



# Random walk centrality in interconnected multilayer networks



Albert Solé-Ribalta, Manlio De Domenico, Sergio Gómez, Alex Arenas\*

Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, 43007 Tarragona, Spain

## HIGHLIGHTS

- We provide analytical expressions for the centrality of random walks in interconnected multilayer networks.
- We check the theoretical results with extensive Monte Carlo simulations of random walkers in different topologies, and achieve an excellent agreement.
- Our results are useful for the ranking of nodes in multi-categorical systems.

## ARTICLE INFO

### Article history:

Received 23 June 2015

Received in revised form

8 January 2016

Accepted 8 January 2016

Available online 2 March 2016

### Keywords:

Multilayer complex networks

Random walks

Centrality

## ABSTRACT

Real-world complex systems exhibit multiple levels of relationships. In many cases they require to be modeled as interconnected multilayer networks, characterizing interactions of several types simultaneously. It is of crucial importance in many fields, from economics to biology and from urban planning to social sciences, to identify the most (or the less) influential nodes in a network using centrality measures. However, defining the centrality of actors in interconnected complex networks is not trivial. In this paper, we rely on the tensorial formalism recently proposed to characterize and investigate this kind of complex topologies, and extend two well known random walk centrality measures, the random walk betweenness and closeness centrality, to interconnected multilayer networks. For each of the measures we provide analytical expressions that completely agree with numerically results.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

It is common practice in many studies involving networks to assume that nodes are connected by a single type of edge that encapsulates all relations between them. In a myriad of applications this assumption oversimplifies the complexity of the system, leading to inaccurate or wrong results. Examples can be found in temporal networks, where neglecting time-dependence washes out the memory of sequences of human contacts in transmission of diseases [1], in co-authorship networks, where neglecting the existence of multiple relationships between actors might alter the topology which may lead to misestimating crucial node's properties [2–7] or in transportation networks where the multilayer topology must be considered to accurately model the dynamics to *a posteriori* predict congested locations [8].

Historically, the term *multiplex* was coined to indicate the presence of more than one relationship between the same actors of a social network [9]. This type of network is well understood in terms of “coloring” (or labeling) the edges corresponding to interactions

of different nature. For instance, in a social network the same individual might have connections to other individuals based on financial interests (e.g., color red) and connections with the same or different individuals based on friendship (e.g., color blue). In other real-world systems, like the transportation network of a city, the same geographical position can be part, for instance, of the network of subway or the network of bus routes, simultaneously. In this specific case, an edge-colored graph would not capture the full structure of the network, since information about the cost to *move* from the subway network to the bus route is missing. This cost can be economic or might account for the time required to physically commute between the two layers. It is in these cases where an interconnected multilayer network provides a better representation of the system. Fig. 2 shows an illustration of an interconnected multilayer (Fig. 2(A)) and the classical representation with an aggregated network (Fig. 2(C)). It is evident that a simple projection of the former – mathematically equivalent to sum up the corresponding adjacency matrices of the individual layers – would provide a network where the information about the relation type is lost. On the other hand, an edge-colored graph (Fig. 2(B)) cannot account for interconnections. For further details about the classification of such multilayer networks we refer to [10] and references therein. In the rest of the paper interconnected multilayer networks will be referred to in short as multilayer networks.

\* Corresponding author. Tel.: +34 977559687.

E-mail address: [alexandre.arenas@urv.cat](mailto:alexandre.arenas@urv.cat) (A. Arenas).

The remainder of this paper is organized as follows. In Section 2 we briefly describe the tensorial notation, defined in [11], adopted overall the paper. In Section 3 we capitalize on this notation to extend some random walk centrality descriptors, well known in the case of single layer networks, to interconnected multilayer networks. Finally, we discuss our findings in Section 4.

## 2. Tensorial notation

Edge-colored graphs can be represented by a set of adjacency matrices [12–15]. However, standard matrices, used to represent networks, are limited in the complexity of the relationships that they can capture, i.e., they do not represent a suitable framework in the case of multilayer networks. This is the case of multiple types of relationships – that can also change in time – between nodes. Such a level of complexity can be characterized by considering tensors and algebras of higher order [11].

A great advantage of tensor formalism developed in [11] relies on its compactness. An adjacency tensor can be written using a more compact notation that is very useful for the generalization of network descriptors to multilayer networks. In this notation, a row vector  $\mathbf{a} \in \mathbb{R}^N$  is given by a covariant vector  $a_\alpha$  ( $\alpha = 1, 2, \dots, N$ ), and the corresponding contravariant vector  $a^\alpha$  (i.e., its dual vector) is a column vector in Euclidean space. A canonical vector is assigned to each node and the corresponding interconnected multi-layer network is represented by a mixed rank-4 adjacency tensor.

However, in the majority of applications, it is not necessary to perform calculations using canonical vectors and tensors explicitly. In this cases, a classical single-layer network can be represented by a rank-2 mixed adjacency tensor  $W_\beta^\alpha$  [11], where the layer information is disregarded. But, in general, systems may exhibit several types of relationships between pairs of nodes and a more general system represented as a multilayer object – in which each type of relationship is represented within a single layer  $\alpha$  ( $\alpha = 1, 2, \dots, L$ ) of the network – is required.<sup>1</sup> In these cases, we use an *intra-layer adjacency tensor* for the 2nd-order tensor  $W_j^i(\alpha)$  that indicates the relationships between nodes within the same layer  $\alpha$  and the 2nd-order *inter-layer adjacency tensor*  $C_j^i(\alpha\beta)$  to encode information about relationships that incorporate multiple layers.

It has been shown that the mathematical object accounting for the whole interconnected multilayer structure is given by a 4th-order (i.e., rank-4) *multilayer adjacency tensor*  $M_{j\beta}^{i\alpha}$ . This tensor might be simply thought as a higher-order matrix with four indices. It is the direct generalization of the adjacency matrix in the case of single layer networks and encodes the intensity of the relationship (which may not be symmetric) from a node  $i$  in layer  $\alpha$  to a node  $j$  in layer  $\beta$  [11].

To reduce the notational complexity in the tensorial equations the Einstein summation convention is adopted. It is applied to repeated indices in operations that involve tensors. For example, we use this convention in the left-hand sides of the following equations:

$$A_i^i \equiv \sum_{i=1}^N A_i^i, \quad A_j^i B_i^j \equiv \sum_{i=1}^N \sum_{j=1}^N A_j^i B_i^j, \\ A_{j\beta}^{i\alpha} B_{i\gamma}^{k\beta} \equiv \sum_{i=1}^N \sum_{\beta=1}^L A_{j\beta}^{i\alpha} B_{i\gamma}^{k\beta},$$

<sup>1</sup> To avoid confusion, in the following we refer to nodes with Latin letters and to layers with Greek letters, allowing us to distinguish indices that correspond to nodes from those that correspond to layers in tensorial equations.

whose right-hand sides include the summation signs explicitly. It is straightforward to use this convention for the product of any number of tensors of any order. In the following, we will use the  $t$ th power of rank-4 tensors, defined by multiple tensor multiplications:

$$(A^t)_{j\beta}^{i\alpha} = (A)_{j_1\beta_1}^{i_1\alpha_1} (A)_{j_2\beta_2}^{i_2\alpha_2} \dots (A)_{j_t\beta_t}^{i_t\alpha_t}. \quad (1)$$

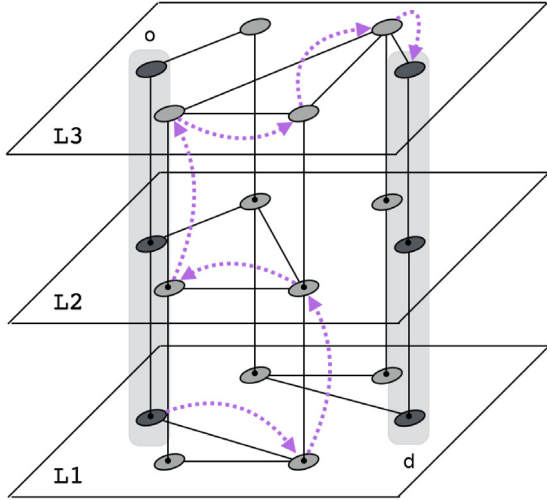
Repeated indices, such that one index is a subscript and the other is a superscript, is equivalent to perform a tensorial operation known as a *contraction*. Moreover, one should be very careful in performing tensorial calculations. For instance, using traditional notation the product  $a^i b^j$  would be a number, i.e., the product of the components of two vectors. However, in our formulation, the same calculation denotes a Kronecker product between two vectors, resulting in a rank-2 tensor, i.e., a matrix.

## 3. Random walk centrality measures in multilayer networks

In practical applications one is often interested in assigning a global measure of importance to each node. If the system we deal with contains several types of relations between actors we expect that the measures, in some way, consider the importance obtained from the different layers. A simple choice could be to combine the centrality of the nodes – obtained from the different layers independently – according to some heuristic choice. This is a viable solution when there is no interconnection between layers, i.e., in the case of edge-colored graphs [16,17]. However, the main drawback of this approach is that it depends on the choice of the heuristics and thus might not evaluate the actual importance of nodes. Our approach accounts for the higher level of complexity of such systems without relying on external assumptions and naturally extends the well-known centrality measures adopted for several decades in the case of single layer networks.

A random walk is one of the simplest dynamical process that can occur on a network, and random walks can be used to approximate other types of diffusion processes [18,19]. Random walks on networks [18,20,19] have attracted considerable interest because they are both important and easy to interpret. They have yielded important insights on a huge variety of applications and can be studied analytically. For example, random walks have been used to rank Web pages [21] and sports teams [22], optimize searches [23], investigate the efficiency of network navigation [24,25], characterize cyclic structures in networks [26], and coarse-grain networks to highlight meso-scale features such as community structure [27–29]. Another interesting application of random walks is to calculate the centrality of actors in complex networks when there is no knowledge about the full network topology but only local information is available. In such cases, centrality descriptors based on shortest-paths, e.g., betweenness and closeness centrality, should be substituted by centrality notions based on random walks [20,30]. In the following we extend these measures to multilayer networks.

First of all, we define a discrete-time random walk, between two individuals  $o$  and  $d$ ,  $o \rightarrow d$ , on a multilayer network consisting of  $L$  layers and  $N$  nodes per layer, as a random sequence of nodes which starts from node  $o$  in any layer and finish in node  $d$  in any layer where each edge's endpoints are the preceding and following vertices in the sequence. The reasoning behind this definition is that the different node replicas in the different layers correspond to the same individual and so anything traveling between them is independent on the starting and ending layer. Fig. 1 shows and example of a random walk between two nodes in a multilayer network where it is evident the introduction of non-trivial effects because of the presence of inter-layer connections that affects its navigation in the networked system [31].



**Fig. 1.** Schematic of a walk (dotted trajectories) between two individuals  $o$  and  $d$  using a multilayer network. A walker can jump between nodes within the same layer, or it might switch to another layer. This illustration evidences how multilayer structure allows a walker to move between nodes that belong to different (disconnected) components on a given layer (L1).

**Random walk occupation centrality.** Let  $T_{j\beta}^{i\alpha}$  denote the tensor of transition probabilities for jumping between pairs of nodes and switching between pairs of layers. Covariant indexes  $i\alpha$  indicate source node and layer and contravariant indexes  $j\beta$  destination node and layer. Similarly to the single layer case, the sum of the probabilities for each outgoing edges of any node adds one. That is,  $u_{i\alpha} = T_{j\beta}^{i\alpha} u_{j\beta}$  where  $u_{i\alpha}$  and  $u_{j\beta}$  are tensors whose components are all ones. In addition, let  $p_{i\alpha}(t)$  be the time-dependent tensor that gives the probability to find a walker at a particular node in a particular layer. Hence, the covariant master equation that governs the discrete-time evolution of the probability from time  $t$  to time  $t + 1$  is  $p_{j\beta}(t + 1) = T_{j\beta}^{i\alpha} p_{i\alpha}(t)$ . The steady-state solution of this equation, i.e., for  $t \rightarrow \infty$ , is given by  $\Pi_{i\alpha}$  which quantifies the probability to find a walker in the node  $i$  of layer  $\alpha$ . In the case of single layer networks, the steady-state solution can be obtained by calculating the leading eigenvector corresponding to the unitary eigenvalue. Similarly, in the case of multilayer networks, the solution can be obtained by calculating the leading eigentensor, solution of the higher-order eigenvalue problem

$$T_{j\beta}^{i\alpha} \Pi_{i\alpha} = \lambda \Pi_{j\beta}. \quad (2)$$

We refer to [Appendix A](#) for the mathematical details to solve this problem. The probability  $\Pi_{j\beta}$ , defined as *random walk occupation centrality*, accounts for the full interconnected structure of the multilayer network. Although different exploration strategies can be adopted to define the transition tensor  $T_{j\beta}^{i\alpha}$  walk in a multilayer network [31], here we only focus on the natural extension of well-known random walks in single layer networks [20]. In this process, the walker in node  $i$  and layer  $\alpha$  might jump to one of its neighbors  $j \neq i$  – within the same layer – or might switch to its counterpart  $i$  in a different interconnected layer  $\beta \neq \alpha$  with equal probability. That is, the inter-layer connection is treated as an edge that can be chosen randomly among all outgoing edges of the node. In the more general case of weighted networks, the jumping probability is usually defined proportional to the weight of the edges. Let us indicate with  $s_{i\alpha}$  the strength of node  $i$  in layer  $\alpha$ , including the inter-layer connections. The multi-strength vector, whose components indicate the strength of each node accounting for the full multilayer structure, is given by summing up its strengths across all layers, i.e., by  $S_i = s_{i\alpha} u^\alpha$ . We indicate with  $D_{j\beta}^{i\alpha}$  the strength tensor whose entries are all zeros, except for  $i = j$  and  $\alpha = \beta$  where the entries are given by  $s_{i\alpha}$ . This tensor represents the multilayer

extension of the well-known diagonal strength matrix in the case of single layer networks. Therefore, the transition tensor is given by  $T_{j\beta}^{i\alpha} = M_{j\beta}^{k\gamma} \tilde{D}_{k\gamma}^{i\alpha}$ , where  $\tilde{D}_{k\gamma}^{i\alpha}$  is the tensor whose entries are the inverse<sup>2</sup> of the non-zero entries of the strength tensor and  $M_{j\beta}^{k\gamma}$  is the weighted adjacency tensor. For this classical random walk, it can be easily shown that  $\Pi_{i\alpha} \propto s_{i\alpha}$  [31].

This centrality, as others in the rest of the paper, assigns a measure of importance to each node in each layer, accounting for the full-interconnected structure of the multilayer network. However, in practical applications one is often interested in assigning a global measure of importance to each node, aggregating the information obtained from the different layers. The choice of the aggregation method is in general not trivial and it strongly influences the final estimation and might lead to wrong results.

However, this is not case for the occupation probability. Since the centrality  $\Pi_{i\alpha}$  is calculated accounting for the full interconnected structure of the whole system we do not require any arbitrary combination of the information from different layers. In our framework, the most intuitive type of aggregation, i.e., summing up over layers, represents the unique and correct choice. Let  $\pi_i = \Pi_{i\alpha} u^\alpha$  be the random walk centrality measure obtained by aggregating over the layers. Here,  $\pi_i$  indicates the probability of finding the walker in node  $i$ , regardless of the layer. It is worth noting that this probability, as well as in single layer networks is, is proportional to  $s_{i\alpha} u^\alpha$ , i.e., the multi-strength of node  $i$ . Therefore, in this specific case, the computation of the centrality by means of the aggregated network would provide the same result of the calculation accounting for the multilayer structure, if inter-layer edges are mapped to self-loops. Unfortunately, this is not the case for the other centrality measures discussed in the rest of this study, where calculating the diagnostics from the aggregate might lead to wrong conclusions.

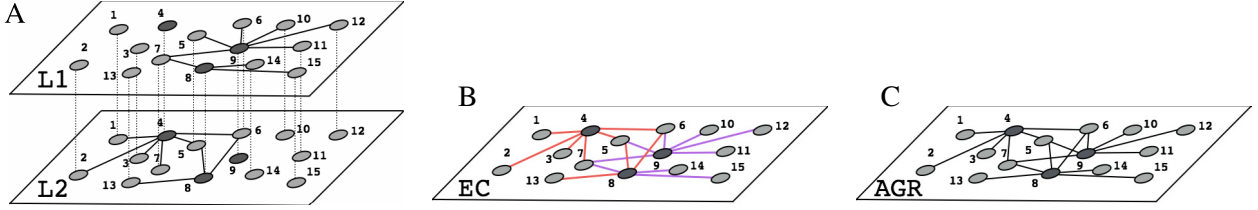
A measure related to random walk occupation centrality is the Page Rank [32] that has been recently extended to interconnected networks [8]. In fact, the Page Rank centrality can be seen as the steady-state solution of the random walk master equation governed by the transition tensor  $R_{j\beta}^{i\alpha}$ , where the walker jumps to a neighbor with rate  $r$  and teleport to any other node in the network with rate  $1 - r$ . This rank-4 tensor is given by

$$R_{j\beta}^{i\alpha} = r T_{j\beta}^{i\alpha} + \frac{(1 - r)}{NL} u_{j\beta}^{i\alpha}, \quad (3)$$

where  $u_{j\beta}^{i\alpha}$  is the rank-4 tensor with all components equal to 1. The steady-state solution of the master equation corresponding to this transition tensor provides the Page Rank centrality for multilayer networks.

**Random walk betweenness centrality.** The betweenness is a measure of network centrality that instead of accounting for topological centrality accounts for the importance of nodes in terms of dynamical processes that run over the network. In particular, the betweenness measures to which extent a node lies in the path between any two other nodes [30]. One can think of packets traveling in internet, in this case the betweenness measures the influence of nodes in the controlling of information. The most common betweenness is the shortest path betweenness [33] where the centrality of a node  $j$  is relative to the number of shortest paths, for any pair  $(o, d)$  of origin and destination nodes, that pass through  $j$ . However, in real networks, entities (rumors, messages or packets over the Internet) that travel the network do not always take the shortest path [34,35]. Consider, for instance, rumors that can

<sup>2</sup> It is worth remarking that, in general, this is different from the inverse of a tensor  $A_{j\beta}^{i\alpha}$ , that is defined as the tensor  $B_{j\beta}^{i\alpha}$  such that  $A_{j\beta}^{i\alpha} B_{j\beta}^{i\alpha} = \delta_{j\beta}^{i\alpha}$ , where  $\delta_{j\beta}^{i\alpha} = \delta_j^i \delta_\beta^\alpha$ .



**Fig. 2.** (A) An interconnected multilayer network representing the same actors exhibiting different relationships on different levels. The cost to move from one layer to the other is represented by dotted vertical lines. (B) Edge-colored graph representing the same actors with the same relations in (A) with two different types of interactions (solid and dashed edges). In this case the representation does not allow modeling the cost to move between layers. (C) Classical approach of representing the different types of relations using an aggregated network. The network represents the same actor and relation in (A) and (B) but disregarding the type of relation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

be wandering around the network or packets trying to avoid overloaded routers. In such cases, the shortest path betweenness is not always a good proxy for the centrality of nodes. For these scenarios the random walk betweenness of a node  $j$  is defined as the amount of random walks between any pair  $(o, d)$  of nodes that pass through  $j$  [30].

To analytically compute the number of random walks visiting a particular node, it is often convenient to use the concept of absorbing random walk, where the absorbing state is selected to be the destination node  $d$  [30,19]. To extend this concept to the case of interconnected multilayer networks, we consider random walks that begin, pass and end in nodes in different layers while accounting for the existence of several replicas of the same node.

Specifically, to extend the concept of random walks to interconnected networks, we define the absorbing transition tensor on a particular node  $d$  by

$$(T_{[d]})_{j\beta}^{\sigma\sigma} = \begin{cases} 0 & j = d \\ T_{j\beta}^{\sigma\sigma} & \text{else.} \end{cases} \quad (4)$$

Random walkers governed by this transition tensor will vanish once they arrive to any absorbing state [19]. Note that  $T_{[d]}$  has one absorbing state for each replica of node  $d$  in different layers. It can be shown (see Appendix B) that the average number of times a random walk (with origin in node  $o$  in layer  $\sigma$  and destination  $d$  independently of the layer) will pass by a node  $j$  in layer  $\beta$ , regardless of the time step, is given by

$$(\tau_{[d]})_{j\beta}^{\sigma\sigma} = [(\delta - T_{[d]})^{-1}]_{j\beta}^{\sigma\sigma}, \quad (5)$$

where  $\delta_{j\beta}^{\sigma\sigma} = \delta_j^{\sigma\sigma}$  and  $\delta$  is the Kronecker delta. Note that the average number of times that the walk will visit node  $j$  still depends on the layer where  $j$  is located and on the originating layer  $\sigma$ . Since we are interested on node properties, regardless of the layer, we average over all possible starting layers  $\sigma$  and aggregate the walks that pass through  $j$  in the different layers,

$$(\tau_{[d]})_j^o = \frac{1}{L} (\tau_{[d]})_{j\beta}^{\sigma\sigma} u^\beta u_\sigma. \quad (6)$$

The overall betweenness centrality is obtained by averaging over all possible origins and destinations:

$$\tau_j = \frac{1}{N(N-1)} \sum_{d=1}^N (\tau_{[d]})_j^o u_o. \quad (7)$$

The comparison between the values of  $\tau_i$  obtained from simulations and theoretical predictions are shown in Fig. 3(A)–(C). As expected, the results are in excellent agreement. It is worth remarking that the equivalence holds regardless of the number of nodes in the network, the topology and the number of layers. Fig. 3(D)–(F) shows the relative difference of the betweenness computed using the aggregated network and the multilayer network. The results show that in general the betweenness computed

on the aggregated network tends to underestimate the betweenness of the multilayer network.

**Random walk closeness centrality.** The distance between two nodes in a network is given by shortest-path which separates them. The farness of an individual is given by the sum of all geodesics from that node to any other node. In general, the inverse of this farness provides a measure of the *closeness* of the node. Such a diagnostic is related to how fast information is expected to spread from a given actor to the others in the network. A variant of the closeness when random walks are considered is given by the random walk closeness centrality. In the case of single layer networks, it has been introduced to quantify how central a node is located regarding its potential to receive information randomly diffusing over the network [20].

We define the random walk closeness centrality of a node  $i$  in a multilayer network as the inverse of the average number of steps that a random walker, starting from any other node in the multilayer network, requires to reach  $i$  for the first time. The computation of the closeness centrality is generally based on the mean first-passage time (MFPT), that is defined as the average number of steps to reach a node  $d$ , starting from a given node  $o$ . The MFPT matrix can be computed analytically by means of Kemeny–Snell fundamental matrix  $Z$  [36,37] or by means of absorbing random walks [38,19]. In this study, we adopt the second approach as for the calculation of random walk betweenness centrality. The following calculation involve the use of the transition tensor  $T$  governing random walks over multilayer networks and the corresponding absorbing transition tensor  $T_{[d]}$ . Hence, the tensor

$$p_{j\beta}^{\sigma\sigma}(t) = (T_{[d]}^t)_{j\beta}^{\sigma\sigma} \quad (8)$$

indicates the probability of visiting node  $j$  in layer  $\beta$ , after  $t$  time steps, considering that the walk originated in node  $o$  in layer  $\sigma$ . This transition tensor is absorbing on node  $d$  regardless of the layer and, consequently, any walker reaching an absorbing state will vanish, i.e.,  $p_{d\beta}^{\sigma\sigma}(t) = 0$  for any  $\beta$  and  $t$ . The probability that the walker is absorbed in some node  $d$  at a time  $h$  equal or smaller than  $t$ , regardless of the layer, is given by

$$(q_{[d]})^{\sigma\sigma}(t) = u^{\sigma\sigma} - (T_{[d]}^t)_{j\beta}^{\sigma\sigma} u^{j\beta}. \quad (9)$$

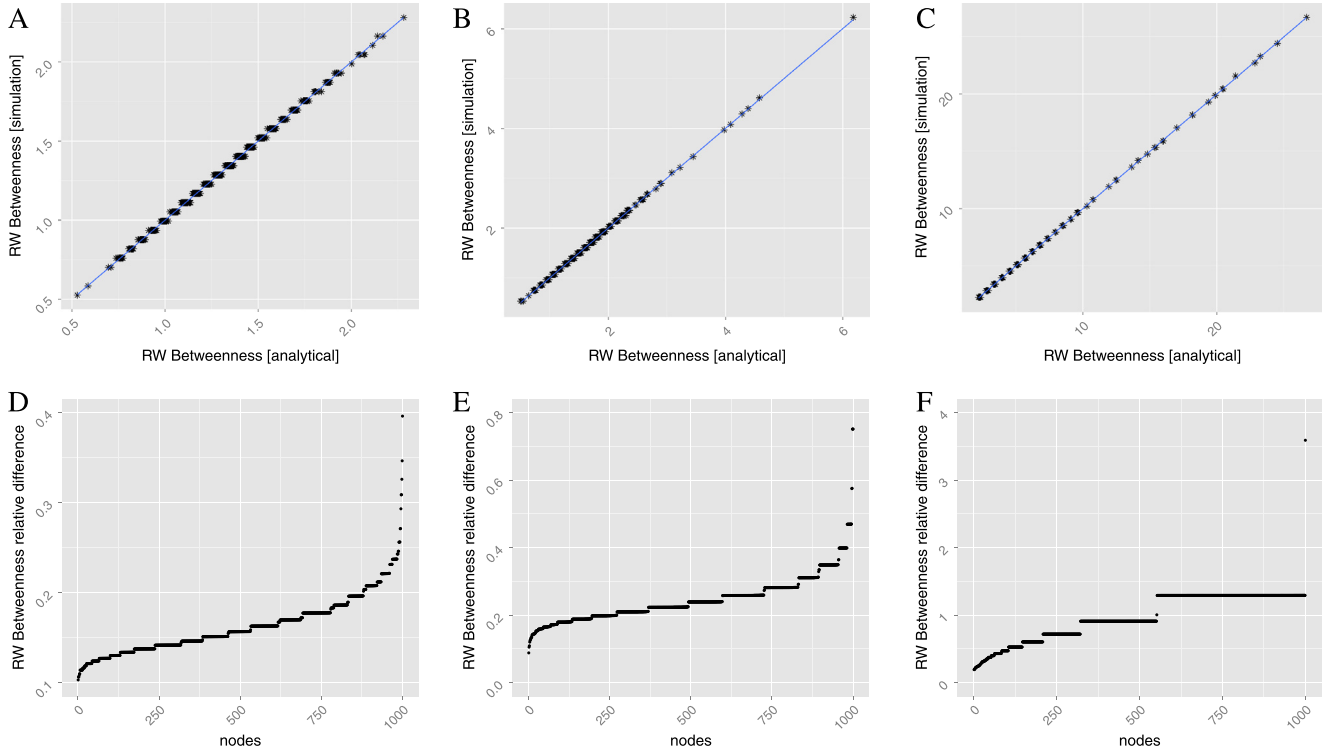
Note that we have a rank-2 tensor  $q$  for each choice of  $d$  and we put in evidence this dependence by means of  $[d]$ . From each tensor  $q$  we can calculate the probability that the first passage time for node  $d$  is exactly  $t$  by

$$\begin{aligned} (q_{[d]})^{\sigma\sigma}(h=t) &= (q_{[d]})^{\sigma\sigma}(t) - (q_{[d]})^{\sigma\sigma}(t-1) \\ &= [(T_{[d]}^{t-1}) - (T_{[d]}^t)]_{j\beta}^{\sigma\sigma} u^{j\beta}. \end{aligned} \quad (10)$$

Considering the walk starts from node  $o$  in layer  $\sigma$ , each tensor encoding the mean first passage time to node  $d$  is obtained from Eq. (10) as

$$(H_{[d]})^{\sigma\sigma} = \sum_{t=0}^{\infty} t (q_{[d]})^{\sigma\sigma}(h=t) = [(\delta - T_{[d]})^{-1}]_{j\beta}^{\sigma\sigma} u^{j\beta}. \quad (11)$$





**Fig. 3.** (A, B, C) Comparison of the random walk betweenness centrality obtained by simulation and by our analytical approach for different multilayer network topologies. (D, E, F) Relative difference of the betweenness,  $(\tau_j - \tau_j^{\text{agr}}) / \tau_j^{\text{agr}}$ , computed using the aggregated network as shown in Fig. 2 and the multilayer network. Nodes in the horizontal axis are sorted according to the betweenness relative difference. In all experiments, each multilayer network is composed of two layers with 1000 nodes per layer. (A, D) Results on a multilayer network with two Erdős–Rényi networks as layers. (B, E) Results on a multilayer network with one Erdős–Rényi network and one Barabási–Albert network as layers. (C, F) Results on a multilayer network with two Barabási–Albert networks as layers.

Note the presence of the individual betweenness contributions as in Eq. (5).

The geometric series in Eq. (11) converges since the maximum eigenvalue of  $T_{[d]}$  is strictly smaller than one, and the sum can be calculated exploiting the self-similarity of the series. The mean first passage time to  $d$  still depends on the origin of the walk, i.e., node  $o$  in layer  $\sigma$ .

The average mean first passage time  $h_{[d]}$  to node  $d$  is obtained by averaging  $(H_{[d]})^{o\sigma}$  over all possible starting nodes and layers as

$$h_{[d]} = \frac{1}{(N-1)L} u_{o\sigma} (H_{[d]})^{o\sigma} + \frac{1}{N} \pi_{[d]}^{-1}, \quad (12)$$

where  $\pi_{[d]}$  is the occupation probability of node  $d$  and the term  $\frac{1}{N} \pi_{[d]}^{-1}$  is included explicitly to account for the average return time, that is not accounted for when using absorbing random walks.

Finally, the random walk closeness centrality of node  $d$  is defined as the inverse of  $h_{[d]}$ . We introduce the vector  $\xi_i$  whose components are given by the inverse of the corresponding values of  $h$ .

The comparison between the values of  $\xi_i$  obtained from simulations and theoretical predictions are shown in Fig. 4(A)–(C). As in the betweenness centrality, the results are in excellent agreement and it is worth remarking that the equivalence holds regardless of the number of nodes in the network, the topology and the number of layers. Fig. 3 panels D–F show the relative difference of the closeness computed using the aggregated network and the multilayer network. The results show that the closeness centrality computed on the aggregated network tends to overestimate the closeness of the multilayer network.

#### 4. Conclusions and discussion

We have extended the main random walk centrality measures to interconnected multilayer networks and gave interpretation

of their meaning. In addition, we have presented analytical approaches, based on the tensorial formalism defined in [11], for their computation. The comparison of the predictions given by our analytical approach with the results obtained by simulations show a perfect agreement, concluding that the presented analytical expressions are ready to be applied to the analysis of real complex networks. We expect that the presented results are useful in many interdisciplinary applications ranging from social sciences to transportation networks.

#### Acknowledgments

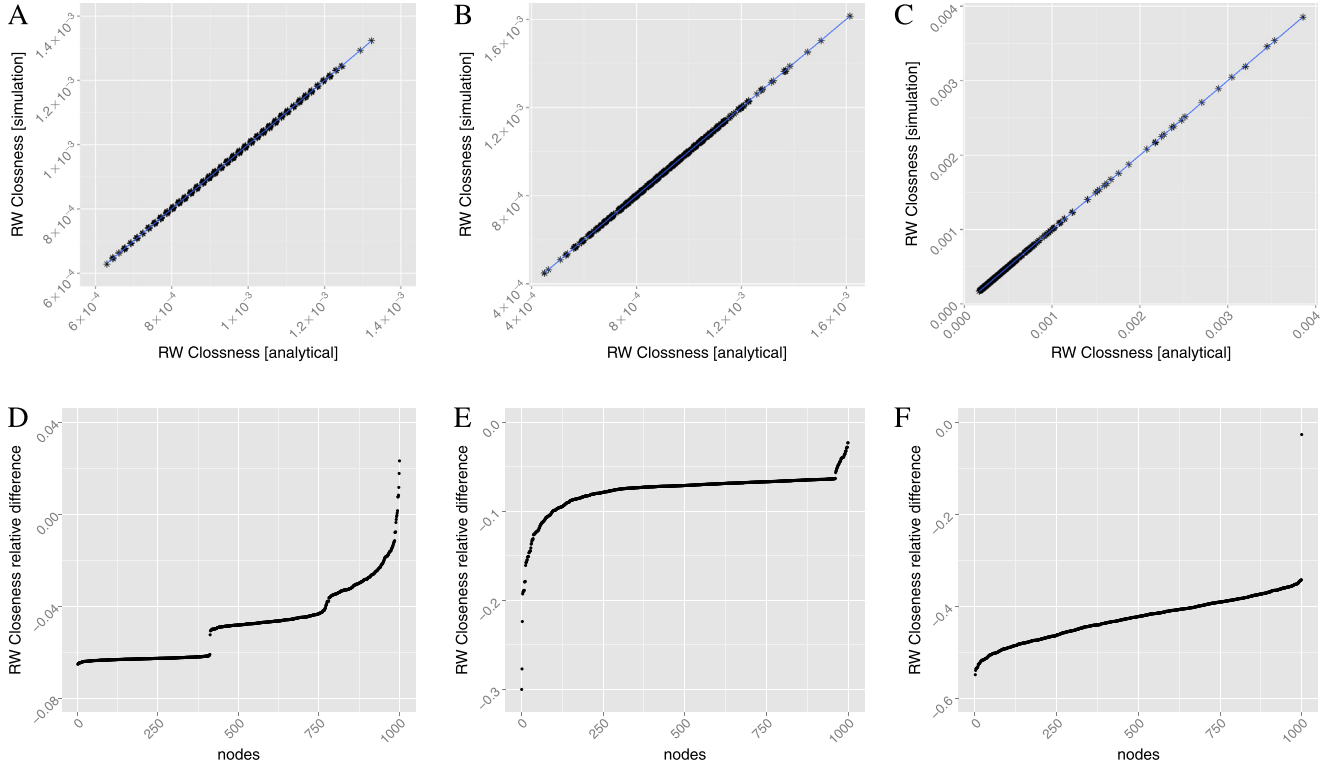
AA, MDD, SG, and AS were supported by the European Commission FET-Proactive project PLEXMATH (Grant No. 317614), the MULTIPLEX (grant 317532) and the Generalitat de Catalunya 2009-SGR-838. AA also acknowledges financial support from the ICREA Academia and the James S. McDonnell Foundation Grant #220020325, and SG and AA were supported by FIS2015-71582-C2-1.

#### Appendix A. Eigenvalue problem with tensors

The eigenvalue problem for a rank-2 tensor, i.e., a standard matrix, is defined by  $W_j^i v_i = \lambda v_j$ . The extension of this problem to rank-4 tensors leads to the equation

$$M_{j\beta}^{i\alpha} V_{i\alpha} = \lambda V_{j\beta}. \quad (\text{A.1})$$

To solve this problem, it is worth noting that any tensor can be *unfolded* to lower rank tensors [39]. For instance, a rank-2 tensor like  $W_j^i$ , with  $N^2$  components, can be flattened to a vector  $w_k$  with  $N^2$  components. In the case of the rank-4 multilayer adjacency tensor  $M_{j\beta}^{i\alpha}$ , although any unfolding is allowed, it is particularly



**Fig. 4.** (A, B, C) Comparison of the random walk closeness centrality obtained by simulation and by our analytical approach for different multilayer network topologies. (D, E, F) Relative difference of the closeness,  $(\xi_i - \xi_i^{agr}) / \xi_i^{agr}$ , computed using the aggregated network as shown in Fig. 2 and the multilayer network. Nodes in the horizontal axis are sorted according to the closeness relative difference. In all experiments, each multilayer network is composed of two layers with 1000 nodes per layer. (A, D), results on a multilayer network with two Erdős–Rényi networks as layers. (B, E), results on a multilayer network with one Erdős–Rényi network and one Barabási–Albert network as layers. (C, F), results on a multilayer network with two Barabási–Albert networks as layers.

useful for some applications to choose the ones flattening to a squared rank-2 tensor  $\tilde{M}_l^k$  with  $NL \times NL$  components, where  $L$  indicates the number of layers [40]. In fact, this unfolding produces as many block adjacency matrices, named *supra-adjacency matrices* in some applications [40,31,41], as the number of permutations of diagonal blocks of size  $N^2$ , i.e.,  $L!$ . However, such unfoldings do not alter the spectral properties of the resulting supra-matrix and can be used to solve the eigenvalue problem for rank-4 tensors. In fact, the solution of the eigenvalue problem

$$\tilde{M}_l^k \tilde{v}_k = \tilde{\lambda}_l \tilde{v}_l, \quad (\text{A.2})$$

is a *supra-vector* with  $NL$  components which corresponds to the unfolding of the eigentensor  $V_{i\alpha}$ .

## Appendix B. Mean number of crossing times

Given  $M$  random walks starting in node  $o$  on layer  $\sigma$  and ending when reaching node  $d$ , regardless of the layer, the expected number of times a random walk will pass by node  $j$  on layer  $\beta$  is given by

$$(\tau_{[d]})_{j\beta}^{o\sigma} = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M \sum_{t=0}^{\infty} z_{j\beta}^{o\sigma}(t, m), \quad (\text{B.1})$$

where  $z_{j\beta}^{o\sigma}(t, m) = 1$  if walk  $m$  was visiting node  $j$  in layer  $\beta$  at time step  $t$  and  $z_{j\beta}^{o\sigma}(t, m) = 0$  otherwise.

Following the frequentist interpretation, the probability of being in node  $j$  in layer  $\beta$  at time step  $t$ , provided that the walk originated in node  $o$  in layer  $\sigma$ , is given by

$$p_{j\beta}^{o\sigma}(t) = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M z_{j\beta}^{o\sigma}(t, m). \quad (\text{B.2})$$

Substituting (B.2) in (B.1) we obtain that

$$\begin{aligned} (\tau_{[d]})_{j\beta}^{o\sigma} &= \sum_{t=0}^{\infty} p_{j\beta}^{o\sigma}(t) = \sum_{t=0}^{\infty} (T_{[d]}^t)_{j\beta}^{o\sigma} \\ &= \left[ (\delta - T_{[d]})^{-1} \right]_{j\beta}^{o\sigma} \end{aligned} \quad (\text{B.3})$$

where  $T_{[d]}$  corresponds to the absorbing transition tensor defined in Eq. (4).

## References

- [1] P. Holme, J. Saramäki, Temporal networks, *Phys. Rep.* 519 (2012) 97–125.
- [2] L.C. Freeman, Centrality in social networks conceptual clarification, *Soc. Networks* 1 (3) (1979) 215–239.
- [3] H. Jeong, S.P. Mason, A.-L. Barabási, Z.N. Oltvai, Lethality and centrality in protein networks, *Nature* 411 (6833) (2001) 41–42.
- [4] R. Guimerà, A. Diaz-Guilera, F. Vega-Redondo, A. Cabrales, A. Arenas, Optimal network topologies for local search with congestion, *Phys. Rev. Lett.* 89 (24) (2002) 248701.
- [5] R. Guimerà, S. Mossa, A. Turtleschi, L.N. Amaral, The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles, *PNAS* 102 (22) (2005) 7794–7799.
- [6] M. Barthélemy, Betweenness centrality in large complex networks, *Eur. Phys. J. B* 38 (2) (2004) 163–168.
- [7] V. Nicosia, R. Criado, M. Romance, G. Russo, V. Latora, Controlling centrality in complex networks, *Sci. Rep.* 2 (2012).
- [8] M. De Domenico, A. Solé-Ribalta, E. Omodei, S. Gómez, A. Arenas, Ranking in interconnected multilayer networks reveals versatile nodes, *Nature Commun.* 6 (2015).
- [9] J.F. Padgett, C.K. Ansell, Robust action and the rise of the medici, 1400–1434, *Am. J. Sociol.* (1993) 1259–1319.
- [10] M. Kivelä, A. Arenas, M. Barthélemy, J.P. Gleeson, Y. Moreno, M.A. Porter, Multilayer networks, *J. Complex Netw.* 2 (3) (2014) 203–271.
- [11] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M.A. Porter, S. Gómez, A. Arenas, Mathematical formulation of multi-layer networks, *Phys. Rev. X* 3 (4) (2013) 041022.

- [12] A. Cardillo, J. Gómez-Gardeñes, M. Zanin, M. Romance, D. Papo, F. del Pozo, S. Boccaletti, Emergence of network features from multiplexity, *Sci. Rep.* 3 (2013).
- [13] V. Nicosia, G. Bianconi, V. Latora, M. Barthélemy, Growing multiplex networks, *Phys. Rev. Lett.* 111 (2013) 058701.
- [14] G. Bianconi, Statistical mechanics of multiplex networks: Entropy and overlap, *Phys. Rev. E* 87 (2013) 062806.
- [15] F. Battiston, V. Nicosia, V. Latora, Structural measures for multiplex networks, *Phys. Rev. E* 89 (3) (2014) 032804.
- [16] L. Sola, M. Romance, R. Criado, J. Flores, A.G. del Amo, S. Boccaletti, Eigenvector centrality of nodes in multiplex networks, *Chaos* 3 (2013) 033131.
- [17] A. Halu, R.J. Mondragon, P. Pansaraza, G. Bianconi, Multiplex pagerank, *PLoS ONE* 8 (10) (2013) e78293.
- [18] F.R.K. Chung, *Spectral Graph Theory*, second ed., American Mathematical Society, Providence, RI, 1997.
- [19] M. Newman, *Networks: An Introduction*, Oxford University Press, Inc., New York, NY, USA, 2010.
- [20] J.D. Noh, H. Rieger, Random walks on complex networks, *Phys. Rev. Lett.* 92 (11) (2004) 118701.
- [21] S. Brin, L. Page, The anatomy of a large-scale hypertextual web search engine, in: *Seventh International World-Wide Web Conference, WWW 1998, 1998*.
- [22] T. Callaghan, P.J. Mucha, M.A. Porter, Random walker ranking for NCAA Division I-A football, *Amer. Math. Monthly* 114 (9) (2007) 761–777.
- [23] G. Viswanathan, S.V. Buldyrev, S. Havlin, M. Da Luz, E. Raposo, H.E. Stanley, Optimizing the success of random searches, *Nature* 401 (6756) (1999) 911–914.
- [24] S.-J. Yang, Exploring complex networks by walking on them, *Phys. Rev. E* 71 (1) (2005) 016107.
- [25] L. da Fontoura Costa, G. Travieso, Exploring complex networks through random walks, *Phys. Rev. E* 75 (1) (2007) 016102.
- [26] H.D. Rozenfeld, J.E. Kirk, E.M. Bollt, D. Ben-Avraham, Statistics of cycles: how loopy is your network? *J. Phys. A* 38 (21) (2005) 4589.
- [27] D. Gfeller, P. De Los Rios, Spectral coarse graining of complex networks, *Phys. Rev. Lett.* 99 (3) (2007) 038701.
- [28] M. Rosvall, C.T. Bergstrom, An information-theoretic framework for resolving community structure in complex networks, *PNAS* 104 (18) (2007) 7327–7331.
- [29] R. Lambiotte, J.-C. Delvenne, M. Barahona, Random walks, Markov processes and the multiscale modular organization of complex networks, *IEEE Trans. Netw. Sci. Eng.* 1 (2) (2014) 76–90.
- [30] M.E. Newman, A measure of betweenness centrality based on random walks, *Soc. Networks* 27 (1) (2005) 39–54.
- [31] M. De Domenico, A. Solé-Ribalta, S. Gómez, A. Arenas, Navigability of interconnected networks under random failures, *Proc. Natl. Acad. Sci.* 111 (23) (2014) 8351–8356.
- [32] S. Brin, L. Page, The anatomy of a large-scale hypertextual web search engine, in: *Seventh International World-Wide Web Conference, WWW 1998, 1998*. URL: <http://ilpubs.stanford.edu:8090/361/>.
- [33] L.C. Freeman, A set of measures of centrality based on betweenness, *Sociometry* (1977) 35–41.
- [34] L.C. Freeman, S.P. Borgatti, D.R. White, Centrality in valued graphs: A measure of betweenness based on network flow.
- [35] K. Stephenson, M. Zelen, Rethinking centrality: Methods and examples, *Social Networks* 11 (1) (1989) 1–37.
- [36] L. Lovász, Random walks on graphs: A survey, in: *Combinatorics Paul Erdos is Eighty 2* (1), 1993, pp. 1–46.
- [37] Z. Zhang, A. Julaiti, B. Hou, H. Zhang, G. Chen, Mean first-passage time for random walks on undirected networks, *Eur. Phys. J. B* 84 (4) (2011) 691–697.
- [38] J.G. Kemeny, J.L. Snell, *Finite Markov Chains*, Van Nostrand Reinhold, 1960.
- [39] T.G. Kolda, B.W. Bader, Tensor decompositions and applications, *SIAM Rev.* 51 (3) (2009) 455–500. <http://dx.doi.org/10.1137/07070111x>.
- [40] S. Gómez, A. Díaz-Guilera, J. Gómez-Gardeñes, C. Pérez-Vicente, Y. Moreno, A. Arenas, Diffusion dynamics on multiplex networks, *Phys. Rev. Lett.* 110 (2) (2013) 028701.
- [41] E. Cozzo, M. Kivelä, M.D. Domenico, A. Solé-Ribalta, A. Arenas, S. Gómez, M.A. Porter, Y. Moreno, Structure of triadic relations in multiplex networks, *New J. Phys.* 17 (7) (2015) 073029. URL <http://stacks.iop.org/1367-2630/17/i=7/a=073029>.