Congestion Induced by the Structure of Multiplex Networks

Albert Solé-Ribalta, Sergio Gómez, and Alex Arenas
Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, 43007 Tarragona, Spain
(Received 16 December 2015; published 10 March 2016)

Multiplex networks are representations of multilayer interconnected complex networks where the nodes are the same at every layer. They turn out to be good abstractions of the intricate connectivity of multimodal transportation networks, among other types of complex systems. One of the most important critical phenomena arising in such networks is the emergence of congestion in transportation flows. Here, we prove analytically that the structure of multiplex networks can induce congestion for flows that otherwise would be decongested if the individual layers were not interconnected. We provide explicit equations for the onset of congestion and approximations that allow us to compute this onset from individual descriptors of the individual layers. The observed cooperative phenomenon is reminiscent of Braess' paradox in which adding extra capacity to a network when the moving entities selfishly choose their route can in some cases reduce overall performance. Similarly, in the multiplex structure, the efficiency in transportation can unbalance the transportation loads resulting in unexpected congestion.

DOI: 10.1103/PhysRevLett.116.108701

Introduction.—Complex networks have become a natural abstraction of the interactions between elements in complex systems [1]. When the type of interaction is essentially identical between any two elements, the theory of complex networks provides us with a wide set of tools and diagnostics that turn out to be very useful to gain insight in the system under study. However, there are particular cases where this classical approach may lead to misleading results, e.g., when the entities under study are related to each other using different types of relations in what is being called multilayer interconnected networks [2–4]. Representative examples are multimodal transportation networks [5,6] where two geographic places may be connected by different transport modes, or social networks [7–10] where users are connected using several platforms or different categorical layers.

Here, we focus our study on the transportation congestion problem in multiplex networks, where each node is univocally represented in each layer so the interconnectivity pattern among layers becomes a one-to-one connection (i.e., each node in one layer is connected to the same node in the rest of the layers, thus allowing travelling elements to switch layers at all nodes). This representation is an excellent proxy of the structure of multimodal transportation systems in geographic areas [6]. The particular topology of each layer is conveniently represented as a spatial network where nodes correspond to a certain coarse grain of the common geography at all layers [11–14].

Transportation dynamics on networks can be, in general, interpreted as the flow of elements from an origin node to a destination node. When the network is facing a number of simultaneous transportation processes, we find that many elements travel through the same node or link. This, in combination with the possible physical constraints of the nodes and links, can lead to network congestion, in which the number of elements in transit on the network grows proportionally with time [15,16]. Usually, to analyze the phenomenon, a discrete abstraction of the transportation dynamics in networks is used [15–21].

Multimodal transportation can also be mathematically abstracted as transportation dynamics on top of a multiplex structure. Note that routings on the multilayer transportation system are substantially different with respect to routings on single layer transportation networks. In the multilayer case, each location of the system (e.g., geographical location) has different replicas that represent each entry point to the system using the different transportation media. Thus, each element with the intention of traveling between locations i and j has the option to choose between the most appropriate media to start and end its traversal. We assume that elements traverse the network using the shortest paths, so each element chooses the starting and ending media that minimize the distance between the starting or ending locations. As we will show in this work, this "selfish" behavior provokes an unbalance in the load of the transportation layers inducing congestion, similarly to what is presented in the classical counterintuitive result of Braess' paradox [22].

Note that in a multiplex network we can have two types of shortest paths: paths that only use a single layer (intralayer paths) and paths that use more than one layer (interlayer paths). Hereafter, we develop the analysis of transportation in multiplex networks, consisting of N locations (nodes per layer) and L layers, and quantify when this structure will induce congestion. To this aim, we describe, with a set of discrete time balance equations (one for each node at each layer), the increment of elements $\Delta q_{i\alpha}$ in the queue of each node i on layer α:

$$\Delta q_{i\alpha} = g_{i\alpha} + \sigma_{i\alpha} - d_{i\alpha}, \quad (1)$$

where $g_{i\alpha}$ is the average number of elements injected at node i in layer α (also called the injection rate, which can be assimilated to an external particle reservoir), $\sigma_{i\alpha}$ is the average number of elements that arrive at node i in layer α from the adjacent links of that node (ingoing rate), and $d_{i\alpha} \in [0, \tau_{i\alpha}]$ corresponds to the average number of elements that finish their traversal in node i in layer α or are forwarded to other neighboring nodes. The control parameter is $g_{i\alpha}$: small values of it correspond to a low density of elements in the network and high values to a high density of elements. A graphical explanation of the variables of the model is shown in Fig. 1.

Before reaching congestion, the number of elements in the queue of each node is constant on average, $\Delta q_{i\alpha} = 0 \quad \forall\, i\alpha$, and consequently $d_{i\alpha} = g_{i\alpha} + \sigma_{i\alpha} < \tau_{i\alpha}$, where $\tau_{i\alpha}$ is the maximum processing rate of the node. A node i on layer α becomes congested when it is requested to process more elements than its maximum processing rate, $d_{i\alpha} > \tau_{i\alpha}$, and therefore its onset of congestion is achieved when $d_{i\alpha} = \tau_{i\alpha}$. We are interested in computing the maximum injection rate $g_{i\alpha}$ for which the network is congestion free. In the noncongested phase, as well as at the onset of congestion, the number of elements ingoing to each node $\sigma_{i\alpha}$ can be obtained in terms of the node's effective betweenness, see Ref. [15]. Our scenario is slightly different since we need to account for the effective betweenness of the multiplex. In addition to the intralayer and interlayer paths, our definition of the dynamics also accounts for the number of shortest paths that start ($s_{i\alpha}$) and end ($e_{i\alpha}$) at node i on layer α (this can be computed using any classical shortest path algorithm [23]). Note that $\sum_\alpha s_{i\alpha} = \sum_\alpha e_{i\alpha} = N - 1$. These factors are essential to understand the unbalance of the loads between layers in the multiplex network, and only depend on the distribution of shortest paths in the full structure.

In the following, we assume a constant injection rate $\sum_\alpha g_{i\alpha} = \rho L$ with ρ being the common injection rate at all locations i. In addition we also suppose, without loss of generality, that the maximum processing rate is the same for all nodes of the multiplex network, $\tau_{i\alpha} = \tau$. These hypotheses simplify the analysis but are not crucial to develop it.

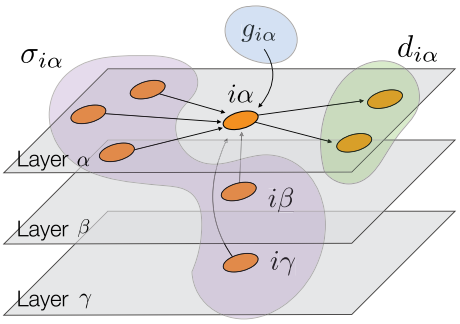

FIG. 1. Contributions to the size of the queue of each node at each layer of a multiplex network in our standardized transportation model. Arrows mark the flow direction of elements in and out of the node.

To obtain the critical injection rate of the multiplex, we require expressions for $g_{i\alpha}$ and $\sigma_{i\alpha}$. The injection rate of node i on layer α can be obtained as the product of the number of elements that enter the network using location i, ρL, and the fraction of multiplex shortest paths that start on node i on layer α, $s_{i\alpha}/(N-1)$:

$$g_{i\alpha} = \rho L \frac{s_{i\alpha}}{N-1}. \quad (2)$$

The ingoing rate of each node $\sigma_{i\alpha}$ depends on the fraction of shortest paths that pass through or end in it [15]. Thus, $\sigma_{i\alpha}$ can be obtained as the number of generated elements over all the network at each time step, ρLN, times the fraction of them that arrive ($e_{i\alpha}$) or traverse it ($B_{i\alpha}$ is the topological betweenness):

$$\sigma_{i\alpha} = \rho L \frac{B_{i\alpha} + e_{i\alpha}}{N-1}. \quad (3)$$

When the network is already congested, Eq. (3) does not generally holds since elements traversing congested paths stack in intermediate nodes, resulting in a cascade effect not captured by the betweenness. Therefore, our analysis only covers the onset of congestion and it cannot be directly applied to the congested regime.

An efficient algorithm to compute the betweenness on multiplex structures can be found in Ref. [24] for shortest path dynamics and in Ref. [25] for random walk dynamics. The computation of $s_{i\alpha}$ and $e_{i\alpha}$ for shortest path dynamics can be obtained by modifying the previously cited algorithm to account for the number of paths that reach the source and destination nodes.

The onset of congestion of the multiplex is attained when a node i in layer α is required to process elements at its maximum processing rate, i.e., $g_{i\alpha} + \sigma_{i\alpha} = \tau$. Therefore, the critical injection rate of the system ρ_c becomes

$$\rho_c = \tau L^{-1} \frac{N-1}{\mathcal{B}^*}, \quad (4)$$

where $\mathcal{B}_{i\alpha} \equiv B_{i\alpha} + s_{i\alpha} + e_{i\alpha}$ and $\mathcal{B}^* \equiv \max_{i\alpha} \mathcal{B}_{i\alpha}$. In the following we call $\mathcal{B}_{i\alpha}$ the interconnected betweenness. Note that $\mathcal{B}_{i\alpha}$ depends on the intralayer paths and interlayer paths, and on the migration of shortest paths between layers (more efficient layers contain a larger proportion of the starting and ending routes). We tested the validity of Eq. (4) against Monte Carlo simulations on top of Erdős-Rényi multiplex networks, see Fig. 2.

In the following, we investigate the role of the topology of the individual layers on the multiplex congestion. First of

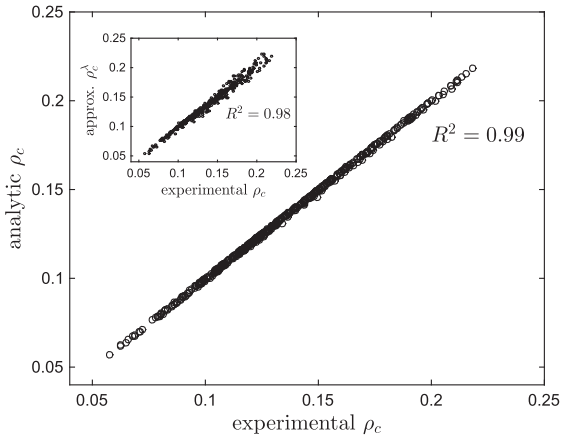

FIG. 2. Accuracy of the analytical value of ρ_c given by Eq. (4) predicting the actual onset of congestion in experimental simulations on 500 random multiplex networks formed by two Erdős-Rényi networks (of 500 nodes) as layers. The inset shows the correlation between the experimentally obtained critical injection rate and the analytical approximation in Eq. (6) where λ is approximated by 1. R^2 is the coefficient of determination for the linear fits.

all, note that in the definition and computation of the multiplex betweenness (see Ref. [24]), the shortest paths (possibly degenerated) between all pairs of multiplex locations $N(N-1)$ are considered. The multiplex structure unbalances, in a highly nonlinear way, the distribution of shortest paths among the layers. However, some approximations are possible to grasp the effect of the different contributions to the onset of congestion in multiplex structures.

As stated before, an important parameter of traffic dynamics in multiplex networks is the fraction of interlayer shortest paths, i.e., the fraction of shortest paths that contain, at least, one interlayer edge. Experiments with multiplex networks composed of two layers, each one being a different random Erdős-Rényi network, show that most of the shortest paths are fully contained within a layer, see the Supplemental Material [[26]]. This effect becomes more evident as the degree of the layers increases. Therefore, the fraction of shortest paths fully contained within layers, λ, is basically 1, and the main factor influencing the traffic dynamics is the migration of shortest paths from the less efficient layer (the one with larger shortest paths) to the most efficient one. Under this situation we can approximate the interconnected betweenness of node i in layer α, $\mathcal{B}^{i\alpha}$, in terms of the betweenness of node i of layer α, $\mathcal{B}^i_{(\alpha)}$, when layer α is considered as a single layer network:

$$\mathcal{B}^{i\alpha} \approx \lambda \mu_\alpha \mathcal{B}^i_{(\alpha)}, \tag{5}$$

where $\mu_\alpha < 1$ is the fraction of shortest paths using only layer α, satisfying $\sum_\alpha \mu_\alpha = 1$. The effect of the product of $\lambda\mu_\alpha$ is to precisely account for the fraction of all shortest paths that traverse only layer α in the multiplex. Note that the approximation in Eq. (5) does not account for the betweenness contribution of the paths that use interlayer edges. However, the high value $\lambda \approx 1$ indicates that they are usually negligible, and we can even further approximate $\mathcal{B}^{i\alpha} \approx \mu_\alpha \mathcal{B}^i_{(\alpha)}$.

Taking advantage of Eq. (5), the critical injection rate of the multiplex can be obtained by rescaling the critical injection rate of the individual layers:

$$\rho_c \approx \tau L^{-1} \frac{N-1}{\lambda \mu_\ell B^*_{(\ell)}} \approx \frac{1}{L\mu_\ell} \rho_c^{(\ell)}, \tag{6}$$

where $\rho_c^{(\ell)}$ is the critical injection rate of the most efficient layer ℓ. Fractions μ_ℓ and λ are genuine properties of the multiplex network structure that can be obtained by means of the multiplex extension of Brandes' betweenness algorithm [24]. Figure 2(a) shows the accuracy of this approximation in the calculation of ρ_c. The high accuracy obtained in the approximation evidences that the critical injection rate of the multiplex crucially depends on the migration of shortest paths between layers, which is captured in μ_ℓ.

As an example, consider a multiplex structure composed of two identical layers. In this case, there are no shortest paths using interlayer edges since they would be longer than the ones fully included in one layer; thus, $\lambda = 1$. Since the paths in both layers are identical, there is a multiplex path degeneration: for each shortest path in layer 1 there is an equivalent shortest path in layer 2. As a consequence, the nodes on the paths only obtain $1/2$ of the betweenness contribution they would obtain if the layers were separated, which results in $\mu_\ell = 1/2$. Eventually, we see that for identical layers the multiplex betweenness is $1/2$ of the betweenness computed on any of the layers.

On the other side, consider a multiplex network in which most of the paths in layer 1 have length 2 and most of the paths in layer 2 have length 3. Again, there are very few shortest paths using interlayer edges since their minimum length is 3 (i.e., one intralayer edge, followed by a change of layers through an interlayer link, and finally another intralayer edge); therefore, $\lambda \approx 1$. Moreover, most of the shortest paths make use of layer 1, where the lengths are shorter, so $\mu_1 \approx 1$ and $\mu_2 \approx 0$. Substitution in Eq. (5) shows that the interconnected betweenness of the multiplex is equivalent to the betweenness of the most efficient layer, which in this case is layer 1.

We can compute the congestion induced by a multiplex as the situation in which a multiplex network reaches congestion with less load than the worst of its layers when operating individually. In a multiplex with two layers, 1 and 2 (layer 2 being the most efficient), this limiting situation is obtained when $\rho_c < \rho_c^{(1)}$, and consequently

$$\frac{1}{L\lambda\mu_2} \lesssim \frac{\mathcal{B}^*_{(2)}}{\mathcal{B}^*_{(1)}}. \quad (7)$$

Figure 3(a) shows the regions where the multiplex structure induces congestion for sets of Erdős-Rényi multiplex networks. In each experiment, two Erdős-Rényi networks with a different mean degree are coupled to form a multiplex network. For each pair of mean degrees we have evaluated 100 random realizations of the multiplex network and for each realization we have computed the onset of congestion of the multiplex network and of the individual layers. We have then obtained the fraction of times that the multiplex network reaches congestion before both layers. The boundaries approximated by Eq. (7) determine accurately the regions where the multiplex induces congestion. As expected, the approximation using only μ works well except when both mean degrees are low since in these cases the number of shortest paths using the multiplex structure is more relevant. Surprisingly, for larger degrees (in the diagonal) the Erdős-Rényi networks generated present small fluctuations on the average degree that eventually make a node in one layer have a maximum degree a little bit larger than in the other layer. This asymmetry, for such dense networks, is enough to provoke a load unbalance that is reflected in the simulations.

We have used homogenous random network multiplexes to demonstrate the use of the analytical approach; however, the theory is general for any other multiplex network structure. To conclude this Letter, we have also used a different type of topology, random geometric graphs, more akin to representing transportation networks [11,27]. To this end, we propose a simple configuration of a random geometric multiplex. We assume that each random geometric multiplex is composed of two types of transportation media: short range (e.g., the bus network) and long range (e.g., the subway), see Fig. 4.

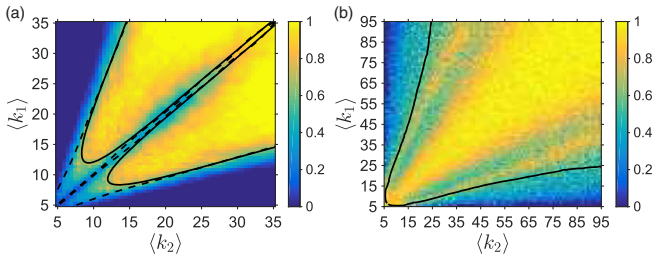

FIG. 3. Probability of obtaining a multiplex configuration that induces congestion when (a) the multiplex is composed of two Erdős-Rényi layers and (b) the topology is a random geometric multiplex. In both network topologies each layer has 500 nodes. The number of simulations points is 50^2, and for each point we generate 10^2 configurations fixing $\langle k_1 \rangle$ and $\langle k_2 \rangle$. The colors indicate the probability of observing that the onset of congestion of the multiplex satisfies $\rho_c < \min(\rho_c^{(1)}, \rho_c^{(2)})$. The lines show the accuracy of Eq. (7) in detecting the region where the multiplex structure induces congestion. The solid lines represent the expression when the real value of λ is used and the dashed lines when we approximate λ by 1.

Our construction method allows us to generate very extreme geometric multiplexes, from configurations where the long range layer only contains some of the longer edges of the short range layer ($R_{L2}^{\mathrm{Max}} \approx R_{L1}$) to those where the long range layer only contains edges larger than the ones in the short range layer ($R_{L2}^{\mathrm{Min}} \approx R_{L1}$). However, we usually obtain configurations where the long range layer has some degree of edge overlap with the short range layer. The test set where we have performed the experiments has been constructed by creating 10^5 random geometric multiplex networks, choosing uniformly at random the parameters of the model. Figure 3(b) shows that Eq. (7) accurately predicts the region where the multiplex structure induces congestion.

In summary, we have analyzed the congestion phenomena on multiplex transportation networks. We developed a standardized model of how elements traverse those networks and we provided an analytical expression for the onset of congestion. We then showed that the multiplex structure induces congestion and derived analytical expressions to determine the network parameters that raise these phenomena. All analytical expressions have been assessed on Erdős-Rényi and geometric multiplex networks, and show perfect agreement with the empirical results. The reason behind this phenomenology is the unbalance of the shortest paths between the layers. The flow follows the shortest path, increasing the load of the most efficient (in terms of the shortest paths) layer, and eventually congesting it. The theory

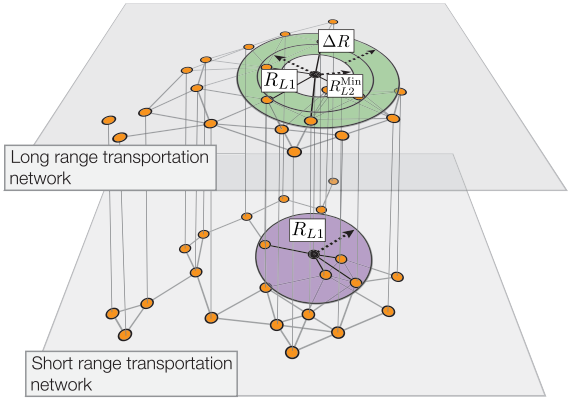

FIG. 4. Sketch of the generation process of a random geometric multiplex. We first choose uniformly at random a set of N points in a bidimensional space, $(x, y) \in [0, 1]^2$; these are our node locations. We then generate the first layer by adding edges between all locations i and j separated by a distance d_{ij} smaller than a certain radius $R_{L1} \in [0, 0.4]$. The second layer is generated by adding edges between all node pairs with distance $R_{L2}^{\mathrm{Min}} < d_{ij} < R_{L2}^{\mathrm{Max}}$. The values of $R_{L2}^{\mathrm{Min}} \in [0, R_{L1}]$ force minimum overlapping between both layers. The value of $R_{L2}^{\mathrm{Max}} = R_{L2}^{\mathrm{Min}} + \Delta R$ with $\Delta R \in [0, R_{L1}]$ ensures the range $[R_{L2}^{\mathrm{Min}}, R_{L2}^{\mathrm{Max}}]$ does not exceed the radius of the first layer.

and experiments developed in this Letter are especially useful to understand the transportation dynamics in multilayer networks and might help in the development of more efficient transportation networks and routing algorithms.

This work has been supported by Ministerio de Economía y Competitividad (Grant No. FIS2015-71582-C2-1) and European Comission FET-Proactive Projects PLEXMATH (Grant No. 317614). A. A. also acknowledges partial financial support from the ICREA Academia and the James S. McDonnell Foundation.

[1] M. Newman, *Networks: An Introduction* (Oxford University Press, New York, 2010).
[2] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M. A Porter, S. Gómez, and A. Arenas, Mathematical Formulation of Multilayer Networks, Phys. Rev. X **3**, 041022 (2013).
[3] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, Multilayer networks, J. Complex Netw. **2**, 203 (2014).
[4] S. Boccaletti, G. Bianconi, R. Criado, C. I. del Genio, J. Gmez-Gardees, M. Romance, I. Sendia-Nadal, Z. Wang, and M. Zanin, The structure and dynamics of multilayer networks, Phys. Rep. **544**, 1 (2014).
[5] M. De Domenico, A. Solé-Ribalta, S. Gómez, and A. Arenas, Navigability of interconnected networks under random failures, Proc. Natl. Acad. Sci. U.S.A. **111**, 8351 (2014).
[6] E. Strano, S. Shai, S. Dobson, and M. Barthelemy, Multiplex networks in metropolitan areas: generic features and local effects, J. R. Soc. Interface **12** (2015).
[7] P. J. Mucha, T. Richardson, K. Macon, M. A Porter, and J.-P. Onnela, Community structure in time-dependent, multiscale, and multiplex networks, Science **328**, 876 (2010).
[8] M. Magnani and L. Rossi, in Advances in Social Networks Analysis and Mining (ASONAM), *2011 International Conference on* (IEEE, New York, 2011), pp. 5–12.
[9] C. Granell, S. Gómez, and A. Arenas, Dynamical Interplay between Awareness and Epidemic Spreading in Multiplex Networks, Phys. Rev. Lett. **111**, 128701 (2013).
[10] M. De Domenico, A. Solé-Ribalta, E. Omodei, S. Gómez, and A. Arenas, Ranking in interconnected multilayer networks reveals versatile nodes, Nat. Commun. **6** (2015).
[11] M. Barthelemy, Spatial networks, Phys. Rep. **499**, 1 (2011).
[12] R. G. Morris and M. Barthelemy, Transport on Coupled Spatial Networks, Phys. Rev. Lett. **109**, 128703 (2012).
[13] R. Louf and M. Barthelemy, Modeling the Polycentric Transition of Cities, Phys. Rev. Lett. **111**, 198702 (2013).
[14] R. Louf and M. Barthelemy, A typology of street patterns, J. R. Soc. Interface **11**, 0924 (2014).
[15] R. Guimerà, A. Diaz-Guilera, F. Vega-Redondo, A. Cabrales, and A. Arenas, Optimal Network Topologies for Local Search with Congestion, Phys. Rev. Lett. **89**, 248701 (2002).
[16] L. Zhao, Y.-C. Lai, K. Park, and N. Ye, Onset of traffic congestion in complex networks, Phys. Rev. E **71**, 026125 (2005).
[17] F. Tan, J. Wu, Y. Xia, and K. Tse Chi, Traffic congestion in interconnected complex networks, Phys. Rev. E **89**, 062813 (2014).
[18] P. Echenique, J. Gómez-Gardeñes, and Y. Moreno, Dynamics of jamming transitions in complex networks, Europhys. Lett. **71**, 325 (2005).
[19] P. Echenique, J. Gómez-Gardeñes, and Y. Moreno, Improved routing strategies for internet traffic delivery, Phys. Rev. E **70**, 056105 (2004).
[20] S. Chen, W. Huang, C. Cattani, and G. Altieri, Traffic dynamics on complex networks: a survey, Math. Probl. Eng. **2012** (2011).
[21] A. Cardillo, M. Zanin, J. Gómez-Gardeñes, M. Romance, A. J. García del Amo, and S. Boccaletti,, Modeling the multi-layer nature of the european air transport network: Resilience and passengers re-scheduling under random failures, Eur. Phys. J. Spec. Top. **215**, 23 (2013).
[22] D. Braess, A. Nagurney, and T. Wakolbinger, On a paradox of traffic planning., Transp. Sci. **39**, 446 (2005).
[23] In the case of shortest path degeneracy (multiple shortest paths between the same two locations), the fractional contributions must be considered. For an efficient computation a modified version of Brandes' algorithm might be used [24].
[24] A. Solé-Ribalta, M. De Domenico, S. Gómez, and A. Arenas, Centrality rankings in multiplex networks, in *Proceedings of the 2014 ACM Conference on Web Science, Bloomington, IN, USA, June 23–26* (ACM, New York, 2014), pp. 149–155.
[25] A. Solé-Ribalta, M. De Domenico, S. Gómez, and A. Arenas, Random walk centrality in interconnected multilayer networks, Phys. D (Amsterdam) (to be published).
[26] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.116.108701 for supplemental experiments showing the fraction of interlayer shortest paths.
[27] D. Taylor, F. Klimm, H. A. Harrington, M. Kramár, K. Mischaikow, M. A. Porter, and P. J. Mucha, Topological data analysis of contagion maps for examining spreading processes on networks, Nat. Commun. **6** (2015).