Growing Multiplex Networks

V. Nicosia,[1] G. Bianconi,[1] V. Latora,[1,2] and M. Barthelemy[3]

[1]*School of Mathematical Sciences, Queen Mary University of London, Mile End Road, E1 4NS London, United Kingdom*
[2]*Dipartimento di Fisica e Astronomia, Università di Catania and INFN, 95123 Catania, Italy*
[3]*Institut de Physique Théorique, CEA, CNRS-URA 2306, F-91191 Gif-sur-Yvette, France*

(Received 28 February 2013; published 31 July 2013)

We propose a modeling framework for growing multiplexes where a node can belong to different networks. We define new measures for multiplexes and we identify a number of relevant ingredients for modeling their evolution such as the coupling between the different layers and the distribution of node arrival times. The topology of the multiplex changes significantly in the different cases under consideration, with effects of the arrival time of nodes on the degree distribution, average shortest path length, and interdependence.

DOI: 10.1103/PhysRevLett.111.058701 PACS numbers: 89.75.Hc, 89.75.−k, 89.75.Fb

Many different physical, biological, and social systems are structured as networks, and their properties are now, after a decade of effort, well understood [1–4]. However, a complex network is rarely isolated, and some of its nodes could be part of many graphs at the same time. Examples include multimodal transportation networks [5,6], climatic systems [7], economic markets [8], energy-supply networks [9], and the human brain [10]. In these cases, each network is part of a larger system in which a set of interdependent networks with different structure and function coexist, interact, and coevolve. So far network scientists have investigated these systems by looking at one type of relationship at a time, e.g., by analyzing collaboration networks and email communications as separate graphs. However, the structural properties of each of these networks and their evolution can depend in a nontrivial way on that of other graphs to which they are interconnected. Consequently, these systems are better represented as multiplexes, i.e., graphs composed by M different layers in which the same set of N nodes can be connected to each other by means of links belonging to M different classes or types. Despite some early attempts in the field of social network analysis [11], the characterization of multiplexes is still in its infancy, mainly due to the lack of multiplex data. However, some recent works have already proposed suitable extensions to multilayer graphs of classic network metrics and models [12–14]. Preliminary results show that multiplexicity has important consequences for the dynamics of processes occurring in real systems, including routing [12,15], diffusion [16], cooperation [17], election models [18], and epidemic spreading [19]. Nowadays, an increasing number of new data sets of multiplex systems, e.g., coming from large online social networks [20,21], trading networks [22] and human neuroimaging techniques [23], are rapidly becoming available and demand for adequate models to understand their structure and evolution.

In this Letter we propose and study a generic model of multiplex growth, inspired by classical models based on preferential attachment, in which the probability for a newly arrived node to establish connections to existing nodes in each of the layers of a multiplex is a function of the degree of other nodes at all layers. We define two new metrics to characterize the structure of multiplexes and we study the effect of different attachment rules and the impact of delays in the arrival of nodes at different layers on the structure of the resulting network. We provide closed forms for both the degree distributions at each layer and the interlayer degree-degree correlations, and we show how different attachment kernels can change the distributions of distances and interdependence.

More precisely, a multiplex is a set of N nodes which are connected to each other by means of edges belonging to M different classes or types. We represent each class of edges as a separate *layer*, and we assume that a node i of the multiplex consists of M replicas, one for each layer. We denote by $V^{[\alpha]}$ the set of the nodes in layer α, and by $E^{[\alpha]}$ the set of all the edges of a given type α. An M-layer multiplex is therefore fully specified by the vector $\mathcal{A} = [A^{[1]}, A^{[2]}, \ldots, A^{[M]}]$, whose elements are the adjacency matrices $A^{[\alpha]} = \{a_{ij}^{[\alpha]}\}$, where $a_{ij}^{[\alpha]} = 1$ if node i and node j are connected by an edge of type α, whereas $a_{ij}^{[\alpha]} = 0$ otherwise. We denote by $k_i^{[\alpha]} = \sum_j a_{ij}^{[\alpha]}$ the degree of node i at layer α, i.e., the number of edges of type α of which i is an endpoint, and by $\boldsymbol{k}_i$ the M-dimensional vector of the degrees of the replicas of i. In general, the degrees of the replicas of i are distinct, and some replicas can also be isolated (i.e., $k_i^{[\alpha]} = 0$ for some value of α). In the following we consider all the edges at all layers to be undirected and unweighted. As in the case of classical "singlex" graphs, we can characterize each layer α of a multiplex by studying the degree distribution $P(k^{[\alpha]})$, and the joint-degree distribution $P(k^{[\alpha]}, k'^{[\alpha]})$. However, we are interested here in the structural properties of the multiplex as a whole, so we propose to quantify the correlations between the degrees of replicas of the same node at two different layers α and α', by constructing the interlayer joint-degree distributions $P(k^{[\alpha]}, k^{[\alpha']})$, or the conditional degree distributions $P(k^{[\alpha']}|k^{[\alpha]})$. In particular, we can look at the projection of the conditional distribution obtained by considering the average degree $\bar{k}^{[\alpha']}$ at layer α' of nodes having degree $k^{[\alpha]}$ at layer α:

$$\bar{k}^{[\alpha']}(k^{[\alpha]}) = \sum_{k^{[\alpha']}} k^{[\alpha']} P(k^{[\alpha']}|k^{[\alpha]}). \quad (1)$$

By plotting this quantity as a function of $k^{[\alpha]}$ we can detect the presence and the sign of degree correlations between the two layers. For a multiplex with no correlations between layers α and α' we expect $\bar{k}^{[\alpha']}(k^{[\alpha]}) = \langle k^{[\alpha']}\rangle$ and $\bar{k}^{[\alpha]}(k^{[\alpha']}) = \langle k^{[\alpha]}\rangle$. If $\bar{k}^{[\alpha']}(k^{[\alpha]})$ increases with $k^{[\alpha]}$ we say that the degrees of the two layers have positive (assortative) correlations, while if $\bar{k}^{[\alpha']}(k^{[\alpha]})$ is a decreasing function of $k^{[\alpha]}$ we say that the degrees on layer α and α' are anticorrelated (or disassortatively correlated). We notice that a similar concept of internetwork assortativity was already defined in Ref. [24] for the case of interdependent graphs, while the authors of Ref. [13] proposed to measure interlayer assortativity by means of the Pearson's linear correlation coefficient of degrees [25].

In addition to the assortativity, we can also characterize the "multiplex reachability" of a node i, e.g., by computing the average distance L_i from i to any other node of the multiplex, and comparing this average distance with that measured on each layer separately. The presence of more than one layer in a multiplex produces an increase in the number of available paths, so that the distance between two nodes of a multiplex will be, in general, smaller than, or at most equal to, that measured on each layer separately. A better measure to quantify the value added by the multiplexicity to the reachability of nodes is the *interdependence* [12], which for a node i is defined by

$$\lambda_i = \sum_{\substack{j\in N \\ j\neq i}} \frac{\psi_{ij}}{\sigma_{ij}}, \quad (2)$$

where ψ_{ij} is the number of shortest paths between node i and node j which use edges lying on more than one layer, while σ_{ij} is the total number of shortest paths between i and j in the multiplex. The interdependence of a multiplex is computed as the average node interdependence $\lambda = 1/N\sum_i \lambda_i$ with $\lambda \in [0, 1]$. If λ is close to zero, then most of the shortest paths among nodes lie on just one layer, while if λ is close to 1 the majority of the shortest paths exploit more than one layer.

The few models of multiplexes proposed so far are based on the juxtaposition of random graphs [13]. However, networks usually result from a growing process consisting of the addition of nodes and edges over time. For this reason, we introduce here a model of growing multiplex networks. Most of the classical growing models for single-layer networks start from an initial connected graph with m_0 nodes and assume that new nodes arrive in the graph one by one, carrying m edge stubs, and connect with other existing nodes according to a prescribed attachment rule. In that case, each node i has a unique *arrival time* t_i, but in multiplexes, instead, each layer can exhibit a different edge-formation dynamics, and in general the edges of the M replicas of a new node are not created at the same time. For instance, a face-to-face interaction relationship is usually established before two individuals become friends, while two locations are usually connected by a road before a direct railway line between them is constructed. Consequently, we assume that a newly arrived node has exactly m stubs on each layer of the multiplex (in [27] we briefly discuss the case where m is a random variable), but the replica of a node i on layer α can connect its m stubs at a different time $t_i^{[\alpha]}$. We denote by $\boldsymbol{t}_i$ the vector of arrival times of the replicas of node i. In order to make the model analytically tractable, we make two simplifying assumptions. The first is that there exists a layer $\bar{\alpha}$ so that $t_i^{[\bar{\alpha}]} \leq t_i^{[\alpha]} \forall i, \forall \alpha \neq \bar{\alpha}$. This is equivalent to saying that a newly arrived node must first create its connections on layer $\bar{\alpha}$ before any of its replicas can create connections on any other layer $\alpha \neq \bar{\alpha}$. We call $\bar{\alpha}$ the *master layer* (in Ref. [27] we briefly discuss the case in which this assumption does not hold, and each node can arrive first on any of the M layers of the multiplex). The second assumption is that nodes arrive one by one on the master layer, at equal discrete time intervals $t = \{1, 2, \ldots\}$. We label the nodes of a growing multiplex according to the ordering induced by their arrival on the master layer. Without loss of generality, in the following we assume that the master layer is the first one, i.e., $\bar{\alpha} = 1$, and that the arrival times of the replicas of node i have the form

$$t_i^{[\alpha]} = T[t_i^{[1]}, \xi^{[\alpha]}(\tau)], \quad (3)$$

where T is a certain function of $t_i^{[1]}$ and of the random variable $\xi^{[\alpha]}(\tau)$. By appropriately choosing T and $\xi^{[\alpha]}(\tau)$ we can model different arrival behaviors, including (i) simultaneous arrival ($T = t_i^{[1]}$), and (ii) power-law delayed arrival [$T = t_i^{[1]} + \xi(\tau)$ and $P(\xi = \tau) = (\beta - 1)\tau^{-\beta}$ for $\tau \geq 1$ and $\beta > 1$]. Upon arrival, the newborn node i connects to m existing nodes in the master layer, according to a certain attachment rule. As in the preferential attachment models [28], we assume that the attachment probability depends on the degree of a node. However, in a multiplex the probability for node i to connect to node j on each layer α can depend not only on $k_j^{[\alpha]}$ but also on the degrees of j's replicas on the other layers

$$\Pi_{i\to j}^{[\alpha]} = \frac{F_j^{[\alpha]}(\boldsymbol{k}_j)}{\sum_l F_l^{[\alpha]}(\boldsymbol{k}_l)}. \quad (4)$$

For the sake of clarity and without loss of generality, we focus in the following on 2-layer multiplexes with

$\alpha = 1, 2$. We begin with the simplest case of *linear attachment* which is the natural extension of the Barabási-Albert model [28]. In this case, we consider that the probability for a newborn node i to connect to an existing node j on layer α is proportional to a linear combination of the degrees of j at all layers. The attachment kernels can then be expressed as

$$\begin{bmatrix} F^{[1]}[k, q] \\ F^{[2]}[k, q] \end{bmatrix} = C \begin{bmatrix} k \\ q \end{bmatrix} = \begin{bmatrix} c^{[1,1]} & c^{[1,2]} \\ c^{[2,1]} & c^{[2,2]} \end{bmatrix} \begin{bmatrix} k \\ q \end{bmatrix}, \quad (5)$$

where we use here and in the following the notations $k^{[1]} = k$ and $k^{[2]} = q$ [29]. The coefficients $c^{[r,s]}$ tune the dependence of the attachment probability at layer r on the degrees of nodes at layer s. In the case of 2-layer multiplexes we can represent the set of coefficients $C = \{c^{[r,s]}\}$ using the compact notation $\{c^{[1,1]}, c^{[1,2]}, c^{[2,1]}, c^{[2,2]}\}$. The dynamics can be easily solved in the mean-field (see [27] for details) and in some specific cases we can fully characterize the degree correlations within the two different layers by analytically solving the master equation. If we denote by $N_{k,q}(t)$ the number of nodes having, at time t, degree k on the first layer and degree q on the second layer, and by $\Pi^{[\alpha]}_{k,q}$ the probability that one of these $N_{k,q}(t)$ nodes acquires one of the m new links on layer α at time $t + 1$, the master equation can be written as [30]

$$N_{k,q}(t + 1) = N_{k,q}(t) + \mathcal{G} - \mathcal{L}, \quad (6)$$

where

$$\mathcal{G} = m[\Pi^{[1]}_{k-1,q} N_{k-1,q}(t) + \Pi^{[2]}_{k,q-1} N_{k,q-1}(t)] + \delta_{k,m}\delta_{q,m},$$
$$\mathcal{L} = m[\Pi^{[1]}_{k,q} + \Pi^{[2]}_{k,q}] N_{k,q}(t),$$

represent, respectively, the expected increase ($\mathcal{G}$) and the expected decrease ($\mathcal{L}$) of $N_{k,q}$ at time $(t + 1)$. Assuming that $N_{k,q} = tP(k, q)$ for large t, the solution of Eq. (6) is obtained by solving the corresponding recursive expression (see [27] for details). In the following we summarize the master-equation solution in some particularly interesting cases. First of all, let us consider simultaneous arrival of the nodes in the two layers. If we set $C = \{1, 0, 0, 1\}$ then the attachment probability at each layer will depend only on the degree of the nodes in the same layer. In this case the degree distribution in the first layer reads [31,32]

$$P(k) = \frac{2m(m + 1)}{k(k + 1)(k + 2)}, \qquad k > m, \quad (7)$$

and the degree distribution in the second layer is identically equal. This distribution goes as $P(k) \sim k^{-\gamma}$ with $\gamma = 3$. If we solve the master equation for the multiplex evolution we obtain the analytical expression for the interlayer joint degree probability $P(k, q)$

$$P(k,q) = \frac{2\Gamma(2+2m)\Gamma(k)\Gamma(q)\Gamma(k+q-2m+1)}{\Gamma(m)\Gamma(m)\Gamma(k+q+3)\Gamma(k-m+1)\Gamma(q-m+1)}. \quad (8)$$

The average degree $\bar{k}(q)$ at layer 1 of nodes having degree q at layer 2 reads

$$\bar{k}(q) = \frac{m(q + 2)}{1 + m}. \quad (9)$$

Notice that even if the two layers grow independently, the simultaneous arrival introduces nontrivial interlayer degree correlations. In fact, in the mean-field approach, the degree of a node on each layer increases over time as $k_i^{[\alpha]}(t) = m(t/t_i^{[\alpha]})^{1/2}$ (see [27] for details), so that the degrees of the two replicas of a node i depend, for large t, only on their arrival time. If both replicas have the same arrival time, i.e., $t_i^{[1]} = t_i^{[2]}$ then the degree of the two replicas will be positively correlated. In Fig. 1 we report the degree distribution and the values of $\bar{k}(q)$ for two coupling patterns, which are in good agreement with the theoretical curves [33]. It is clear from the figure that in the synchronous arrival case the shape of the coupling matrix is actually not very relevant and that the value of the degree distribution exponent and strong assortativity are robust features of these multiplexes.

If we consider a power-law delayed arrival time on the second layer, the results are significantly different. In Fig. 2 we illustrate how the exponent of the delay distribution β affects the structure of the obtained multiplex. The bulk of the degree distributions are still power laws $P(k) \sim k^{-\gamma}$ with $\gamma = 3$, but the shape of the far tail depends now on β: for small β, a few nodes are predominant and become super-hubs (as also shown in Fig. S-1 in [27]). The average shortest path and the interdependence are also significantly affected, as shown in Figs. 2(c) and 2(d). In particular, when β is closer to one the presence of more predominant

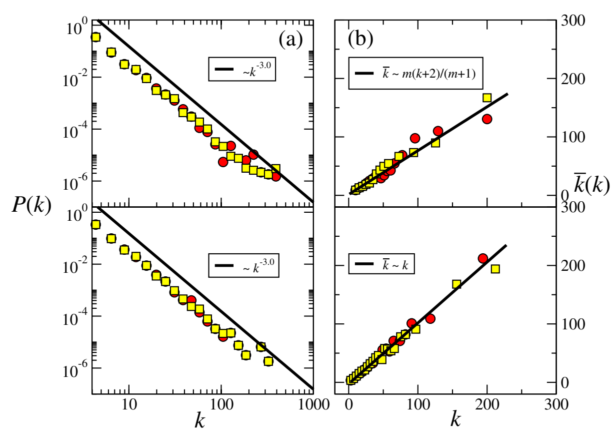

FIG. 1 (color online). (a),(b): the degree distribution $P(k)$ (left) and the projection $\bar{k}(k)$ of the interlayer degree correlations (right) closely follow the theoretical curves (solid black lines) and are relatively insensitive to the coupling matrix.

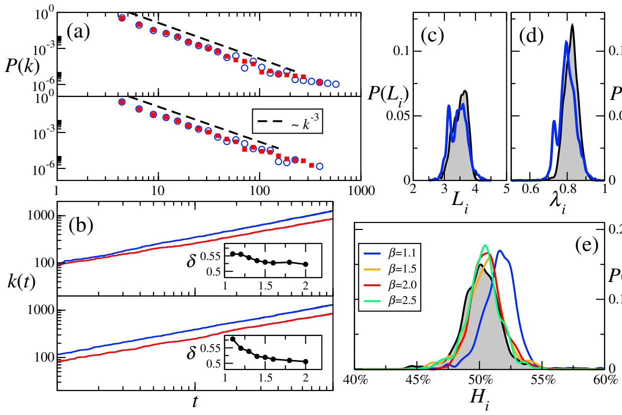

FIG. 2 (color online). (a) Degree distribution on the first layer. When β is close to 1, superhubs appear. (b) The degree $k(t)$ of the largest hub of the first layer as a function of time scales as $(t/t_0)^{\delta}$, where δ approaches 0.5 when β increases (insets). The value of β tunes the shape of the distribution of average shortest path lengths (c) and node interdependence (d). (e) The percentage of times H_i that the maximal-degree node in a shortest path from node i belongs to the first layer. When β is small, superhubs in the first layer are more abundant in the shortest paths.

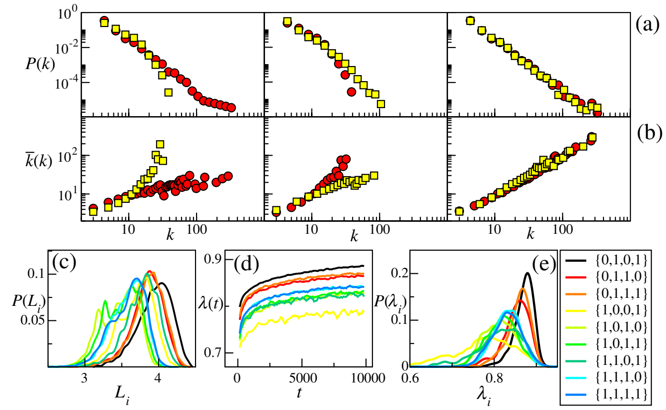

FIG. 3 (color online). Degree distributions (a) and interlayer degree correlations (b) for semilinear attachment. (c) The distribution of the average shortest path length from one node to all the other nodes heavily depends on the coupling pattern. Similarly, the interdependence of a node $\lambda(t)$ is always a sublinear function of the arrival time t but its shape depends on the coupling pattern at work (d). In general, older nodes have smaller interdependence. (e) The coupling pattern also affects the distribution of node interdependence. The smallest average interdependence is observed when the two layers are independent (yellow curve).

"old" hubs lowers the average shortest path and the interlayer assortativity. Moreover, broader delays cause a lower participation of hubs of the second layer in shortest paths, as shown in Fig. 2(e).

So far we have considered the case of two scale-free growing networks, but it would be interesting to construct multiplexes in which a scale-free network is coupled to a network with a peaked degree distribution. In this respect, we introduce a *semilinear attachment kernel* which allows to grow multiplexes in which the two layers have different topological structures. The model is defined as follows:

$$\begin{bmatrix} F^{[1]}[k,q] \\ F^{[2]}[k,q] \end{bmatrix} = C\begin{bmatrix} k \\ 1 \end{bmatrix}, \tag{10}$$

where C is still a 2×2 matrix of coefficients, as in the linear model. In this case, the degree of a node on any of the two layers could depend only on its degree on layer 1 and does not ever depend on its degree on layer 2. If we set $C = \{1, 0, 0, 1\}$ we can analytically solve the master equation and the degree distributions of the two layers read

$$P^{[1]}(k) \sim k^{-3}, \qquad P^{[2]}(q) \sim e^{-q}, \tag{11}$$

while the interlayer joint degree distribution is equal to

$$P(k,q) = a(k)\sum_{n=0}^{k-m}\binom{k-m}{n}\left(\frac{2m}{2+2m+k-n}\right)^{q-m+1} \times (-1)^{k-m+n}, \tag{12}$$

where $a(k) = \Gamma(k)/[\Gamma(m+1)\Gamma(k-m+1)]$. The function $\bar{k}(q)$ is given by

$$\bar{k}(q) = m\left(\frac{2(m+1)}{1+2m}\right)^{q-m+1}. \tag{13}$$

Similar relations can be derived for the other coupling patterns. In panel (a) and (b) of Fig. 3 we report the degree distribution and the value of $\bar{k}(q)$ for three different coupling patterns, which are in good agreement with the theoretical curves. In the semilinear model the coupling pattern has a dramatic impact on other structural properties of the multiplex such as the distribution of the average shortest path length from each node and the distribution of node interdependence. In particular, the interdependence is smaller for older nodes, and grows sublinearly with time. This implies that navigation for old nodes is easier within a single layer while younger nodes will have to resort to the different layers to reach a target. In addition, a sublinear growth implies that the system performance increases very slowly.

V. N. and V. L. acknowledge support from the Project LASAGNE, Contract No. 318132 (STREP), funded by the European Commission. M. B. thanks Queen Mary University of London for its warm welcome at the start of this project, and is supported by the FET-Proactive project PLEXMATH (FP7-ICT-2011-8; Grant No. 317614) funded by the European Commission.

[1] R. Albert and A.-L. Barabasi, Rev. Mod. Phys. **74**, 47 (2002).
[2] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
[3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez and D.-U. Hwang, Phys. Rep. **424**, 175 (2006).

[4] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, England, 2008).
[5] M. Kurant and P. Thiran, Phys. Rev. Lett. **96**, 138701 (2006).
[6] S.-R. Zou, T. Zhou, A.-F. Liu, X.-L. Xu, and D.-R. He, Phys. Lett. A **374**, 4406 (2010).
[7] J. Donges, H. Schultz, N. Marwan, Y. Zou, and J. Kurths, Eur. Phys. J. B **84**, 635 (2011).
[8] J. Yang, W. Wang, and G. Chen, Physica (Amsterdam) **388A**, 2435 (2009).
[9] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, Nature (London) **464**, 1025 (2010).
[10] E. Bullmore and O. Sporns, Nat. Rev. Neurosci. **10**, 186 (2009).
[11] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications* (Cambridge University Press, Cambridge, England, 1994), Vol. 8.
[12] R. G. Morris and M. Barthelemy, Phys. Rev. Lett. **109**, 128703 (2012).
[13] K.-M. Lee, J. Y. Kim, W. kuk Cho, K.-I. Goh, and I.-M. Kim, New J. Phys. **14**, 033027 (2012).
[14] G. Bianconi, Phys. Rev. E **87**, 062806 (2013).
[15] Y. Zhuo, Y. Peng, C. Liu, Y. Liu, and K. Long, Physica (Amsterdam) **390A**, 2401 (2011).
[16] S. Gómez, A. Díaz-Guilera, J. Gómez-Gardeñes, C. J. Pérez-Vicente, Y. Moreno, and A. Arenas, Phys. Rev. Lett. **110**, 028701 (2013).
[17] J. Gómez-Gardeñes, I. Reinares, A. Arenas, and L. M. Floría, Sci. Rep. **2**, 620 (2012).
[18] A. Halu, K. Zhao, A. Baronchelli, and G. Bianconi, Europhys. Lett. **102**, 16 002 (2013).
[19] A. Saumell-Mendiola, M. Á. Serrano, and M. Boguñá, Phys. Rev. E **86**, 026106 (2012).
[20] M. Szell, R. Lambiotte, and S. Thurner, Proc. Natl. Acad. Sci. U.S.A. **107**, 13 636 (2010).
[21] M. Magnani, B. Micenkova, and L. Rossi, arXiv:1303.4986.
[22] M. Barigozzi, G. Fagiolo, and D. Garlaschelli, Phys. Rev. E **81**, 046104 (2010).
[23] V. Nicosia, M. Valencia, M. Chavez, A. Díaz-Guilera, and V. Latora, Phys. Rev. Lett. **110**, 174102 (2013).
[24] R. Parshani, C. Rozenblat, D. Ietri, C. Ducruet, and S. Havlin, Europhys. Lett. **92**, 68 002 (2010).
[25] Notice that, as shown in [26], the Pearson's coefficient becomes infinitesimal in N when degree fluctuations are unbounded (e.g., in the case of power-law distributions $P(k) \sim k^{-\gamma}$ with $\gamma < 3$), so that it should not be employed to accurately quantify degree correlations in a generic scale-free network.
[26] N. Litvak and R. van der Hofstad, Phys. Rev. E **87**, 022801 (2013).
[27] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.111.058701 for more results and details about mean-field approaches, exact calculations, and other variants of the model.
[28] R. Albert, H. Jeong, and A.-L. Barabasi, Nature (London) **401**, 130 (1999).
[29] It is easy to show that the time required to simulate the growth of a multiplex with N nodes and M layers increases as $\mathcal{O}(N^2 M)$. See [27] for additional details.
[30] Notice that for $t \gg 1$ we have $\Pi_{k,q}^{[\alpha]} \ll 1$ $\forall k, q$; consequently, aymptotically in time we can neglect the probability that the same node acquires, at time t, a link in layer 1 and in layer 2.
[31] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. Lett. **85**, 4633 (2000).
[32] P. L. Krapivsky and S. Redner, Phys. Rev. E **63**, 066123 (2001).
[33] All the results shown in the figures correspond to 2-layer multiplexes with $N = 10\,000$ nodes, $m = 3$, $m_0 = 3$, and are averaged over 50 realizations.