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 (UK)*
[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*

We propose a modelling 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.

Many different physical, biological and social systems are structured as networks, and their properties are now, after a decade of efforts, 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 non-trivial 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 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 multi-layer 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 inter-layer 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 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 inter-layer 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:

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

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 inter-network assortativity was already defined in Ref. [24] for the case of interdependent graphs, while the authors of Ref. [13] proposed to measure inter-layer assortativity by means of the Pearson's linear correlation coefficient of degrees [25].
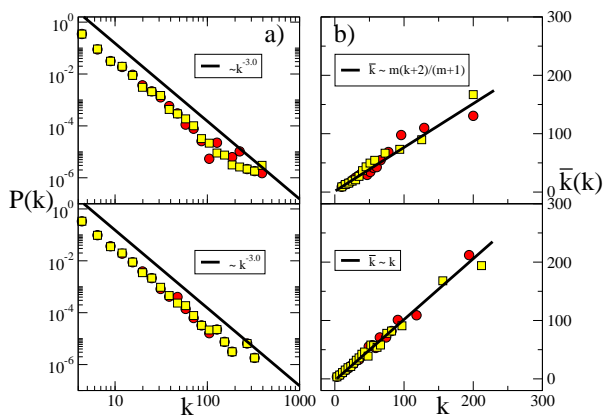

FIG. 1. (color online) **Synchronous linear attachment.** Panel a-b: the degree distribution $P(k)$ (left) and the projection $\bar{k}(k)$ of the inter-layer degree correlations (right) closely follow the theoretical curves (solid black lines) and are relatively insensitive to the coupling matrix.

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

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

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 in 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 friend, 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 Appendix 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 Appendix 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 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

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

where T is a certain function of $t_i^{[1]}$ and of the random variable $\xi^{[\alpha]}(\tau)$. By appropriately choosing T and $\xi^{[\alpha]}(\tau)$
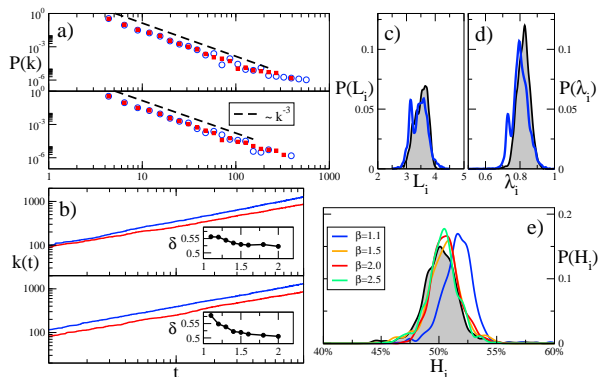
FIG. 2. (color online) **Delayed linear attachment.** (a) Degree distribution on the first layer. When β is close to 1, super-hubs appear. (b) The degree $k(t)$ of the largest hub of the first layer as a function of time scales as $(t/t_0)^\delta$, where δ approaches 0.5 when β increases (insets). The 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, super-hubs in the first layer are more abundant in the shortest paths.

we can model different arrival behaviors, including *a) simultaneous arrival* $(T = t_i^{[1]})$; *b) 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 [27], 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

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

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 [27]. 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

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

where we use here and in the following the notations $k^{[1]} = k$ and $k^{[2]} = q$ [28]. 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 mean-field (see Appendix 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 $\Pi_{k,q}^{[\alpha]}$ 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 [29]

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

where

$$\mathcal{G} = m \left[\Pi_{k-1,q}^{[1]} N_{k-1,q}(t) + \Pi_{k,q-1}^{[2]} N_{k,q-1}(t) \right] + \delta_{k,m}\delta_{q,m}$$

$$\mathcal{L} = m \left[\Pi_{k,q}^{[1]} + \Pi_{k,q}^{[2]} \right] N_{k,q}(t)$$

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 Appendix 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 [30, 31]

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

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 inter-layer joint degree probability $P(k, q)$

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

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

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

Notice that even if the two layers grow independently, the simultaneous arrival introduces non-trivial inter-layer degree correlations. In fact, in the mean-field approach, the degree of a node on each layer increases over time as $k_i^{[\alpha]}(t) = m \left(t/t_i^{[\alpha]} \right)^{1/2}$ (see Appendix for details), so
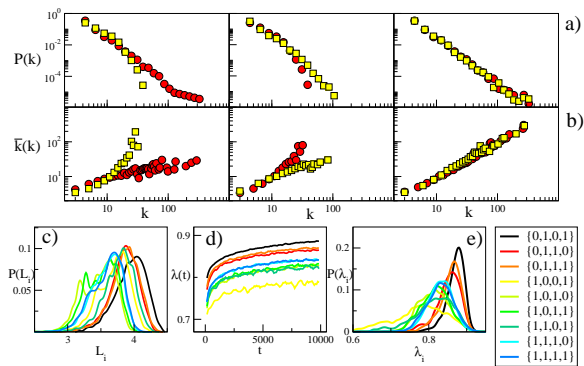
FIG. 3. (color online) **Semi-linear attachment.** Degree
distributions (a) and inter-layer degree correlations (b) for
semi-linear attachment. (c) The distribution of the average
shortest path length from one node to all the other nodes
heavily depends on the coupling pattern. Similarly, the in-
terdependence 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 indepen-
dent (yellow curve).

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 [32]. 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 distri-
bution 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 predom-
inant and become super-hubs (as also shown in Fig. S-1
in Appendix). The average shortest path and the inter-
dependence are also significantly affected, as shown in
Fig. 2c-d. In particular, when β is closer to one the pres-
ence of more predominant 'old' hubs lowers the average
shortest path and the inter-layer assortativity. Moreover,
broader delays cause a lower participation of hubs of the
second layer in shortest paths, as shown in Fig. 2e.

So far we have considered the case of two scale-free
growing networks, but it would be interesting to con-
struct multiplexes in which a scale-free network is cou-
pled to a network with a peaked degree distribution. In
this respect, we introduce a *semi-linear attachment ker-*

nel which allows to grow multiplexes in which the two
layers have different topological structures. The model is
defined as follows

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

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:

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

while the inter-layer joint degree distribution is equal to

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

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

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

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 dif-
ferent coupling patterns, which are in good agreement
with the theoretical curves. In the semi-linear model the
coupling pattern has a dramatic impact on other struc-
tural 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.

[1] R. Albert and A.-L. Barabasi, *Rev. Mod. Phys.* **74**, 47
(2002).
[2] M. E. J. Newman, *SIAM Review* **45**, 167–256 (2003).
[3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez and D.-
U. Hwang, *Phys. Rep.* **424**, 175–308 (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**, 4406–4410 (2010).

[7] J. Donges, H. Schultz, N. Marwan, Y. Zou and J. Kurths, *Eur. Phys. J. B* **84**, 635–651 (2011).

[8] J. Yang, W. Wang and G. Chen, *Physica* **388A**, 2435–2449 (2009).

[9] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley and S. Havlin, *Nature (London)* **464**, 1025–1028 (2010).

[10] E. Bullmore and O. Sporns, *Nat. Rev. Neurosci.* **10**, 186–198 (2009).

[11] S. Wasserman and K. Faust *Social Network Analysis* (Cambridge University Press, Cambridge, 1994)

[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 A* **390**, 2401–2407 (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, *EPL–Europhys. Lett.* **102**, 16002 (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. USA* **107**, 13636 (2010).

[21] M. Magnani, B. Micenkova and L. Rossi, *preprint, arXiv:1303.4986* (2013).

[22] M. Barigozzi, G. Fagiolo and D. Garlaschelli, *Phys. Rev. E* **81**, 046104 (2010).

[23] V. Nicosia, M. Valencia, M. Chavez, A. Díaz-Guilera, V. Latora, *Phys. Rev. Lett.* **110**, 174102 (2013).

[24] R. Parshani, C. Rozenblat, D. Ietri, C. Ducruet and S. Havlin, *EPL–Europhys. Lett.* **92**, 68002 (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] R. Albert, H. Jeong and A.-L. Barabasi, *Nature* **401**, 130–131 (1999).

[28] 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 Appendix for additional details

[29] Notice that for $t \gg 1$ we have $\Pi_{k,q}^{[\alpha]} \ll 1 \quad \forall k, q$; 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.

[30] S. N. Dorogovtsev, J. F. F. Mendes and A. N. Samukhin, *Phys. Rev. Lett.* **85**, 4633 (2000).

[31] P. L. Krapivsky and S. Redner, *Phys. Rev. E* **63**, 066123 (2001).

[32] All the results shown in the figures correspond to 2-layer multiplexes with $N = 10000$ nodes, $m = 3$, $m_0 = 3$, and are averaged over 50 realizations.

Mean-field theory

In this section, using a mean-field approach, we discuss the time evolution of the degree of the nodes of a multiplex in the different layers and we derive long-time expressions for the degree distribution at each layer and for the inter-layer degree-degree correlations. We first consider the linear attachment case (A) and then proceed to the semi-linear case (B). We also provide a concise discussion of the case in which new nodes bring a random number of new edges (C).

Linear attachment kernel on both layers

According to the growth model discussed in the main text, the probability that a newly arrived node i in the multiplex creates a link to node j on layer α can be written as

$$\Pi_{i \to j}^{[\alpha]} = \frac{F_j^{[\alpha]}(\boldsymbol{k}_j)}{\sum_l F_l^{[\alpha]}(\boldsymbol{k}_l)} \tag{S-1}$$

where $F_j^{[\alpha]}(\boldsymbol{k}_j)$ is a certain function of the degrees of the replicas of node j. If $F_j^{[\alpha]}(\boldsymbol{k}_j)$ is a linear function of $\boldsymbol{k}_j$ $\forall \alpha$, and all the replicas of the new node arrive at the same time, the temporal evolution of the degree of a node on each layer is governed by the equations

$$\begin{bmatrix} \frac{dk^{[1]}}{dt} \\ \frac{dk^{[2]}}{dt} \end{bmatrix} = \frac{1}{2t} C \begin{bmatrix} k^{[1]} \\ k^{[2]} \end{bmatrix} = \frac{1}{2t} \begin{bmatrix} c^{[1,1]} & c^{[1,2]} \\ c^{[2,1]} & c^{[2,2]} \end{bmatrix} \begin{bmatrix} k^{[1]} \\ k^{[2]} \end{bmatrix} \tag{S-2}$$

with the constraints $c^{[1,1]} + c^{[1,2]} = 1$ and $c^{[2,1]} + c^{[2,2]} = 1$. Since the matrix elements are real and non-zeros, the maximal eigenvalue is real. Moreover since we have just two layers then both eigenvalues λ_1 and λ_2 are real. We notice that if we impose that each row of matrix C must sum to 1, and that the coefficients $c^{[r,s]}$ are non-negative (to ensure that $\Pi_{i \to j}^{[\alpha]}$ is a probability distribution $\forall \alpha$), then we can write

$$C = \begin{bmatrix} a & 1-a \\ b & 1-b \end{bmatrix}, \quad 0 \le a \le 1, 0 \le b \le 1$$

It is easy to verify that if $b - a + 1 \neq 0$ the matrix C has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = (a - b) \neq 1$, with eigenvectors $\boldsymbol{u^1} = [1, 1]$ and $\boldsymbol{u^2} = [1, -b/(1-a)]$ (the degenerate case $b - a + 1 = 0$ is considered below). Since the eigenvalues are distinct, then C is diagonalizable, i.e. it is similar to the diagonal matrix

$$\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & a-b \end{bmatrix}, \quad 0 \le a \le 1, \ 0 \le b \le 1$$

whose non-zero elements are the eigenvalues of C. The system in Eq. (S-2) can be also written in the form

$$\dot{\boldsymbol{k}}(t) = A(t)\boldsymbol{k}(t) = \alpha(t)C\boldsymbol{k}(t) \tag{S-3}$$

where $\alpha(t) = \frac{1}{2t}$ is a scalar function, and C is a constant matrix. This is a homogeneous time-varying linear dynamical system, whose temporal evolution is fully determined by the initial state $\boldsymbol{k}_s = m\boldsymbol{u^1}$ and by the state transition matrix $\Phi(t, s)$, where s is the time at which a node is added to the graph

$$\boldsymbol{k}(t) = \Phi(t, s)\boldsymbol{k}_s = m\Phi(t, s)\mathbf{u^1}. \tag{S-4}$$

Since C is diagonalizable, then the transition matrix can be written as

$$\Phi(t, s) = e^{\int_s^t A(\tau) \, d\tau} = e^{C \int_s^t \frac{d\tau}{2\tau}} = e^{C\sigma}. \tag{S-5}$$

To compute the transition matrix $\Phi(t, s)$ we first compute the exponential matrix $e^{C\sigma}$ and then we substitute σ with $\int_s^t \frac{d\tau}{2\tau}$. Since the eigenvalues of C are distinct, the corresponding eigenvectors form a base of $\mathbb{R}^2$, so we have

$$e^{C\sigma} = Ve^{\Lambda}V^{-1} \tag{S-6}$$

where V is the matrix whose columns are the eigenvectors of C and Λ is the diagonal matrix of the eigenvalues of C. After some simple algebra we obtain

$$e^{C\sigma} = \frac{1}{1-(a-b)} \begin{bmatrix} be^{\sigma} + (1-a)e^{(a-b)\sigma} & (1-a)e^{\sigma} - (1-a)e^{(a-b)\sigma} \\ be^{\sigma} + \frac{b(1-a)}{a-b}e^{(a-b)\sigma} & (1-a)e^{\sigma} - \frac{b(1-a)}{(a-b)}e^{(a-b)\sigma} \end{bmatrix}. \tag{S-7}$$

By substituting σ with $\int_s^t \frac{d\tau}{2\tau} = \frac{1}{2}(\log t - \log s)$, we get:

$$\Phi(t,s) = \frac{1}{1-(a-b)} \begin{bmatrix} b(\frac{t}{s})^{\frac{1}{2}} + (1-a)(\frac{t}{s})^{\frac{a-b}{2}} & (1-a)(\frac{t}{s})^{\frac{1}{2}} - (1-a)(\frac{t}{s})^{\frac{a-b}{2}} \\ b(\frac{t}{s})^{\frac{1}{2}} + \frac{b(1-a)}{a-b}(\frac{t}{s})^{\frac{a-b}{2}} & (1-a)(\frac{t}{s})^{\frac{1}{2}} - \frac{b(1-a)}{a-b}(\frac{t}{s})^{\frac{a-b}{2}} \end{bmatrix} \tag{S-8}$$

and the solution of the system reads

$$\boldsymbol{k}_s(t) = m\Phi(t,s)\mathbf{u^1} = m\left(\frac{t}{s}\right)^{\frac{1}{2}}\mathbf{u^1}. \tag{S-9}$$

In this case we have

$$\langle k^{[1]}|k^{[2]}\rangle \propto k^{[2]}, \tag{S-10}$$

$$\langle k^{[2]}|k^{[1]}\rangle \propto k^{[1]}. \tag{S-11}$$

Let us now consider the degenerate case $b - a + 1 = 0$. Since we imposed that each row of the matrix C has to sum to 1, then $b - a + 1 = 0$ only if $a = 1$ and $b = 0$. In this case the two layers evolve independently, the matrix C is diagonal and the time evolution of the degree on each layer reads

$$k_s^{[\alpha]}(t) = m\left(\frac{t}{s}\right)^{\frac{1}{2}}, \tag{S-12}$$

with $\alpha = 1, 2$. This means that in the case of linear attachment kernel on both layers without delay the degree distribution of each layer is a power-law $P(k) \sim k^{-\gamma}$ with exponent $\gamma = 3$ and we have

$$\langle k^{[1]}|k^{[2]}\rangle \propto k^{[2]}, \tag{S-13}$$

$$\langle k^{[2]}|k^{[1]}\rangle \propto k^{[1]}. \tag{S-14}$$

Semi-linear attachment kernel

For the semi-linear attachment kernel we have

$$\begin{bmatrix} \frac{dk^{[1]}}{dt} \\ \frac{dk^{[2]}}{dt} \end{bmatrix} = \frac{1}{t}\begin{bmatrix} \frac{am}{2am+1-a} & 0 \\ \frac{bm}{2bm+1-b} & 0 \end{bmatrix}\begin{bmatrix} k^{[1]} \\ k^{[2]} \end{bmatrix} + \frac{1}{t}\begin{bmatrix} \frac{m(1-a)}{2am+1-a} & 0 \\ 0 & \frac{m(1-b)}{2bm+1-b} \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{S-15}$$

which is in the form

$$\dot{\boldsymbol{k}}(t) = A(t)\boldsymbol{k}(t) + B(t)\boldsymbol{w}(t) \tag{S-16}$$

where

$$A(t) = \begin{bmatrix} \frac{am}{t(2am+1-a)} & 0 \\ \frac{bm}{t(2am+1-a)} & 0 \end{bmatrix}, \quad B(t) = \begin{bmatrix} \frac{m(1-a)}{t(2am+1-a)} & 0 \\ 0 & \frac{m(1-b)}{t(2bm+1-b)} \end{bmatrix}, \quad \boldsymbol{w}(t) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The system in Eq. (S-16) is a non-homogeneous time-varying linear dynamical system where $\boldsymbol{w}(t)$ represents an external forcing function. If we call $\Phi(t,s)$ the state transition matrix of the corresponding homogeneous system $\dot{\boldsymbol{k}}(t) = A(t)\boldsymbol{k}(t)$, it is possible to show that the unique solution of Eq. (S-16) is given by

$$\boldsymbol{k}_s(t) = \Phi(t,s)\boldsymbol{k}_s + \int_s^t d\sigma\, \Phi(t,\sigma)B(\sigma)w(\sigma). \tag{S-17}$$

The form of the transition matrix $\Phi(t, s)$ associated to the homogeneous system depends on the value of a. When $a \neq 0$ then $\Phi(t, s)$ reads

$$\Phi(t, s) = \begin{bmatrix} \left(\frac{t}{s}\right)^\beta & 0 \\ \frac{(2abm+(1-a)b)}{(2abm-ab+a)}\left(\frac{t}{s}\right)^\beta + \frac{((a-1)b-2abm)}{(2abm-ab+a)} & 1 \end{bmatrix}, \qquad \text{where} \quad \beta = \frac{am}{2am-a+1}. \qquad \text{(S-18)}$$

By plugging Eq. (S-18) into Eq. (S-17) one obtains the mean-field temporal evolution of $k_s^{[1]}$ and $k_s^{[2]}$:

$$k_s^{[1]}(t) = \frac{am-a+1}{a}\left(\frac{t}{s}\right)^\beta + \frac{a-1}{a},$$

$$k_s^{[2]}(t) = \delta \left(\frac{t}{s}\right)^\beta + \eta\left(\log(t) - \log(s)\right) + \epsilon, \qquad \text{(S-19)}$$

where

$$\delta = \frac{2a^2bm^2 + (3a - 3a^2)bm + (a+1)^2 b}{2a^2bm - a^2b + a^2},$$

$$\eta = \frac{(a^2 - ab)m}{2a^2bm - a^2b + a^2},$$

$$\epsilon = \frac{((2a^2 - 3a)b + a^2)m - (a-1)^2 b}{2a^2bm - a^2b + a^2}.$$

In general, if $b \neq 0$ then for $t \to \infty$ we have $\left(\frac{t}{s}\right)^\beta \gg (\log(t) - \log(s))$. Consequently, Eqs. (S-19) can be written as

$$k_s^{[1]}(t) \simeq \left(m + \frac{1-a}{a}\right)\left(\frac{t}{s}\right)^\beta, \qquad \text{(S-20)}$$

$$k_s^{[2]}(t) \simeq \delta \left(\frac{t}{s}\right)^\beta \qquad \text{(S-21)}$$

so that the degree distribution on both layers reads

$$P(k^{[\ell]}) \sim k^{-(\frac{1}{\beta}+1)} = k^{-(3+\frac{1-a}{am})} \qquad \text{(S-22)}$$

and we have

$$\langle k^{[1]} | k^{[2]} \rangle \propto k^{[2]},$$

$$\langle k^{[2]} | k^{[1]} \rangle \propto k^{[1]}. \qquad \text{(S-23)}$$

Instead, if $b = 0$ the solution for the degree of nodes on the second layer reads

$$k_s^{[2]}(t) = \eta(\log(t) - \log(s)) + \epsilon, \qquad \text{(S-24)}$$

while $k_s^{[1]}(t)$ is expressed by Eq. (S-20). In this case, the degree distribution on the first layer is the same as in Eq. (S-22), while for the second layer we have

$$P(k^{[2]}) \sim e^{-\frac{k}{\eta}} \qquad \text{(S-25)}$$

and in the limit of large $k^{[1]}(t), k^{[2]}(t)$ we obtain

$$\langle k^{[1]} | k^{[2]} \rangle \propto e^{\frac{\beta k^{[2]}}{\eta}},$$

$$\langle k^{[2]} | k^{[1]} \rangle \propto \log(k^{[1]}). \qquad \text{(S-26)}$$

Eqs. (S-18—S-24) are valid when $a \neq 0$. When $a = 0$ the state transition matrix reads

$$\Phi(t, s) = \begin{bmatrix} 1 & 0 \\ \frac{bm(\log t - \log s)}{2bm - b + 1} & 1 \end{bmatrix} \qquad \text{(S-27)}$$

and the generic solutions for $k_s^{[1]}(t)$ and $k_s^{[2]}(t)$ are

$$k_s^{[1]}(t) = m\left(\log(t) - \log(s)\right) + m,$$

$$k_s^{[2]}(t) = \frac{bm^2\left(\log(t) - \log(s)\right)^2 + \left((2-2b)m + 2bm^2\right)\left(\log(t) - \log(s)\right) + 4bm^2 + (2-2b)m}{4bm - 2b + 2}. \tag{S-28}$$

In this case, the degree distribution on the first layer is exponential $P(k^{[1]}) \sim e^{-\frac{k}{m}}$. On the second layer, the functional form of the degree distribution depends on the value of b. It is easy to verify that when $b = 0$ then $P(k^{[2]}) \sim e^{-\frac{k}{m}}$, and we have in the limit of large $k^{[1]}(t), k^{[2]}(t)$

$$\langle k^{[1]} | k^{[2]} \rangle \propto k^{[2]},$$

$$\langle k^{[2]} | k^{[1]} \rangle \propto k^{[1]}. \tag{S-29}$$

Conversely, when $b > 0$ the degree distribution on the second layer is

$$P(k^{[2]}) \sim \frac{e^{-\sqrt{\mu k + \nu}}}{\sqrt{\mu k + \nu}} \tag{S-30}$$

where

$$\mu = m\left[4b^2 m + \left(2b - 2b^2\right)\right],$$

$$\nu = m^2\left[b^2 m^2 + \left(2b - 6b^2\right)m + \left(3b^2 - 4b + 1\right)\right].$$

In this case, in the limit of large $k^{[1]}(t), k^{[2]}(t)$, we have

$$\langle k^{[1]} | k^{[2]} \rangle \propto \sqrt{k^{[2]}},$$

$$\langle k^{[2]} | k^{[1]} \rangle \propto \left(k^{[1]}\right)^2. \tag{S-31}$$

Fluctuations in the number of edges

In principle, the mean-field approach could be also applied to the case in which the number of edges brought on layer α by each new-born node is not fixed but is a random variable $\xi^{[\alpha]}$ drawn from a given distribution $P(\xi^{[\alpha]})$. In this case we should solve the system of stochastic differential equations:

$$\begin{bmatrix} \frac{dk^{[1]}}{dt} \\ \frac{dk^{[2]}}{dt} \end{bmatrix} = \frac{1}{2\left(\kappa^{[1]}(t) + \kappa^{[2]}(t)\right)} \begin{bmatrix} \xi^{[1]}(t) & 0 \\ 0 & \xi^{[2]}(t) \end{bmatrix} \begin{bmatrix} a & 1-a \\ b & 1-b \end{bmatrix} \begin{bmatrix} k^{[1]} \\ k^{[2]} \end{bmatrix} \tag{S-32}$$

where:

$$\kappa^{[\alpha]}(t) = \int_1^t d\tau \ \xi^{[\alpha]}(\tau)$$

The random variable $\xi^{[\alpha]}$ is a positive integer with average $\langle \xi^{[\alpha]} \rangle$ and the dominant term at large times of $\kappa^{[\alpha]}$ is then given by

$$\kappa^{[\alpha]}(t) \simeq t\langle \xi^{[\alpha]} \rangle + o(t)$$

This implies in particular that at large times, the effect of randomness in the number of edges is negligible, and the behavior of the system is governed by the average number of edges $\langle \xi^{[\alpha]} \rangle$ added in each layer.

Master Equation approach for the model without delay

We provide here the derivation of exact expressions of $P(k)$ and $P(k,q)$ starting from the master equation of the system. We denote by $N_{k,q}(t)$ the average number of nodes that at time t have degree k in layer 1 and degree q in

layer 2. We start from a small connected network and at each time we add a node which brings, at same time, m new edges in layer 1 and m new edges in layer 2. We assume that, when we add the new node i to the network, the expected number of new links in layer 1 attached to a node j of degree k in layer 1 and degree q in layer 2 is given by $m\Pi_{i\to j}^{[1]} = \frac{A_{k,q}}{t}$. Similarly, the expected number of new links in layer 2 attached to a node j of degree k in layer 1 and degree q in layer 2 is given by $m\Pi_{i\to j}^{[2]} = \frac{B_{k,q}}{t}$. In addition to that, we work in the hypothesis that in the large t limit, $t \gg 1$, we have $A_{k,q}/t \ll 1$ and $B_{k,q}/t \ll 1$ so that we can neglect the probability that a node acquires at the same time a link in both layers. In this hypothesis the master equation for evolving multiplex network is given by

$$N_{k,q}(t+1) = N_{k,q}(t) + \frac{A_{k-1,q}}{t}N_{k-1,q}(t) + \frac{B_{k,q-1}}{t}N_{k,q-1}(t) - \left[\frac{A_{k,q}}{t} + \frac{B_{k,q}}{t}\right]N_{k,q}(t) + \delta_{k,m}\delta_{q,m} \tag{S-33}$$

for $k \geq m$ and $q \geq m$, as long as $N_{m-1,q}(t) = N_{k,m-1}(t) = 0$. Assuming that $N_{k,q} = tP(k,q)$ is valid in the large time limit $t \gg 1$, we can solve for the combined degree distribution $P(k,q)$ indicating the probability that a node has at the same time degree k in layer 1 and degree q in layer 2. We get the master equations

$$P(m,q) = \left(\prod_{j=m+1}^{q} \frac{B_{m,j-1}}{1 + A_{m,j} + B_{m,j}}\right) P(m,m),$$

$$P(k,q) = \sum_{r=m}^{q} \left(\prod_{j=r+1}^{q} \frac{B_{k,j-1}}{1 + A_{k,j} + B_{k,j}}\right) \frac{A_{k-1,r}}{1 + A_{k,r} + B_{k,r}} P(k-1,r) \tag{S-34}$$

Solution of the master equation in three simple cases

(i) Linear attachment kernel – Let us first consider a linear preferential attachment kernel, in which $c^{[1,1]} = c^{[2,2]} = 1$ and $c^{[1,2]} = c^{[2,1]} = 0$. In this case we have

$$A_{k,q} = \frac{k}{2},$$
$$B_{k,q} = \frac{q}{2}. \tag{S-35}$$

The recursive Eqs. (S-34) read

$$P(m,q) = \frac{\Gamma(q)\Gamma(3+2m)}{\Gamma(m)\Gamma(3+q+m)}P(m,m),$$

$$P(k,q) = \sum_{r=m}^{q} \left(\frac{\Gamma(q)\Gamma(3+r+k)}{\Gamma(r)\Gamma(3+q+k)}\right) \frac{k-1}{2+k+q}P(k-1,r) \tag{S-36}$$

where $P(m,m)$ is fixed by the normalization condition $\sum_{k=m}^{\infty}\sum_{q=m}^{\infty}P(k,q) = 1$. Using the relation

$$\sum_{r=m}^{q} \frac{\Gamma(k+r-2m)}{\Gamma(r-m+1)\Gamma(k-m)} = \frac{\Gamma(k+q-2m+1)}{\Gamma(k-m+1)\Gamma(q-m+1)} \tag{S-37}$$

it can be proved recursively that $P(k,q)$ takes the following expression

$$P(k,q) = \frac{2\Gamma(2+2m)}{\Gamma(m)\Gamma(m)}\frac{\Gamma(k+q-2m+1)}{\Gamma(k+q+3)}\frac{\Gamma(q)}{\Gamma(q-m+1)}\frac{\Gamma(k)}{\Gamma(k-m+1)} \tag{S-38}$$

Summing over the degree in layer 2 we can find the degree distribution $P(k)$ in layer 1, i.e. $P(k) = \sum_{q=m}^{\infty}P(k,q)$ obtaining the known result for a single layer,

$$P(k) = \frac{2m(1+m)}{k(k+1)(k+2)} \tag{S-39}$$

The function $\langle k(q)\rangle$ is given by

$$\langle k(q)\rangle = \frac{\sum_{k=m}^{\infty}kP(k,q)}{\sum_{k=m}^{\infty}P(k,q)} = \frac{m}{1+m}(q+2) \tag{S-40}$$

Similar expressions are obtained for $P(q)$ and $\langle q(k) \rangle$, by summing Eq. (S-38) over k.

(ii) Uniform attachment kernel – Let us now consider a uniform attachment kernel, in which every target node j is chosen with probability $\Pi^{[1]}_{i \to j} = \frac{1}{t}$ in layer 1 and with probability $\Pi^{[2]}_{i \to j} = \frac{1}{t}$ in layer 2, so that

$$A_{k,q} = m$$
$$B_{k,q} = m. \tag{S-41}$$

In this case the recursive Eqs. (S-34) read

$$P(m,q) = \left(\frac{m}{1+2m} \right)^{q-m} P(m,m),$$

$$P(k,q) = \sum_{r=m}^{q} \left(\frac{m}{1+2m} \right)^{q-r} \frac{m}{1+2m} P(k-1,r) \tag{S-42}$$

where $P(m,m)$ is again fixed by the normalization condition $\sum_{k=m}^{\infty} \sum_{q=m}^{\infty} P(k,q) = 1$. Using again the relation provided in Eq. (S-37) it is easy to prove recursively that $P(k,q)$ is given in this case by

$$P(k,q) = \frac{1}{m} \left(\frac{m}{1+2m} \right)^{k+q-2m+1} \frac{\Gamma(k+q-2m+1)}{\Gamma(k-m+1)\Gamma(q-m+1)} \tag{S-43}$$

Moreover the degree distribution $P(k) = \sum_{q=m}^{\infty} P(k,q)$ and $P(q) = \sum_{k=m}^{\infty} P(k,q)$ of a single network are given by

$$P(k) = \frac{1}{1+m} \left(\frac{m}{1+m} \right)^{k-m} \quad \text{and} \quad P(q) = \frac{1}{1+m} \left(\frac{m}{1+m} \right)^{q-m} \tag{S-44}$$

while the function $\langle k(q) \rangle = \sum_{k=m}^{\infty} kP(k,q)/P(q)$ is given by

$$\langle k(q) \rangle = \frac{(q+2)m}{1+m} \tag{S-45}$$

(iii) Semi-linear attachment kernel – Finally we analyze the case of semi-linear attachment with $c^{[1,1]} = c^{[2,2]} = 1$ and $c^{[1,2]} = c^{[2,1]} = 0$. We have:

$$A_{k,q} = \frac{k}{2}$$
$$B_{k,q} = m. \tag{S-46}$$

In this case the Eqs. (S-34) read as

$$P(m,q) = \left(\frac{2m}{2+3m} \right)^{q-m} P(m,m),$$

$$P(k,q) = \sum_{r=m}^{q} \left(\frac{2m}{2+k+2m} \right)^{q-r} \frac{k-1}{2+k+2m} P(k-1,r) \tag{S-47}$$

where $P(m,m)$ is fixed by the normalization condition $\sum_{q=m}^{\infty} \sum_{k=m}^{\infty} P(k,q) = 1$. It can be shown recursively that these equations have the following solution,

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

with the associated degree distributions $P(k) = \sum_{q=m}^{\infty} P(k,q)$ and $P(q) = \sum_{k=m}^{\infty} P(k,q)$ given by

$$P(k) = \frac{2m(1+m)}{k(k+1)(k+2)},$$

$$P(q) = \frac{1}{\Gamma(1+m)} \sum_{k=m}^{\infty} \sum_{n=0}^{k-m} \binom{k-m}{n} \left(\frac{2m}{2+2m+k-n} \right)^{q-m+1} (-1)^{k-m+n} = \left(\frac{m}{1+m} \right)^{q-m} \frac{1}{1+m}$$

Finally the function $\langle k(q) \rangle = \sum_{k=m}^{\infty} kP(k,q)/P(q)$ is given by

$$\langle k(q) \rangle = \left(\frac{2(m+1)}{1+2m} \right)^{q-m+1} m$$

<div align="right">(S-48)</div>

Role of β in the delayed arrival

In Fig. S-1 we show the time evolution of the maximum degree $k_M(t)$ on the first layer, for different values of β. Notice that $k_M(t) \sim (t/s)^{\delta}$. The effect of the exponent β tuning the width of the delay distribution is evident: the larger the value of β, the closer δ is to 0.5, the value observed in the case of synchronous arrival. Consequently, the rightmost part of the degree distribution is broader when β is close to 1 and becomes more similar to $P(k) \sim k^{-3}$ when β increases.

Finite size effects

It is interesting to investigate how the properties of the multiplexes generated using the model we propose depend on the number of nodes N. For instance, most of the mean-field predictions for the degree distributions and inter-layer degree correlations are valid in the limit of large N. However, as shown in Fig. S-2 and in Fig. S-3, the properties of the degree distributions and of inter-layer degree-degree correlations are similar to those predicted for large N even for relatively small multiplexes, e.g. with $N = 1000$.

Time complexity

The most efficient algorithm for the construction of a simplex networks based on preferential attachment takes advantage of random sampling with rejection and runs in $\mathcal{O}(Nm^2)$. However, in the case of a multiplex the procedure to sample a candidate neighbour j of a newly-arrived node i is a bit more complicated. Let us first consider a two-layer multiplex described by Eq. (S-2). When we sample the candidate neighbours of node i at layer 1 at time t, each node j should be sampled with a probability proportional to $ak_j^{[1]} + (1-a)k_j^{[2]}$. The simplest way to implement such sampling is to construct a vector $\mathcal{S}$ whose n-th entry is equal to $\sum_{j=1}^{n} ak_j^{[1]} + (1-a)k_j^{[2]}$ (the first element $\mathcal{S}[0]$ of the array is set equal to zero); then we sample a real number ζ in the interval $(0, \mathcal{S}[t]]$ and we choose the node j such that $\zeta \leq \mathcal{S}[j]$ and $\zeta > \mathcal{S}[j-1]$. The construction of the vector $\mathcal{S}$ at each time t requires $\mathcal{O}(t)$ operations while the sampling of a single node j can be efficiently implemented by binary search, requiring at most $\mathcal{O}(\log(t))$ operations per edge, so that the sampling of m edges requires at most $\mathcal{O}(m\log(t))$ steps. Thus, the total number of operations needed to sample a layer of a multiplex is:

$$\sum_{j=m_0+1}^{N} \mathcal{O}(t) + \mathcal{O}(m\log(t)) = \mathcal{O}(N^2) + \mathcal{O}(mN\log(N)) = \mathcal{O}(N^2 + mN\log(N))$$

<div align="right">(S-49)</div>

It is easy to verify that the construction of a M-layer multiplex requires a number of steps

$$\mathcal{O}(M(N^2 + mN\log(N))$$

<div align="right">(S-50)</div>

which, for m fixed, is dominated by $\mathcal{O}(MN^2)$. Therefore, the time complexity of this algorithm is linear in the number of layers and quadratic in the number of nodes. We notice that in principle it is possible to construct better algorithms to sample growing multiplexes by implementing a smart policy to update the array $\mathcal{S}$.

Randomly-chosen master layer

In the main text we made the simplifying assumption that each node arrives first on the master layer and then on the other layers, after a certain delay. We call this assumption "Equal master layer" (EML). In this Section we briefly comment on the case in which this assumption does not hold, i.e. when a node first arrives either on the first
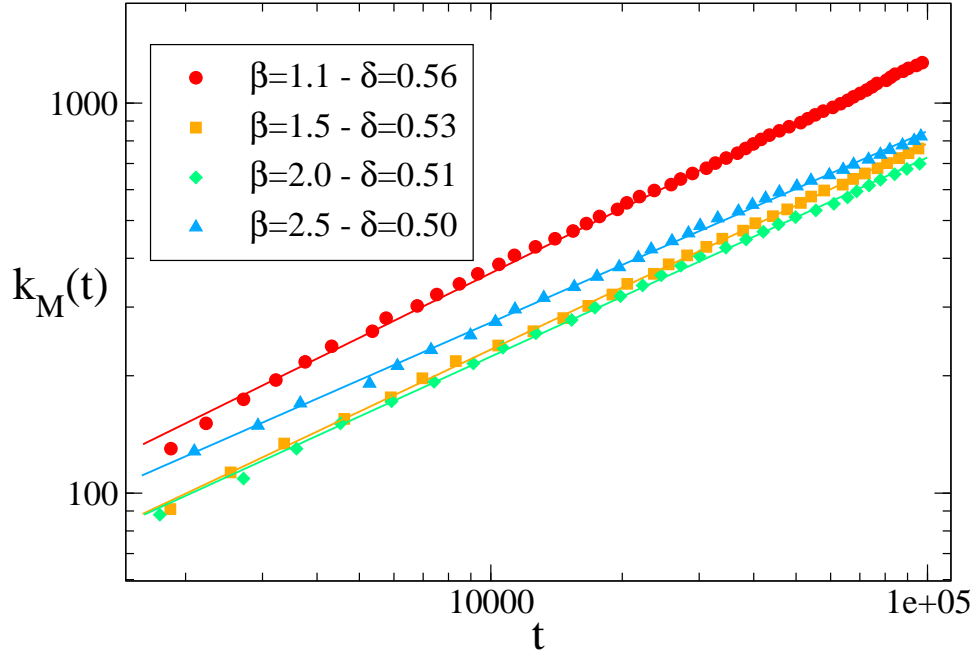
FIG. S-1. **Effect of β on the maximum degree of a layer.** When the exponent β of the delay distribution is close to 1, the maximum degree of a layer scales as $(t/s)^{\delta}$, with $\delta > 1/2$. Consequently, the rightmost side of the degree distribution is broader when β is close to 1 and converges to $P(k) \sim k^{-3}$ as β increases.
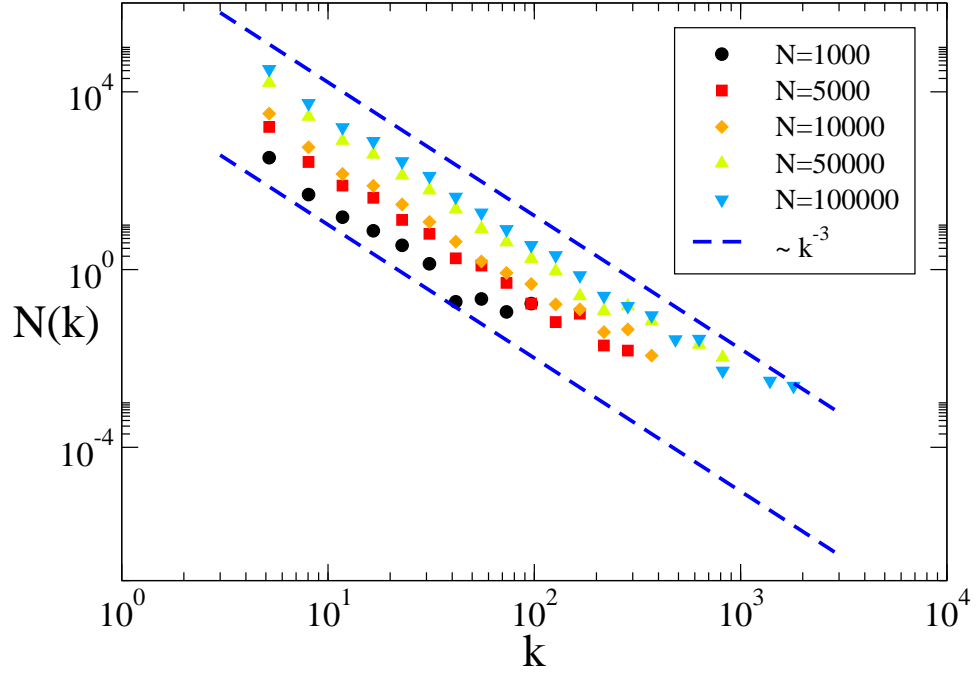

FIG. S-2. **Degree distribution on the first layer for different values of** N. Even for relatively small multiplexes, i.e. $N = 1000$, the exponent of the degree distribution is almost equal to $\gamma = 3$.

or on the second layer, and then arrives on the other layer after a power-law distributed delay. We call this case "Randomly-chosen master layer" (RML), to stress the fact that the master layer of each node is chosen at random among the M layers of the multiplex. In particular, we are interested in the case in which a newly arrived node
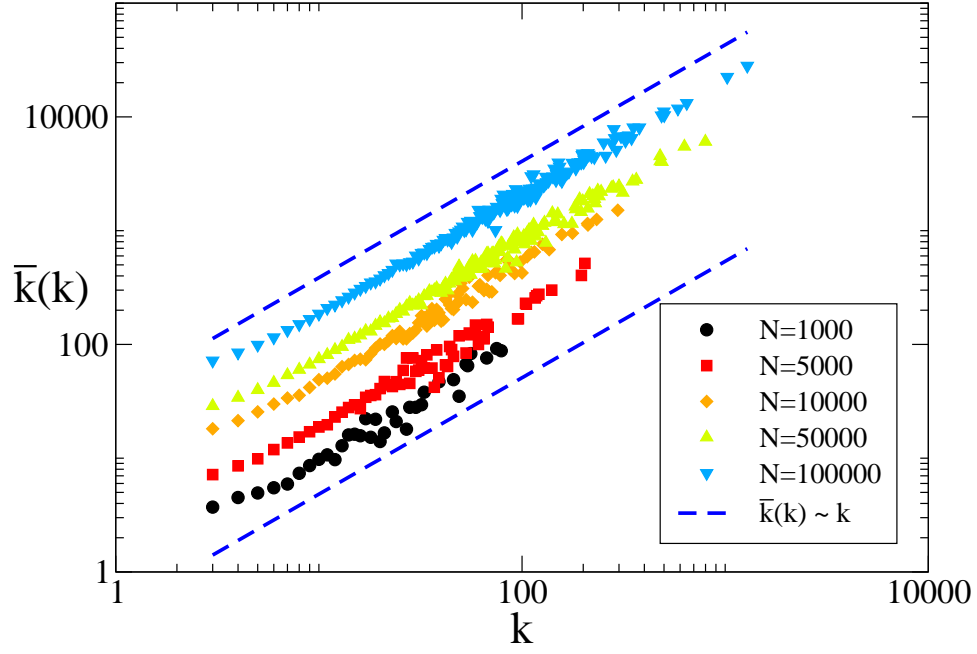
FIG. S-3. **Inter-layer degree correlation as a function of** N. The shape of the inter-layer degree-degree correlations is already evident even for small values of N. The curves have been vertically displaced to facilitate visual comparison.

selects one of the M layers of the multiples as its master layer with uniform probability $p = 1/M$. In Fig. S-4, S-5 and S-6 we report, respectively, the degree distributions, the temporal scaling of the degree of the largest hub and the distribution of shortest path lengths and node interdependence for RML with $M = 2$. The plots suggest that the random choice of the master layer produces a more balanced distribution of super hubs between the two layers, which has a relevant impact on the distribution of shortest path lengths and node interdependence (Fig. S-6). Conversely, the degree distributions and the temporal scaling of the degree of the largest hub are practically indistinguishable from those observed in EML (Fig. S-4 and S-5).
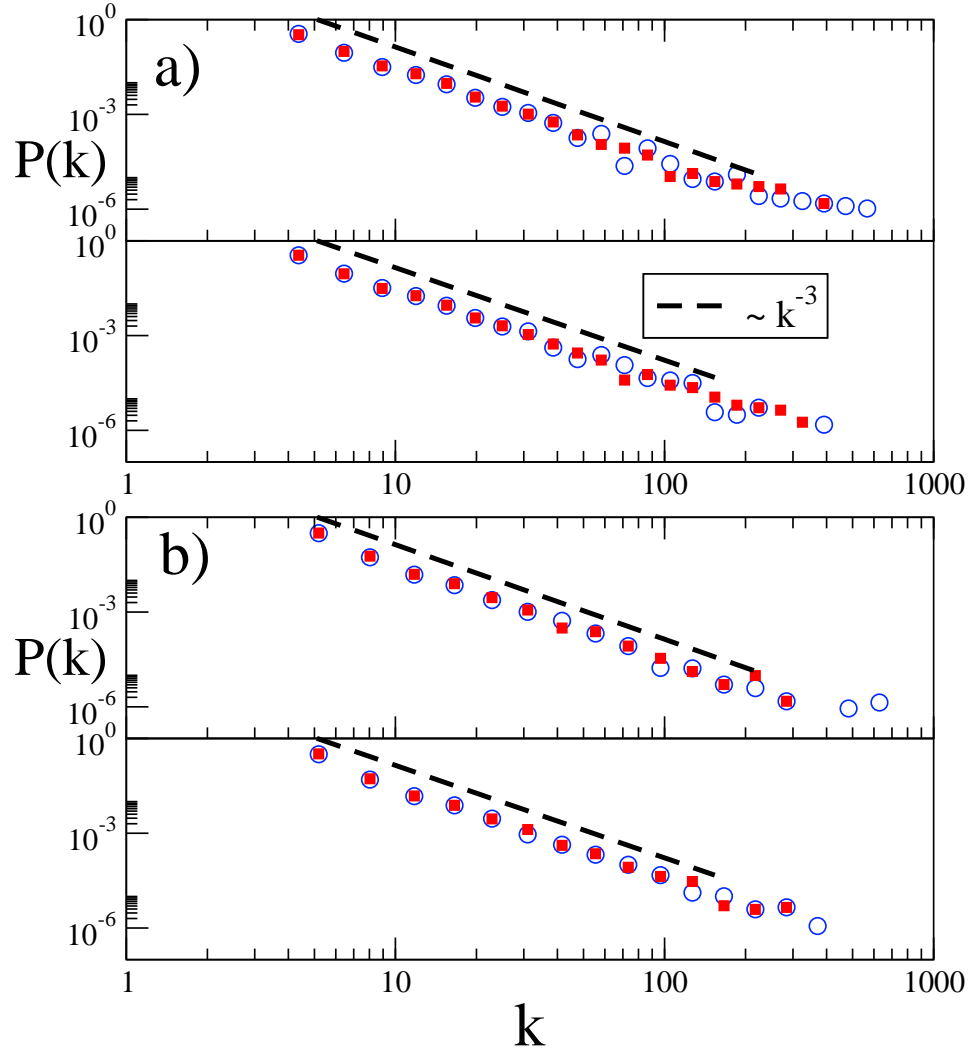
FIG. S-4. **Degree distribution of mixing master layers**. The degree distribution of the first layer for the EML (panel a) and the RML (panel b) for $\beta = 1.1$ (blue circles) and $\beta = 2.0$ (red squares). The random choice of the master layer for each node does not sensibly change the shape of the degree distribution.
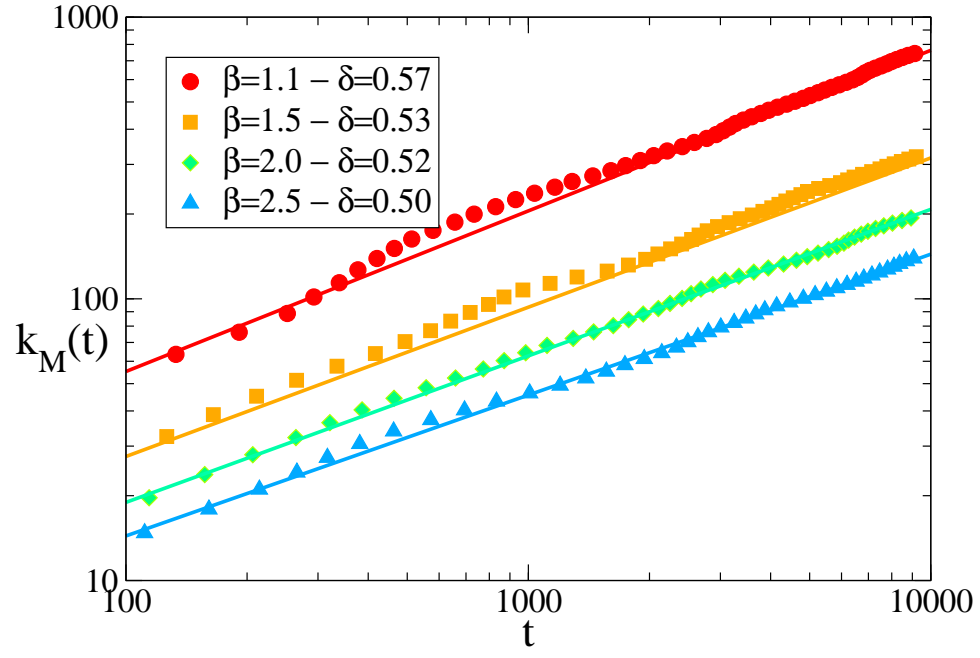
FIG. S-5. **Temporal scaling of degree for RML**. The degree k_M of the largest hub in the first layer scales as $(t/s)^\delta$, where δ depends on the exponent β of the delay distribution and approaches 0.5 as β increases. Notice that the values of δ are pretty similar to those observed in the case of EML, reportes in Fig. S-1.
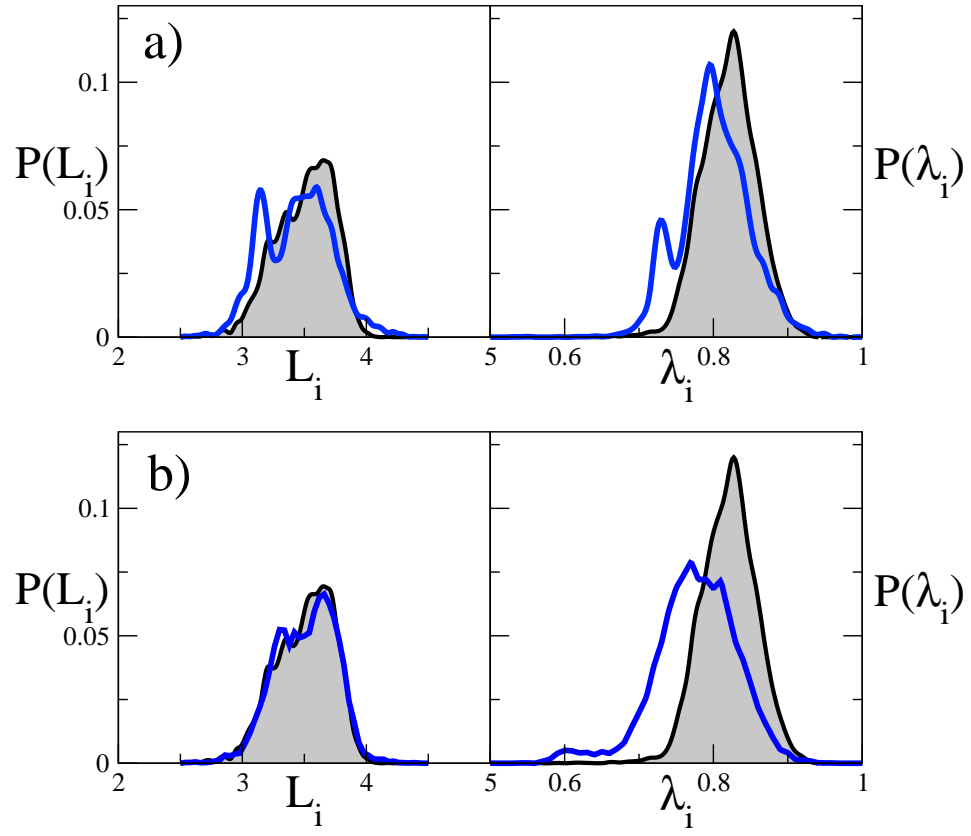
FIG. S-6. **Distribution of shortest path lengths and node interdependence for RML**. When the master layer of each node is chosen at random, the distribution of shortest path lengths (panel b, left) is more similar to that obtained for synchronous arrival (shaded curve) than that obtained in EML (panel a, left). Conversely, the distribution of node interdependence for RML (panel b, right) sensibly deviates from that obtained for synchronous arrival (shaded curve) and differs from that observed in EML (panel a, right). These results can be explained by a more balanced distribution of super-hubs among the two layers of the multiplex.