



PAPER

Structure of triadic relations in multiplex networks

OPEN ACCESS

RECEIVED
5 March 2015REVISED
25 May 2015ACCEPTED FOR PUBLICATION
9 June 2015PUBLISHED
31 July 2015

Content from this work
may be used under the
terms of the [Creative
Commons Attribution 3.0
licence](#).

Any further distribution of
this work must maintain
attribution to the
author(s) and the title of
the work, journal citation
and DOI.

Emanuele Cozzo¹, Mikko Kivelä², Manlio De Domenico³, Albert Solé-Ribalta³, Alex Arenas³, Sergio Gómez³,
Mason A Porter^{2,4} and Yamir Moreno^{1,5,6}¹ Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, E-50018 Zaragoza, Spain² Oxford Centre for Industrial and Applied Mathematics, Mathematical Institute, University of Oxford, Oxford, OX2 6GG, UK³ Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, E-43007 Tarragona, Spain⁴ CABDyN Complexity Centre, University of Oxford, Oxford, OX1 1HP, UK⁵ Department of Theoretical Physics, University of Zaragoza, E-50009 Zaragoza, Spain⁶ Complex Networks and Systems Lagrange Lab, Institute for Scientific Interchange, Turin, ItalyE-mail: yamir.moreno@gmail.com**Keywords:** multilayer networks, transitivity, complex systems**Abstract**

Recent advances in the study of networked systems have highlighted that our interconnected world is composed of networks that are coupled to each other through different 'layers' that each represent one of many possible subsystems or types of interactions. Nevertheless, it is traditional to aggregate multilayer networks into a single weighted network in order to take advantage of existing tools. This is admittedly convenient, but it is also extremely problematic, as important information can be lost as a result. It is therefore important to develop multilayer generalizations of network concepts. In this paper, we analyze triadic relations and generalize the idea of transitivity to multiplex networks. By focusing on triadic relations, which yield the simplest type of transitivity, we generalize the concept and computation of clustering coefficients to multiplex networks. We show how the layered structure of such networks introduces a new degree of freedom that has a fundamental effect on transitivity. We compute multiplex clustering coefficients for several real multiplex networks and illustrate why one must take great care when generalizing standard network concepts to multiplex networks. We also derive analytical expressions for our clustering coefficients for ensemble averages of networks in a family of random multiplex networks. Our analysis illustrates that social networks have a strong tendency to promote redundancy by closing triads at every layer and that they thereby have a different type of multiplex transitivity from transportation networks, which do not exhibit such a tendency. These insights are invisible if one only studies aggregated networks.

1. Introduction

The quantitative study of networks is fundamental for investigations of complex systems throughout the biological, social, information, engineering, and physical sciences [1–3]. The broad applicability of networks, and their success in providing insights into the structure and function of both natural and designed systems, has generated considerable excitement across myriad scientific disciplines. Numerous tools have been developed to study networks, and the realization that several common features arise in a diverse variety of networks has facilitated the development of theoretical tools to study them. For example, many networks constructed from empirical data have heavy-tailed degree distributions, satisfy the small-world property, and/or possess modular structures. Such structural features can have important implications for information dissemination, robustness against component failure, and more.

Traditional studies of networks generally assume that nodes are adjacent to each other by a single type of static edge that encapsulates all connections between them. This assumption is almost always a gross oversimplification, and it can lead to misleading results and even the fundamental inability to address certain problems. Most real systems have multilayer structures [4, 5], as there are almost always multiple types of ties or interactions that can occur between nodes, and it is crucial to take them into account. For example,

transportation systems include multiple modes of travel, biological systems include multiple signaling channels that operate in parallel, and social networks include multiple types of relationships and multiple modes of communication. We will represent such systems using the formalism of *multiplex networks*, which allow one to incorporate multiple types of edges between nodes.

The notion of multiplexity was introduced years ago in fields such as engineering [6, 7] and sociology [1, 8–10], but the discussions included few analytical tools to accompany them. This situation arose for a simple reason: although many aspects of single-layer networks are well understood, it is challenging to properly generalize even the simplest concepts to multiplex networks. Theoretical developments on multilayer networks (including both multiplex networks and interconnected networks) have gained steam only in the last few years [11–20], and even basic notions like centrality and diffusion have barely been studied in multiplex settings [21–30]. New degrees of freedom arise from the multilayer structure of multiplex networks, and this brings both new challenges [4, 31] and new phenomena. The new phenomena include multiplexity-induced correlations [19], new types of dynamical feedbacks [26], and ‘costs’ of inter-layer connections [32]. For reviews about networks with multiple layers, see [4, 5].

In the present article, we focus on one of the most important structural properties of networks: triadic relations, which are used to describe the simplest and most fundamental type of transitivity in networks [1, 3, 33–35]. We develop multiplex generalizations of clustering coefficients, which can be done in myriad ways, and (as we will illustrate) the most appropriate generalization depends on the application under study. Such considerations are crucial when developing multiplex generalizations of any single-layer (i.e., ‘monoplex’) network diagnostic. There have been several attempts to define multiplex clustering coefficients [36–40], but there are significant shortcomings in these definitions. For example, some of them do not reduce to the standard single-layer clustering coefficient or are not properly normalized (see appendix B).

The fact that existing definitions of multiplex clustering coefficients are mostly ad hoc makes them difficult to interpret. In our definitions, we start from the basic concepts of walks and cycles to obtain a transparent and general definition of transitivity. This approach also guarantees that our clustering coefficients are always properly normalized. It reduces to a weighted clustering coefficient [41] of an aggregated network for particular values of the parameters; this allows comparison with existing single-layer diagnostics. We also address two additional, very important issues: (1) multiplex networks have many types of connections, and our multiplex clustering coefficients are (by construction) decomposable, so that the contribution of each type of connection is explicit; (2) because our notion of multiplex clustering coefficients builds on walks and cycles, we do not require every node to be present in all layers, which removes a major (and very unrealistic) simplification that is used in existing definitions.

Using the example of clustering coefficients, we illustrate how the new degrees of freedom that result from the existence of multiple layers in a multiplex network yield rich new phenomena and subtle differences in how one should define key network diagnostics. As an illustration of such phenomena, we derive analytical expressions for the expected values of clustering coefficients on multiplex networks in which each layer is an independent Erdős–Rényi (ER) graph. We find that the clustering coefficients depend on the intra-layer densities in a nontrivial way if the probabilities for an edge to exist are heterogeneous across the layers. We thereby demonstrate for multiplex networks that it is insufficient to generalize existing diagnostics in a naïve manner and that one must instead construct their generalizations from first principles (e.g., as walks and cycles in this case).

2. Methods

2.1. Mathematical representation

We represent a multiplex network using a finite sequence of graphs $\{G^\alpha\}$, with $G^\alpha = (V^\alpha, E^\alpha)$, where $\alpha \in L$ is the set of layers. Without loss of generality, we let $L = \{1, \dots, b\}$ and $V^\alpha \subseteq \{1, \dots, n\}$. For simplicity, we examine unweighted and undirected multiplex networks. We define the intra-layer supra-graph as $G_A = (V, E_A)$, where the set of nodes is $V = \bigcup_\alpha \{(u, \alpha) : u \in V^\alpha\}$ and the set of edges is $E_A = \bigcup_\alpha \{[(u, \alpha), (v, \alpha)] : (u, v) \in E^\alpha\}$. We also define the coupling supra-graph $G_C = (V, E_C)$ using the same sets of nodes and the edge set $E_C = \bigcup_{\alpha, \kappa} \{[(u, \alpha), (u, \kappa)] : u \in V^\alpha, u \in V^\kappa, \alpha \neq \kappa\}$ and its associated adjacency matrix C . If $[(u, \alpha), (u, \kappa)] \in E_C$, we say that (u, α) and (u, κ) are ‘interconnected’. The nonzero entries of the matrix $C = C^T$ indicates the connections between corresponding nodes (i.e., between the same entity) on different layers. We say that a multiplex network is ‘node-aligned’ [4] if all layers share the same set of nodes (i.e., if $V^\alpha = V^\kappa$ for all α and κ). The supra-graph is $\bar{G} = (V, \bar{E})$, where $\bar{E} = E_A \cup E_C$. The corresponding adjacency matrix is the supra-adjacency matrix \bar{A} .

Supra-adjacency matrices satisfy $\bar{A} = \mathcal{A} + C$ and $\mathcal{A} = \bigoplus_\alpha \mathbf{A}^{(\alpha)}$, where $\mathbf{A}^{(\alpha)}$ is the adjacency matrix of layer α (i.e., the adjacency matrix associated to G^α) and \bigoplus denotes the direct sum of the matrices. We consider undirected networks, so $\mathcal{A} = \mathcal{A}^T$. For clarity, we denote nodes in a given layer and in monoplex networks using the symbols u, v, w ; and we denote indices in a supra-adjacency matrix using the symbols i, j, h . We also define

$l(u) = \{(u, \alpha) \in V \mid \alpha \in L\}$ to be the set of supra-adjacency matrix indices that correspond to node u , and we refer to the nodes (u, α) of a supra-graph as a node-layer pair. An entity u corresponds to a ‘physical node.’

The local clustering coefficient C_u of node u in an unweighted monoplex network is the number of triangles (i.e., triads) that include node u divided by the number of connected triples (i.e., either 2-stars or triangles) with node u in the center [3, 34]. The local clustering coefficient is a measure of transitivity [33], and it can be interpreted as the density of a focal node’s neighborhood. For our purposes, it is convenient to define the local clustering coefficient C_u as the number of three-cycles t_u that start and end at the focal node u divided by the number of three-cycles d_u such that the second step of the cycle occurs in a complete graph (i.e., assuming that the neighborhood of the focal node is as dense as possible)⁷. In mathematical terms, $t_u = (\mathbf{A}^3)_{uu}$ and $d_u = (\mathbf{AFA})_{uu}$, where \mathbf{A} is the adjacency matrix of the graph and \mathbf{F} is the adjacency matrix of a complete graph with no self-edges. (In other words, $\mathbf{F} = \mathbf{J} - \mathbf{I}$, where \mathbf{J} is a complete square matrix of 1s and \mathbf{I} is the identity matrix.)

The *local clustering coefficient* for node u is thus given by the formula $C_u = t_u/d_u$. This is equivalent to the usual definition of the local clustering coefficient: $C_u = t_u/(k_u(k_u - 1))$, where $k_u \geq 2$ is the degree of node u . (The local clustering coefficient is often set to 0 for nodes of degree 0 and 1, and another option is to state that it is not defined in such cases.) One can calculate a single global clustering coefficient for a monoplex network either by averaging C_u over all nodes or by computing $C = \frac{\sum_u t_u}{\sum_u d_u}$. Henceforth, we will use the term *global clustering coefficient* for the latter quantity.

2.2. Triads on multiplex networks

In addition to three-cycles (i.e., triads) that occur within a single layer, multiplex networks also contain cycles that incorporate more than one layer but still have three intra-layer steps. Such cycles are important for the analysis of transitivity in multiplex networks. In social networks, for example, transitivity involves social ties across multiple social environments [1, 42]. In transportation networks, there typically exist several means of transport to return to one’s starting location, and different combinations of transportation modes are important in different cities [43]. For dynamical processes on multiplex networks, it is important to consider three-cycles that traverse different numbers of layers, so one needs to take them into account when defining a multiplex clustering coefficient. We define a *supra-walk* as a walk on a multiplex network in which, either before or after each intra-layer step, a walk can either continue on the same layer or change to an adjacent layer. We represent this choice using the following matrix:

$$\hat{\mathbf{C}} = \beta \mathbf{I} + \gamma \mathbf{C}, \quad (1)$$

where \mathbf{I} is the $|V| \times |V|$ identity matrix, $|V|$ is the number of node-layer pairs, the parameter β is a weight that accounts for the walk staying in the current layer, and γ is a weight that accounts for the walk stepping to another layer. In a supra-walk, a *supra-step* consists either of only a single intra-layer step or of a step that includes both an intra-layer step and an inter-layer step, in which one changes from one layer to another (either before or after the intra-layer step). In the latter type of supra-step, note that we are disallowing two consecutive inter-layer steps. The number of three-cycles for node i is then

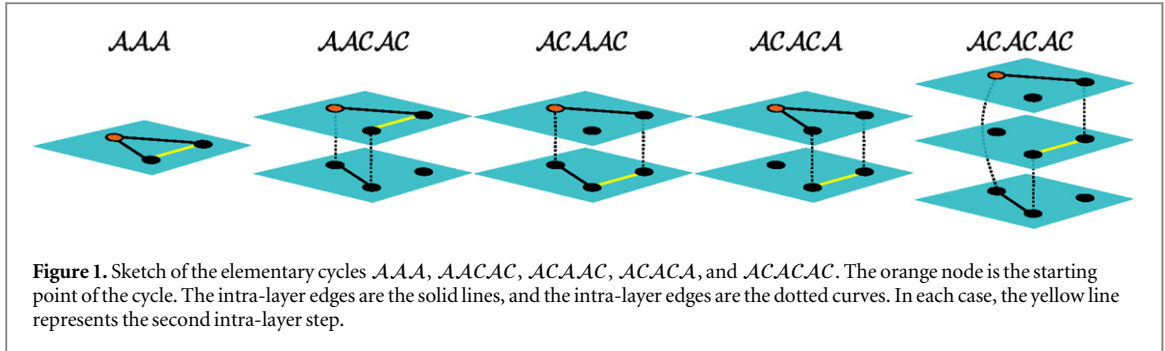
$$t_{M,i} = \left[(\mathcal{A}\hat{\mathbf{C}})^3 + (\hat{\mathbf{C}}\mathcal{A})^3 \right]_{ii}, \quad (2)$$

where the first term corresponds to cycles in which the inter-layer step is taken after an intra-layer one and the second term corresponds to cycles in which the inter-layer step is taken before an intra-layer one. The subscript M refers to the particular way that we define a supra-walk in a multiplex network through the supra-matrices $\mathcal{A}\hat{\mathbf{C}}$ and $\hat{\mathbf{C}}\mathcal{A}$. However, one can also define other types of supra-walks (see appendices C and D), and we will use different subscripts when we refer to them. We can simplify equation (2) by exploiting the fact that both \mathcal{A} and $\hat{\mathbf{C}}$ are symmetric. This yields

$$t_{M,i} = 2 \left[(\mathcal{A}\hat{\mathbf{C}})^3 \right]_{ii}. \quad (3)$$

It is useful to decompose multiplex clustering coefficients that are defined in terms of multilayer cycles into so-called *elementary cycles* by expanding equation (3) and writing it in terms of the matrices \mathcal{A} and \mathbf{C} . That is, we write $t_{M,i} = \sum_{\mathcal{E} \in \mathcal{E}} w_{\mathcal{E}}(\mathcal{E})_{ii}$, where \mathcal{E} denotes the set of elementary cycles and $w_{\mathcal{E}}$ are weights of different elementary cycles. We can use symmetries in our definition of cycles and thereby express all of the elementary cycles in a standard form with terms from the set $\mathcal{E} = \{AAA, AACAC, ACAAC, ACACA, ACACAC\}$. See figure 1 for an illustration of elementary cycles and appendix E for details on deriving the elementary cycles.

⁷ Note that we use the term ‘cycle’ to refer to a walk that starts and ends at the same physical node u . As we will discuss later, in a multiplex network, it is permissible (and relevant) to return to the same node via a different layer from the one that was used originally to leave the node.



Note that some of the alternative definitions of a three-cycle—which we discuss in appendix C—lead to more elementary cycles than the ones that we just enumerated.

To define multiplex clustering coefficients, we need both the number $t_{*,i}$ of cycles and a normalization $d_{*,i}$. The symbol $*$ stands for any type of cycle: the three-cycle that we define in the main text, an elementary cycle, or the alternatives definition of three-cycles that we give in appendix C. Choosing a particular definition coincides to a given way to calculate the associated expression for $t_{*,i}$. To determine the normalization, it is natural to follow the same procedure as with monoplex clustering coefficients and use a complete multiplex network $\mathcal{F} = \bigoplus_{\alpha} \mathbf{F}^{(\alpha)}$, where $\mathbf{F}^{(\alpha)} = \mathbf{J}^{(\alpha)} - \mathbf{I}^{(\alpha)}$ is the adjacency matrix for a complete graph on layer α . We can then proceed from any definition of $t_{*,i}$ to $d_{*,i}$ by replacing the second intra-layer step with a step in the complete multiplex network. For example, we obtain $d_{M,i} = 2(\mathcal{A}\hat{\mathcal{C}}\mathcal{F}\hat{\mathcal{C}}\mathcal{A})_{ii}$ for $t_{M,i} = 2[(\mathcal{A}\hat{\mathcal{C}})^3]_{ii}$. Similarly, one can use any other definition of a cycle (e.g., any of the elementary cycles or the cycles that we discuss in appendix C) as a starting point for defining a multiplex clustering coefficient.

The above formulation allows us to define local and global clustering coefficients for multiplex networks analogously to their definition in monoplex networks. We can calculate a natural multiplex analog to the usual monoplex local clustering coefficient for any node i of the supra-graph. Additionally, in a multiplex network, a node u of an intra-layer network allows an intermediate description for clustering that lies between local and the global clustering coefficients. We define

$$c_{*,i} = \frac{t_{*,i}}{d_{*,i}}, \quad (4)$$

$$C_{*,u} = \frac{\sum_{i \in l(u)} t_{*,i}}{\sum_{i \in l(u)} d_{*,i}}, \quad (5)$$

$$C_* = \frac{\sum_i t_{*,i}}{\sum_i d_{*,i}}, \quad (6)$$

where $l(u)$ is as defined before. Note that we refer to clustering coefficients defined either by equation (4) or by equation (5) as local clustering coefficients. In equations (4)–(6), and in our subsequent formulas for clustering coefficients, we are of course requiring denominators to be nonzero (as in the monoplex case). In situations in which a denominator vanishes, we set the value of the associated clustering coefficient to 0.

We can decompose the expression in equation (6) in terms of the contributions from cycles that traverse exactly one, two, and three layers (where $m = 1, 2, 3$ indicates the number of layers) to give

$$t_{*,i} = t_{*,1,i}\beta^3 + t_{*,2,i}\beta\gamma^2 + t_{*,3,i}\gamma^3, \quad (7)$$

$$d_{*,i} = d_{*,1,i}\beta^3 + d_{*,2,i}\beta\gamma^2 + d_{*,3,i}\gamma^3, \quad (8)$$

$$C_*^{(m)} = \frac{\sum_i t_{*,m,i}}{\sum_i d_{*,m,i}}. \quad (9)$$

We can similarly decompose equations (4) and (5). Using the decomposition in equation (7) yields an alternative way to average over contributions from the three types of cycles:

$$C_*(\omega_1, \omega_2, \omega_3) = \sum_m^3 \omega_m C_*^{(m)}, \quad (10)$$

where $\vec{\omega}$ is a vector that gives the relative weights of the different contributions. We use the term *layer-decomposed clustering coefficients* for $C_*^{(1)}$, $C_*^{(2)}$, and $C_*^{(3)}$. There are also analogs of equation (10) for the clustering coefficients defined in equations (4) and (5). Each of the clustering coefficients in equations (4)–(6)

depends on the values of the parameters β and γ , but the dependence vanishes if $\beta = \gamma$. Unless we explicitly indicate otherwise, we assume in our calculations that $\beta = \gamma$.

2.3. Clustering coefficients for aggregated networks

A common way to study multiplex networks is to aggregate layers to obtain either multi-graphs or weighted networks, where the number of edges or the weight of an edge is the number of different types of edges between a pair of nodes [4]. One can then use any of the numerous ways to define clustering coefficients for weighted monoplex networks [44, 45] to calculate clustering coefficients for the aggregated network.

One of the weighted clustering coefficients is a special case of our multiplex clustering coefficient (for others, see appendix A). References [41, 46, 47] calculated a weighted clustering coefficient as

$$C_{Z,u} = \frac{\sum_{vw} W_{uv} W_{vw} W_{wu}}{w_{\max} \sum_{v \neq w} W_{uv} W_{uw}} = \frac{(\mathbf{W}^3)_{uu}}{(\mathbf{W}(w_{\max} \mathbf{F})\mathbf{W})_{uu}}, \quad (11)$$

where $W_{uv} = \sum_{i \in I(u), j \in I(v)} \mathcal{A}_{ij}$ is an element of the weighted adjacency matrix \mathbf{W} . The elements of \mathbf{W} are the weights of the edges, the quantity $w_{\max} = \max_{u,v} W_{uv}$ is the maximum weight in \mathbf{W} , and \mathbf{F} is the adjacency matrix of the complete unweighted graph. We can define the global version C_Z of $C_{Z,u}$ by summing over all of the nodes in the numerator and the denominator of equation (11) (analogously to equation (6)).

For node-aligned multiplex networks, the clustering coefficients $C_{Z,u}$ and C_Z are related to our multiplex clustering coefficients $C_{M,u}$ and C_M . Letting $\beta = \gamma = 1$ and summing over all layers yields $\sum_{i \in I(u)} ((\mathcal{A}\hat{\mathcal{C}})^3)_{ii} = (\mathbf{W}^3)_{uu}$. That is, in this special case, the weighted clustering coefficients $C_{Z,u}$ and C_Z are equivalent to the corresponding multiplex clustering coefficients $C_{M,u}$ and C_M . In particular, $C_{M,u}(\beta = \gamma) = \frac{w_{\max}}{b} C_{Z,u}$ and $C_M(\beta = \gamma) = \frac{w_{\max}}{b} C_Z$. We need the term w_{\max}/b to match the normalizations because aggregation removes the information about the number of layers b , so the normalization must be based on the maximum weight instead of the number of layers. That is, a step in the complete weighted network is described by using $w_{\max} \mathbf{F}$ in equation (11) instead of using $b\mathbf{F}$.

Note that this relationship between our multiplex clustering coefficient and the weighted clustering coefficient in equation (11) is only true for node-aligned multiplex networks. If some nodes are not shared among all layers, then the normalization of our multiplex clustering coefficient depends on how many nodes are present in the local neighborhood of the focal node. This contrasts with the ‘global’ normalization by w_{\max} used by the weighted clustering coefficient in equation (11).

2.4. Clustering coefficients in Erdős–Rényi networks

Almost all real networks contain some amount of transitivity, and it is often desirable to know if a network contains more transitivity than would be expected by chance. In order to examine this question, one typically compares clustering-coefficient values of a network to what would be expected from some random network that acts as a null model. The simplest random network to use is an Erdős–Rényi (ER) network. In this section, we give formulas for expected clustering coefficients in node-aligned multiplex networks in which each intra-layer network is an ER network that is created independently of other intra-layer networks and the inter-layer connections are created as described in section 2.1.

The expected value of the local clustering coefficient in an unweighted monoplex ER network is equal to the probability p of an edge to exist. That is, the density of the neighborhood of a node, measured by the local clustering coefficient, has the same expectation as the density of the entire network for an ensemble of ER networks. In multiplex networks with ER intra-layer graphs with connection probabilities p_α , the same result holds only when all of the layers are statistically identical (i.e., $p_\alpha = p$ for all α). Note that this is true even if the network is not node-aligned. However, heterogeneity among layers complicates the behavior of clustering coefficients. If the layers have different connection probabilities, then the expected value of the mean clustering coefficient is a nontrivial function of the connection probabilities. In particular, it is not always equal to the mean of the connection probabilities. For example, the formulas for the expected global layer-decomposed clustering coefficients are

$$\langle C_M^{(1)} \rangle = \frac{\sum_\alpha p_\alpha^3}{\sum_\alpha p_\alpha^2} \equiv \frac{\overline{p^3}}{p^2}, \quad (12)$$

$$\langle C_M^{(2)} \rangle = \frac{3 \sum_{\alpha \neq \kappa} p_\alpha p_\kappa^2}{(b-1) \sum_\alpha p_\alpha^2 + 2 \sum_{\alpha \neq \kappa} p_\alpha p_\kappa}, \quad (13)$$

$$\langle C_M^{(3)} \rangle = \frac{\sum_{\alpha \neq \kappa, \kappa \neq \mu, \mu \neq \alpha} p_\alpha p_\kappa p_\mu}{(b-2) \sum_{\alpha \neq \kappa} p_\alpha p_\kappa}. \quad (14)$$

Table 1. Clustering coefficients C_M , $C_M^{(1)}$, $C_M^{(2)}$, and $C_M^{(3)}$ that correspond, respectively, to the global, one-layer, two-layer, and three-layer clustering coefficients for various multiplex networks. ‘Tailor Shop’: Kapferer tailor-shop network ($n = 39$, $b = 4$) [48]. ‘Management’: Krackhardt office cognitive social structure ($n = 21$, $b = 21$) [49]. ‘Families’: Padgett Florentine families social network ($n = 16$, $b = 2$) [50]. ‘Bank’: Roethlisberger and Dickson bank wiring-room social network ($n = 14$, $b = 6$) [51]. ‘Tube’: the London Underground (i.e., ‘the Tube’) transportation network ($n = 314$, $b = 14$) [52]. ‘Airline’: network of flights between cities, in which each layer corresponds to a single airline ($n = 3108$, $b = 530$) [53]. The rows labeled ‘orig.’ give the clustering coefficients for the original networks, and the rows labeled ‘ER’ give the expected value and the standard deviation of the clustering coefficient in an ER random network with exactly as many edges in each layer as in the original network. For the original values, we perform a two-tailed Z-test to examine whether the observed clustering coefficients could have been produced by the ER networks. We designate the p -values as follows: *: $p < 0.05$, **: $p < 0.01$ for Bonferroni-corrected tests with 24 hypothesis; ’: $p < 0.05$, ’’: $p < 0.01$ for uncorrected tests. We do not use any symbols for values that are not significant. We symmetrize directed networks by considering two nodes to be adjacent if there is at least one edge between them. The social networks in this table are node-aligned multiplex graphs, but the transportation networks are not node-aligned. We report values that are means over different numbers of realizations: 1.5×10^5 for Tailor Shop, 1.5×10^3 for Airline, 1.5×10^4 for Management, 1.5×10^5 for Families, 1.5×10^4 for Tube, and 1.5×10^5 for Bank.

		Tailor Shop	Management	Families	Bank	Tube	Airline
C_M	orig.	0.319**	0.206**	0.223’	0.293**	0.056	0.101**
	ER	0.186 ± 0.003	0.124 ± 0.001	0.138 ± 0.035	0.195 ± 0.009	0.053 ± 0.011	0.038 ± 0.000
$C_M^{(1)}$	orig.	0.406**	0.436**	0.289’	0.537**	0.013’’	0.100**
	ER	0.244 ± 0.010	0.196 ± 0.015	0.135 ± 0.066	0.227 ± 0.038	0.053 ± 0.013	0.064 ± 0.001
$C_M^{(2)}$	orig.	0.327**	0.273**	0.198	0.349**	0.043*	0.150**
	ER	0.191 ± 0.004	0.147 ± 0.002	0.138 ± 0.040	0.203 ± 0.011	0.053 ± 0.020	0.041 ± 0.000
$C_M^{(3)}$	orig.	0.288**	0.192**	—	0.227**	0.314**	0.086**
	ER	0.165 ± 0.004	0.120 ± 0.001	—	0.186 ± 0.010	0.051 ± 0.043	0.037 ± 0.000

See appendix G for analogous formulas for the local multiplex clustering coefficients and for the numerical validation of our theoretical results.

3. Results and discussions

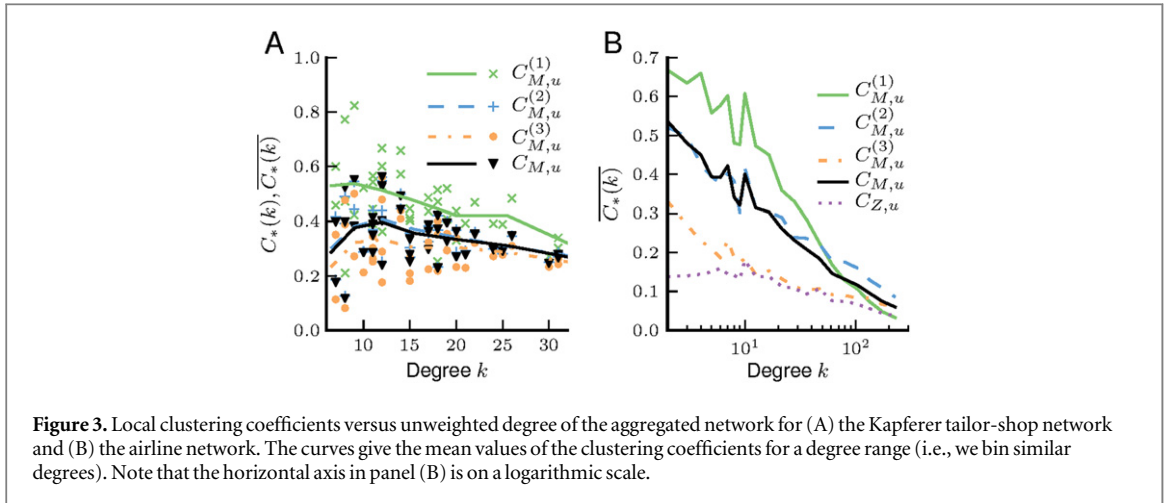
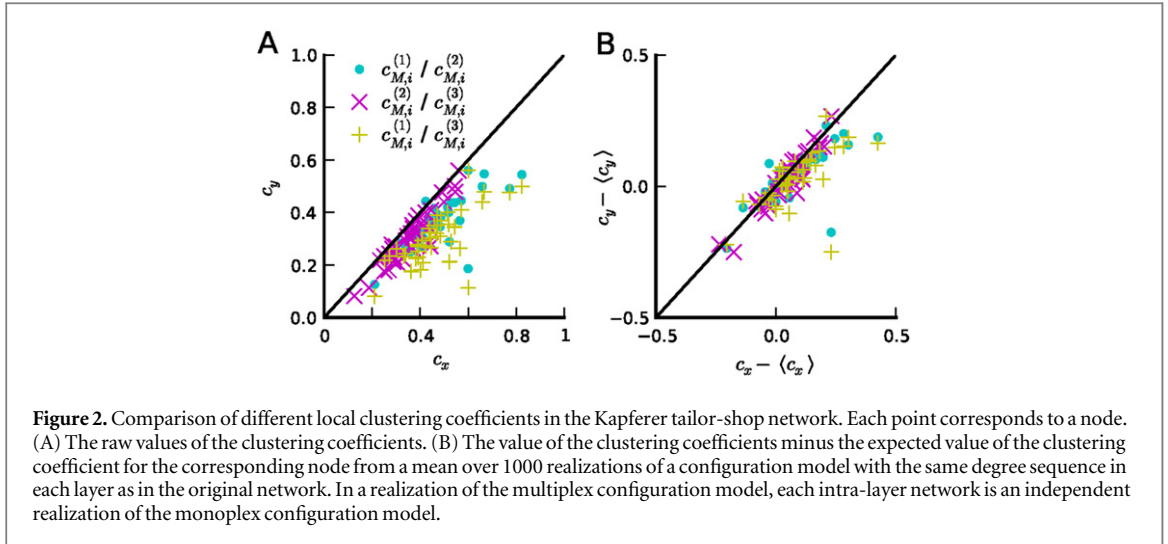
We investigate transitivity in empirical multiplex networks by calculating clustering coefficients. In table 1, we give the values of layer-decomposed global clustering coefficients for multiplex networks (four social networks and two transportation networks) constructed from real data. Note that the two transportation networks have different numbers of nodes in different layers (i.e., they are not node-aligned [4]). To help give context to the values, the table also includes the clustering-coefficient values that we obtain for ER networks with matching edge densities in each layer. See appendix H for a similar table that uses an alternative null model in which we shuffle the inter-layer connections.

As we will now discuss, multiplex clustering coefficients give insights that are impossible to infer by calculating weighted clustering coefficients for aggregated networks or by calculating clustering coefficients separately for each layer of a multiplex network.

For each social network in table 1, note that $C_M < C_M^{(1)}$ and $C_M^{(1)} > C_M^{(2)} > C_M^{(3)}$. Consequently, the primary contribution to the triadic structure of these multiplex networks arises from three-cycles that stay within a given layer. To check that the ordering of the different clustering coefficients is not an artifact of the heterogeneity of densities of the different layers, we also calculate the expected values of the clustering coefficients in ER networks with identical edge densities to the data. We observe that all clustering coefficients exhibit larger inter-layer transivities than would be expected in corresponding ER networks with identical edge densities, although the same ordering relationship (i.e. $C_M^{(1)} > C_M^{(2)} > C_M^{(3)}$) holds. From our results in table 1, it seems that triadic-closure mechanisms in social networks cannot be considered purely at the aggregated network level; these mechanisms appear to be more effective inside of layers than between layers. For example, if there is a connection between individuals u and v and also a connection between v and w in the same layer, then it is more likely that u and w ‘meet’ in the same layer than in some other layer.

The transportation networks that we examine exhibit the opposite pattern from the social networks. For example, for the London Underground (‘Tube’) network, in which each layer corresponds to a line, we observe that $C_M^{(3)} > C_M^{(2)} > C_M^{(1)}$. This reflects the fact that single lines in the Tube are designed to avoid redundant connections. A single-layer triangle would require a line to make a loop among three stations. Two-layer triangles, which are a bit more frequent than single-layer ones, entail that two lines run in almost parallel directions and that one line jumps over a single station. For three-layer triangles, the geographical constraints do not matter because one can construct a triangle with three straight lines.

We also analyze the local triadic closure of the Kapferer tailor-shop social network by examining the local clustering-coefficient values. In figure 2(a), we show a comparison of the layer-decomposed local clustering



coefficients (also see figure 6(a) of [40]). Observe that the condition $c_{M,i}^{(1)} > c_{M,i}^{(2)} > c_{M,i}^{(3)}$ holds for most of the nodes. In figure 2(b), we subtract the expected values of the clustering coefficients of nodes in a network generated with the configuration model⁸ from the corresponding clustering-coefficient values observed in the data to discern whether we should also expect to observe the relative order of the local clustering coefficients in an associated random network (with the same layer densities and degree sequences as the data). Similar to our results for global clustering coefficients, we see that taking a null model into account lessens—but does not remove—the difference between the coefficients that count different numbers of layers.

We investigate the dependence of local triadic structure on degree for one social network and one transportation network. In figure 3(a), we show how the different multiplex clustering coefficients depend on the unweighted degrees of the nodes in the aggregated network for the Kapferer tailor shop. Note that the relative ordering of the values of the mean clustering coefficient does not depend on degree. In figure 3(b), we illustrate that the aggregated network for the airline transportation network exhibits a non-constant difference between the curves of $C_{M,u}$ and the weighted clustering coefficient $C_{Z,u}$. Using a global normalization (see the discussion in section 2.3) reduces the clustering-coefficient values for the small airports much more than it does for the large airports. This, in turn, introduces a bias.

The airline network is organized differently from the London Tube network. When comparing these networks, note that each layer in the former encompasses flights from a single airline. For the airline network (see figure 3(b)), we observe that the two-layer local clustering coefficient is larger than the single-layer one for hubs (i.e., high-degree nodes), but it is smaller for small airports (i.e., low-degree nodes). However, the global clustering coefficient counts the total number of three-cycles and connected triplets, and it thus gives more weight to high-degree nodes than to low-degree nodes. We thus find that the global clustering coefficients for the

⁸ We use the configuration model instead of an ER network as a null model because the local clustering-coefficient values are typically correlated with node degree in monoplex networks [3], and an ER-network null model does not preserve degree sequence.

airline network satisfy $C_M^{(2)} > C_M^{(1)} > C_M^{(3)}$. The intra-airline clustering coefficients have small values, presumably because it is not in the interest of an airline to introduce new flights between two airports that can already be reached by two flights via the same airline through some major airport. The two-layer cycles correspond to cases in which an airline has a connection from an airport to two other airports and a second airline has a direct connection between those latter two airports. Completing a three-layer cycle requires using three distinct airlines, and this type of congregation of airlines to the same area is not frequent in the data. Three-layer cycles are more common than single-layer cycles only for a few of the largest airports.

4. Conclusions

We derived measurements of transitivity for multiplex networks by developing multiplex generalizations of triadic relationships and clustering coefficients. By using examples from empirical data in diverse settings, we showed that different notions of multiplex transitivity are important in different situations. For example, the balance between intra-layer versus inter-layer clustering is different in social networks versus transportation networks (and even in different types of networks within each category, as we illustrated explicitly for transportation networks), reflecting the fact that multilayer transitivity can arise from different mechanisms. Such differences are rooted in the new degrees of freedom that arise from inter-layer connections and are invisible to calculations of clustering coefficients on single-layer networks obtained via aggregation. In other words, transitivity is inherently a multilayer phenomenon: all of these diverse flavors of transitivity reduce to the same description when one throws away the multilayer information. Generalizing clustering coefficients for multiplex networks makes it possible to explore such phenomena and to gain deeper insights into different types of transitivity in networks. The existence of multiple types of transitivity also has important implications for multiplex network motifs and multiplex community structure. In particular, our work on multiplex clustering coefficients demonstrates that the definition of any clustering notion for multiplex networks needs to be able to consider diverse forms of transitivity.

Acknowledgments

All authors were supported by the European Commission FET-Proactive project PLEXMATH (Grant No. 317614). AA also acknowledges financial support from the ICREA Academia, Generalitat de Catalunya (2009-SGR-838), and the James S.McDonnell Foundation; and SG and AA were supported by FIS2012-38266. YM was also supported by MINECO through Grants FIS2011-25167 and by DGA (Spain). MAP acknowledges a grant (EP/J001759/1) from the EPSRC. We thank David Krackhardt and the anonymous referees for useful comments.

Appendix A. Weighted clustering coefficients

There are two primary weighted clustering coefficients for monoplex networks that provide alternatives to the one that we discussed in the main text [54, 55]. They are

$$C_{O,u} = \frac{1}{w_{\max} k_u (k_u - 1)} \sum_{v,w} (W_{uv} W_{vw} W_{uw})^{1/3}, \quad (\text{A.1})$$

$$C_{\text{Ba},u} = \frac{1}{s_u (k_u - 1)} \sum_{v,w} \frac{(W_{uv} + W_{uw})}{2} A_{uv} A_{uw} A_{vw}, \quad (\text{A.2})$$

where \mathbf{A} is the unweighted adjacency matrix associated with the weighted adjacency matrix \mathbf{W} , the degree of node u is $k_u = \sum_v A_{uv}$, the strength of u is $s_u = \sum_v W_{uv}$, and the quantity $w_{\max} = \max_{u,v} W_{uv}$ is the maximum weight in \mathbf{W} . When using equations (A.1) and (A.2), one also has $C_{O,u} = 0$ and $C_{\text{Ba},u} = 0$ for nodes of degree $k_u = 0$ and $k_u = 1$.

Appendix B. Multiplex clustering coefficients in the literature

Let $\mathbf{A}^{(\alpha)}$ denote the intra-layer adjacency matrix for layer α . For a weighted multiplex network, we use $\mathbf{W}^{(\alpha)}$ to denote the intra-layer weight matrix (i.e., the weighted intra-layer adjacency matrix) for layer α . We use \mathbf{W} to denote the weight matrix of the aggregated network. (See section 2.3 in the main text.) The clustering coefficient that was defined in [36] for node-aligned multiplex networks is

$$C_{\text{Be},u} = \frac{\sum_{v,w} \sum_{\alpha} A_{uv}^{(\alpha)} \sum_{\kappa} A_{uw}^{(\kappa)} \sum_{\mu} A_{vw}^{(\mu)}}{\sum_{v,w} \sum_{\kappa} A_{uv}^{(\kappa)} \sum_{\alpha} \max(A_{uw}^{(\alpha)}, A_{vw}^{(\alpha)})}, \quad (\text{B.1})$$

which can be expressed in terms of the aggregated network as

$$C_{\text{Be},u} = \frac{\sum_{v,w} W_{uv} W_{uw} W_{vw}}{\sum_{v,w} W_{uv} \sum_{\alpha} \max(A_{uw}^{(\alpha)}, A_{vw}^{(\alpha)})}. \quad (\text{B.2})$$

The numerator of equation (B.2) is the same as the numerator of the weighted clustering coefficient $C_{Z,u}$, but the denominator is different. Because of the denominator in equation (B.2), the values of the clustering coefficient $C_{\text{Be},u}$ do not have to lie in the interval $[0, 1]$. For example, $C_{\text{Be},u} = (n - 2)b/n$ for a complete multiplex network (where n is the number of nodes in the multiplex network), so $C_{\text{Be},u} > 1$ when $b > \frac{n}{n-2}$.

References [37, 38] defined a family of local clustering coefficients for directed and weighted multiplex networks:

$$C_{\text{Br},u,t} = \frac{\sum_{\alpha \in L} \sum_{v,w \in N(u,t)} (W_{vw}^{(\alpha)} + W_{vw}^{(\alpha)})}{2|N(u,t)|b}, \quad (\text{B.3})$$

where $N(u,t) = \{v : |\{\alpha : A_{uv}^{(\alpha)} = 1 \text{ and } A_{vu}^{(\alpha)} = 1\}| \geq t\}$, t is a threshold, and we recall that $L = \{1, \dots, b\}$ is the set of layers. The clustering coefficient (B.3) does not yield the ordinary monoplex local clustering coefficient for unweighted (i.e., networks with binary weights) and undirected networks when it is calculated for the special case of a monoplex network (i.e., a multiplex network with $b = 1$ layer). Furthermore, its values are not normalized to lie between 0 and 1. For example, consider a complete multiplex network with n nodes and an arbitrary number of layers. In this case, the clustering coefficient (B.3) takes the value of $n - 2$ for each node. If a multiplex network is undirected (and unweighted), then $C_{\text{Br},u,t}$ can always be calculated when one is only given an aggregated network and the total number of layers in the multiplex network. As an example, for the threshold value $t = 1$, one obtains

$$C_{\text{Br},u,1} = \frac{1}{k_u b} \sum_{v,w} \frac{W_{vw}}{2} A_{uv} A_{uw} A_{vw}, \quad (\text{B.4})$$

where \mathbf{A} is the binary adjacency matrix corresponding to the weighted adjacency matrix \mathbf{W} and $k_u = \sum_v A_{uv}$ is the degree of node u .

Reference [39] defined a clustering coefficient for multiplex networks that are not necessarily node-aligned as

$$C_{\text{Cr},u} = \frac{2 \sum_{\alpha \in L} |\bar{E}_{\alpha}(u)|}{\sum_{\alpha} |\Gamma_{\alpha}(u)| (|\Gamma_{\alpha}(u)| - 1)}, \quad (\text{B.5})$$

where $L = \{1, \dots, b\}$ is again the set of layers, $\Gamma_{\alpha}(u) = \Gamma(u) \cap V_{\alpha}$, the quantity $\Gamma(u)$ is the set of neighbors of node u in the aggregated network, V_{α} is the set of nodes in layer α , and $\bar{E}_{\alpha}(u)$ is the set of edges in the subgraph induced by $\Gamma_{\alpha}(u)$ in the aggregated network. For a node-aligned multiplex network, $V_{\alpha} = V$ and $\Gamma_{\alpha}(u) = \Gamma(u)$, so one can write

$$C_{\text{Cr},u} = \frac{\sum_{v \neq w} A_{uv} W_{vw} A_{wu}}{b \sum_{v \neq w} A_{uv} A_{wu}}, \quad (\text{B.6})$$

which is a local clustering coefficient for the aggregated network.

Battiston *et al* [40] defined two versions of clustering coefficients for node-aligned multiplex networks:

$$C_{\text{Bat1},u} = \frac{\sum_{\alpha} \sum_{\kappa \neq \alpha} \sum_{v \neq u, w \neq u} A_{uv}^{(\alpha)} A_{vw}^{(\kappa)} A_{wu}^{(\alpha)}}{(b-1) \sum_{\alpha} \sum_{v \neq u, w \neq u} A_{uv}^{(\alpha)} A_{wu}^{(\alpha)}}, \quad (\text{B.7})$$

$$C_{\text{Bat2},u} = \frac{\sum_{\alpha} \sum_{\kappa \neq \alpha} \sum_{\mu \neq \alpha, \kappa} \sum_{v \neq u, w \neq u} A_{uv}^{(\alpha)} A_{vw}^{(\mu)} A_{wu}^{(\kappa)}}{(b-2) \sum_{\alpha} \sum_{\kappa \neq \alpha} \sum_{v \neq u, w \neq u} A_{uv}^{(\alpha)} A_{wu}^{(\kappa)}}. \quad (\text{B.8})$$

The first definition, $C_{\text{Bat1},u}$, counts the number of \mathcal{ACACA} -type elementary cycles; and the second definition, $C_{\text{Bat2},u}$, counts the three-layer elementary cycles \mathcal{ACACAC} . In both of these definitions, the sums in the denominators allow terms in which $v = w$, so a complete multiplex network has a local clustering coefficient of $(n-1)/(n-2)$ for every node.

Table B1. Summary of the properties of the different multiplex clustering coefficients. The notation $C_{*,(u)}$ means that the property holds for both the global version and the local version of the associated clustering coefficient (C.C.).

Property	$C_{M,(u)}$	$C_{Be,u}$	$C_{Z,(u)}$	$C_{Ba,u}$	$C_{O,u}$	$C_{Br,u}$	$C_{Cr,u}$	$C_{Bat(1,2),u}$
(1) Reduces to monoplex C.C.	✓	✓	✓	✓	✓		✓	
(2) $C_* \leq 1$	✓		✓	✓	✓		✓	✓
(3) $C_* = p$ in multiplex ER graph	✓						✓	
(4) Monoplex C.C. for copied layers	✓		✓	✓	✓		✓	
(5) Defined for node-layer pairs	✓							
(6) Defined for non-node-aligned	✓						✓	

Reference [31] proposed definitions for global clustering coefficients using a tensorial formalism for multilayer networks; when representing a multiplex network as a third-order tensor, the formulas in [31] reduce to the clustering coefficients that we propose in the present article. (See equation (6) of the main text.)

Parshani *et al* [56] defined an ‘inter-clustering coefficient’ for two-layer interdependent networks that can be interpreted as multiplex networks [4, 57–59]. Their definition is similar to edge ‘overlap’ [40]; in our framework, it corresponds to counting two-cycles of type $(\mathcal{A}\mathcal{C})^2$. A few other scholars [60, 61] have also defined generalizations of clustering coefficients for multilayer networks that cannot be interpreted as multiplex networks [4].

In table B1, we show a summary of the properties satisfied by several different (local and global) multiplex clustering coefficients. In particular, we check the following properties. (1) The value of the clustering coefficient reduces to the values of the associated monoplex clustering coefficient for a single-layer network. (2) The value of the clustering coefficient is normalized so that it takes values that are less than or equal to 1. (All of the clustering coefficients are non-negative.) (3) The clustering coefficient has a value of p in a large (i.e., when the number of nodes $n \rightarrow \infty$) node-aligned multiplex network in which each layer is an independent ER network with an edge probability of p in each layer. (4) Suppose that we construct a multiplex network by replicating the same given monoplex network in each layer. We indicate whether the clustering coefficient for the multiplex network has the same value as for the monoplex network. (5) There exists a version of the clustering coefficient that is defined for each node-layer pair separately. (6) The clustering coefficient is defined for multiplex networks that are not node-aligned.

Appendix C. Other possible definitions of cycles

There are many possible ways to define cycles in multiplex networks. If we relax the condition of disallowing two consecutive inter-layer steps, then we can write

$$t_{SM,i} = \left[(\hat{\mathcal{C}}\mathcal{A}\hat{\mathcal{C}})^3 \right]_{ii}, \quad (\text{C.1})$$

$$t_{SM',i} = \left[(\hat{\mathcal{C}}'\mathcal{A} + \mathcal{A}\hat{\mathcal{C}}')^3 \right]_{ii}, \quad (\text{C.2})$$

where $\hat{\mathcal{C}}' = \frac{1}{2}\beta\mathbf{I} + \gamma\mathcal{C}$. (Our discussion in appendix D is helpful for understanding the factor of 1/2.) Unlike the matrix $\mathcal{A}\hat{\mathcal{C}}$ in equation (3) in the main text, the matrices $\hat{\mathcal{C}}\mathcal{A}\hat{\mathcal{C}}$ and $\hat{\mathcal{C}}'\mathcal{A} + \mathcal{A}\hat{\mathcal{C}}'$ are symmetric. We can thus interpret them as weighted adjacency matrices of symmetric supra-graphs, and we can then calculate cycles and clustering coefficients in these supra-graphs (see appendix D).

It is sometimes desirable to forbid the option of staying inside of a layer in the first step of the second term of equation (C.2). In this case, one can write

$$t_{M',i} = \left[(\mathcal{A}\hat{\mathcal{C}})^3 + \gamma\mathcal{C}\mathcal{A}(\hat{\mathcal{C}}\mathcal{A})^2 \right]_{ii}. \quad (\text{C.3})$$

With this restriction, cycles that traverse two edges of the focal node i are only calculated two times instead of four times. We simplify equation (C.3) to obtain

$$t_{M',i} = \left[2(\mathcal{A}\hat{\mathcal{C}})^2\mathcal{A}\hat{\mathcal{C}}' \right]_{ii}, \quad (\text{C.4})$$

which is similar to equation (3) in the main text. In table C1, we show the values of the clustering coefficients that we calculate using this last definition of cycle for the empirical networks that we studied in the main text.

Table C1. Clustering coefficients (rows) for the same empirical networks (columns) from table 1 in the main text. For the Tube and the Airline networks, we only calculate clustering coefficients for non-node-aligned networks. For local clustering coefficients, we average over all nodes to obtain $\overline{C_{*,u}} = \frac{1}{n} \sum_u C_{*,u}$.

CC	Families	Bank	Tailor Shop	Management	Tube	Airline
C_M	0.218	0.289	0.320	0.206	0.070	0.102
$C_M^{(1)}$	0.289	0.537	0.406	0.436	0.013	0.100
$C_M^{(2)}$	0.202	0.368	0.338	0.297	0.041	0.173
$C_M^{(3)}$	—	0.227	0.288	0.192	0.314	0.086
$C_M^{(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})}$	0.164	0.377	0.344	0.309	0.123	0.120
$\overline{C_{Cr,u}}$	0.342	0.254	0.308	0.150	0.038	0.329
$\overline{C_{Ba,u}}$	0.195	0.811	0.612	2.019	—	—
$\overline{C_{Br,u}}$	0.674	1.761	4.289	1.636	—	—
$\overline{C_{O,u}}$	0.303	0.268	0.260	0.133	—	—
$\overline{C_{Be,u}}$	0.486	0.775	0.629	0.715	—	—
$\overline{C_{Bat1,u}}$	0.159	0.199	0.271	0.169	—	—
$\overline{C_{Bat2,u}}$	—	0.190	0.282	0.179	—	—

Appendix D. Defining multiplex clustering coefficients using auxiliary networks

An elegant way to generalize clustering coefficients for multiplex networks is to define a new (possibly weighted) auxiliary supra-graph G_M so that one can define cycles of interest as weighted three-cycles in G_M . Once we have a function that produces the auxiliary supra-adjacency matrix $\mathcal{M} = \mathcal{M}(\mathcal{A}, C)$, we can define the auxiliary complete supra-adjacency matrix $\mathcal{M}^F = \mathcal{M}(\mathcal{F}, C)$. One can then define a local clustering coefficient for node-layer pair i with the formula

$$c_i = \frac{(\mathcal{M}^3)_{ii}}{(\mathcal{M}\mathcal{M}^F\mathcal{M})_{ii}}. \quad (\text{D.1})$$

As with a monoplex network, the denominator written in terms of the complete matrix \mathcal{M}^F is equivalent to the usual one written in terms of connectivity. We thereby consider the connectivity of a node in the supra-graph induced by the matrix \mathcal{M} . We refer to the matrix \mathcal{M} as the *multiplex walk matrix* because it encodes the permissible steps in a multiplex network. When \mathcal{M} is equal to $\mathcal{A}\hat{C}$ or to $\hat{C}\mathcal{A}$, the induced supra-graph is directed, so one needs to distinguish between in-degrees and out-degrees.

A key advantage of defining clustering coefficients using an auxiliary supra-graph is that one can then use it to calculate other diagnostics (e.g., degree or strength) for nodes. One can thereby investigate correlations between clustering-coefficient values and the size of the multiplex neighborhood of a node. (The size of the neighborhood is the number of nodes that are reachable in a single step via connections defined by the matrix \mathcal{M} .)

We can write the symmetric multiplex walk matrices in equations (C.1) and (C.2) as

$$\mathcal{M}_{SM} = \hat{C}\mathcal{A}\hat{C}, \quad (\text{D.2})$$

$$\mathcal{M}_{SM'} = (\hat{C}'\mathcal{A} + \mathcal{A}\hat{C}'). \quad (\text{D.3})$$

To avoid double-counting intra-layer steps in the definition of $\mathcal{M}_{SM'}$, we need to rescale either the intra-layer weight parameter β (i.e., we can write $\hat{C}' = \beta'\mathbf{I} + \gamma C = \frac{1}{2}\beta'\mathbf{I} + \gamma C$) or the inter-layer weight parameter γ (i.e., we can write $\hat{C}' = \beta\mathbf{I} + \gamma C = \beta\mathbf{I} + 2\gamma C$ and also define $\mathcal{M}_{SM'} = \frac{1}{2}(\mathcal{A}\hat{C}' + \hat{C}'\mathcal{A})$).

Consider a supra-graph induced by a multiplex walk matrix. The distinction between the matrices \mathcal{M}_{SM} and $\mathcal{M}_{SM'}$ is that \mathcal{M}_{SM} also includes terms of the form $C\mathcal{A}C$ that take into account walks that have an inter-layer step (C) followed by an intra-layer step (\mathcal{A}) and then another inter-layer step (C). Therefore, in the supra-graph induced by \mathcal{M}_{SM} , two nodes in the same layer that are not adjacent in that layer are nevertheless adjacent if the same physical nodes are adjacent in another layer.

The matrix \hat{C} sums the contributions of all node-layer pairs that correspond to the same physical node when $\beta = \gamma = 1$. In other words, if we associate a vector of the canonical basis e_i to each node-layer pair i and let $\Gamma_C((u, \alpha)) = \{(u, \kappa) | \kappa \in L\}$ denote all node-layer pairs that correspond to the same physical node, then

$$\hat{C}e_i = \sum_{j \in \Gamma_C(i)} e_j \quad (\text{D.4})$$

produces a vector whose entries are equal to 1 for nodes that belong to the basis vector and which are equal to 0 for nodes that do not belong to that vector. Consequently, \mathcal{M}_{SM} is related to the weighted adjacency matrix of

the aggregated graph for $\beta = \gamma = 1$. To be precise, we obtain the following relation:

$$\left(\widehat{\mathcal{C}}\mathcal{A}\widehat{\mathcal{C}}\right)_{ij} = W_{uv}, \quad \text{for any } i \in l(u), j \in l(v). \quad (\text{D.5})$$

One can also write the multiplex clustering coefficient induced by equation (2) in terms of the auxiliary supra-adjacency matrix by considering equation (3), which is a simplified version of the equation that counts cycles only in one direction. This yields

$$\mathcal{M}_M = \sqrt[3]{2} \mathcal{A}\widehat{\mathcal{C}}. \quad (\text{D.6})$$

The matrix \mathcal{M}_M is not symmetric, which implies that the associated graph is a directed supra-graph. Nevertheless, the clustering coefficient induced by \mathcal{M}_M is the same as that induced by its transpose \mathcal{M}_M^T if \mathcal{A} is symmetric.

Appendix E. Expressing clustering coefficients using elementary three-cycles

We now give a detailed explanation of the process of decomposing any of our walk-based clustering coefficients into elementary cycles. An *elementary cycle* is a term that consists of products of the matrices \mathcal{A} and \mathcal{C} (i.e., there are no sums) after one expands the expression for a cycle (which is a weighted sum of such terms). Because we are only interested in the diagonal elements of the terms and we consider only undirected intra-layer supra-graphs and coupling supra-graphs, we can transpose the terms and still write them in terms of the matrices \mathcal{A} and \mathcal{C} rather than also using their transposes. There are also multiple ways of writing non-symmetric elementary cycles (e.g., $(\mathcal{A}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C})_{ii} = (\mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{A})_{ii}$).

We adopt a convention in which we transpose all elementary cycles so that we select the one in which the first element is \mathcal{A} rather than \mathcal{C} when comparing the two versions of the term from left to right. That is, for two equivalent terms, we choose the one that comes first in alphabetical order. To calculate the clustering coefficients that we defined in the appendix (see appendices C and D), we also need to include elementary cycles that start and end in an inter-layer step. The set of elementary three-cycles is thus $\mathcal{E} = \{\mathcal{A}\mathcal{A}\mathcal{A}, \mathcal{A}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}, \mathcal{A}\mathcal{C}\mathcal{A}\mathcal{A}\mathcal{C}, \mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A}, \mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}, \mathcal{C}\mathcal{A}\mathcal{A}\mathcal{A}\mathcal{C}, \mathcal{C}\mathcal{A}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}, \mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}\}$.

We now write our clustering coefficients using elementary three-cycles. We obtain the normalization formulas by using the elementary three-cycles and then replacing the second \mathcal{A} term with \mathcal{F} . This yields a standard form for any of our multiplex clustering coefficients. For example,

$$c_{*,i} = \frac{t_{*,i}}{d_{*,i}}, \quad (\text{E.1})$$

where

$$t_{*,i} = \left[w_{\mathcal{A}\mathcal{A}\mathcal{A}}\mathcal{A}\mathcal{A}\mathcal{A} + w_{\mathcal{A}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}}\mathcal{A}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C} + w_{\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{A}\mathcal{C}}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{A}\mathcal{C} \right. \\ \left. + w_{\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A}}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A} + w_{\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C} \right. \\ \left. + w_{\mathcal{C}\mathcal{A}\mathcal{A}\mathcal{A}\mathcal{C}}\mathcal{C}\mathcal{A}\mathcal{A}\mathcal{A}\mathcal{C} + w_{\mathcal{C}\mathcal{A}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}}\mathcal{C}\mathcal{A}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C} \right. \\ \left. + w_{\mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}}\mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C} \right]_{ii} \quad (\text{E.2})$$

$$d_{*,i} = \left[w_{\mathcal{A}\mathcal{A}\mathcal{A}}\mathcal{A}\mathcal{F}\mathcal{A} + w_{\mathcal{A}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}}\mathcal{A}\mathcal{F}\mathcal{C}\mathcal{A}\mathcal{C} + w_{\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{A}\mathcal{C}}\mathcal{A}\mathcal{F}\mathcal{C}\mathcal{A}\mathcal{C} \right. \\ \left. + w_{\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A}}\mathcal{A}\mathcal{F}\mathcal{C}\mathcal{A} + w_{\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}}\mathcal{A}\mathcal{F}\mathcal{C}\mathcal{A}\mathcal{C} \right. \\ \left. + w_{\mathcal{C}\mathcal{A}\mathcal{A}\mathcal{A}\mathcal{C}}\mathcal{C}\mathcal{A}\mathcal{F}\mathcal{A}\mathcal{C} + w_{\mathcal{C}\mathcal{A}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}}\mathcal{C}\mathcal{A}\mathcal{F}\mathcal{C}\mathcal{A}\mathcal{C} \right. \\ \left. + w_{\mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}\mathcal{A}\mathcal{C}}\mathcal{C}\mathcal{A}\mathcal{F}\mathcal{C}\mathcal{A}\mathcal{C} \right]_{ii}, \quad (\text{E.3})$$

where i is a node-layer pair and the $w_{\mathcal{E}}$ coefficients are scalars that correspond to the weights for each type of elementary cycle. These weights are different for different types of clustering coefficients; one can choose whatever is appropriate for a given problem. Note that we have absorbed the parameters β and γ into these coefficients (see below and table E1). We illustrate the possible elementary cycles in figure 1 of the main text and in figure E1.

One can even express the cycles that include two consecutive inter-layer steps in the standard form of equations (E.2)–(E.3) for node-aligned multiplex networks, because $\mathcal{C}^2 = (b-1)\mathcal{I} + (b-2)\mathcal{C}$ in this case. Without the assumption that $\beta = \gamma = 1$, the expansion for the coefficient $c_{\text{SM},i}$ is cumbersome because it includes coefficients $\beta^k\gamma^h$ with all possible combinations of k and h such that $k+h=6$ and $h \neq 1$. Furthermore, in the general case, it is also not possible to infer the number of layers in which a walk traverses an intra-layer edge based on the exponents of β and γ for $c_{\text{SM},i}$ and $c_{\text{SM}',i}$. For example, in $c_{\text{SM}',i}$ the intra-layer elementary triangle $\mathcal{A}\mathcal{A}\mathcal{A}$ includes a contribution from both β^3 (i.e., the walk stays in the original layer) and

Table E1. Coefficients of elementary multiplex three-cycle terms $w_{\mathcal{E}}$ (see equations (E.2) and (E.3)) for different multiplex clustering coefficients. For example, w_{AAA} for type- M clustering coefficients (i.e., C_M , $C_{M,u}$, and $C_{M,i}$) is equal to $2\beta^3$. For type-SM' and type-SM clustering coefficients, we calculate the expansions only for node-aligned multiplex networks.

C.C.	$\beta^h\gamma^k$	AAA	$AACAC$	$ACAAC$	$ACACA$	$ACACAC$	$CAAAC$	$CAACAC$	$CACACAC$
M	β^3	2		2	2				
	$\beta\gamma^2$		2						
	γ^3					2			
M'	β^3	1		2	1				
	$\beta\gamma^2$		2						
	γ^3					2			
SM'	β^3	1		4	3		1		
	$\beta\gamma^2$	$2(b-1)$	2						
	γ^3			$2(b-2)$	$2(b-2)$	2		2	
SM	β^6	1							
	$\beta^4\gamma^2$	$2(b-1)$	4	4	4		1		
	$\beta^3\gamma^3$		$2(b-2)$	$2(b-2)$	$4(b-2)$	8		4	
	$\beta^2\gamma^4$	$(b-1)^2$	$4(b-1)$	$4(b-1)$	$(b-2)^2$	$8(b-2)$	$2(b-1)$	$2(b-2)$	4
	$\beta^1\gamma^5$		$2(b-2)(b-1)$	$2(b-2)(b-1)$		$2(b-2)^2$		$4(b-1)$	$4(b-2)$
	γ^6						$(b-1)^2$	$2(b-2)(b-1)$	$(b-2)^2$

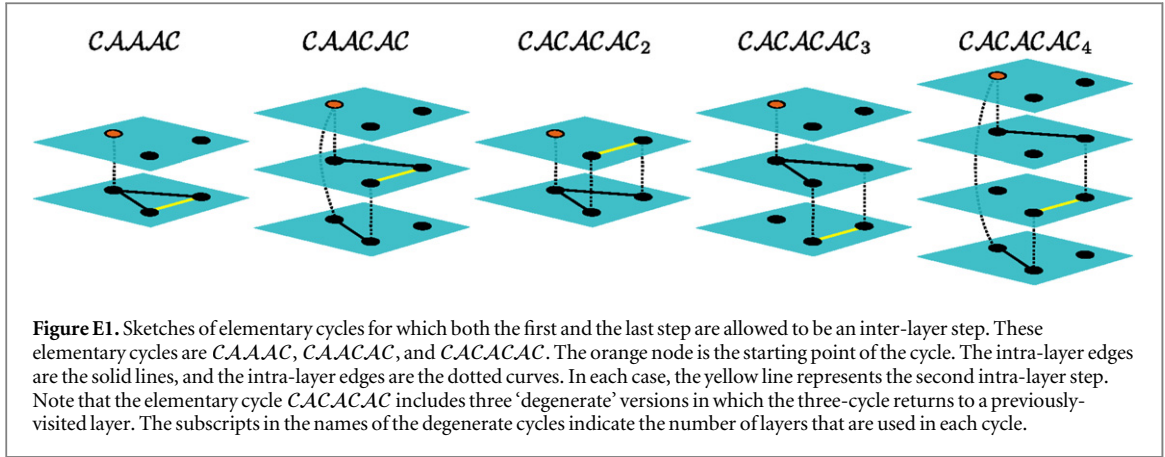
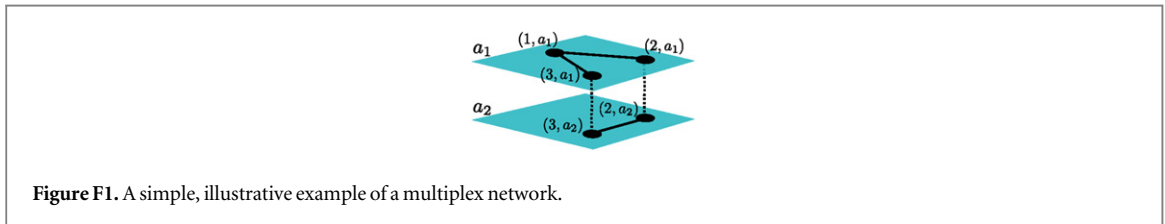


Table E2. Coefficients of the elementary multiplex three-cycle terms $w_{\mathcal{E}}$ (see equations (E.2) and (E.3)) for different multiplex clustering coefficients when $\beta = \gamma = 1$. For type-SM' and type-SM clustering coefficients, we calculate the expansions only for node-aligned multiplex networks.

C. C.	AAA	$AACAC$	$ACAAC$	$ACACA$	$ACACAC$	$CAAAAC$	$CAACAC$	$CACACAC$
M	2	2	2	2	2	0	0	0
M'	1	2	2	1	2	0	0	0
SM	$1b^2$	$2b^2$	$2b^2$	b^2	$2b^2$	b^2	$2b^2$	b^2
SM'	$2b - 1$	2	$2b$	$2b - 1$	2	1	2	0



$\beta\gamma^2$ (i.e., the walk visits some other layer but then comes back to the original layer without traversing any intra-layer edges while it was gone). Moreover, all of the terms with b arise from a walk moving to a new layer and then coming right back to the original layer in the next step. Because there are $b - 1$ other layers from which to choose, the influence of cycles with such transient layer visits is amplified by the total number of layers in a network. That is, adding more layers (even ones that do not contain any edges) changes the relative importance of different types of elementary cycles.

In table E1, we show the values of the coefficients $w_{\mathcal{E}}$ for the different ways that we define three-cycles in multiplex networks. In table E2, we show their corresponding expansions in terms of elementary cycles for the case $\beta = \gamma = 1$. These cycle decompositions illuminate the difference between $c_{M,i}$, $c_{M',i}$, $c_{SM,i}$ and $c_{SM',i}$. The clustering coefficient $c_{M,i}$ gives equal weight to each elementary cycle, whereas $c_{M',i}$ gives half of the weight to AAA and $ACACA$ cycles (i.e., the elementary cycles that include an implicit double-counting) as compared to the other cycles. The matrices that correspond to elementary cycles with such double-counting are symmetric, and the same cycle is thus counted in two different directions.

Appendix F. A simple example

We now use a simple example (see figure F1) to illustrate the differences between the various notions of a multiplex clustering coefficient. Consider a two-layer multiplex network with three nodes in layer a_1 and two nodes in layer a_2 . The three node-layer pairs in layer a_1 form a 2-star, the two node-layer pairs that are not connected directly to each other on layer a_1 are each adjacent via an inter-layer edge to a counterpart node-layer pair in layer a_2 , and the two node-layer pairs on layer a_2 are adjacent to each other.

The adjacency matrix \mathcal{A} for the intra-layer supra-graph is

$$\mathcal{A} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad (\text{F.1})$$

and the adjacency matrix \mathcal{C} of the coupling supra-graph is

$$\mathcal{C} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}. \quad (\text{F.2})$$

Therefore, the supra-adjacency matrix is

$$\bar{\mathcal{A}} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}. \quad (\text{F.3})$$

The multiplex walk matrix \mathcal{M}_M is

$$\mathcal{M}_M = \sqrt[3]{2} \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix}, \quad (\text{F.4})$$

and we note that it is not symmetric. For example, node-layer pair $(2, a_2)$ is reachable from $(1, a_1)$, but node-layer pair $(1, a_1)$ is not reachable from $(2, a_2)$. The edge $[(1, a_1), (2, a_2)]$ in this supra-graph represents the walk $\{(1, a_1), (2, a_1), (2, a_2)\}$ in the multiplex network. The symmetric walk matrix $\mathcal{M}_{SM'}$ is

$$\mathcal{M}_{SM'} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix}. \quad (\text{F.5})$$

The matrix $\mathcal{M}_{SM'}$ is the sum of \mathcal{M}_M and \mathcal{M}_M^T with rescaled diagonal blocks in order to not double-count the edges $[(1, a_1), (2, a_1)]$ and $[(1, a_1), (3, a_1)]$. Additionally,

$$\mathcal{M}_{SM} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix}, \quad (\text{F.6})$$

which differs from $\mathcal{M}_{SM'}$ in the fact that node-layer pairs $(2, a_1)$ and $(3, a_1)$ are connected through the multiplex walk $\{(2, a_1), (2, a_2), (3, a_2), (3, a_1)\}$.

The adjacency matrix of the aggregated graph is

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}. \quad (\text{F.7})$$

That is, it is a complete graph without self-edges.

We now calculate $c_{*,i}$ using different definitions of a multiplex clustering coefficient. To calculate $c_{M,i}$, we need to compute the auxiliary complete supra-adjacency matrix \mathcal{M}_M^F according to equation (D.6). We obtain

$$\mathcal{M}_M^F = \sqrt[3]{2} \mathcal{F}\hat{\mathcal{C}} = \sqrt[3]{2} \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix}. \quad (\text{F.8})$$

The clustering coefficient of node-layer pair $(1, a_1)$, which is part of two triangles that are reachable along the directions of the edges, is

$$c_{M,(1,a_1)} = \frac{1}{2}. \quad (\text{F.9})$$

For node-layer pair $(2, a_1)$, we get

$$c_{M,(2,a_1)} = 1, \quad (\text{F.10})$$

which is the same as the clustering-coefficient values of the remaining node-layer pairs.

To calculate $c_{SM',i}$, we need to compute $\mathcal{M}_{SM'}^F$, which we obtain using equation (D.3). We thus obtain

$$\mathcal{M}_{SM'}^F = \mathcal{F}\hat{\mathcal{C}} + \hat{\mathcal{C}}\mathcal{F} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix}. \quad (\text{F.11})$$

In the supra-graph associated with the supra-adjacency matrix $\mathcal{F}\hat{\mathcal{C}} + \hat{\mathcal{C}}\mathcal{F}$, all node-layer pairs are adjacent to all other node-layer pairs except those that correspond to the same physical nodes. The clustering coefficient of node-layer pair $(1, a_1)$, which is part of six triangles, is

$$c_{SM',(1,a_1)} = \frac{1}{2} = c_{M,(1,a_1)}. \quad (\text{F.12})$$

The clustering coefficient of node-layer pair $(2, a_1)$, which is part of one triangle, is

$$c_{SM',(2,a_1)} = 1. \quad (\text{F.13})$$

To calculate $c_{SM,i}$, we compute \mathcal{M}_{SM}^F using equation (D.2). We thus obtain

$$\mathcal{M}_{SM}^F = \hat{\mathcal{C}}\mathcal{F}\hat{\mathcal{C}} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 0 & 2 \\ 1 & 2 & 0 & 2 & 0 \\ 1 & 0 & 2 & 0 & 2 \\ 1 & 2 & 0 & 2 & 0 \end{pmatrix}. \quad (\text{F.14})$$

The only difference between the graphs associated with the matrices $\hat{\mathcal{C}}\mathcal{F}\hat{\mathcal{C}}$ and $\mathcal{F}\hat{\mathcal{C}} + \hat{\mathcal{C}}\mathcal{F}$ is the weight of the edges in $\hat{\mathcal{C}}\mathcal{F}\hat{\mathcal{C}}$ that take into account the fact that intra-layer edges might be repeated in the two layers.

The clustering coefficient of node-layer pair $(1, a_1)$, which is part of eight triangles, is

$$c_{SM,(1,a_1)} = \frac{8}{12} = \frac{2}{3}. \quad (\text{F.15})$$

The clustering coefficient of node-layer pair $(2, a_1)$, which is part of four triangles, is

$$c_{SM,(2,a_1)} = \frac{4}{6} = \frac{2}{3}. \quad (\text{F.16})$$

Because we are weighting edges based on the number of times an edge between two nodes is repeated in different layers among a given pair of physical nodes in the normalization, none of the node-layer pairs has a clustering coefficient equal to 1. By contrast, all nodes have clustering coefficients with the same value in the aggregated network, for which the layer information has been lost. In particular, they each have a clustering-coefficient value of $\frac{2}{3}$, independent of the definition of the multiplex clustering coefficient.

Appendix G. Further discussion of clustering coefficients in multiplex ER networks

The expected values of the local clustering coefficients in node-aligned multiplex ER networks are

$$\langle c_{AAAA,i} \rangle = \frac{1}{b} \sum_{\alpha \in L} p_\alpha \equiv \bar{p}, \quad (\text{G.1})$$

$$\langle c_{AACAC,i} \rangle = \frac{1}{b} \sum_{\alpha \in L} p_\alpha \equiv \bar{p}, \quad (\text{G.2})$$

$$\langle c_{ACAAC,i} \rangle = \frac{1}{b} \sum_{\alpha \in L} \frac{\sum_{\kappa \neq \alpha} p_\kappa^2}{\sum_{\kappa \neq \alpha} p_\kappa}, \quad (\text{G.3})$$

$$\langle c_{ACACA,i} \rangle = \frac{1}{b} \sum_{\alpha \in L} p_\alpha \equiv \bar{p}, \quad (\text{G.4})$$

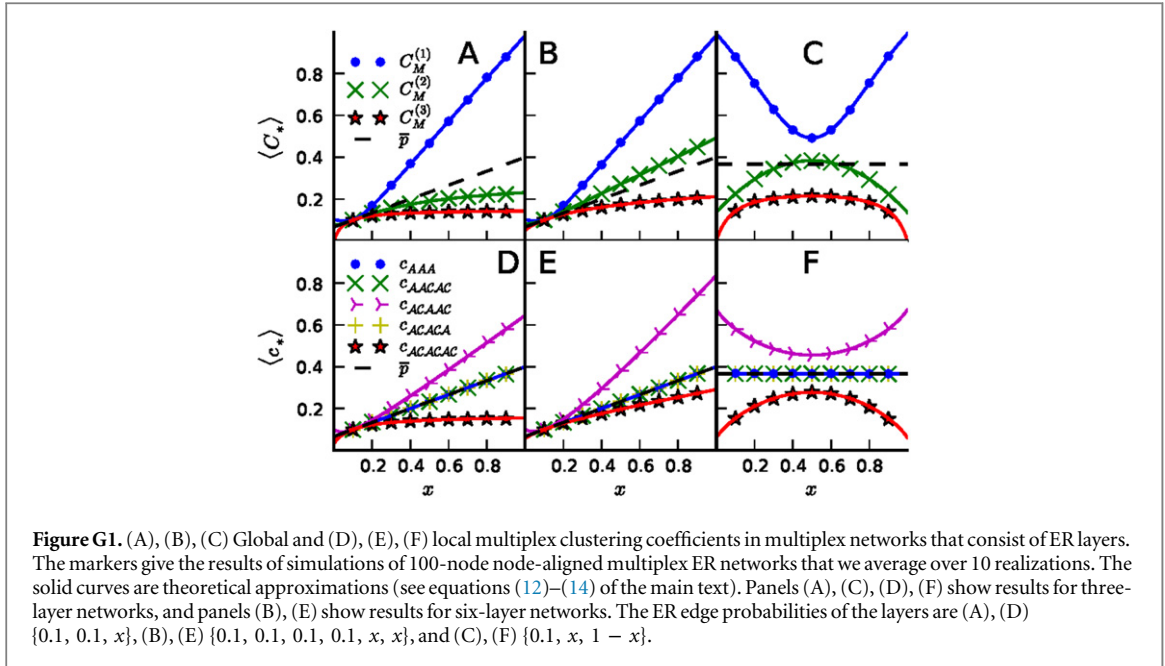


Figure G1. (A), (B), (C) Global and (D), (E), (F) local multiplex clustering coefficients in multiplex networks that consist of ER layers. The markers give the results of simulations of 100-node node-aligned multiplex ER networks that we average over 10 realizations. The solid curves are theoretical approximations (see equations (12)–(14) of the main text). Panels (A), (C), (D), (F) show results for three-layer networks, and panels (B), (E) show results for six-layer networks. The ER edge probabilities of the layers are (A), (D) {0.1, 0.1, x }, (B), (E) {0.1, 0.1, 0.1, 0.1, x , x }, and (C), (F) {0.1, x , $1 - x$ }.

$$\langle c_{ACACAC,i} \rangle = \frac{1}{b(b-1)} \sum_{\alpha \in L} \frac{\sum_{\kappa \neq \alpha; \mu \neq \kappa, \alpha} p_{\kappa} p_{\mu}}{\sum_{\kappa \neq \alpha} p_{\kappa}}. \quad (\text{G.5})$$

Note that $c_{M,i}^{(1)} = c_{AAA,i}$ and $c_{M,i}^{(3)} = c_{ACACAC,i}$, but the two-layer clustering coefficient $c_{M,i}^{(2)}$ arises from a weighted sum of contributions from three different elementary cycles.

In figure G1, we illustrate the behavior of the global and local clustering coefficients in multiplex networks in which the layers consist of ER networks with varying amounts of heterogeneity in the intra-layer edge densities. Although the global and mean local clustering coefficients are equal to each other when averaged over ensembles of monoplex ER networks, we do not obtain a similar result for multiplex networks with ER layers unless the layers have the same value of the parameter p . The global clustering coefficients give more weight than the mean local clustering coefficients to denser layers. This is evident for the intra-layer clustering coefficients $c_{M,i}^{(1)}$ and $C_M^{(1)}$, for which the ensemble average of the mean of the local clustering coefficient $c_{M,i}^{(1)}$ is always equal to the mean edge density, whereas the ensemble average of the global clustering coefficient $C_M^{(1)}$ has values that are greater than or equal to the mean edge density. This situation is a good example of a case in which transitivity in multiplex networks differs from the results and intuition from monoplex networks.

In particular, failing to take into account the heterogeneity of edge densities in multiplex networks can lead to incorrect or misleading results when trying to distinguish among values of a clustering coefficient that are what one would expect from an ER random network versus those that are a signature of a triadic-closure process (see figure G1).

Appendix H. Null model for shuffling inter-layer connections

In table H1, we compare empirical values of layer-decomposed global clustering coefficients with clustering-coefficient values for a null model in which we preserve the topology of each intra-layer network but independently shuffle the labels of the nodes inside of each layer. That is, for each intra-layer network $G^\alpha = (V^\alpha, E^\alpha)$, we choose a permutation $\pi : V^\alpha \mapsto V^\alpha$ uniformly at random and construct a new multiplex network starting from $\{\pi(G^\alpha)\}$, where $\pi(G^\alpha) = (\pi(V^\alpha), \pi(E^\alpha))$ and $\pi(E^\alpha) = \{(\pi(u), \pi(v)) | (u, v) \in E^\alpha\}$. In this way, we effectively randomize inter-layer edges but preserve both the structure of intra-layer networks and the number of inter-layer edges between each pair of layers. For our comparisons using this null model, most of the clustering coefficients take values that are significant for our data sets (see table H1). Because of the way that we construct the null model, the global single-layer clustering coefficients are exactly the same for the original data and the null model.

Table H1. Clustering coefficients C_M , $C_M^{(1)}$, $C_M^{(2)}$, $C_M^{(3)}$ that correspond, respectively, to the global, one-layer, two-layer, and three-layer clustering coefficients for various multiplex networks. ‘Tailor Shop’: Kapferer tailor-shop network ($n = 39$, $b = 4$) [48]. ‘Management’: Krackhardt office cognitive social structure ($n = 21$, $b = 21$) [49]. ‘Families’: Padgett Florentine families social network ($n = 16$, $b = 2$) [50]. ‘Bank’: Roethlisberger and Dickson bank wiring-room social network ($n = 14$, $b = 6$) [51]. ‘Tube’: the London Underground (i.e., ‘the Tube’) transportation network ($n = 314$, $b = 14$) [52]. ‘Airline’: network of flights between cities, in which each layer corresponds to a single airline ($n = 3108$, $b = 530$) [53]. The rows labeled ‘orig.’ give the clustering coefficients for the original networks, and the rows labeled ‘NM’ give the expected value and the standard deviation of the clustering coefficient in a null-model network in which we preserve the topology of the intra-layer networks but separately (and independently) shuffle the node labels for each intra-layer network. For the original values, we perform a two-tailed Z-test to examine whether the observed clustering coefficients could have been produced by the null model. We designate the p -values as follows: * : $p < 0.05$, ** : $p < 0.01$ for Bonferroni-corrected tests with 18 hypothesis; † : $p < 0.05$, †† : $p < 0.01$ for uncorrected tests. We do not use any symbols for values that are not significant. We symmetrize directed networks by considering two nodes to be adjacent if there is at least one edge between them. The social networks in this table are node-aligned multiplex graphs, but the transportation networks are not node-aligned. We report values that are means over different numbers of realizations: 10^5 for Tailor Shop, 10^3 for Airline, 10^4 for Management, 10^5 for Families, 10^4 for Tube, and 10^5 for Bank.

		Tailor Shop	Management	Families	Bank	Tube	Airline
C_M	orig.	0.319**	0.206**	0.223	0.293**	0.056**	0.101**
	NM	0.218 ± 0.007	0.131 ± 0.003	0.194 ± 0.029	0.223 ± 0.014	0.025 ± 0.008	0.054 ± 0.001
$C_M^{(1)}$	orig.	0.406	0.436	0.289	0.537	0.013	0.100
	NM	0.406	0.436	0.289	0.537	0.013	0.100
$C_M^{(2)}$	orig.	0.327**	0.273**	0.198	0.349**	0.043	0.150**
	NM	0.220 ± 0.009	0.176 ± 0.003	0.158 ± 0.041	0.240 ± 0.018	0.035 ± 0.018	0.082 ± 0.002
$C_M^{(3)}$	orig.	0.288**	0.192**	—	0.227†	0.314**	0.086**
	NM	0.165 ± 0.010	0.120 ± 0.003	—	0.186 ± 0.016	0.053 ± 0.043	0.037 ± 0.002

References

- [1] Wasserman S and Faust K 1994 *Social Network Analysis: Methods and Applications Structural Analysis in the Social Sciences* (Cambridge: Cambridge University Press)
- [2] Boccaletti S, Latora V, Moreno Y, Chavez M and Hwang D U 2006 *Phys. Rep.* **424** 175–308
- [3] Newman M E J 2010 *Networks: An Introduction* (Oxford: Oxford University Press)
- [4] Kivela M, Arenas A, Barthelemy M, Gleeson J P, Moreno Y and Porter M A 2014 *J. Complex Netw.* **2** 203–71
- [5] Boccaletti S, Bianconi G, Criado R, Del Genio C I, Gómez-Gardeñes J, Romance M, Sendiña-Nadal I, Wang Z and Zanin M 2014 *Phys. Rep.* **544** 1–122
- [6] Chang S E, Seligson H A and Eguchi R T 1996 Estimation of the economic impact of multiple lifeline disruption: memphis light, gas, and water division case study *Technical Report No. NCEER-96-0011* Multidisciplinary Center for Earthquake Engineering Research (MCEER), Buffalo, NY
- [7] Little R G 2002 *J. Urban Tech.* **9** 109–23
- [8] Mitchell J C 1969 *Social Networks in Urban Situations: Analyses of Personal Relationships in Central African Towns* (Manchester: Manchester University Press)
- [9] Verbrugge L M 1979 *Soc. Forces* **57** 1286–309
- [10] Coleman J S 1988 *Am. J. Sociol.* **94**(Suppl) 95–120
- [11] Leicht E A and D’Souza R M 2009 Percolation on interacting networks arXiv: 0907.0894
- [12] Mucha P J, Richardson T, Macon K, Porter M A and Onnela J-P 2010 *Science* **328** 876–8
- [13] Criado R, Flores J, García del Amo A, Gómez-Gardeñes J and Romance M 2010 *Int. J. Bifurcation Chaos* **20** 877–83
- [14] Buldyrev S V, Parshani R, Paul G, Stanley H E and Havlin S 2010 *Nature* **464** 1025–8
- [15] Gao J, Buldyrev S V, Havlin S and Stanley H E 2011 *Phys. Rev. Lett.* **107** 195701
- [16] Brummitt C D, D’Souza R M and Leicht E A 2012 *Proc. Natl. Sci. USA* **109** 680–9
- [17] Yağan O and Gligor V 2012 *Phys. Rev. E* **86** 036103
- [18] Gao J, Buldyrev S V, Stanley H E and Havlin S 2012 *Nat. Phys.* **8** 40–48
- [19] Lee K M, Kim J Y, Cho W K, Goh K-I and Kim I M 2012 *New J. Phys.* **14** 033027
- [20] Brummitt C D, Lee K M and Goh K-I 2012 *Phys. Rev. E* **85** 045102 (R)
- [21] Gómez S, Díaz-Guilera A, Gómez-Gardeñes J, Pérez-Vicente C, Moreno Y and Arenas A 2013 *Phys. Rev. Lett.* **110** 028701
- [22] De Domenico M, Solé-Ribalta A, Gómez S and Arenas A 2014 *Proc. Natl. Acad. Sci. USA* **111** 8351–6
- [23] Bianconi G 2013 *Phys. Rev. E* **87** 062806
- [24] Solé L, Romance M, Criado R, Flores J, García del Amo A and Boccaletti S 2013 *Chaos* **23** 033131
- [25] Halu A, Mondragón R J, Panzarasa P and Bianconi G 2013 *PLOS ONE* **8** e78293
- [26] Cozzo E, Arenas A and Moreno Y 2012 *Phys. Rev. E* **86** 036115
- [27] Cozzo E, Baños R A, Meloni S and Moreno Y 2013 *Phys. Rev. E* **88** 050801
- [28] Granel C, Gómez S and Arenas A 2013 *Phys. Rev. Lett.* **111** 128701
- [29] Solé-Ribalta A, De Domenico M, Kouvaris N E, Díaz-Guilera A, Gómez S and Arenas A 2013 *Phys. Rev. E* **88** 032807
- [30] De Domenico M, Solé-Ribalta A, Omodei E, Gómez S and Arenas A 2015 *Nat. Commun.* **6** 6868
- [31] De Domenico M, Solé-Ribalta A, Cozzo E, Kivela M, Moreno Y, Porter M A, Gómez S and Arenas A 2013 *Phys. Rev. X* **3** 041022
- [32] Min B and Goh K-I 2013 arXiv:1307.2967
- [33] Luce R D and Perry A D 1949 *Psychometrika* **14** 95–116
- [34] Watts D and Strogatz S H 1998 *Nature* **393** 440–2
- [35] Karlberg M 1997 *Soc. Netw.* **19** 325–43

- [36] Barrett L, Henzi S P and Lusseau D 2012 *Phil. Trans. R. Soc. London B* **367** 2108–18
- [37] Bródka P, Musiał K and Kazienko P 2010 A method for group extraction in complex social networks *Knowledge Management, Information Systems, E-Learning, and Sustainability Research (Communications in Computer and Information Science vol 111)* ed M D Lytras, P Ordóñez De Pablos, A Ziderman, A Roulstone, H Maurer and J B Imber (Berlin: Springer) pp 238–47
- [38] Bródka P, Kazienko P, Musiał K and Skibicki K 2012 *Int. J. Comput. Intell. Syst.* **5** 582–96
- [39] Criado R, Flores J, García del Amo A, Gómez-Gardeñes J and Romance M 2011 *Int. J. Comput. Math.* **89** 291–309
- [40] Battiston F, Nicosia V and Latora V 2014 *Phys. Rev. E* **89** 032804
- [41] Zhang B and Horvath S 2005 *Stat. Appl. Genetics Mol. Biol.* **4** 17
- [42] Szell M, Lambiotte R and Thurner S 2010 *Proc. Natl Sci. USA* **107** 13636–41
- [43] Gallotti R and Barthelemy M 2014 *Sci. Rep.* **4** 6911
- [44] Saramäki J, Kivelä M, Onnela J-P, Kaski K and Kertész J 2007 *Phys. Rev. E* **75** 027105
- [45] Opsahl T and Panzarasa P 2009 *Soc. Netw.* **31** 155–63
- [46] Ahnert S E, Garlaschelli D, Fink T M A and Caldarelli G 2007 *Phys. Rev. E* **76** 016101
- [47] Grindrod P 2002 *Phys. Rev. E* **66** 066702
- [48] Kapferer B 1972 *Strategy and Transaction in an African Factory: African Workers and Indian Management in a Zambian Town* (Manchester: Manchester University Press)
- [49] Krackhardt D 1987 *Soc. Netw.* **9** 109–34
- [50] Breiger R L and Pattison P E 1986 *Soc. Netw.* **8** 215–256
- [51] Roethlisberger F and Dickson W 1939 *Management and the Worker* (Cambridge: Cambridge University Press)
- [52] Rombach M P, Porter M A, Fowler J H and Mucha P J 2014 *SIAM J. Appl. Math.* **74** 167–90
- [53] <http://openflights.org/data.html> (accessed on 2013-07-03)
- [54] Onnela J-P, Saramäki J, Kertész J and Kaski K 2005 *Phys. Rev. E* **71** 065103
- [55] Barrat A, Barthélemy M, Pastor-Satorras R and Vespignani A 2004 *Proc. Natl Acad. Sci. USA* **101** 3747–52
- [56] Parshani R, Rozenblat C, Ietri D, Ducruet C and Havlin S 2010 *Europhys. Lett.* **92** 68002
- [57] Son S W, Grassberger P and Paczuski M 2011 *Phys. Rev. Lett.* **107** 195702
- [58] Son S W, Bizhani G, Christensen C, Grassberger P and Paczuski M 2012 *Europhys. Lett.* **97** 16006
- [59] Baxter G J, Dorogovtsev S N, Goltsev A V and Mendes J F F 2012 *Phys. Rev. Lett.* **109** 248701
- [60] Donges J F, Schultz H C H, Marwan N, Zou Y and Kurths J 2011 *Eur. Phys. J. B* **84** 635–51
- [61] Podobnik B, Horvatić D, Dickison M and Stanley H E 2012 *Europhys. Lett.* **100** 50004