Multiplexity and multireciprocity in directed multiplexes

Valerio Gemmetto,[1] Tiziano Squartini,[2] Francesco Picciolo,[3] Franco Ruzzenenti,[4,5,6] and Diego Garlaschelli[1]

[1]*Instituut-Lorentz for Theoretical Physics, Leiden Institute of Physics, University of Leiden, The Netherlands*
[2]*IMT School for Advanced Studies, Lucca, Italy*
[3]*Department of Biotechnology, Chemistry and Pharmacy, University of Siena, Italy*
[4]*Department of Economics and Statistics, University of Siena, Italy*
[5]*Institute of Sociology, Jagiellonian University, Krakow, Poland*
[6]*Department of Management, University of Naples "Parthenope," Naples, Italy*

Real-world multilayer networks feature nontrivial dependencies among links of different layers. Here we argue that if links are directed, then dependencies are twofold. Besides the ordinary tendency of links of different layers to align as the result of "multiplexity," there is also a tendency to antialign as a result of what we call "multireciprocity," i.e., the fact that links in one layer can be reciprocated by *opposite* links in a different layer. Multireciprocity generalizes the scalar definition of single-layer reciprocity to that of a square matrix involving all pairs of layers. We introduce multiplexity and multireciprocity matrices for both binary and weighted multiplexes and validate their statistical significance against maximum-entropy null models that filter out the effects of node heterogeneity. We then perform a detailed empirical analysis of the world trade multiplex (WTM), representing the import-export relationships between world countries in different commodities. We show that the WTM exhibits strong multiplexity and multireciprocity, an effect which is, however, largely encoded into the degree or strength sequences of individual layers. The residual effects are still significant and allow us to classify pairs of commodities according to their tendency to be traded together in the same direction and/or in opposite ones. We also find that the multireciprocity of the WTM is significantly lower than the usual reciprocity measured on the aggregate network. Moreover, layers with low (high) internal reciprocity are embedded within sets of layers with comparably low (high) mutual multireciprocity. This suggests that, in the WTM, reciprocity is inherent to groups of related commodities rather than to individual commodities. We discuss the implications for international trade research focusing on product taxonomies, the product space, and fitness and complexity metrics.

I. INTRODUCTION

Several real-world systems are composed of intricately interconnected units, thus exhibiting a nontrivial network structure. The behavior and dynamics of such systems are strongly dependent on how information can propagate throughout the network. Both the directionality and the intensity of connections crucially affect this process and should possibly be incorporated in the network description. For instance, most of the communication relations among individuals, such as exchanges of letters, e-mails, or texts, are intrinsically directional and are therefore best represented as directed networks [1]. Furthermore, such interactions typically have heterogeneous intensities, calling for a description in terms of weighted networks [2].

Recently, it has been realized that many real-world systems often require an even more detailed representation, because a given set of units can be connected by different kinds of relations. This property can be abstractly captured in terms of so-called *edge-colored graphs* (where links of different colors are allowed among the same set of nodes) or, equivalently, *multilayer* or *multiplex networks* (where the same set of nodes is replicated in multiple layers, each of which is an ordinary network) [3,4]. The nontrivial properties of these systems, with respect to ordinary single-layer ("monochromatic" or "monoplex") networks, arise from the facts that the various layers are interdependent and the presence of a link in one layer can influence the presence of a link in a different layer. A clear example is represented by the different kinds

of relationships existing between employees in a university department [5], where individuals can be connected by coauthorship, common leisure activities, online social networks, etc. The interdependence of layers implies that the topological properties usually defined for monoplex networks admit nontrivial generalizations to multiplex networks and that some properties which are uninteresting, or even undefined, for single-layer networks become relevant for multiplexes.

This paper introduces metrics characterizing the dependencies among layers in multiplexes *with directed links*. While various measures of interlayer overlap for multiplexes have already been introduced [6,7], they suffer from two main limitations. First, most definitions are available only for multiplexes with undirected links, and their straightforward generalization to the directed case would overshadow important properties that are inherent to directed networks, most importantly the reciprocity (which is one of our main focuses here). Second, even in "trivial" multiplexes where there is no dependence among layers (i.e., in independent superpositions of single-layer networks with the same set of nodes), a certain degree of interlayer overlap can be created entirely by chance. This effect becomes more pronounced as the density of the single-layer networks increases and as the correlation among single-node properties (like degrees or strengths) across layers increases. For instance, if a node is a hub in multiple layers, then there is an increased chance of overlap among these layers, even if the presence of links in one layer is assumed not to influence the presence of links in another layer.

The two limitations discussed above highlight the need to define metrics that appropriately filter out both global (network-wide) and local (node-specific) density effects. Correlation-based measures of interlayer overlap have been proposed with this aim in mind [8]. However, as recently pointed out [9], correlation-based metrics for multiplexes are not a correct solution in general, because they implicitly assume that edges observed between different pairs of nodes are sampled from the same probability distribution. This assumption is strongly violated in real-world networks, whose markedly heterogeneous topology is a signature of very different probabilities for edges emanating from different nodes, e.g., the probability of links being found around more important nodes clearly differs from the probability of links being found around less important nodes.

The above considerations motivate us to introduce new multiplexity metrics that explicitly take the directionality of links into account and appropriately filter out the spurious effects of chance, while controlling for the extreme heterogeneity of empirical node-specific properties. In this paper we carry out this program by extending recent "filtered" definitions of multiplexity [9], originally defined for undirected links, to the case of directed links. Although this might seem a straightforward procedure at first, we will in fact show that it requires different null models, triggers concepts, and leads to quantities that are undefined in the undirected case. Indeed, while in the undirected case there is only one possible notion of dependency among links in different layers, in the directed case there are two possibilities, depending on whether links are "aligned" or "anti-aligned."

Aligned links between two layers are observed when a directed link from node i to node j exists in both layers. This situation is the straightforward analog of what can happen in undirected multiplex networks and is a signature of the fact that the connection from i to j is relevant for multiple layers. We will denote this effect simply as *multiplexity*, in analogy with the undirected case [9], and will study it in the general case of an arbitrary number of layers. By contrast, antialigned links form between two layers when a link from node i to node j in one layer is *reciprocated* by an opposite link from node j to node i in the other layer. This situation does not have a counterpart in the case of undirected multiplexes and leads us to the definition of the concept of *multireciprocity*, i.e., the generalization of the popular concept of reciprocity to the case of multiplex networks.

In monoplex networks—either binary [10] or weighted [11]—reciprocity is defined as the tendency of vertex pairs to form mutual connections. This property, which is one of the best-studied properties of single-layer directed networks, can crucially affect various dynamical processes such as diffusion [12], percolation [13], and growth [14,15]. For instance, the presence of directed, reciprocal connections can lead to the establishment of functional communities and hierarchies of groups of neurons in the cerebral cortex [16].

In binary graphs, a simple measure of reciprocity is the ratio of the number of reciprocated links (i.e., realized links for which the link pointing in the opposite direction between the same two nodes is also realized) to the total number of directed links. However, it has been shown [10] that this measure is not *per se* informative about the actual tendency

towards reciprocation, because even in a random network a certain number of reciprocated links will appear. So the number of observed mutual interactions has to be compared with the expected number obtained for a given random null model if one wants to understand whether mutual links are present in the real network significantly more (or less) often than in the random benchmark [17]. It is therefore crucial to make use of proper null models for networks. Since in most real-world directed networks the distribution of the number of in-coming and out-going links (i.e., the in-degree and out-degree) of nodes is very broad, an appropriate null model should fix the in- and out-degrees of all nodes equal to their observed values. The null model of directed networks with given in- and out-degrees often goes under the name of the directed binary configuration model (DBCM) [18]. The rationale underlying the DBCM is the consideration that the in- and out-degree of a node might reflect some intrinsic "size," or other characteristic, of that node; therefore a null model tailored for a specific network should preserve the observed degree heterogeneity. Conveniently, the DBCM is also the correct null model to use when measuring the *multiplexity* among layers of a multiplex with directed links. Indeed, the DBCM is the directed generalization of the undirected binary configuration model used in Ref. [9] for the definition of appropriately filtered, undirected multiplexity metrics. This nicely implies that we can use the DBCM as a single null model in our analysis of both multiplexity and multireciprocity.

Recently, the definition of reciprocity has been extended to weighted networks [11]. A simple measure of weighted reciprocity is the ratio of "total reciprocated link weight" to total link weight, where the reciprocated link weight is defined, for any two reciprocated links, as the minimum weight of the two links. Similarly to the binary case, some level of weighted reciprocity can be generated purely by chance. So the empirical measure has to be compared to its expected value under a proper null model, represented in this case by a random weighted network where each node has the same in-strength and out-strength (i.e., total in-coming link weight and total out-going link weight, respectively) as in the real network. This null model is sometimes called the directed weighted configuration model (DWCM) [19] and, conveniently, is also the relevant null model (generalizing its undirected counterpart [9]) to study the multiplexity in presence of weighted directed links.

We stress that the concept of reciprocity has not been generalized to multiplex networks yet. Our definition of multireciprocity represents the first step in this direction and captures the tendency of a directed link in one layer of a multiplex to be reciprocated by an opposite link in a possibly *different* layer. While ordinary reciprocity can be quantified by a scalar quantity, multireciprocity requires a square matrix where all the possible pairs of layers are considered. Similarly, the multiplexity also requires a square matrix. Together, the multiplexity matrix and the multireciprocity matrix represent the two "directed" extensions of the undirected multiplexity matrix that has recently been introduced [9] to characterize undirected (either binary or weighted) multiplexes.

The rest of the paper is organized as follows. In Sec. II we introduce our methods, null models, and main definitions for both binary and weighted multiplexes. In Sec. III we apply our

techniques to the analysis of the world trade multiplex (WTM), a directed weighted multiplex representing the import-export relations between countries of the world in different products. We identify a number of empirical properties of the WTM that are impossible to access via the usual aggregate (monoplex) analysis of the network of total international trade. We finally conclude the paper in Sec. IV, where we discuss some important implications of our results, both for the general study of multiplex networks and for more specific research questions in international trade economics. Several necessary technical details are given in the Appendices.

II. MULTIPLEXITY AND MULTIRECIPROCITY METRICS

In this section we give definitions of multiplexity and multireciprocity metrics for both binary and weighted multiplexes. These definitions require, as a preliminary step, the introduction of appropriate null models. In turn, null models require the choice of a convenient notation. We address these points in the resulting order.

We represent a directed multiplex $\vec{G} = (G^1, \ldots, G^M)$ as the superposition of M-directed networks (layers) G^α ($\alpha = 1, \ldots, M$), all sharing the same set of N nodes [3]. Links can be either binary or weighted. In the binary case, each layer α is represented by a $N \times N$ binary adjacency matrix $G^\alpha = (a_{ij}^\alpha)_{i,j=1}^N$, where $a_{ij}^\alpha = 0, 1$ depending on whether a directed link from node i to node j is absent or present, respectively. In the weighted case, each layer α is represented by a $N \times N$ non-negative integer adjacency matrix $G^\alpha = (w_{ij}^\alpha)_{i,j=1}^N$, where $w_{ij}^\alpha = 0, 1, \ldots, \infty$ is the weight of the directed link from node i to node j ($w_{ij}^\alpha = 0$ indicating the absence of such link). We denote by $\mathcal{G}_N$ the set of all (binary or weighted) single-layer graphs with N nodes and by $\mathcal{G}_N^M \equiv (\mathcal{G}_N)^M$ the set of all (binary or weighted) M-layer multiplexes with N nodes.

A. Null models of multiplex networks: maximum entropy and maximum likelihood

Since our purpose is that of measuring correlations between directed links (possibly, in opposite directions) in different layers, we define independent reference models for each layer of the multiplex, thus creating an uncorrelated null model for the entire multiplex [7,9]. This means that if $\mathcal{P}(\vec{G} \mid \vec{\theta})$ denotes the joint probability of the entire multiplex $\vec{G} \in \mathcal{G}_N^M$ (given a set of constraints enforced via the vector $\vec{\theta}$ of parameters, see Appendix A) and

$$P^\alpha(G^\alpha \mid \vec{\theta^\alpha}) \equiv \underbrace{\sum_{G^1 \in \mathcal{G}_N} \cdots \sum_{G^\beta \in \mathcal{G}_N} \cdots \sum_{G^M \in \mathcal{G}_N}}_{\beta \neq \alpha} \mathcal{P}(\vec{G} \mid \vec{\theta}) \quad (1)$$

denotes the (marginal) probability for the single-layer graph $G^\alpha \in \mathcal{G}_N$ (given a set of layer-specific constraints enforced via the partial vector $\vec{\theta^\alpha}$), then we require the null model to obey the factorization property

$$\mathcal{P}(\vec{G} \mid \vec{\theta}) = \prod_{\alpha=1}^{M} P^\alpha(G^\alpha \mid \vec{\theta^\alpha}). \quad (2)$$

The above property ensures that the definition of the null model for the entire multiplex reduces to the definition of independent null models for each layer separately (see Appendix A for a rigorous derivation).

In the case of binary multiplexes, the null model we want to use to control for the heterogeneity of nodes in each layer is, as we have already mentioned, the DBCM [20,21], defined as the ensemble of binary networks with given in-degree and out-degree sequences. At this point, we have to make a major decision, since the DBCM can be implemented either microcanonically or canonically.

In the microcanonical approach, node degrees are "hard," i.e., enforced sharply on each realization. The most popular microcanonical implementation of the DBCM is based on the random degree-preserving rewiring of links [18] (or the local rewiring algorithm), which, unfortunately, introduces a bias. This bias arises because, if the degree distribution is sufficiently broad (as in most real-world cases), the randomization process explores the space of possible network configurations not uniformly, giving higher probability to the configurations that are "closer" to the initial one [22] (more details are given in Appendix B). Another possible microcanonical implementation, based on the random matching of "edge stubs" (half links) to the nodes, creates undesired self-loops and multiple edges [18,23]. Besides these limitations, microcanonical approaches are computationally demanding. Indeed, in order to measure the expected value of any quantity of interest, it is necessary to generate several randomized networks, on each of which the quantity needs to be calculated. This sampling method is per se very costly, and even more so in the case of multiplex networks, due to the presence of several layers requiring a further multiplication of iterations (see Appendix B).

By contrast, in the canonical implementation [20,21] of the DBCM the in- and out-degrees are "soft," i.e., preserved only on average. The resulting probability distribution over the ensemble of possible graphs is obtained analytically by maximizing the entropy subject to the enforced constraints [20,24–26] (see Appendix A for details). This procedure leads to the class of models also known as exponential random graphs or $p^\star$ models [27–29]. In order to fit such exponential random graphs to real-world networks, we adopt an exact, unbiased, and fast method [20,21] based on the maximum likelihood principle [30]. The method is summarized in Appendix B and implemented in our analysis using the so-called maximize-and-sample (MAX&SAM) algorithm [21]. The latter yields the exact probabilities of occurrence of any graph in the ensemble and the explicit expectation values of the quantities of interest. This has the enormous advantage that an explicit sampling of graphs is not required: expectation values are calculated analytically and not as sample averages. In particular, the probability p_{ij}^α that a link from node i to node j is realized in layer α ($a_{ij}^\alpha = 1$) can be easily calculated. From the set of all such probabilities, the expected value of the multireciprocity can be computed analytically and directly compared with the empirical value in order to obtain a filtered measure.

We now come to the case of multiplexes with weighted links. In this case, we want the enforced constraints to be the in-strength and out-strength sequences of the real network, separately for each layer. The corresponding model is

sometimes referred to as the DWCM [11]. As for the binary case, we want to build the null model canonically as a maximum-entropy ensemble of weighted networks, leading to a weighted exponential random graph model [11,20]. The implementation we use is again based on the MAX&SAM algorithm [21], which in this case calculates the exact probability that, in the null model, the weight of the directed link connecting node i to node j in layer α has a particular value w_{ij}^α for each pair of nodes and each layer. From this probability, the expected weighted multireciprocity can be computed analytically and compared with the empirical one, thus producing a filtered value that, in this case as well, does not require the explicit sampling of graphs.

B. Binary multiplexity and multireciprocity

Our first set of main definitions are specific for multiplexes with binary links. Consider a directed and binary multiplex $\vec{G}$ with M layers. We quantify the similarity and reciprocity between any two layers α and β by defining the binary *multiplexity* $m_b^{\alpha,\beta}$ and *multireciprocity* $r_b^{\alpha,\beta}$ as follows:

$$m_b^{\alpha,\beta} = \frac{2\sum_i \sum_{j\neq i} \min\{a_{ij}^\alpha, a_{ij}^\beta\}}{L^\alpha + L^\beta} = \frac{2L^{\alpha\rightrightarrows\beta}}{L^\alpha + L^\beta}, \quad (3a)$$

$$r_b^{\alpha,\beta} = \frac{2\sum_i \sum_{j\neq i} \min\{a_{ij}^\alpha, a_{ji}^\beta\}}{L^\alpha + L^\beta} = \frac{2L^{\alpha\rightleftarrows\beta}}{L^\alpha + L^\beta}, \quad (3b)$$

where $L^\alpha = \sum_i \sum_{j\neq i} a_{ij}^\alpha$ represents the total number of directed links in layer α (analogously for layer β), $L^{\alpha\rightrightarrows\beta} = \sum_i \sum_{j\neq i} \min\{a_{ij}^\alpha, a_{ij}^\beta\}$ is the number of links of layer α that are multiplexed in layer β (clearly, $L^{\alpha\rightrightarrows\beta} = L^{\beta\rightrightarrows\alpha}$), and $L^{\alpha\rightleftarrows\beta} = \sum_i \sum_{j\neq i} \min\{a_{ij}^\alpha, a_{ji}^\beta\}$ is the number of links of layer α that are reciprocated in layer β (clearly, $L^{\alpha\rightleftarrows\beta} = L^{\beta\rightleftarrows\alpha}$). Note that possible self-loops (terms of the type a_{ii}^α) are deliberately ignored because they are indistinguishable from links pointing in the opposite direction, thus making their contribution to either multiplexity or multireciprocity undefined.

Equations (3) can be regarded as defining the entries of two $M \times M$ matrices, which we will call the *binary multiplexity matrix* $\mathbf{M}_b = (m_b^{\alpha,\beta})_{\alpha,\beta}$ and the *binary multireciprocity matrix* $\mathbf{R}_b = (r_b^{\alpha,\beta})_{\alpha,\beta}$, respectively. The matrices $\mathbf{M}_b$ and $\mathbf{R}_b$ represent the two natural extensions, to the case of directed multiplexes, of the single binary multiplexity matrix introduced in Ref. [9] for undirected binary multiplexes. Both matrices provide information about the "overlap" between directed links connecting pairs of nodes in different layers. Their entries range in $[0,1]$ and are maximal only when layers α and β are, respectively, identical (i.e., $a_{ij}^\alpha = a_{ij}^\beta$ for all $i \neq j$) and fully "multireciprocated" (i.e., $a_{ij}^\alpha = a_{ji}^\beta$ for all $i \neq j$). The matrix $\mathbf{M}_b$ has by construction a unit diagonal, since the intralayer multiplexity trivially has the maximum value $m_b^{\alpha,\alpha} = 1$ for all α. By contrast, the diagonal of $\mathbf{R}_b$ is nontrivial and of special significance, as the intralayer multireciprocity $r_b^{\alpha,\alpha}$ reduces to the ordinary definition of binary reciprocity for monoplex networks [10].

For "trivial," uncorrelated multiplexes made of sparse noninteracting layers with narrow degree distributions, the matrix $\mathbf{M}_b$ would asymptotically (i.e., in the limit of large

N but not necessarily large M) be the $M \times M$ identity matrix, and the matrix $\mathbf{R}_b$ would asymptotically be a $M \times M$ diagonal matrix. This is because, in the presence of sparse uncorrelated layers without hubs, the chance of a link in one layer "overlapping" with a (mutual) link in a different layer is negligible. For finite and/or dense networks and/or broad degree distributions, however, positive values of $m_b^{\alpha,\beta}$ and $r_b^{\alpha,\beta}$ (with $\alpha \neq \beta$) can be produced entirely by chance even in a multiplex with no dependencies among layers. For instance, if the same node is a hub in multiple layers, then the chance of a large overlap of links among all pairs of such layers is very high, even if the layers are noninteracting.

The above considerations imply that, in order to extract statistically significant information about the tendency towards multiplexity and multireciprocity in a real-world multiplex, it becomes necessary to compare the empirical values of $m_b^{\alpha,\beta}$ and $r_b^{\alpha,\beta}$ with the corresponding expected values calculated under the chosen null model of independent multiplexes with given degrees (i.e., the DBCM). Hence, we introduce the *transformed* binary multiplexity and multireciprocity matrices with entries

$$\mu_b^{\alpha,\beta} = \frac{m_b^{\alpha,\beta} - \langle m_b^{\alpha,\beta}\rangle_{\text{DBCM}}}{1 - \langle m_b^{\alpha,\beta}\rangle_{\text{DBCM}}} \quad (\alpha \neq \beta), \qquad (4a)$$

$$\rho_b^{\alpha,\beta} = \frac{r_b^{\alpha,\beta} - \langle r_b^{\alpha,\beta}\rangle_{\text{DBCM}}}{1 - \langle r_b^{\alpha,\beta}\rangle_{\text{DBCM}}}, \qquad (4b)$$

where $\langle\cdot\rangle_{\text{DBCM}}$ denotes the expected value under the DBCM. Note that, since $\langle m_b^{\alpha,\alpha}\rangle_{\text{DBCM}} = m_b^{\alpha,\alpha} = 1$ for all α, we formally set the diagonal terms $\mu_b^{\alpha,\alpha} \equiv 1$, as the definition (4a) would produce an indeterminate expression if extended to $\alpha = \beta$. The explicit calculation of the above expected values is provided in Appendix C and more details are provided later in this section.

The filtered quantities (4) are directly informative about the presence of dependencies between layers. Positive values represent higher-than-expected multiplexity or multireciprocity (correlated or "attractive" pairs of layers), while negative values represent lower-than-expected quantities (anticorrelated or "repulsive" pairs of layers). Pairs of uncorrelated ("noninteracting") layers are characterized by multiplexity and multireciprocity values comparable with 0. In principle, a layer that is uncorrelated with all other layers can be separated from the multiplex and analzed separately from it.

The choice of the denominator of (4a) and (4b), *a priori* not obvious, guarantees that the maximum value for the transformed multiplexity and multireciprocity is 1. Moreover, it ensures that $\rho_b^{\alpha,\alpha}$ reduces to the rescaled reciprocity ρ_b defined for single-layer networks [10]. It should also be noted that the multiplexity defined in (3a) is just the normalized version of the interlayer overlap introduced in Refs. [6] and [7], extended to directed multiplex networks. In this context, the contribution that we give is the comparison with a null model. Indeed, while (3a) only provides information about the raw similarity of the layers, which is strongly density dependent, the transformed measure (4a) is mapped to a universal interval. In combination with the z scores that we introduce later, it can be used to consistently compare the statistical significance of the multiplexity of different systems. The quantity defined in

(3b), which focuses explicitly on the reciprocity properties of the multiplex, has never been introduced before, along with its transformed quantity defined in (4b). The latter can be used for a consistent comparison of the multireciprocity of mutiplexes with different densities.

The calculation of the expected values of $m_b^{\alpha,\beta}$ and $r_b^{\alpha,\beta}$ under the DBCM can be carried out analytically using the MAX&SAM method [21], with no need to actually randomize the empirical network or numerically sample the null model ensemble. Ultimately, the calculation requires the computation of the expected value of the minimum between two binary random variables (see Appendix C). If $p_{ij}^{\alpha} \equiv \langle a_{ij}^{\alpha} \rangle_{\text{DBCM}}$ denotes the probability that, under the DBCM, a directed link is realized from node i to node j in layer α, then the adjacency matrix entry a_{ij}^{α} is described by the Bernoulli mass probability function,

$$P\left(a_{ij}^{\alpha}\right) = \left(p_{ij}^{\alpha}\right)^{a_{ij}^{\alpha}}\left(1 - p_{ij}^{\alpha}\right)^{\left(1 - a_{ij}^{\alpha}\right)}. \qquad (5)$$

Using the above equation, and given the explicit expression for p_{ij}^{α}, it is possible to calculate $\mu_b^{\alpha,\beta}$ and $\rho_b^{\alpha,\beta}$ analytically as reported in Appendix C.

It is instructive to compare the multivariate quantities measured on the multiplex with the corresponding scalar quantities defined on the aggregate monoplex network obtained by combining all layers together. This comparison can highlight the gain of information resulting from the multiplex representation, with respect to the ordinary monoplex projection where all the distinct types of links are treated as equivalent. The binary aggregate monoplex can be defined in terms of the adjacency matrix with entries

$$a_{ij}^{\text{mono}} = 1 - \prod_{\alpha=1}^{M}\left(1 - a_{ij}^{\alpha}\right) = \begin{cases} 1 & \text{if } \exists \alpha : a_{ij}^{\alpha} = 1 \\ 0 & \text{otherwise} \end{cases}. \qquad (6)$$

For the quantities we defined so far, the only meaningful comparison between the multiplex and the aggregate network can be done in terms of the reciprocity, because the multiplexity of the aggregate is $m_b^{\text{mono}} = 1$ by construction. The single, global reciprocity of the aggregated monoplex network is given by

$$r_b^{\text{mono}} = \frac{\sum_i \sum_{j \neq i} \min\left\{a_{ij}^{\text{mono}}, a_{ji}^{\text{mono}}\right\}}{L^{\text{mono}}}, \qquad (7)$$

where $L^{\text{mono}} = \sum_i \sum_{j \neq i} a_{ij}^{\text{mono}}$. Similarly, it is possible to define the corresponding filtered quantity ρ_b^{mono}, in analogy with (4b).

The transformed quantities $\mu_b^{\alpha,\beta}$ and $\rho_b^{\alpha,\beta}$ defined in (4) capture the similarity and reciprocity between layers of a multiplex via a comparison of the empirical values with the expected values under a null model. However, those quantities do not consider any information about the variances of the values of multiplexity and multireciprocity under the null model, thus giving no direct information about statistical significance. In particular, even multiplexes sampled from the null model with independent layers would be characterized by small, but in general nonzero, values of $\mu_b^{\alpha,\beta}$ and $\rho_b^{\alpha,\beta}$. This makes it difficult to disentangle, for an observed real-world multiplex, weak interlayer dependencies from pure noise. Moreover, the random fluctuations around the expectation values will in general differ for different pairs of layers,

potentially making the comparison of the values of $\mu_b^{\alpha,\beta}$ and $\rho_b^{\alpha,\beta}$ for different pairs of layers misleading. To overcome these limitations, we define the z scores associated to $m_b^{\alpha,\beta}$ and $r_b^{\alpha,\beta}$ as:

$$z\left(m_b^{\alpha,\beta}\right) = \frac{m_b^{\alpha,\beta} - \langle m_b^{\alpha,\beta} \rangle_{\text{DBCM}}}{\sqrt{\langle (m_b^{\alpha,\beta})^2 \rangle_{\text{DBCM}} - \langle m_b^{\alpha,\beta} \rangle_{\text{DBCM}}^2}} \quad (\alpha \neq \beta), \quad (8a)$$

$$z\left(r_b^{\alpha,\beta}\right) = \frac{r_b^{\alpha,\beta} - \langle r_b^{\alpha,\beta} \rangle_{\text{DBCM}}}{\sqrt{\langle (r_b^{\alpha,\beta})^2 \rangle_{\text{DBCM}} - \langle r_b^{\alpha,\beta} \rangle_{\text{DBCM}}^2}}. \quad (8b)$$

As for the quantities defined in (4), it is possible to obtain an analytical expression for the z scores as well. This is shown in detail in Appendix C.

Each z score in (8) has the same sign as the corresponding quantity in (4), since the numerator is the same and both have positive denominators. However, except for the common sign, the two sets of quantities can have *a priori* very different values. In particular, the z scores count the number of standard deviations by which the observed raw quantities deviate from their expected values under the null model. As such, they are useful in order to understand whether small measured values of $\mu_b^{\alpha,\beta}$ or $\rho_b^{\alpha,\beta}$ are actually consistent with zero within a small number of standard deviations, in which case we can consider the layers α and β as uncorrelated. We point out that, in general, z scores have a clear statistical interpretation only if their distribution is Gaussian under repeated realizations of the model. In our case, although the quantities $m_b^{\alpha,\beta}$ and $r_b^{\alpha,\beta}$ are not truly normally distributed under the null model, they are defined as the sum of many independent 0/1 random variables (of the type $\min\{a_{ij}^{\alpha}, a_{ij}^{\beta}\}$ or $\min\{a_{ij}^{\alpha}, a_{ji}^{\beta}\}$, respectively), which all have variance in the interval $(0, 1/4]$ and are thus approximately described by a central limit theorem ensuring an asymptotic convergence to the normal distribution. We can therefore consider as statistically significant all the z scores having an absolute value larger than a given threshold, which we set at $z_c = 2$. This selects the observed pairs of layers with values of multiplexity and/or multireciprocity that differ from their expectation values by more than 2 standard deviations, i.e., with $|z| > z_c$.

C. Weighted multiplexity and multireciprocity

We now move to our second set of definitions, valid for weighted multiplexes. In analogy with (3), we define the *weighted* multiplexity and multireciprocity matrices $\mathbf{M}_w$ and $\mathbf{R}_w$ having entries

$$m_w^{\alpha,\beta} = \frac{2\sum_i \sum_{j \neq i} \min\left\{w_{ij}^{\alpha}, w_{ij}^{\beta}\right\}}{W^{\alpha} + W^{\beta}} = \frac{2W^{\alpha \Rightarrow \beta}}{W^{\alpha} + W^{\beta}}, \quad (9a)$$

$$r_w^{\alpha,\beta} = \frac{2\sum_i \sum_{j \neq i} \min\left\{w_{ij}^{\alpha}, w_{ji}^{\beta}\right\}}{W^{\alpha} + W^{\beta}} = \frac{2W^{\alpha \rightleftarrows \beta}}{W^{\alpha} + W^{\beta}}, \quad (9b)$$

where $W^{\alpha} = \sum_i \sum_{j \neq i} w_{ij}^{\alpha}$ is the total weight of the links in layer α (analogously for layer β), $W^{\alpha \Rightarrow \beta} = \sum_i \sum_{j \neq i} \min\{w_{ij}^{\alpha}, w_{ij}^{\beta}\}$ is the total link weight of layer α that is multiplexed in layer β (clearly, $W^{\alpha \Rightarrow \beta} = W^{\beta \Rightarrow \alpha}$), and

$W^{\alpha \rightleftarrows \beta} = \sum_i \sum_{j \neq i} \min\{w_{ij}^{\alpha}, w_{ji}^{\beta}\}$ is the total link weight of layer α that is reciprocated in layer β (clearly, $W^{\alpha \rightleftarrows \beta} = W^{\beta \rightleftarrows \alpha}$). The matrices $\mathbf{M}_w$ and $\mathbf{R}_w$ represent the two generalizations, for directed multiplexes, of the weighted multiplexity matrix introduced in Ref. [9] for undirected weighted multiplexes. Like their binary counterparts, both matrices have entries in the range [0,1], the maximum value being attained by identical ($w_{ij}^{\alpha} = w_{ij}^{\beta}$ for all i,j) and fully "multireciprocated" ($w_{ij}^{\alpha} = w_{ji}^{\beta}$ for all i,j) layers, respectively. In analogy with the corresponding binary case, the diagonal of $\mathbf{M}_w$ has all unit entries while that of $\mathbf{R}_w$ has entries that coincide with the recent definition of reciprocity for weighted monoplex networks [11].

In this case as well, for trivial multiplexes with sparse noninteracting layers and narrow strength distributions, the two matrices are expected to be asymptotically diagonal. However, this is no longer true in the presence of dense layers and/or for broad strength distributions, and we therefore need a comparison of the raw quantities with their expected value under a null model (now the DWCM). This consideration leads us to introduce the transformed weighted multiplexity and multireciprocity matrices with entries

$$\mu_w^{\alpha,\beta} = \frac{m_w^{\alpha,\beta} - \langle m_w^{\alpha,\beta} \rangle_{\text{DWCM}}}{1 - \langle m_w^{\alpha,\beta} \rangle_{\text{DWCM}}} \quad (\alpha \neq \beta), \tag{10a}$$

$$\rho_w^{\alpha,\beta} = \frac{r_w^{\alpha,\beta} - \langle r_w^{\alpha,\beta} \rangle_{\text{DWCM}}}{1 - \langle r_w^{\alpha,\beta} \rangle_{\text{DWCM}}}, \tag{10b}$$

where $\langle \cdot \rangle_{\text{DWCM}}$ denotes the expected value under the DWCM. As in the binary case, we can derive an analytical expression for the expected values that ultimately requires the expectation of the minimum of w_{ij}^{α} and w_{ij}^{β} (or w_{ji}^{β}). This is done in Appendix D. It turns out that, under the DWCM, the distribution of link weights is geometrical [11,21]:

$$P(w_{ij}^{\alpha}) = (p_{ij}^{\alpha})^{w_{ij}^{\alpha}} (1 - p_{ij}^{\alpha}), \tag{11}$$

where p_{ij}^{α} denotes again the probability that a directed link (of any positive weight) from node i to node j is realized in layer α. The above probability can be used to calculate $\mu_w^{\alpha,\beta}$ and $\rho_w^{\alpha,\beta}$ analytically as discussed in Appendix D.

The weighted multireciprocity of the multiplex can be conveniently compared with the weighted reciprocity of the aggregated monoplex network. The link weights of the latter are defined by

$$w_{ij}^{\text{mono}} = \sum_{\alpha=1}^{M} w_{ij}^{\alpha}, \tag{12}$$

and the associated aggregate weighted reciprocity [11] is

$$r_w^{\text{mono}} = \frac{\sum_i \sum_{j \neq i} \min\{w_{ij}^{\text{mono}}, w_{ji}^{\text{mono}}\}}{W^{\text{mono}}} \tag{13}$$

(where $W^{\text{mono}} = \sum_i \sum_{j \neq i} w_{ij}^{\text{mono}}$). The corresponding filtered value ρ_w^{mono} can be defined as in (10b).

In analogy with the binary case, it is possible to define the z scores associated to $m_w^{\alpha,\beta}$ and $r_w^{\alpha,\beta}$ as follows:

$$z(m_w^{\alpha,\beta}) = \frac{m_w^{\alpha,\beta} - \langle m_w^{\alpha,\beta} \rangle_{\text{DWCM}}}{\sqrt{\langle (m_w^{\alpha,\beta})^2 \rangle_{\text{DWCM}} - \langle m_w^{\alpha,\beta} \rangle_{\text{DWCM}}^2}} \quad (\alpha \neq \beta), \tag{14a}$$

$$z(r_w^{\alpha,\beta}) = \frac{r_w^{\alpha,\beta} - \langle r_w^{\alpha,\beta} \rangle_{\text{DWCM}}}{\sqrt{\langle (r_w^{\alpha,\beta})^2 \rangle_{\text{DWCM}} - \langle r_w^{\alpha,\beta} \rangle_{\text{DWCM}}^2}}. \tag{14b}$$

The explicit analytical expressions for these z scores are calculated in Appendix D. Again, the z scores (14) have the same signs as the corresponding quantities (10), but in addition they allow us to test for statistical significance using, e.g., a threshold of $z_c = 2$.

III. EMPIRICAL ANALYSIS OF THE WORLD TRADE MULTIPLEX

In this section, we apply the framework defined so far to the analysis of a real-world system. This system is the WTM, defined as the multilayer network representing the directed trade relations between world countries in different commodities. At both the binary and the weighted levels, the structure of the aggregate (monoplex) version of this network is well studied [31–33], as well as that of many of its layers separately [8,34]. However, much less is known about the interlayer dependencies in the WTM. In particular, an assessment of the interlayer couplings that are not simply explained by the local topological properties of the WTM has been carried out only for the undirected version of the network [9]. Given the importance of the directionality of trade flows, especially at the disaggregated level of individual commodities, it is therefore important to carry out a directed analysis of the WTM. The tools we have introduced in the previous section allow us to make this step and arrive at a novel characterization of the WTM where the undirected multiplexity properties documented in Ref. [9] are resolved into their two directed components, namely multiplexity and multireciprocity. These results have important potential implications for problems related to research on international trade, such as the definition of trade-based "product taxonomies" [8], the construction of the "product space" [35], and the calculation of "fitness and complexity" metrics [36]. These points are discussed later in Sec. IV.

A. Data

We use the BACI-Comtrade data set [37] where international trade flows among all countries of the world are disaggregated into different commodity classes at the two-digit resolution level, defined as in the standard HS1996 classification [38] of traded goods. It is possible to represent this data set as a multiplex as in Refs. [8,9,34]. In particular, we will consider a multilayer representation defined by $N = 207$ nodes (countries) and $M = 96$ layers (commodities) for the year 2011. Since each trade exchange is reported by both the importer and the exporter (and the two values may in general differ), the data set uses a reconciliation procedure to get a unique value for each flow (see Ref. [37] for details). All the resulting trade volumes are expressed in thousands of dollars in the data set. Since our approach works for integer link
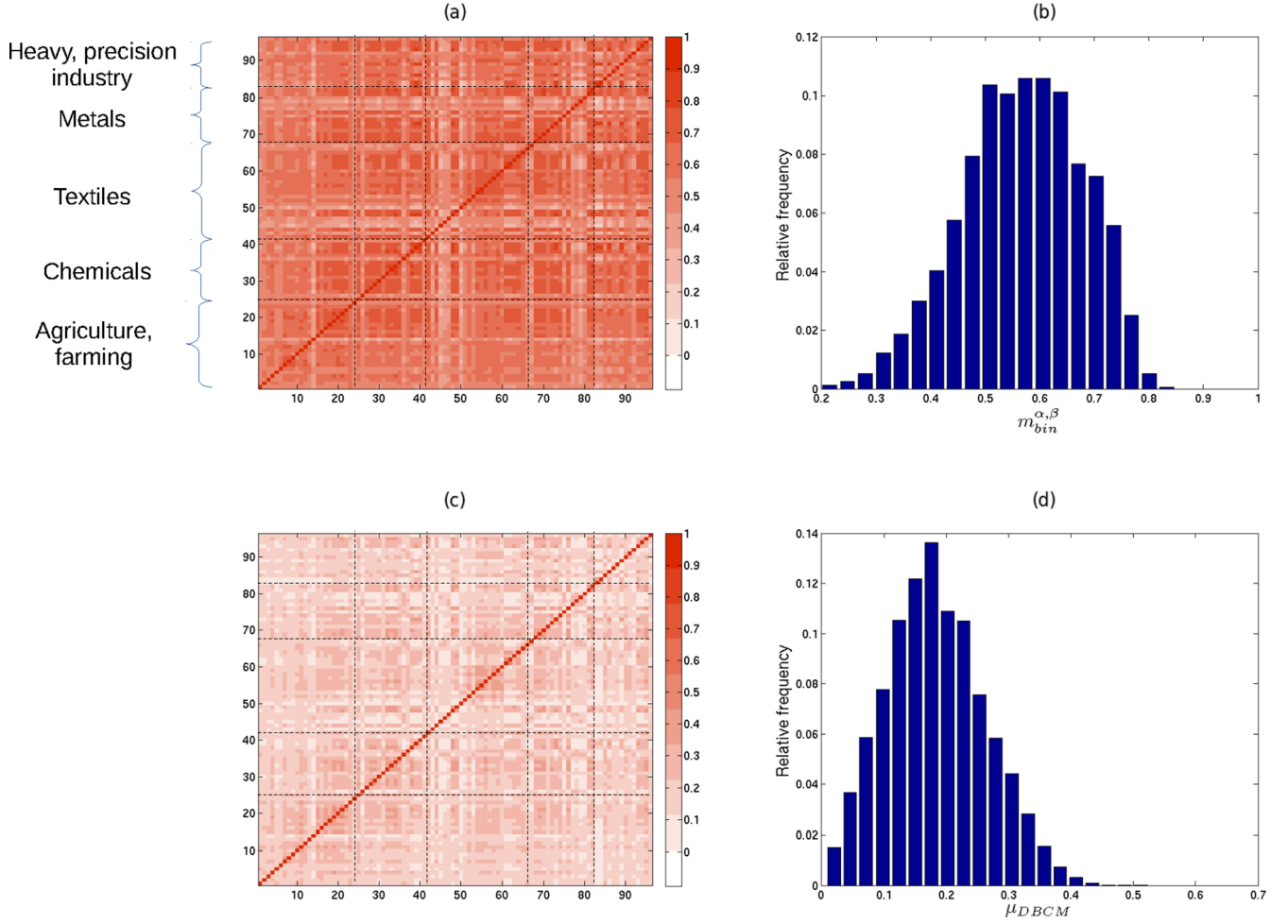
FIG. 1. Top panels: Color-coded binary multiplexity matrix $\mathbf{M}_b$ (a) and corresponding distribution of off-diagonal multiplexity values $m_b^{\alpha,\beta}$ (with $\alpha \neq \beta$) (b). Bottom panels: Same as for the top panels but with raw binary multiplexity $m_b^{\alpha,\beta}$ replaced by rescaled binary multiplexity $\mu_b^{\alpha,\beta}$.

weights, all the reported trade values have been rescaled by first dividing by 10 and then rounding to the closest integer. This defines our integer link weights $\{w_{ij}^\alpha\}$ for all layers. For each entry w_{ij}^α, we then define $a_{ij}^\alpha = 1$ if $w_{ij}^\alpha > 0$ and $a_{ij}^\alpha = 0$ otherwise. We point out that the rounding procedure does not significantly affect the structure of the system under study, as the percentage of original links which are lost (i.e., rounded to zero) is negligible.

From the multiplex trade flows we also compute the aggregate binary and weighted links a_{ij}^{mono} and w_{ij}^{mono} between any two countries i and j in the collapsed monoplex trade network, as in (6) and (12), respectively. This allows us to compare the multiplex structure of trade with the aggregate one and highlight relevant information that is lost in the aggregation procedure. For instance, for both the binary and the weighted representation of the system, we can compare the values of the multireciprocity matrix measured on the commodity-resolved multiplex with the usual scalar reciprocity measured on the monoplex aggregate trade network.

B. Binary analysis

We start with a binary analysis of the WTM, thus taking into account only the topology of the various layers while disregarding the information about trade volumes. In Fig. 1(a) we show the color-coded binary multiplexity matrix $\mathbf{M}_b$. Next to it, in Fig. 1(b), we show the corresponding frequency distribution of off-diagonal matrix entries $m_b^{\alpha,\beta}$ (with $\alpha \neq \beta$). In calculating the frequencies, we discard the diagonal entries because they trivially evaluate to $m_b^{\alpha,\alpha} = 1$, as discussed above. High values of multiplexity are observed for most of the pairs of commodities. This result is in agreement with what has been reported in Ref. [9] on the basis of an undirected analysis of the WTM where imports and exports between any two countries were combined together into a single trade link.

As we mentioned, the multiplexity matrix $\mathbf{M}_b$ would be asymptotically diagonal for trivial multiplexes with sparse noninteracting layers and narrow degree distributions. However, since the layers of the WTM are very dense and their degree distributions significantly broad [8,9,34], this system is an ideal case study requiring the use of a null model in order to assess the presence of a genuine coupling among layers. In Fig. 1(c), we show the color-coded matrix of rescaled multiplexity values $\mu_b^{\alpha,\beta}$, which control for the effects of the heterogeneity of the layer-specific in- and out-degree sequences. Similarly, in Fig. 1(d), we show the corresponding distribution of off-diagonal entries. We find
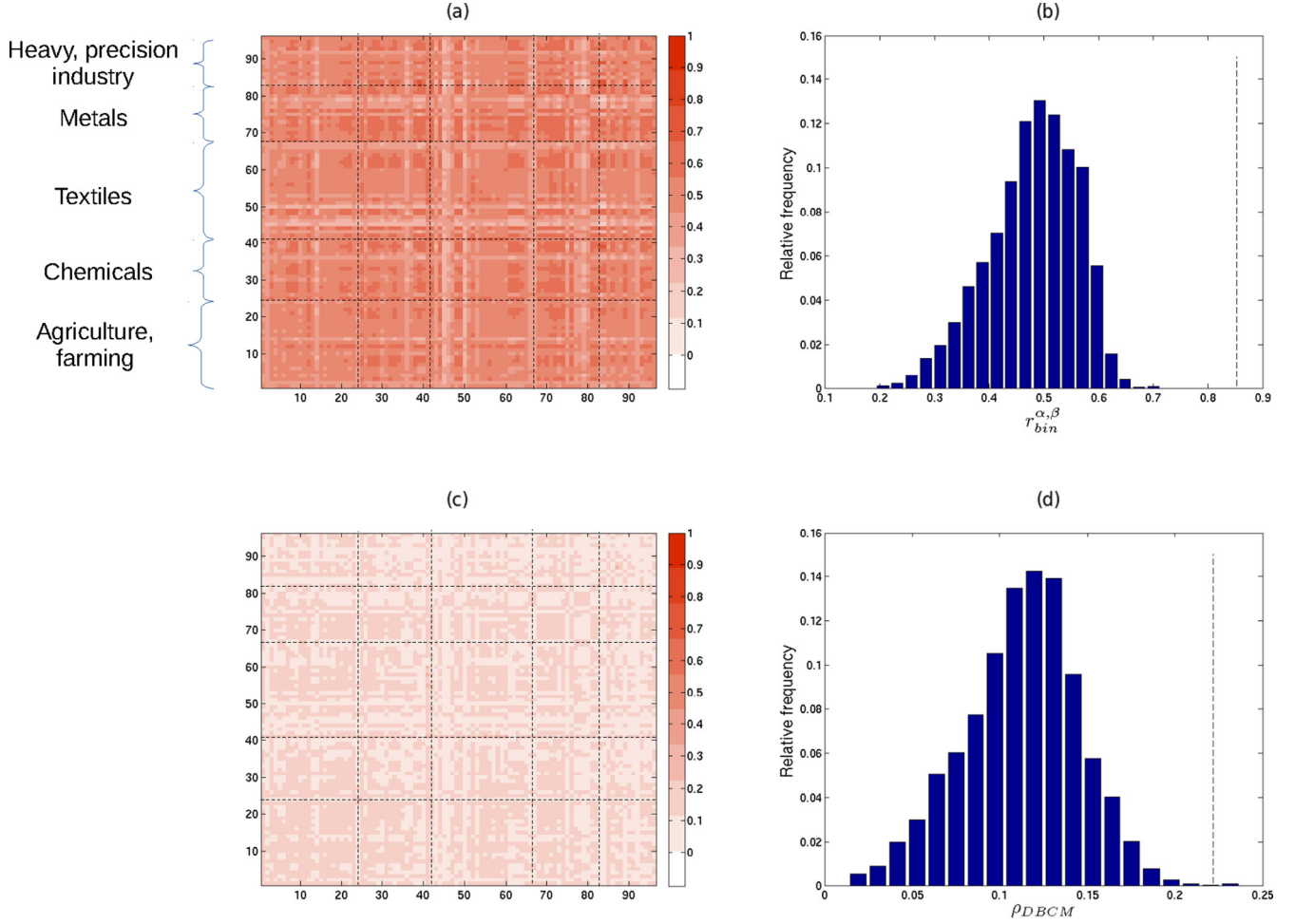
FIG. 2. Top panels: Color-coded binary multireciprocity matrix $\mathbf{R}_b$ (a) and corresponding distribution of off-diagonal multireciprocity values $r_b^{\alpha,\beta}$ (with $\alpha \neq \beta$) (b). Bottom panels: Same as for the top panels but with raw binary multireciprocity $r_b^{\alpha,\beta}$ replaced by rescaled binary multireciprocity $\rho_b^{\alpha,\beta}$. The dashed lines represent the value of (raw and rescaled) binary reciprocity r_b^{mono} and ρ_b^{mono} of the aggregated monoplex network.

that, after controlling for the degrees, a significant amount of correlation is destroyed. However, all the values are still strictly positive, indicating a tendency of all pairs of commodities to be "traded together." The statistical significance of this result is discussed later in terms of z scores.

We now move to the analysis of multireciprocity. It is known that when the aggregate trade in all commodities is considered, the binary monoplex representation of the world trade network exhibits a high level of reciprocity [10,39,40]. It is interesting to see whether such property is preserved also at the multiplex level and how the values compare with the aggregate case. Figure 2(a) shows the color-coded binary multireciprocity matrix $\mathbf{R}_b$ and Fig. 2(b) the corresponding distribution of off-diagonal entries [41], with a superimposed δ function indicating the value of the binary reciprocity r_b^{mono} of the aggregate monoplex network as a comparison. The results are comparable with those found above for the multiplexity. Also in this case, the high multireciprocity values are consistent with the high multiplexity values found for the undirected representation of the WTM [9] (where pairs of reciprocated links in each layer are merged into single undirected links). However, for the multireciprocity this result is much less

trivial than for the multiplexity, given the chosen level of disaggregation into many commodity classes. Indeed, one would expect that, at such a relatively high resolution, it should be not very likely (at least not as likely as in the undirected representation) that the same commodity is traded "back and forth," i.e., both ways between the same two countries. In any case, we do find, in accordance with what we expect, that for all pairs of commodities the multireciprocity is significantly smaller than the reciprocity r_b^{mono} of the aggregate monoplex. This means that, as layers are aggregated, there is a bigger relative increment (with respect to individual layers) in the overall number of reciprocated links than in the total number of links.

As an interesting result, the intralayer reciprocity values $r_b^{\alpha,\alpha}$ lying along the diagonal of the multireciprocity matrix are found to be very similar to the values of the matrix entries $r_b^{\alpha,\beta}$ lying close to the diagonal. Indeed, in the matrix plot of Fig. 2(a) the diagonal is visually indistinguishable from the entries of the matrix that are "nearby." Given the order of the commodities in the matrix (as shown for instance in Ref. [9]), these nearby entries represent the multireciprocity between pairs of similar commodities. This
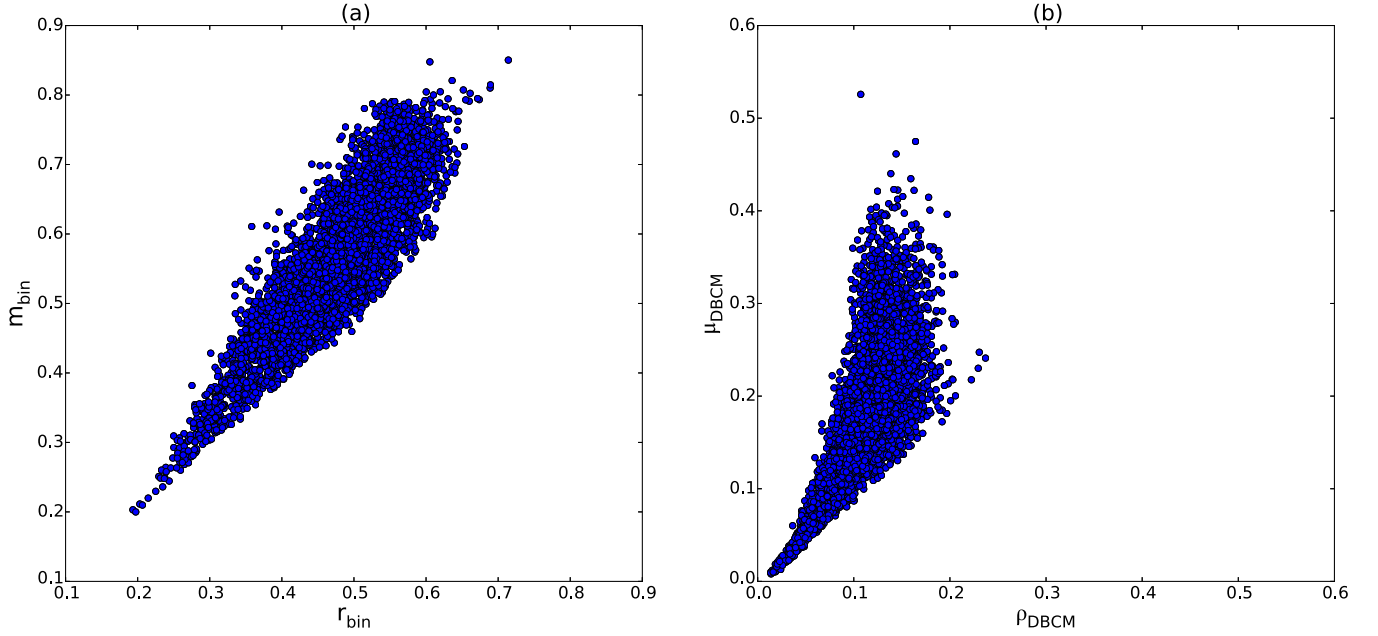
FIG. 3. Scatter plots of off-diagonal binary multireciprocity values versus off-diagonal binary directed multiplexity values. Left: Raw values ($r_b^{\alpha,\beta}$ vs. $m_b^{\alpha,\beta}$); right: rescaled values ($\rho_b^{\alpha,\beta}$ vs. $\mu_b^{\alpha,\beta}$).

result means that the high reciprocity of the aggregate trade monoplex does *not* arise from the superposition of layers with high internal reciprocity and low mutual multireciprocity (as would be the case in presence of an approximately diagonal multireciprocity matrix). Rather, we find that a trade flow in one commodity α tends to be reciprocated by comparable trade flows in several different commodities, including (but not dominated by) the same commodity α and many other related commodities. Specifically, it can be seen from Fig. 2(a) that layers characterized by low (high) values of internal reciprocity are embedded within groups of layers with low (high) mutual multireciprocity. This suggests that the level of reciprocity in international trade is not an intrinsic property of individual commodities but rather a property of whole groups of mutually reciprocated commodities with comparable multireciprocity values.

In Figs. 2(c) and 2(d) we show the color-coded binary rescaled multireciprocity matrix and the corresponding distribution of off-diagonal entries $\rho_b^{\alpha,\beta}$ (with $\alpha \neq \beta$). The relatively small values (with respect to the non-rescaled quantities) indicate that, in analogy with what we found for the multiplexity, the apparent correlation between the topology of pairs of layers is largely encoded in the relatedness of the degree sequences of such pairs. For the vast majority of pairs of commodities the multireciprocity is still lower than that measured on the aggregate network. However, all pairs of layers preserve a positive residual multireciprocity, the statistical significance of which is studied later in our z-score analysis.

When we look at the multiplexity matrix in Fig. 1(a) and the corresponding multireciprocity matrix in Fig. 2(a), we see the appearance of similar patterns. Such similarity is further investigated in Fig. 3(a), where we report the scatter plots of pairwise multireciprocity values versus the corresponding multiplexity values. We observe a roughly linear trend, which

is, however, lost when we look at the filtered values, as shown in Fig. 3(b). We see that, in the latter case, the relationship between $\rho_b^{\alpha,\beta}$ and $\mu_b^{\alpha,\beta}$ is nonlinear and significantly scattered. Although the presence of a nonlinear relation may be related to the particular choice of normalization adopted in (4), we point out that the entity of the scatter is so big that it is not possible to retrieve the value of multiplexity from the multireciprocity, and vice versa. This illustrates that the two quantities convey different pieces of information that are irreducible to each other.

Similar considerations apply to the z scores. In Figs. 4(a) and 4(b) we show the empirical relation between the transformed multiplexity and multireciprocity and their corresponding z scores. It is worth recalling that the information provided by these two quantities can differ *a priori*, given the lack of information about the standard deviation in the rescaled multiplexity and multireciprocity metrics. Empirically, we, however, find a strong correlation between these quantities, indicating that large values of binary multiplexity or multireciprocity correspond to large z scores, and vice versa. Moreover, even the smallest z scores (those found for the pairs of layers showing very low multiplexity or multireciprocity) are still quite high (i.e., positive and larger than $z_c = 2$) in terms of statistical significance. This means that even the pairs of layers with smallest multiplexity or multireciprocity should be considered as significantly and positively correlated. We therefore conclude that, at a binary level, every commodity of the WTM tends to be traded together with all other commodities, both in the same and in the opposite directions. As we show below, this is no longer the case when the weighted version of the multiplex is considered.

Figure 4(c) shows the relation existing between $z(r_b^{\alpha,\beta})$ and $z(m_b^{\alpha,\beta})$ for each pair of layers. If we compare this figure with Fig. 3, then we see that in this case the trend is more linear, although the scatter is again quite large. This confirms that it
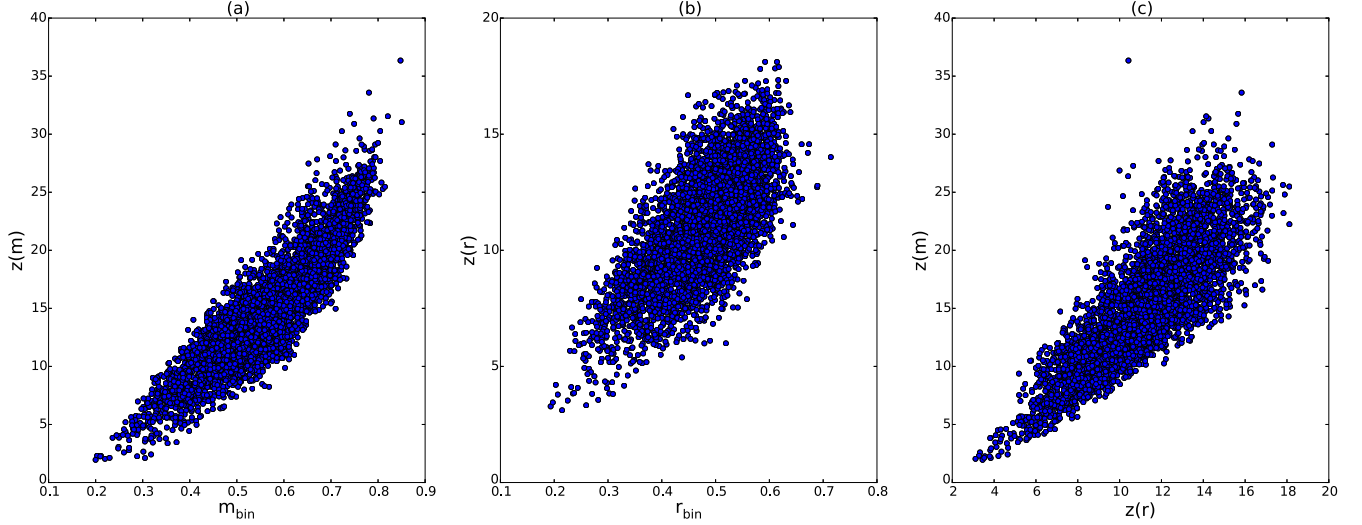
FIG. 4. Left: Binary transformed multiplexity $\mu_b^{\alpha,\beta}$ versus its corresponding z score $z(m_b^{\alpha,\beta})$; center: binary transformed multireciprocity $\rho_b^{\alpha,\beta}$ versus its corresponding z score $z(r_b^{\alpha,\beta})$; right: $z(r_b^{\alpha,\beta})$ vs. $z(m_b^{\alpha,\beta})$. Only off-diagonal values are reported.

is not possible to recover the values of multiplexity from those of multireciprocity and vice versa.

C. Weighted analysis

We now perform a weighted analysis of the world trade multiplex by taking into account the value of imports and exports observed between countries.

In Figs. 5(a) and 5(b) we show the color-coded weighted directed multiplexity matrix $\mathbf{M}_w$ and the distribution of its off-diagonal entries. We clearly see that, even though several pairs of commodities are still strongly overlapping, the multiplexity distribution is concentrated over a range of significantly smaller values with respect to the corresponding binary distribution. Indeed, the notion of weighted multiplexity, by involving the minimum of the weights of two reciprocated links, provides a stricter criterion with respect to the unweighted case. In particular, for any pair of nodes and any pair of layers, it is more unlikely to achieve the maximum weighted value $\min\{w_{ij}^\alpha, w_{ij}^\beta\}$ than the maximum binary value $\min\{a_{ij}^\alpha, a_{ij}^\beta\}$. Lower values of multiplexity with respect to Fig. 1(a) are therefore expected. We also expect to find a similar reduction for the multireciprocity later.

In Figs. 5(c) and 5(d) we report the color-coded weighted rescaled multiplexity matrix and the corresponding distribution of off-diagonal entries $\mu_w^{\alpha,\beta}$. The fact that many values are now mapped to zero means that a significant component of the overlap between commodities can be explained simply in terms of the correlated strength sequences of the various layers. Importantly, we see that some pairs of layers actually exhibit negative rescaled multiplexity, even though the distribution is far from symmetric. This result, which is only visible in the weighted analysis, means that there are pairs of commodities for which the observed trade multiplicity is actually *lower* than expected under the null model: these commodities prefer "not to be traded together."

We then analyze the weighted multireciprocity of the WTM. Recently, it has been shown that the aggregated version

of the network has a strong weighted reciprocity [11], a result that we can now complement with the analysis of the disaggregated multiplex. In Figs. 6(a) and 6(b) we report the color-coded weighted multireciprocity matrix $\mathbf{R}_w$, along with the distribution of its off-diagonal entries. In analogy with the binary case, we see that the aggregated network exhibits a reciprocity that is significantly higher than the multireciprocity associated to any individual pair of layers. Yet several pairs of commodities are characterized by a substantial level of multireciprocity. In Figs. 6(c) and 6(d) we show the corresponding results for the rescaled weighted multireciprocity $\rho_w^{\alpha,\beta}$. We see that many values become close to zero and some become negative, in analogy with the behavior of the multiplexity. The identification of pairs of layers with negative rescaled multireciprocity indicates that the corresponding commodities "prefer not to be traded in opposite directions," in contrast with the results we found in the binary analysis.

In Fig. 7 we compare the weighted multireciprocity and the weighted multiplexity. When we consider the raw values (a), we observe a clear linear trend (although more scattered than in the corresponding unweighted case). The trend becomes even more robust, and less noisy, for the filtered values, as shown in (b). In both panels, the most significant commodities (both in terms of trade volumes and economic relevance) mainly lie along the diagonal, while the outliers represent less relevant products (for instance, some textiles or less traded craft goods). We also see pairs of commodities whose multireciprocity is similar to the reciprocity of the aggregate trade network. These commodities, such as cereals and heavy industry products, are not necessarily the most traded ones; still, they better represent the reciprocity patterns of total trade among countries, possibly because they give the main contribution to the reciprocity of the aggregated network.

Quantitatively, another important difference between the binary and the weighted approach lies in the statistical significance of the values of multiplexity and multireciprocity, as we can see from the analysis of the z scores (Fig. 8). Indeed, in the unweighted case we found that even the smallest values of $\mu_b^{\alpha,\beta}$
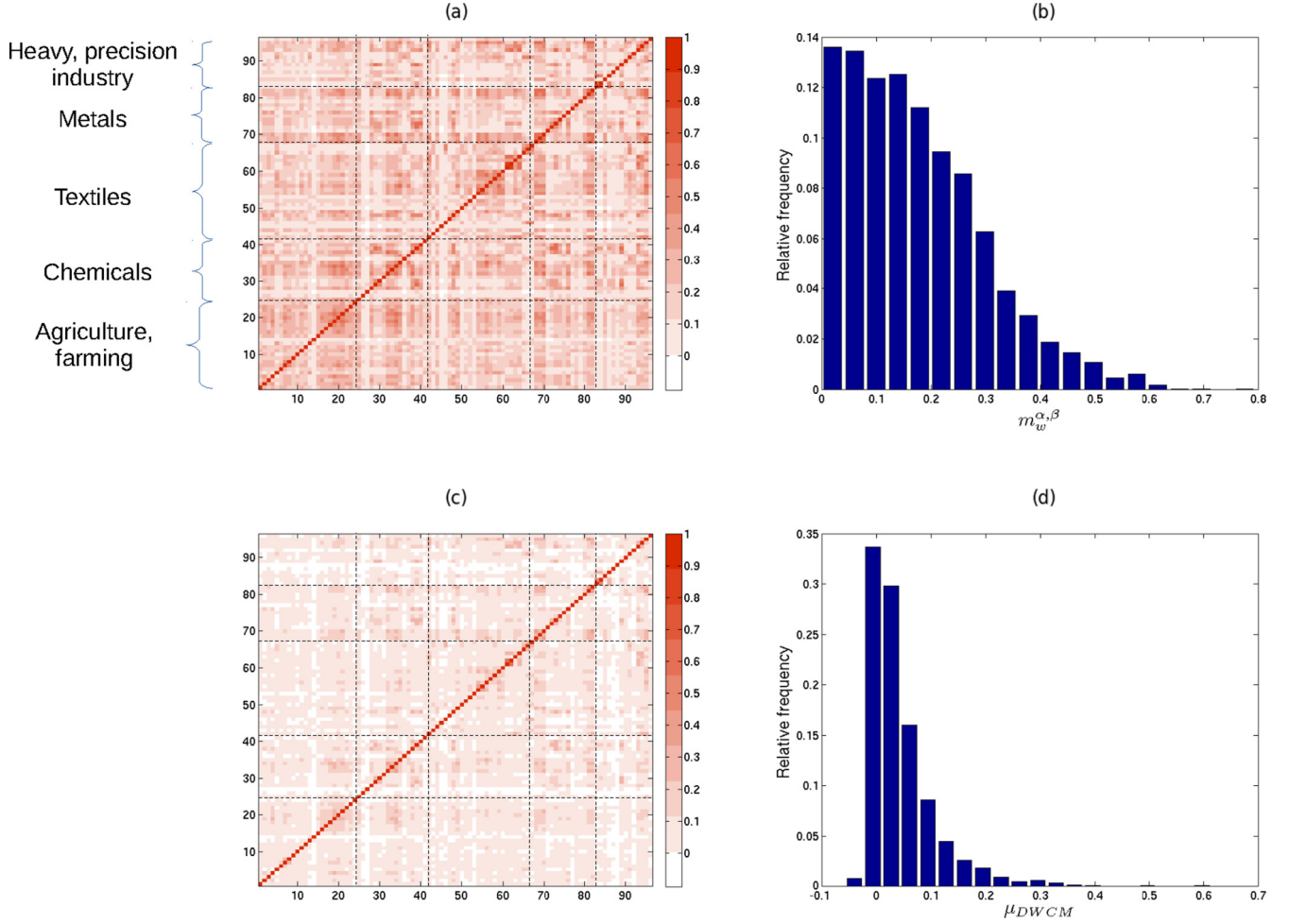
FIG. 5. Top panels: Color-coded weighted multiplexity matrix $\mathbf{M}_w$ (a) and corresponding distribution of off-diagonal multiplexity values $m_w^{\alpha,\beta}$ (with $\alpha \neq \beta$) (b). Bottom panels: Same as for the top panels but with raw weighted multiplexity $m_w^{\alpha,\beta}$ replaced by rescaled weighted multiplexity $\mu_w^{\alpha,\beta}$. Note that, in panel (c), white entries represent negative values.

and $\rho_b^{\alpha,\beta}$ are significant, as the corresponding z scores are larger than the critical value z_c. Instead, here we observe almost no correlation (except for the aforementioned sign concordance) between weighted multiplexity or multireciprocity and the corresponding z scores [see Figs. 8(a) and 8(b), respectively]. Indeed, the same value of $\mu_b^{\alpha,\beta}$ or $\rho_b^{\alpha,\beta}$ may even correspond to z scores with different orders of magnitude. This means that, even for two pairs of layers with the same observed value of weighted multiplexity or multireciprocity, the statistical significance of the interlayer coupling can differ considerably. Moreover, the absolute value of many weighted z scores is found below the significance threshold $z_c = 2$, identifying pairs of uncorrelated layers (a result that is unobserved in the binary case). Finally, many pairs of commodities have a negative z score below $-z_c$ for the multiplexity and/or multireciprocity. For these pairs, the tendency *not* to be traded in the same direction and/or in opposite direction is statistically validated and confirms a difference with respect to the binary case.

As a final result, in Fig. 8(c) we show the relation existing between $z(m_w^{\alpha,\beta})$ and $z(r_w^{\alpha,\beta})$. We find an overall level of correlation which, however, leaves room for a significant scatter of points around the identity line. This scatter is big

enough to imply that, for a given significance threshold z_c, the pairs of commodities can be partitioned in the following five classes:

(1) a few pairs of commodities that tend to be traded in the same direction [$z(m_w^{\alpha,\beta}) > z_c$] but not in opposite directions [$z(r_w^{\alpha,\beta}) < -z_c$]: examples are apparel articles vs. ships and boats; food industry residues, prepared animal feed vs. ores, slag, and ash;

(2) a few pairs of commodities that tend to be traded in opposite directions [$z(r_w^{\alpha,\beta}) > z_c$] but not in the same direction [$z(m_w^{\alpha,\beta}) < -z_c$]: examples are ores, slag, and ash vs. footwear and gaiters; apparel articles vs. ores, slag, and ash;

(3) a moderately sized group of pairs of commodities that tend to be traded neither in the same direction [$z(m_w^{\alpha,\beta}) < -z_c$] nor in opposite ones [$z(r_w^{\alpha,\beta}) < -z_c$]: examples are raw hides and skins vs. arms and ammunitions; tobacco vs. ships and boats;

(4) a large group of pairs of commodities for which there is no statistically significant tendency in at least one of the two directions [$|z(m_w^{\alpha,\beta})| < z_c$ and/or $|z(r_w^{\alpha,\beta})| < z_c$]: examples are tobacco vs. inorganic chemicals; explosives, pyrotechnic products vs. vehicles (note that this class can be
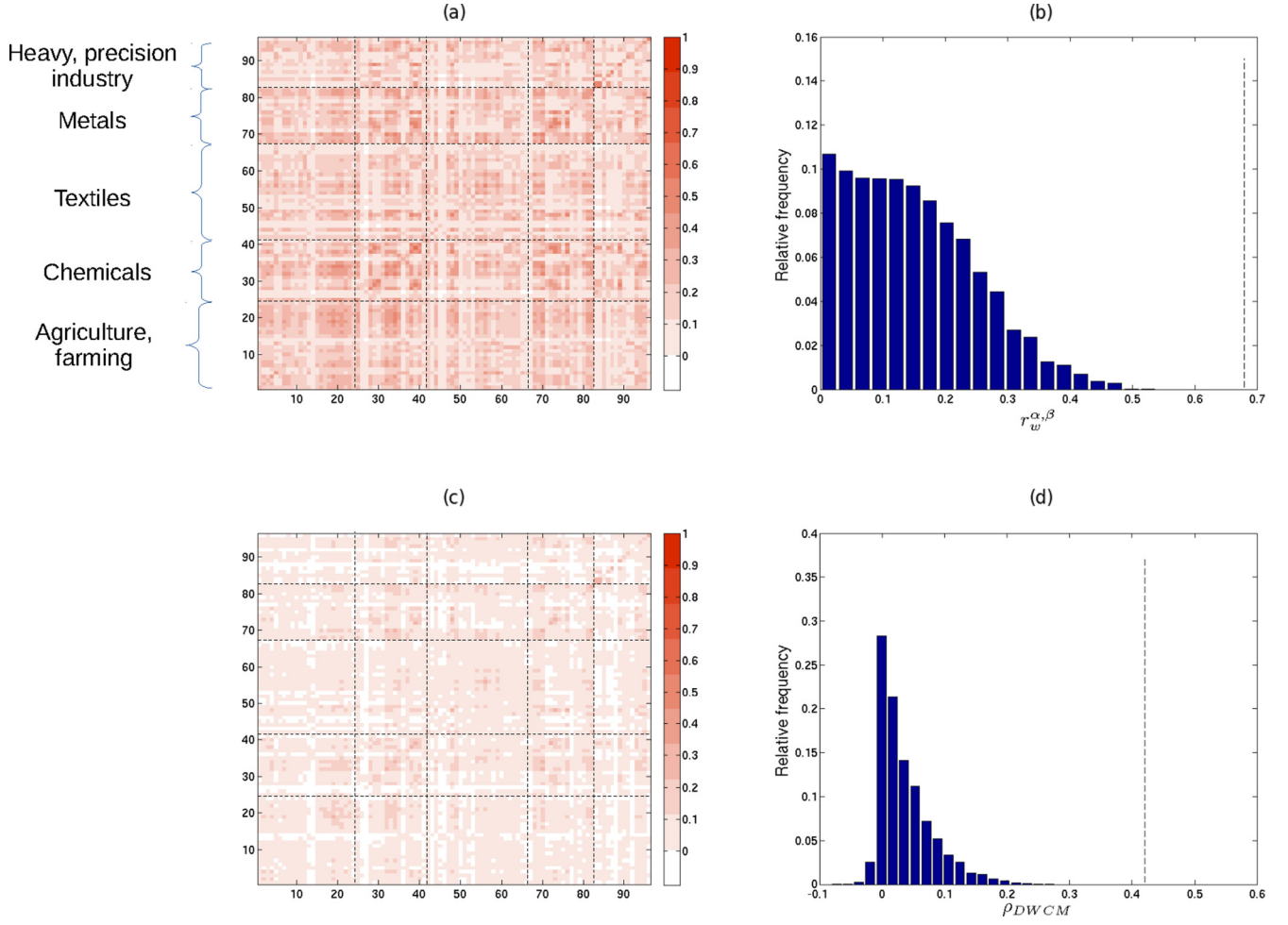
FIG. 6. Top panels: Color-coded weighted multireciprocity matrix $\mathbf{R}_w$ (a) and corresponding distribution of off-diagonal multireciprocity values $r_w^{\alpha,\beta}$ (with $\alpha \neq \beta$) (b). Bottom panels: Same as for the top panels but with raw weighted multireciprocity $r_w^{\alpha,\beta}$ replaced by rescaled weighted multireciprocity $\rho^{\alpha,\beta}$. The dashed lines represent the value of (raw and rescaled) weighted reciprocity r_w^{mono} and ρ_w^{mono} of the aggregated monoplex network. Note that, in panel (c), white entries represent negative values.

further split in subclasses where commodities are uncorrelated in one direction but correlated in different ways in the other direction);

(5) a very large group of pairs of commodities that tend to be traded both in the same direction $[z(m_w^{\alpha,\beta}) > z_c]$ and in opposite ones $[z(r_w^{\alpha,\beta}) > z_c]$: examples are sugar vs. cocoa; soap, waxes, candles vs. sugar.

It should be noted that, in contrast with the above classification, the binary analysis concluded that all pairs of commodities belong to the last class only.

IV. DISCUSSION AND CONCLUSIONS

The study of multilayer networks has received substantial attention in the past few years, leading to the introduction of several quantities characterizing the structure of multiplexes as well as the behavior of several dynamical processes taking place on them. The aim of all these studies is that of highlighting the role of the interlayer couplings, the latter being the ultimate reason why layers of a multiplex should be analyzed together in the first place rather than separately. In this paper we have argued that even the *simplest* definitions of

interlayer coupling, based merely on the structural overlap of links across layers, are strongly biased by the density, finiteness, and heterogeneity of the network. We have shown that controlling for the above effects requires a quite elaborate statistical treatment. Focusing on multiplexes with (binary or weighted) directed links, we have introduced maximum-entropy multiplex ensembles with given node properties as the unbiased null models serving as a benchmark for the empirically observed properties. We have then defined multiplexity and multireciprocity metrics, respectively quantifying the tendency of pairs of links to "align" and/or "anti-align" across each pair of layers of a real-world directed multiplex. Since links can exist in both directions in every layer, the possible tendencies of forming aligned (multiplexed) and antialigned (multireciprocated) links do not conflict with each other and can actually coexist. Both multiplexity and multireciprocity are matrix valued, as they represent the possible couplings among all pairs of layers. While multiplexity is a natural extension of the corresponding definition for undirected multiplexes, multireciprocity is a novel concept representing a nontrivial extension of the notion of single-layer reciprocity to multilayer networks.
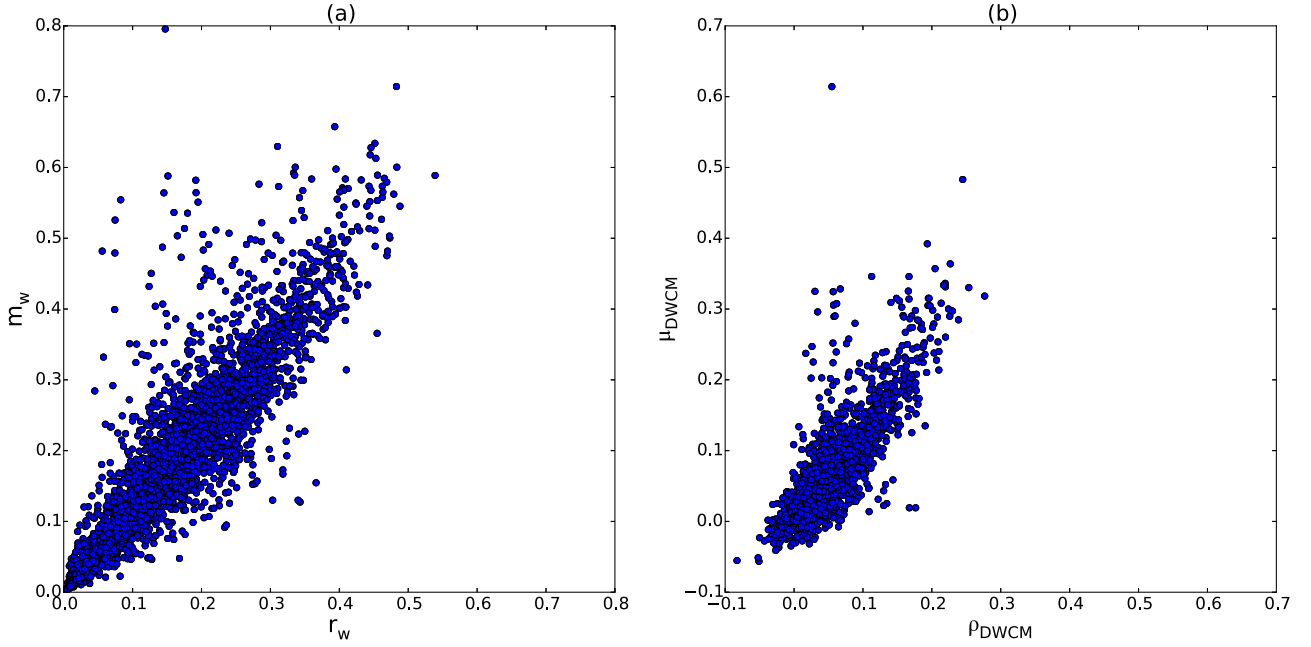
FIG. 7. Scatter plots of off-diagonal weighted multireciprocity values versus off-diagonal weighted directed multiplexity values. Left: Raw values ($r_w^{\alpha,\beta}$ vs. $m_w^{\alpha,\beta}$); right: rescaled values ($\rho_w^{\alpha,\beta}$ vs. $\mu_w^{\alpha,\beta}$).

We believe that our results can be of value for several applications. For instance, they provide a statistically rigorous way to identify possible (groups of) layers that are uncorrelated from the other layers, thus allowing us to simplify the whole multiplex into mutually independent subsystems with smaller numbers of layers. This problem has received significant attention recently [42,43]. Our finding of a strong influence of the local node properties on the overall level of interlayer coupling suggests that many of the results found with alternative techniques that do not control for these effects might be subject to an uncontrolled level of bias.

Other more specific applications are relevant for the case study of the WTM. Our detailed analysis of this system confirmed that its multiplex structure contains much more information than does the aggregated network of total trade. At a binary level, we found that all pairs of commodities tend to be traded together between countries, both in the same direction (high multiplexity) and in opposite directions (high multireciprocity). At a weighted level, this result only holds for a subset of pairs of commodities. Other commodity pairs are not correlated and others even tend to avoid being traded together in the same direction and/or in opposite ones. The
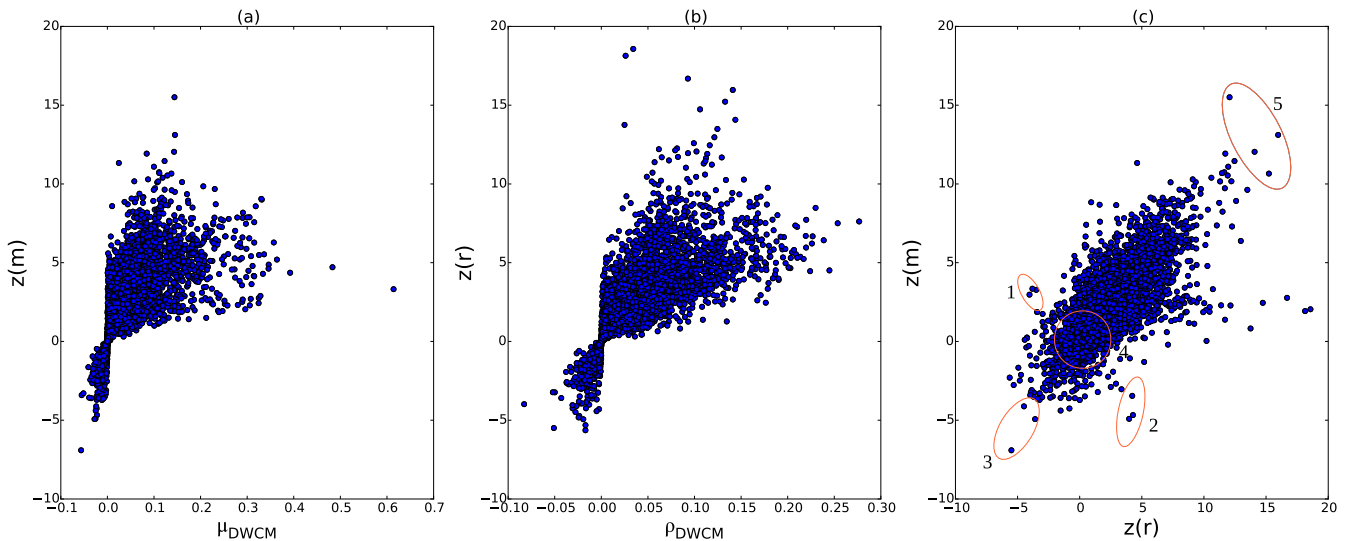
FIG. 8. Left: Weighted transformed multiplexity $\mu_w^{\alpha,\beta}$ versus its corresponding z score $z(m_w^{\alpha,\beta})$; center: weighted transformed multireciprocity $\rho_w^{\alpha,\beta}$ versus its corresponding z score $z(r_w^{\alpha,\beta})$; right: $z(r_w^{\alpha,\beta})$ vs. $z(m_w^{\alpha,\beta})$. In panel (c), numbered circles correspond to the bullet points reported in Sec. III C. Only off-diagonal values are reported.

multireciprocity structure of the WTM highlights a tendency of groups of commodities to have a comparably high mutual reciprocity of the same entity of the internal single-layer reciprocity of these commodities. When aggregated into the monoplex network of total international trade, the WTM has a resulting reciprocity that is much bigger than the multireciprocity among its constituent layers.

In the light of the above results, our approach has implications relevant to various directions in international trade research. In particular, it indicates concrete ways to refine existing measures of intercommodity correlation or similarity that are widely used to construct, among others, "product taxonomies" [8], the "product space" [35], and "fitness and complexity" metrics [36]. All these applications are briefly explained below.

Intercommodity correlation metrics have been introduced to quantify the coupling among layers of the WTM [8], with the goal of constructing "product taxonomies" that reflect empirical trade similarities, as opposed to predefined product categories. However, as already pointed out in Ref. [9], correlation metrics make an implicit and totally unrealistic assumption of structural homogeneity of the network by interpreting all the edges of a layer as independent observations drawn from the same probability distribution. Our results provide alternative metrics of interlayer coupling that replace the homogeneity assumption with a much more realistic null model that accurately controls for the observed degree of node heterogeneity in each layer. The use of our metrics is likely to change the structure of correlation-based product taxonomies significantly.

The "product space" is defined as a network of commodities connected by links whose weight quantifies the tendency of a pair of commodities to be traded together (in the same direction) between the same two countries [35]. Our results clearly indicate that, to be statistically reliable, such an analysis should include a way to filter out the strong empirical heterogeneity of node degrees and/or node strengths. Moreover, they highlight a second layer of information that should be relevant for the product space construction, namely the fact that, besides the tendency of pairs of commodities to be traded together in the same direction (multiplexity), there can be a substantial tendency of being traded in the opposite direction (multireciprocity). We found that these two effects have a comparable magnitude. We also found that pairs of commodities with approximately the same multiplexity can be characterized by very different levels of multireciprocity. This suggests that neglecting multireciprocity in the construction of the product space can represent a substantial loss of information.

Finally, the "fitness and complexity" approach focuses on the bipartite network of countries and their exported products and uses the structure of this network to recursively define metrics of product complexity and country competitiveness (fitness) [36]. This method can reveal the "hidden" potential of countries that is not (yet) reflected in their current GDP levels. Clearly, the output of this approach entirely depends on how the bipartite country-product matrix is constructed. This matrix is ultimately a projection of the WTM but is generally filtered using a null model based on the concept of "revealed comparative advantage" [44], which, however, operates at the

aggregate country-product level and not at the level of the underlying multiplex. As such, it does not control for the size of importers. Our approach provides a way to enforce a more accurate null model on the original WTM and obtain an alternative bipartite country-product projection.

We believe that all the research directions outlined above deserve future explorations and we expect the results reported in this paper to be of use.

APPENDIX A: MAXIMUM ENTROPY METHOD

We define null models of multiplexes as canonical maximum-entropy ensembles satisfying a given set $\vec{\mathcal{C}}$ of $\mathcal{K}$ constraints on average. If $G^{\alpha} \in \mathcal{G}_N$ denotes the graph realized in layer α of the multiplex (recall that $\mathcal{G}_N$ is the set of all directed monoplex graphs with N nodes), and if $\vec{G} \in \mathcal{G}_N^M$ denotes the entire multiplex (where $\mathcal{G}_N^M$ is the set of all directed multiplex graphs with N nodes and M layers), then we write $\vec{G} = (G^{\alpha})_{\alpha=1}^M$. Now let $\vec{\mathcal{C}}$ denote a vector-valued function on $\mathcal{G}_N^M$, evaluating to $\vec{\mathcal{C}}(\vec{G})$ on the particular multiplex $\vec{G}$. The vector $\vec{\mathcal{C}}(\vec{G})$ is to be regarded as a set of structural properties measured on $\vec{G}$.

A canonical ensemble of binary (weighted) directed multiplex networks with the *soft constraint* $\vec{\mathcal{C}}$ is specified by a probability distribution $\mathcal{P}(\vec{G}|\vec{\theta})$ on $\mathcal{G}_N^M$, where $\vec{\theta}$ is a vector of Lagrange multipliers required to enforce a desired expected value

$$\langle \vec{\mathcal{C}} \rangle_{\vec{\theta}} = \sum_{\vec{G} \in \mathcal{G}_N^M} \mathcal{P}(\vec{G}|\vec{\theta}) \vec{\mathcal{C}}(\vec{G}) \qquad (A1)$$

of $\vec{\mathcal{C}}$. Note that both $\vec{\theta}$ and $\vec{\mathcal{C}}$ are vectors of numbers with the same (but model-dependent) dimension $\mathcal{K}$, while $\vec{G}$ is always an M-dimensional vector of graphs. Obviously, an additional constraint on the probability is the normalization condition

$$\sum_{\vec{G} \in \mathcal{G}_N^M} \mathcal{P}(\vec{G}|\vec{\theta}) = 1 \qquad \forall \vec{\theta}. \qquad (A2)$$

We want our ensembles to produce multiplexes with independent layers, as defined in Eq. (2). This requirement corresponds to the enforcement of separate constraints on the different layers, i.e., $\vec{\mathcal{C}} = (\vec{\mathcal{C}^{\alpha}})_{\alpha=1}^M$, where $\vec{\mathcal{C}^{\alpha}}$ is a K^{α}-dimensional vector of structural properties of the network in layer α only, evaluating to $\vec{\mathcal{C}^{\alpha}}(G^{\alpha})$ on the particular single-layer graph G^{α}. This leads to a separation in the corresponding Lagrange multipliers, i.e., $\vec{\theta} = (\vec{\theta^{\alpha}})_{\alpha=1}^M$. K^{α} is the dimension of both $\vec{\mathcal{C}^{\alpha}}$ and $\vec{\theta^{\alpha}}$, and we must have $\sum_{\alpha=1}^M K^{\alpha} = \mathcal{K}$. Consequently, we

can express the entropy of the ensemble of multiplex networks as

$$
\mathcal{S}(\overrightarrow{\theta}) \equiv - \sum_{\overrightarrow{G} \in \mathcal{G}_N^M} \mathcal{P}(\overrightarrow{G}|\overrightarrow{\theta}) \ln \mathcal{P}(\overrightarrow{G}|\overrightarrow{\theta})
$$

$$
= \sum_{\alpha=1}^{M} S^{\alpha}(\overrightarrow{\theta^{\alpha}}), \qquad (A3)
$$

where

$$
S^{\alpha}(\overrightarrow{\theta^{\alpha}}) \equiv - \sum_{G^{\alpha} \in \mathcal{G}_N} P^{\alpha}(G^{\alpha}|\overrightarrow{\theta^{\alpha}}) \ln P^{\alpha}(G^{\alpha}|\overrightarrow{\theta^{\alpha}}) \qquad (A4)
$$

is the entropy of the ensemble of monoplex graphs for the individual layer α, with $P^{\alpha}(G^{\alpha}|\overrightarrow{\theta^{\alpha}})$ defined as in (1) and subject to the normalization condition

$$
\sum_{G^{\alpha} \in \mathcal{G}_N} P^{\alpha}(G^{\alpha}|\overrightarrow{\theta^{\alpha}}) = 1 \quad \forall \overrightarrow{\theta^{\alpha}} \quad \alpha = 1, M. \qquad (A5)
$$

At this point, we want to maximize the entropy $\mathcal{S}(\overrightarrow{\theta})$, subject to the soft constraint $\overrightarrow{\mathcal{C}}$, to find the functional form of $\mathcal{P}(\overrightarrow{G}|\overrightarrow{\theta})$ we are looking for. Equation (A3) ensures that the maximization of $\mathcal{S}(\overrightarrow{\theta})$, subject to (A1), reduces to the maximization of each single-layer entropy $S^{\alpha}(\overrightarrow{\theta^{\alpha}})$, subject to

$$
\langle \overrightarrow{C^{\alpha}} \rangle_{\overrightarrow{\theta^{\alpha}}} = \sum_{G^{\alpha} \in \mathcal{G}_N} P^{\alpha}(G^{\alpha}|\overrightarrow{\theta^{\alpha}}) \overrightarrow{C^{\alpha}}(G^{\alpha}), \qquad (A6)
$$

separately. Therefore the probability $\mathcal{P}(\overrightarrow{G}|\overrightarrow{\theta})$ maximizing $\mathcal{S}(\overrightarrow{\theta})$ reduces, via (2), to the product of all single-layer probability distributions of the type $P^{\alpha}(G^{\alpha}|\overrightarrow{\theta^{\alpha}})$, each of which should separately maximize the corresponding entropy $S^{\alpha}(\overrightarrow{\theta^{\alpha}})$.

The general solution to the problem of maximizing $S^{\alpha}(\overrightarrow{\theta^{\alpha}})$, subject to (A6), for single-layer networks is extensively discussed in Ref. [20] and, in our notation here, leads to the probability distribution

$$
P^{\alpha}(G^{\alpha}|\overrightarrow{\theta^{\alpha}}) = \frac{e^{-H^{\alpha}(G^{\alpha}|\overrightarrow{\theta^{\alpha}})}}{Z(\overrightarrow{\theta^{\alpha}})}, \qquad (A7)
$$

where

$$
H^{\alpha}(G^{\alpha}|\overrightarrow{\theta^{\alpha}}) = \overrightarrow{\theta^{\alpha}} \cdot \overrightarrow{C^{\alpha}}(G^{\alpha}) \qquad (A8)
$$

is the *graph Hamiltonian* (the dot indicating a scalar product, i.e., a linear combination of the enforced constraints) and

$$
Z(\overrightarrow{\theta^{\alpha}}) = \sum_{G^{\alpha} \in \mathcal{G}_N} e^{-H^{\alpha}(G^{\alpha}|\overrightarrow{\theta^{\alpha}})} \qquad (A9)
$$

is the *partition function* (representing the normalizing constant for the probability).

Equation (A9), and, consequently, (A7), lead to different explicit functional forms depending on the choice of the constraint(s), i.e., depending on the functional form of $\overrightarrow{C^{\alpha}}(G^{\alpha})$. In Appendices C and D we explicitly discuss the cases of the

directed binary configuration model (where the constraints are the in- and out-degrees of all nodes in each layer α) and of the directed weighted configuration model (where the constraints are the in- and out-strengths of all nodes in each layer α), respectively.

Once an explicit expression for each $P^{\alpha}(G^{\alpha}|\overrightarrow{\theta^{\alpha}})$ is found, we can use (2) to find the final expression for the whole multiplex probability in the null model:

$$
\mathcal{P}(\overrightarrow{G}|\overrightarrow{\theta}) = \prod_{\alpha=1}^{M} \frac{e^{-H^{\alpha}(G^{\alpha}|\overrightarrow{\theta^{\alpha}})}}{Z(\overrightarrow{\theta^{\alpha}})} = \frac{e^{-\mathcal{H}(\overrightarrow{G}|\overrightarrow{\theta})}}{\mathcal{Z}(\overrightarrow{\theta})}, \qquad (A10)
$$

where

$$
\mathcal{H}(\overrightarrow{G}|\overrightarrow{\theta}) \equiv \sum_{\alpha=1}^{M} H^{\alpha}(G^{\alpha}|\overrightarrow{\theta^{\alpha}}) \qquad (A11)
$$

and

$$
\mathcal{Z}(\overrightarrow{\theta}) \equiv \prod_{\alpha=1}^{M} Z(\overrightarrow{\theta^{\alpha}}). \qquad (A12)
$$

The last three equations rephrase the independence of all layers explicitly.

APPENDIX B: MAXIMUM LIKELIHOOD METHOD

The maximization of the entropy is a constrained, *functional* maximization of $\mathcal{S}(\overrightarrow{\theta})$ in the space of probability distributions. As such, its result is the *functional form* of the maximum-entropy distribution $\mathcal{P}(\overrightarrow{G}|\overrightarrow{\theta})$, given by (A10), but not its *numerical values*. In fact, the distribution depends on the whole vector of parameters $\overrightarrow{\theta}$, and any expectation value calculated analytically using the explicit expression of $\mathcal{P}(\overrightarrow{G}|\overrightarrow{\theta})$ can only be evaluated numerically after a value of $\overrightarrow{\theta}$ is specified. This leads to the problem of choosing $\overrightarrow{\theta}$. Since we are interested in the case where all layers of the multiplex are independent, choosing a value of $\overrightarrow{\theta}$ reduces to the problem of choosing $\overrightarrow{\theta^{\alpha}}$ separately for each layer.

The problem of finding the parameter values of a maximum-entropy model of single-layer networks is solved in the general case in Ref. [30] using the maximum likelihood principle. In our notation here, this solution can be restated as follows. Let G_*^{α} denote, among all graphs $G^{\alpha} \in \mathcal{G}_N$, the particular *empirical* network realized in layer α of the multiplex. Given G_*^{α}, the log-likelihood function

$$
\mathcal{L}^{\alpha}(\overrightarrow{\theta^{\alpha}}) \equiv \ln P(G_*^{\alpha}|\overrightarrow{\theta^{\alpha}}) \qquad (B1)
$$

represents the log of the probability to generate the empirical graph G_*^{α}, given a value of $\overrightarrow{\theta^{\alpha}}$. The maximum likelihood principle [30] states that the optimal choice for $\overrightarrow{\theta^{\alpha}}$ is the one that maximizes the chances to obtain G_*^{α} from the model, i.e., the one that maximizes $\mathcal{L}^{\alpha}(\overrightarrow{\theta^{\alpha}})$. Let this parameter choice be denoted by $\overrightarrow{\theta_*^{\alpha}}$, where

$$
\overrightarrow{\theta_*^{\alpha}} \equiv \arg\max_{\overrightarrow{\theta^{\alpha}}} \mathcal{L}^{\alpha}(\overrightarrow{\theta^{\alpha}}). \qquad (B2)
$$

As a general result [30], the value $\overrightarrow{\theta_*^\alpha}$ defined above is such that

$$\langle \overrightarrow{C^\alpha} \rangle_{\overrightarrow{\theta_*^\alpha}} = \overrightarrow{C^\alpha}(G_*^\alpha), \tag{B3}$$

i.e., the expectation value of each constraint coincides with the empirical value measured on the empirical network G_*^α. This is precisely the outcome we desire, given that our ultimate goal is the construction of ensembles of networks with the same numerical value of the constraints as in the real network.

From a practical point of view, Eqs. (B2) and (B3) represent two equivalent ways to determine $\overrightarrow{\theta_*^\alpha}$. The former requires the maximization of a scalar function over a K^α-dimensional space, while the latter requires the solution of a system of K^α nonlinear coupled equations. For various choices of the graph ensemble $\mathcal{G}_N$ and of the constraints $\overrightarrow{C^\alpha}$ (including those required for our analysis), both approaches are implemented in the MAX&SAM algorithm [21]. More details are given in Appendices C and D. Once the value $\overrightarrow{\theta_*^\alpha}$ is found, it is used to find the *numerical value* $P(G^\alpha | \overrightarrow{\theta_*^\alpha})$ of the probability of any graph $G^\alpha \in \mathcal{G}_N$. So, while the maximization of the entropy generates the functional form of the graph probability, the maximization of the likelihood fixes its numerical values. If X^α denotes any single-layer structural property X of interest, then the above procedure allows us to evaluate the expected value

$$\langle X^\alpha \rangle \equiv \langle X^\alpha \rangle_{\overrightarrow{\theta_*^\alpha}} = \sum_{G^\alpha \in \mathcal{G}_N} P(G^\alpha | \overrightarrow{\theta_*^\alpha}) X^\alpha(G^\alpha) \tag{B4}$$

(and similarly the standard deviation) of X^α explicitly over the desired ensemble [20,21,30]. For many properties of interest, the expected value (B4) can be calculated analytically given the explicit expression of $P(G^\alpha | \overrightarrow{\theta_*^\alpha})$, without the need to sample the graph ensemble explicitly [20]. For more complicated properties, one can instead use the knowledge of $P(G^\alpha | \overrightarrow{\theta_*^\alpha})$ to sample graphs from the ensemble in an unbiased way and then calculate expectations as sample averages [21].

The multiplexity and multireciprocity metrics introduced in Sec. II are not single-layer properties like X^α, as they require measurements on multiple layers simultaneously. Using Eq. (2), we therefore need to generalize Eq. (B4) to the case of an arbitrary multiplex quantity $\mathcal{X}$, evaluating to $\mathcal{X}(\overrightarrow{G})$ on a specific multiplex $\overrightarrow{G} \in \mathcal{G}_N^M$, as follows:

$$\langle \mathcal{X} \rangle \equiv \langle \mathcal{X} \rangle_{\overrightarrow{\theta_*}} = \sum_{\overrightarrow{G} \in \mathcal{G}_N^M} \mathcal{P}(\overrightarrow{G} | \overrightarrow{\theta_*}) \mathcal{X}(\overrightarrow{G}), \tag{B5}$$

where $\overrightarrow{\theta_*} = (\overrightarrow{\theta_*^\alpha})_{\alpha=1}^M$ contains the Langrange multipliers (B2) for all layers and $\overrightarrow{G_*} = (G_*^\alpha)_{\alpha=1}^M \in \mathcal{G}_N^M$ denotes the whole empirical multiplex. Both the expected values and the standard deviations of multiplexity and multireciprocity can be calculated explicitly, and we will therefore follow the analytical approach, which is exact and faster than the sampling approach (see Appendices C and D).

From a computational point of view, the above canonical approach based on soft constraints has many benefits with respect to the microcanonical approach with hard constraints [20,21]. Indeed, the microcanonical approach cannot be

controlled analytically and necessarily requires sampling many randomized multiplexes explicitly from the ensemble. Generating even only a single randomized multiplex requires the iteration of many random constraint-preserving "rewiring moves," which is computationally costly. Such a procedure must be repeated several times to produce a large sample of R randomized multiplexes, on each of which any topological property X of interest has to be calculated. Finally, a sample average should be performed to obtain an estimate of $\langle X \rangle$.

For instance, on single-layer networks with constrained degree sequence, one should iterate the "local rewiring algorithm" [18] that preserves the degrees while randomizing the network. On a monoplex network with L links, the above approach would require a computational time of order $O(L)$, only to generate a single realization of the randomized network. On such a realization, one would then need to measure X (for instance, the monoplex reciprocity), which would require a certain time T_X. The total time needed for a single realization would therefore be $T_X + O(L)$ and for all realizations $R \cdot T_X + O(R \cdot L)$.

In a multiplex network with M layers, the corresponding time required to generate a single randomized multiplex would in principle be of order $O(\sum_{\alpha=1}^M L^\alpha)$, where L^α is the number of links in the αth layer. However, if layers are independent in the null model, then the randomization could (if computational resource allows) be run in parallel on the different layers, thus reducing the above time to $O(\bar{L})$ where $\bar{L}$ is the average number of links per layer, which does not scale with M. However, the calculation of multiplex quantities X (e.g., the multireciprocity) which would require a time T_X for a single layer (e.g., the monoplex reciprocity) would now need to be iterated for each pair of layers, thus requiring a time $O(T_X M^2)$. In total, this means that the total microcanonical computational time for a multiplex is $T_{\text{mic}} = O(R T_X M^2) + O(R \cdot \bar{L})$, before carrying out the final sample averages.

By contrast, our canonical approach does not require the sampling of any multiplex. For individual layers, the calculation of the expected value of most properties of interest basically requires replacing the adjacency matrix of the network with the corresponding expected matrix (or more complicated replacements that in any case require a comparable calculation time). Therefore calculating the expected value $\langle X \rangle$ takes the same time T_X that it would take for the empirical property X to be calculated on the real system. The same holds true for the entire multiplex. Therefore the total canonical time needed is $T_{\text{can}} = O(T_X M^2) + T_{\mathcal{L}}$, where $T_{\mathcal{L}}$ is the one-off time required to preliminary maximize the likelihood (possibly of each layer in parallel) defined in (B1).

As already mentioned above, the time $T_{\mathcal{L}}$ required to maximize the likelihood function can be proxied by the time required to solve a system of coupled, nonlinear equations ($2N$ equations in the case of directed networks, as shown below). However, since such systems can be further simplified by rewriting them only in terms of the sequences of distinct directed degrees or strengths (which are always less than $2N$), the computational time drops to the order of seconds or minutes (depending on the chosen constraints) for each layer. Moreover, further analyses on synthetic networks have shown that this time scales roughly quadratically with the number of nodes; it is in any case considerably shorter than

the corresponding total microcanonical time T_{mic} estimated above.

Besides the computational advantages described above, the canonical approach has the statistical advantage of being a truly unbiased method [21], in the sense that its maximum-entropy nature implies that no preference is given to specific graph configurations other than on the basis of the enforced constraints. So unbiasedness is ensured by the maximum degree of randomness encoded in the graph probability, given the constraints. By constrast, microcanonical approaches are not guaranteed to ensure the same property. In the microcanonical case, unbiasedness means that the realizations of the network should be sampled uniformly (i.e., with exactly the same probability) from the whole set of configurations compatible with the constraints. Ensuring uniform sampling is highly nontrivial and often impossible. For instance, in the case of graphs with fixed degree sequence, it can be proved that the local rewiring algorithm is biased, as it preferentially samples configurations that are "close" to the empirical one [22]. It is in principle possible to remove this bias by calculating the "mobility" function (which is a quantity that depends on the current configuration being randomized) and accepting the "next" randomized configurations with a probability that depends on the mobility itself. This requirement further increases the already heavy computational requirements of the microcanonical approach, because the mobility should be continuously recalculated during the randomization process.

APPENDIX C: DIRECTED BINARY CONFIGURATION MODEL

In this Appendix we explicitly discuss the DBCM model [20,21], obtained through the maximum entropy and maximum likelihood methods in the specific case where $\mathcal{G}_N$ contains all binary directed graphs with N nodes and $\overrightarrow{C^\alpha}$ is a vector of dimension $K^\alpha = 2N$ containing the out-degree k_i^{out} and the in-degree k_i^{in} of all nodes $(i = 1,N)$. Correspondingly, the $2N$-dimensional vector $\overrightarrow{\theta^\alpha}$ contains the associated Lagrange multipliers ϕ_i^α and χ_i^α for all nodes. Note that we enforce the in- and out-degree sequences on all layers, which means that, as a function, $\overrightarrow{C^\alpha} = (\overrightarrow{k^{\text{out}}}, \overrightarrow{k^{\text{in}}})$ is the same for all α. However, the numerical values of the degrees in different layers will in general differ, i.e., $\overrightarrow{C}(G^\alpha) \neq \overrightarrow{C}(G^\beta)$ for $\alpha \neq \beta$, thus $\overrightarrow{\theta^\alpha} = (\overrightarrow{\phi^\alpha}, \overrightarrow{\chi^\alpha})$ must still depend on α explictly.

For single-layer networks, this model is discussed, e.g., in Refs. [20,21]. Here we simply summarize the main steps leading to the final expressions for the expected binary multiplexity and binary reciprocity. Using the notation introduced in Sec. II and in Appendix A, the single-layer Hamiltonian (A8) reads

$$
\begin{aligned}
H(G^\alpha | \overrightarrow{\phi^\alpha}, \overrightarrow{\chi^\alpha}) &= \overrightarrow{\phi^\alpha} \cdot \overrightarrow{k^{\text{out}}}(G^\alpha) + \overrightarrow{\chi^\alpha} \cdot \overrightarrow{k^{\text{in}}}(G^\alpha) \\
&= \sum_{i=1}^{N} \left[\phi_i^\alpha k_i^{\text{out}}(G^\alpha) + \chi_i^\alpha k_i^{\text{in}}(G^\alpha) \right] \\
&= \sum_{i=1}^{N} \sum_{j \neq i} (\phi_i^\alpha + \chi_j^\alpha) a_{ij}^\alpha \quad \text{(C1)}
\end{aligned}
$$

and the partition function (A9) can be calculated as:

$$
\begin{aligned}
Z(\overrightarrow{\phi^\alpha}, \overrightarrow{\chi^\alpha}) &= \prod_{i=1}^{N} \prod_{j \neq i} \left(1 + e^{-\phi_i^\alpha - \chi_j^\alpha}\right) \\
&= \prod_{i=1}^{N} \prod_{j \neq i} \left(1 + x_i^\alpha y_j^\alpha\right), \quad \text{(C2)}
\end{aligned}
$$

where we have set $x_i^\alpha \equiv e^{-\phi_i^\alpha}$ and $y_i^\alpha \equiv e^{-\chi_i^\alpha}$. This implies that the probability (A7) can be written explicitly as

$$
\begin{aligned}
P^\alpha(G^\alpha | \overrightarrow{\phi^\alpha}, \overrightarrow{\chi^\alpha}) &= \prod_{i=1}^{N} \prod_{j \neq i} \frac{\left(x_i^\alpha y_j^\alpha\right)^{a_{ij}^\alpha}}{1 + x_i^\alpha y_j^\alpha} \\
&= \prod_{i=1}^{N} \prod_{j \neq i} \left(p_{ij}^\alpha\right)^{a_{ij}^\alpha} \left(1 - p_{ij}^\alpha\right)^{1 - a_{ij}^\alpha}, \quad \text{(C3)}
\end{aligned}
$$

where

$$
p_{ij}^\alpha = \frac{x_i^\alpha y_j^\alpha}{1 + x_i^\alpha y_j^\alpha} \quad \text{(C4)}
$$

is the probability of a directed link from i to j in layer α. Equation (C3) shows that the random variable a_{ij}^α is drawn, for all $i \neq j$, from a Bernoulli distribution with success probability p_{ij}^α, thus leading to Eq. (5).

Given the real-world multiplex $\overrightarrow{G}_* = (G_*^\alpha)_{\alpha=1}^M$, the single-layer log-likelihood function (B1) to be maximized is then given by

$$
\begin{aligned}
\mathcal{L}(\overrightarrow{x^\alpha}, \overrightarrow{y^\alpha}) &= \sum_{i=1}^{N} \left[k_i^{\text{out}}(G_*^\alpha) \ln x_i^\alpha + k_i^{\text{in}}(G_*^\alpha) \ln y_i^\alpha \right] + \\
&\quad - \sum_{i=1}^{N} \sum_{j \neq i} \ln \left(1 + x_i^\alpha y_j^\alpha\right), \quad \text{(C5)}
\end{aligned}
$$

and the equivalent set of $2N$ coupled nonlinear equations (B3) to be solved is

$$
\sum_{j \neq i} \frac{x_i^\alpha y_j^\alpha}{1 + x_i^\alpha y_j^\alpha} = k_i^{\text{out}}(G_*^\alpha) \quad \forall i = 1,N \quad \text{(C6)}
$$

$$
\sum_{j \neq i} \frac{x_j^\alpha y_i^\alpha}{1 + x_j^\alpha y_i^\alpha} = k_i^{\text{in}}(G_*^\alpha) \quad \forall i = 1,N. \quad \text{(C7)}
$$

Once found, the values of $\{x_i^\alpha\}$ and $\{y_i^\alpha\}$ providing the unique solution to the above problem can be put back into Eqs. (C3) and (C4), allowing us to analytically calculate the expected values $\langle \cdot \rangle_{\text{DBCM}}$ of the quantities of interest via the corresponding probabilities p_{ij}^α (where for simplicity we drop the asterisk indicating that p_{ij}^α is evaluated at the specific values that maximize the likelihood).

In particular, we can calculate the rescaled metrics defined in Eqs. (4) as follows. First, since in the DBCM the in- and out-degrees of all nodes in all layers are equal to their expected values, we necessarily have $\langle L^\alpha \rangle_{\text{DBCM}} = L^\alpha$ for all α, where $L_*^\alpha \equiv L^\alpha(G_*^\alpha)$ is the number of links of the observed, layer-specific graph G_*^α. This means that L^α is a constrained quantity, and we therefore expect the denominators of Eqs. (3) to fluctuate around their expected values $L_*^\alpha + L_*^\beta$ much less

than how the numerators fluctuate around the corresponding expected values. We therefore approximate the expected values of Eqs. (3) as follows:

$$\langle m_b^{\alpha,\beta} \rangle_{\text{DBCM}} = \frac{2\langle L^{\alpha \Rightarrow \beta} \rangle_{\text{DBCM}}}{L_*^\alpha + L_*^\beta} \quad (\alpha \neq \beta), \quad \text{(C8a)}$$

$$\langle r_b^{\alpha,\beta} \rangle_{\text{DBCM}} = \frac{2\langle L^{\alpha \rightleftarrows \beta} \rangle_{\text{DBCM}}}{L_*^\alpha + L_*^\beta}. \quad \text{(C8b)}$$

Consequently,

$$\mu_b^{\alpha,\beta} = \frac{2L_*^{\alpha \Rightarrow \beta} - 2\langle L^{\alpha \Rightarrow \beta} \rangle_{\text{DBCM}}}{L_*^\alpha + L_*^\beta - 2\langle L^{\alpha \Rightarrow \beta} \rangle_{\text{DBCM}}} \quad (\alpha \neq \beta),$$

$$\rho_b^{\alpha,\beta} = \frac{2L_*^{\alpha \rightleftarrows \beta} - 2\langle L^{\alpha \rightleftarrows \beta} \rangle_{\text{DBCM}}}{L_*^\alpha + L_*^\beta - 2\langle L^{\alpha \rightleftarrows \beta} \rangle_{\text{DBCM}}}.$$

Since a_{ij}^α and a_{ij}^β (for $\beta \neq \alpha$), and similarly a_{ij}^α and a_{ji}^β (for any β), are independently drawn from two Bernoulli distributions, the expected values of $\min\{a_{ij}^\alpha, a_{ij}^\beta\}$ (with $\beta \neq \alpha$) and $\min\{a_{ij}^\alpha, a_{ji}^\beta\}$ are easily calculated as

$$\langle \min\{a_{ij}^\alpha, a_{ij}^\beta\} \rangle_{\text{DBCM}} = p_{ij}^\alpha p_{ij}^\beta \quad (\alpha \neq \beta), \quad \text{(C9a)}$$

$$\langle \min\{a_{ij}^\alpha, a_{ji}^\beta\} \rangle_{\text{DBCM}} = p_{ij}^\alpha p_{ji}^\beta. \quad \text{(C9b)}$$

Therefore the final expressions for the transformed multiplexity and multireciprocity are

$$\mu_b^{\alpha,\beta} = \frac{2L_*^{\alpha \Rightarrow \beta} - 2\sum_i \sum_{j \neq i} p_{ij}^\alpha p_{ij}^\beta}{L_*^\alpha + L_*^\beta - 2\sum_i \sum_{j \neq i} p_{ij}^\alpha p_{ij}^\beta} \quad (\alpha \neq \beta) \quad \text{(C10a)}$$

$$\rho_b^{\alpha,\beta} = \frac{2L_*^{\alpha \rightleftarrows \beta} - 2\sum_i \sum_{j \neq i} p_{ij}^\alpha p_{ji}^\beta}{L_*^\alpha + L_*^\beta - 2\sum_i \sum_{j \neq i} p_{ij}^\alpha p_{ji}^\beta}, \quad \text{(C10b)}$$

where the probabilities are defined according to Eq. (C4).

Similarly, we need to calculate the z scores defined in Eqs. (8). To do this, we need to calculate the standard deviations of $m_b^{\alpha,\beta}$ and $r_b^{\alpha,\beta}$ at the denominator of the z scores. Neglecting again the fluctuations of the constrained quantities L^α and L^β around their average values (with respect to the fluctuations of the unconstrained quantities), and since all pairs of nodes are independent, we calculate the variances of $m_b^{\alpha,\beta}$ and $r_b^{\alpha,\beta}$ in a way similar to what we did for the expressions in Eq. (C11):

$$\text{Var}[m_b^{\alpha,\beta}] = \frac{4\sum_i \sum_{j \neq i} \text{Var}[\min\{a_{ij}^\alpha, \ a_{ij}^\beta\}]}{(L_*^\alpha + L_*^\beta)^2} \quad (\alpha \neq \beta),$$

$$\text{Var}[r_b^{\alpha,\beta}] = \frac{4\sum_i \sum_{j \neq i} \text{Var}[\min\{a_{ij}^\alpha, \ a_{ji}^\beta\}]}{(L_*^\alpha + L_*^\beta)^2}. \quad \text{(C11)}$$

Now we note that the minimum of two 0/1 quantities is also a 0/1 quantity. This implies that the square of the minimum is equal to the minimum itself and that the expected square of the minimum is equal to the expected value of the minimum.

In formulas:

$$\langle \min^2\{a_{ij}^\alpha, a_{ij}^\beta\} \rangle_{\text{DBCM}} = p_{ij}^\alpha p_{ij}^\beta \quad (\alpha \neq \beta), \quad \text{(C12)}$$

$$\langle \min^2\{a_{ij}^\alpha, a_{ji}^\beta\} \rangle_{\text{DBCM}} = p_{ij}^\alpha p_{ji}^\beta. \quad \text{(C13)}$$

It then follows that the variance of the minimum is

$$\text{Var}[\min\{a_{ij}^\alpha, a_{ij}^\beta\}] = p_{ij}^\alpha p_{ij}^\beta (1 - p_{ij}^\alpha p_{ij}^\beta) \quad (\alpha \neq \beta),$$

$$\text{Var}[\min\{a_{ij}^\alpha, a_{ji}^\beta\}] = p_{ij}^\alpha p_{ji}^\beta (1 - p_{ij}^\alpha p_{ji}^\beta).$$

Putting these expressions into those for $\text{Var}[m_b^{\alpha,\beta}]$ and $\text{Var}[r_b^{\alpha,\beta}]$, and taking the square root to obtain the standard deviations, we finally arrive at the explicit calculation of the z scores:

$$z(m_b^{\alpha,\beta}) = \frac{L_*^{\alpha \Rightarrow \beta} - \sum_i \sum_{j \neq i} p_{ij}^\alpha p_{ij}^\beta}{\sqrt{\sum_i \sum_{j \neq i} p_{ij}^\alpha p_{ij}^\beta (1 - p_{ij}^\alpha p_{ij}^\beta)}} \quad (\alpha \neq \beta)$$

$$z(r_b^{\alpha,\beta}) = \frac{L_*^{\alpha \rightleftarrows \beta} - \sum_i \sum_{j \neq i} p_{ij}^\alpha p_{ji}^\beta}{\sqrt{\sum_i \sum_{j \neq i} p_{ij}^\alpha p_{ji}^\beta (1 - p_{ij}^\alpha p_{ji}^\beta)}}.$$

From a direct comparison between the above equations and Eqs. (C10), we immediately observe the sign concordance reported in the main text.

APPENDIX D: DIRECTED WEIGHTED CONFIGURATION MODEL

Here we consider the DWCM model [20,21], obtained when $\mathcal{G}_N$ contains all weighted directed graphs (with non-negative integer edge weights) with N nodes and $\overrightarrow{C^\alpha}$ is a vector of dimension $K^\alpha = 2N$ containing the out-strength s_i^{out} and the in-strength s_i^{in} of all nodes ($i = 1, N$). The $2N$-dimensional vector $\overrightarrow{\theta^\alpha}$ contains the associated Lagrange multipliers ϕ_i^α and χ_i^α. As for the DBCM, $\overrightarrow{C^\alpha} = (\overrightarrow{s^{\text{out}}}, \overrightarrow{s^{\text{in}}})$ is the same function for all α. However, the numerical values $\overrightarrow{\theta^\alpha} = (\overrightarrow{\phi^\alpha}, \overrightarrow{\chi^\alpha})$ still depend on α.

For single-layer networks (see, e.g., discussion in Refs. [20,21]), the Hamiltonian (A8) reads

$$H(G^\alpha | \overrightarrow{\phi^\alpha}, \overrightarrow{\chi^\alpha}) = \overrightarrow{\phi^\alpha} \cdot \overrightarrow{s^{\text{out}}}(G^\alpha) + \overrightarrow{\chi^\alpha} \cdot \overrightarrow{s^{\text{in}}}(G^\alpha)$$

$$= \sum_{i=1}^N \left[\phi_i^\alpha s_i^{\text{out}}(G^\alpha) + \chi_i^\alpha s_i^{\text{in}}(G^\alpha) \right]$$

$$= \sum_{i=1}^N \sum_{j \neq i} (\phi_i^\alpha + \chi_j^\alpha) w_{ij}^\alpha \quad \text{(D1)}$$

and the partition function (A9) can be calculated as

$$Z(\overrightarrow{\phi^\alpha}, \overrightarrow{\chi^\alpha}) = \prod_{i=1}^N \prod_{j \neq i} (1 - e^{-\phi_i^\alpha - \chi_j^\alpha})^{-1}$$

$$= \prod_{i=1}^N \prod_{j \neq i} (1 - x_i^\alpha y_j^\alpha)^{-1}, \quad \text{(D2)}$$

where we have set $x_i^\alpha \equiv e^{-\phi_i^\alpha}$ and $y_i^\alpha \equiv e^{-\chi_i^\alpha}$. This implies that the probability (A7) can be written as

$$
\begin{aligned}
P^\alpha(G^\alpha | \vec{\phi^\alpha}, \vec{\chi^\alpha}) &= \prod_{i=1}^{N} \prod_{j \neq i} \left(x_i^\alpha y_j^\alpha\right)^{w_{ij}^\alpha} \left(1 - x_i^\alpha y_j^\alpha\right) \\
&= \prod_{i=1}^{N} \prod_{j \neq i} \left(p_{ij}^\alpha\right)^{w_{ij}^\alpha} \left(1 - p_{ij}^\alpha\right), \qquad \text{(D3)}
\end{aligned}
$$

where

$$
p_{ij}^\alpha = x_i^\alpha y_j^\alpha \tag{D4}
$$

denotes again the probability that a directed link (of any positive weight) from node i to node j is realized in layer α. Equation (D3) gives the interpretation of w_{ij}^α as a geometrically distributed variable, constructed as the iteration of many random events, each defined as incrementing w_{ij}^α by 1, starting from $w_{ij}^\alpha = 0$. In this interpretation, p_{ij}^α is the elementary probability of a "success" event, and the probability that $w_{ij}^\alpha = w$ coincides with the probability $(p_{ij}^\alpha)^w (1 - p_{ij}^\alpha)$ of having w consecutive successes followed by one failure. This leads to the geometric distribution in Eq. (11).

The single-layer log-likelihood function (B1) to be maximized is now given by

$$
\begin{aligned}
\mathcal{L}(\vec{x^\alpha}, \vec{y^\alpha}) &= \sum_{i=1}^{N} \left[s_i^{\text{out}}(G_*^\alpha) \ln x_i^\alpha + s_i^{\text{in}}(G_*^\alpha) \ln y_i^\alpha \right] \\
&\quad + \sum_{i=1}^{N} \sum_{j \neq i} \ln \left(1 - x_i^\alpha y_j^\alpha\right),
\end{aligned} \tag{D5}
$$

and the corresponding equations (B3) are

$$
\sum_{j \neq i} \frac{x_i^\alpha y_j^\alpha}{1 - x_i^\alpha y_j^\alpha} = s_i^{\text{out}}(G_*^\alpha) \quad \forall i = 1, N, \tag{D6}
$$

$$
\sum_{j \neq i} \frac{x_j^\alpha y_i^\alpha}{1 - x_j^\alpha y_i^\alpha} = s_i^{\text{in}}(G_*^\alpha) \quad \forall i = 1, N. \tag{D7}
$$

The expected values $\langle \cdot \rangle_{\text{DWCM}}$ of the relevant quantities can be found through Eqs. (D3) and (D4), evaluated at the values of $\{x_i^\alpha\}$ and $\{y_i^\alpha\}$ that solve the above problem (again, in what follows we drop the asterisk indicating that p_{ij}^α is evaluated at the specific values that maximize the likelihood).

We start with the calculation of the expected values of the multiplexity and multireciprocity metrics defined in Eq. (9). In analogy with what we did for the DBCM, we expect the (constrained) denominators of Eqs. (9) to fluctuate much less than the (unconstrained) numerators and we therefore replace the denominators with their expected values $W_*^\alpha + W_*^\beta$. We therefore write

$$
\langle m_w^{\alpha,\beta} \rangle_{\text{DWCM}} = \frac{2 \langle W^{\alpha \Rightarrow \beta} \rangle_{\text{DWCM}}}{W_*^\alpha + W_*^\beta}, \tag{D8a}
$$

$$
\langle r_w^{\alpha,\beta} \rangle_{\text{DWCM}} = \frac{2 \langle W^{\alpha \rightleftarrows \beta} \rangle_{\text{DWCM}}}{W_*^\alpha + W_*^\beta}, \tag{D8b}
$$

and

$$
\mu_w^{\alpha,\beta} = \frac{2 W_*^{\alpha \Rightarrow \beta} - 2 \langle W^{\alpha \Rightarrow \beta} \rangle_{\text{DWCM}}}{W_*^\alpha + W_*^\beta - 2 \langle W^{\alpha \Rightarrow \beta} \rangle_{\text{DWCM}}} \quad (\alpha \neq \beta),
$$

$$
\rho_w^{\alpha,\beta} = \frac{2 W_*^{\alpha \rightleftarrows \beta} - 2 \langle W^{\alpha \rightleftarrows \beta} \rangle_{\text{DWCM}}}{W_*^\alpha + W_*^\beta - 2 \langle W^{\alpha \rightleftarrows \beta} \rangle_{\text{DWCM}}}.
$$

Since w_{ij}^α and w_{ij}^β (for $\beta \neq \alpha$), and similarly w_{ij}^α and w_{ji}^β (for any β), are independently drawn from two geometric distributions, the expected values of $\min\{w_{ij}^\alpha, w_{ij}^\beta\}$ (with $\beta \neq \alpha$) and $\min\{w_{ij}^\alpha, w_{ji}^\beta\}$ are easily calculated as

$$
\langle \min\{a_{ij}^\alpha, a_{ij}^\beta\} \rangle_{\text{DBCM}} = \frac{p_{ij}^\alpha p_{ij}^\beta}{1 - p_{ij}^\alpha p_{ij}^\beta} \quad (\alpha \neq \beta), \tag{D9a}
$$

$$
\langle \min\{a_{ij}^\alpha, a_{ji}^\beta\} \rangle_{\text{DBCM}} = \frac{p_{ij}^\alpha p_{ji}^\beta}{1 - p_{ij}^\alpha p_{ji}^\beta}. \tag{D9b}
$$

Therefore the transformed multiplexity and multireciprocity read

$$
\mu_w^{\alpha,\beta} = \frac{2 W_*^{\alpha \Rightarrow \beta} - 2 \sum_i \sum_{j \neq i} \frac{p_{ij}^\alpha p_{ij}^\beta}{1 - p_{ij}^\alpha p_{ij}^\beta}}{W_*^\alpha + W_*^\beta - 2 \sum_i \sum_{j \neq i} \frac{p_{ij}^\alpha p_{ij}^\beta}{1 - p_{ij}^\alpha p_{ij}^\beta}} \quad (\alpha \neq \beta) \tag{D10a}
$$

$$
\rho_w^{\alpha,\beta} = \frac{2 W_*^{\alpha \rightleftarrows \beta} - 2 \sum_i \sum_{j \neq i} \frac{p_{ij}^\alpha p_{ji}^\beta}{1 - p_{ij}^\alpha p_{ji}^\beta}}{W_*^\alpha + W_*^\beta - 2 \sum_i \sum_{j \neq i} \frac{p_{ij}^\alpha p_{ji}^\beta}{1 - p_{ij}^\alpha p_{ji}^\beta}}, \tag{D10b}
$$

where the probabilities are defined in Eq. (D4).

We then calculate the z scores. Following an argument similar to the binary case, we write

$$
\text{Var}[m_w^{\alpha,\beta}] = \frac{4 \sum_i \sum_{j \neq i} \text{Var}[\min\{w_{ij}^\alpha, w_{ij}^\beta\}]}{(W_*^\alpha + W_*^\beta)^2} \quad (\alpha \neq \beta),
$$

$$
\text{Var}[r_w^{\alpha,\beta}] = \frac{4 \sum_i \sum_{j \neq i} \text{Var}[\min\{w_{ij}^\alpha, w_{ji}^\beta\}]}{(W_*^\alpha + W_*^\beta)^2}.
$$

After calculating the variance of the minimum of two geometrically distributed random variables, we get

$$
\text{Var}[\min\{w_{ij}^\alpha, w_{ij}^\beta\}] = \frac{p_{ij}^\alpha p_{ij}^\beta}{(1 - p_{ij}^\alpha p_{ij}^\beta)^2} \quad (\alpha \neq \beta),
$$

$$
\text{Var}[\min\{a_{ij}^\alpha, a_{ji}^\beta\}] = \frac{p_{ij}^\alpha p_{ji}^\beta}{(1 - p_{ij}^\alpha p_{ji}^\beta)^2}.
$$

Combining all the relevant expressions together, we get for the z scores:

$$z\left(m_w^{\alpha,\beta}\right) = \frac{W_*^{\alpha \rightrightarrows \beta} - \sum_i \sum_{j \neq i} \frac{p_{ij}^\alpha p_{ij}^\beta}{1 - p_{ij}^\alpha p_{ij}^\beta}}{\sqrt{\sum_i \sum_{j \neq i} \frac{p_{ij}^\alpha p_{ij}^\beta}{\left(1 - p_{ij}^\alpha p_{ij}^\beta\right)^2}}} \quad (\alpha \neq \beta)$$

$$z\left(r_w^{\alpha,\beta}\right) = \frac{W_*^{\alpha \rightleftarrows \beta} - \sum_i \sum_{j \neq i} \frac{p_{ij}^\alpha p_{ji}^\beta}{1 - p_{ij}^\alpha p_{ji}^\beta}}{\sqrt{\sum_i \sum_{j \neq i} \frac{p_{ij}^\alpha p_{ji}^\beta}{\left(1 - p_{ij}^\alpha p_{ji}^\beta\right)^2}}}$$

Comparing with Eqs. (D10), we confirm the concordance of the signs.

---

[1] H. Ebel, L.-I. Mielsch, and S. Bornholdt, Phys. Rev. E **66**, 035103 (2002).

[2] M. E. J. Newman, Phys. Rev. E **70**, 056131 (2004).

[3] S. Boccaletti, G. Bianconi, R. Criado, C. I. del Genio, J. Gomez-Gardeñes, M. Romance, I. Sendiña Nadal, Z. Wang, and M. Zanin, Phys. Rep. **544**, 1 (2014).

[4] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, J. Complex Netw. **2**, 203 (2014).

[5] M. Magnani, B. Micenková, and L. Rossi, arXiv:1303.4986, (2013).

[6] F. Battiston, V. Nicosia, and V. Latora, Phys. Rev. E **89**, 032804 (2014).

[7] G. Bianconi, Phys. Rev. E **87**, 062806 (2013).

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

[9] V. Gemmetto and D. Garlaschelli, Sci. Rep. **5**, 9120 (2015).

[10] D. Garlaschelli and M. I. Loffredo, Phys. Rev. Lett. **93**, 268701 (2004).

[11] T. Squartini, F. Picciolo, F. Ruzzenenti, and D. Garlaschelli, Sci. Rep. **3**, 2729 (2013).

[12] L. A. Meyers, M. E. J. Newman, and B. Pourbohloul, J. Theor. Biol. **240**, 400 (2006).

[13] M. Boguñá and M. A. Serrano, Phys. Rev. E **72**, 016106 (2005).

[14] M. Schnegg, Int. J. Mod. Phys. C **17**, 1067 (2006).

[15] D. Garlaschelli and M. I. Loffredo, Physica A **355**, 138 (2005).

[16] C. Zhou, L. Zemanová, G. Zamora, C. C. Hilgetag, and J. Kurths, Phys. Rev. Lett. **97**, 238103 (2006).

[17] M. E. J. Newman, S. Forrest, and J. Balthrop, Phys. Rev. E **66**, 035101 (2002).

[18] S. Maslov and K. Sneppen, Science **296**, 910 (2002).

[19] M. A. Serrano and M. Boguñá, AIP Conf. Proc. **776**, 101 (2005).

[20] T. Squartini and D. Garlaschelli, New J. Phys. **13**, 083001 (2011).

[21] T. Squartini, R. Mastrandrea, and D. Garlaschelli, New J. Phys. **17**, 023052 (2015).

[22] E. S. Roberts and A. C. C. Coolen, Phys. Rev. E **85**, 046103 (2012).

[23] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).

[24] G. L. Robins, P. E. Pattison, Y. Kalish, and D. Lusher, Soc. Netw. **29**, 173 (2007).

[25] J. Park and M. E. J. Newman, Phys. Rev. E **70**, 066117 (2004).

[26] J. Park and M. E. J. Newman, Phys. Rev. E **68**, 026112 (2003).

[27] P. W. Holland and S. Leinhardt, J. Am. Stat. Assoc. **76**, 33 (1981).

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

[29] T. A. B. Snijders, P. E. Pattison, G. L. Robins, and M. S. Handcock, Sociol. Methodol. **36**, 99 (2006).

[30] D. Garlaschelli and M. I. Loffredo, Phys. Rev. E **78**, 015101 (2008).

[31] D. Garlaschelli and M. I. Loffredo, Phys. Rev. Lett. **93**, 188701 (2004).

[32] T. Squartini, G. Fagiolo, and D. Garlaschelli, Phys. Rev. E **84**, 046117 (2011).

[33] T. Squartini, G. Fagiolo, and D. Garlaschelli, Phys. Rev. E **84**, 046118 (2011).

[34] R. Mastrandrea, T. Squartini, G. Fagiolo, and D. Garlaschelli, Phys. Rev. E **90**, 062804 (2014).

[35] C. A. Hidalgo, B. Klinger, A.-L. Barabási, and R. Hausmann, Science **317**, 482 (2007).

[36] A. Tacchella, M. Cristelli, G. Caldarelli, A. Gabrielli, and L. Pietronero, Sci. Rep. **2**, 723 (2012).

[37] G. Gaulier and S. Zignago, CEPII Working Paper 23 (2010).

[38] http://www.wcoomd.org.

[39] F. Ruzzenenti, D. Garlaschelli, and R. Basosi, Symmetry **2**, 1710 (2010).

[40] F. Picciolo, T. Squartini, F. Ruzzenenti, R. Basosi, and D. Garlaschelli, in *Proceedings of the Eighth International Conference on Signal-Image Technology & Internet-Based Systems (SITIS 2012)* (IEEE, Piscataway, NJ, 2013), pp. 784–792.

[41] We discard the diagonal entries in order to make the distribution compatible with the corresponding distribution for the multiplexity shown above; in any case, if the diagonal entries are included, then the distribution looks very similar.

[42] M. De Domenico, V. Nicosia, A. Arenas, and V. Latora, Nat. Commun. **6**, 6864 (2015).

[43] J. Iacovacci, Z. Wu, and G. Bianconi, Phys. Rev. E **92**, 042806 (2015).

[44] B. Balassa, Manchest. Sch. **33**, 99 (1965).