Multiplexity and multireciprocity in directed multiplexes

V Gemmetto[1], T Squartini[2], F Picciolo[3], F Ruzzenenti[3,4] and D Garlaschelli[1]

[1] Instituut-Lorentz for Theoretical Physics, Leiden Institute of Physics, University of Leiden, Niels Bohrweg 2, 2333 CA, Leiden, The Netherlands

[2] Institute for Complex Systems UOS Sapienza, "Sapienza" University of Rome, Italy

[3] Department of Biotechnology, Chemistry and Pharmacy, University of Siena, Italy

[4] Department of Economics and Statistics, University of Siena, Italy

E-mail: gemmetto@lorentz.leidenuniv.nl

Abstract. In recent years, the study of multi-layer networks has received significant attention. In this work, we provide new measures of dependency between directed links across different layers of multiplex networks. We show that this operation requires more than a straightforward extension of the corresponding multiplexity measures that have been developed for undirected multiplexes. In particular, one should take into account the effects of reciprocity, i.e. the tendency of pairs of vertices to establish mutual connections. In single-layer networks, reciprocity is a crucial property affecting several dynamical processes. Here we extend this quantity to multiplexes and introduce the notion of multireciprocity, defined as the tendency of links in one layer to be reciprocated by links in a different layer. While ordinary reciprocity reduces to a scalar quantity, multireciprocity requires a square matrix generated by all the possible pairs of layers. We introduce multireciprocity metrics valid for both binary and weighted networks and then measure these quantities on the World Trade Multiplex (WTM), representing the import-export relationships between world countries in different products. We show that several pairs of layers exhibit strong multiplexity, an effect which can however be largely encoded into the degree or strength sequences of individual layers. We also find that most pairs of commodities are characterised by positive multireciprocity, and that such values are significantly lower than the usual reciprocity measured on the aggregated network. Moreover, layers with low (high) internal reciprocity are embedded within groups of layers with low (high) mutual multireciprocity. We finally identify robust empirical patterns that allow us to use the multireciprocity matrix to retrieve the two-layer reciprocated degree (strength) of a node from the ordinary in-degree (in-strength) in a single layer and to reconstruct joint multi-layer connection probabilities from marginal ones, hence bridging the gap between single-layer properties and truly multiplex information.

Submitted to: *New J. Phys.*

1. Introduction

Several real-world systems exhibit a network structure, being composed by interconnected elementary units. The study of networks has received much attention in the last two decades. In this perspective, both the intensity and the symmetry of the interactions between nodes have been analysed, thus introducing the concepts of, respectively, weighted and directed graphs. For instance, most of the communication relations among individuals, such as exchanging letters, e-mails or texting, can be suitably represented by directed networks, thus pointing out the directionality of these interactions [1]; furthermore, such interactions can be weighted, in order to quantify the strength of such social connections [2].

Nevertheless, a more detailed representation of such systems is often required, since generally a given set of units can actually be connected by different kinds of relations (as in the so-called edge-colored graphs), therefore yielding interdependent networks where the various layers may influence each other [3, 4]. 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 co-authorship, common leisure activities, on-line social networks etc.

In this context, all the considerations about the intensity and the symmetry of the connections made for single-layer graphs can be applied to multiplexes, as they can be seen as the superposition of several (possibly directed or weighted) networks. Indeed, detailed information about intensity and directionality may sometimes be crucial for a deeper understanding of such system, as it has been observed [6] that weighted multi-level networks showing non-trivial correlations between topology and weights actually exist.

In particular, one of the most well-studied properties of single-layer directed networks - either binary or weighted - is their reciprocity, i.e. the tendency of vertex pairs to form mutual connections. This property can be indeed crucial for dynamical processes taking place on networks, such as diffusion [7], percolation [8] and growth [9, 10]; 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 [11].

It has been shown [12] that the measure of the number of mutual interactions has to be compared with the expected reciprocity obtained for a given reference model, in order to understand whether such observed mutual links are actually present in the real network significantly more (or less) often than in the randomized benchmark [13]. In this context, it is therefore crucial to make use of proper null models for networks; in particular, for unweighted networks the directed (binary) configuration model (DBCM) has been widely used, in order to take into account the heterogeneity of the degree sequence within the null model [14].

This quantity has been extended to weighted networks by means of the definition of the weighted reciprocity [15]. Similarly to the binary case, this measure has to be compared to the expected value under a proper null model, represented for instance by the directed weighted configuration model (DWCM) [16].

The concept of reciprocity has not been generalized to multi-layer networks yet. Here we argue that, in presence of multiple types of connections, one should extend the notion of reciprocity to that of *multireciprocity*, which we define as the tendency of a directed link in one layer of the multiplex to be reciprocated by an opposite link in a *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. Furthermore, we investigate the correlations between directed layers of a multiplex through the directed multiplexity. We introduce multiplexity and multireciprocity matrices for both binary and weighted multiplexes and we then validate our metrics by measuring them on the World Trade Multiplex (WTM), a directed weighted multiplex representing the import-export relations between countries of the world in different products.

2. Methods

2.1. Multiplex approach, null models and Maximum Likelihood Method

We represent our multiplex as the superposition of M networks, each of them sharing the same set of N nodes. Since our purpose is precisely 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 [23, 17].

In this context, we make use of the concept of canonical network ensemble [18], represented by the the family of networks satisfying a set of constraints on average (to be fixed, for instance, based on some properties of the considered real-world network). Such randomized graphs preserve only some of the features of the system under study and are completely random otherwise; therefore, they represent suitable reference models for networks.

In particular, in order to take into consideration the heterogeneity of the degree sequence within the null models, in the unweighted case we make use of the directed (binary) configuration model (DBCM) [14], namely the ensemble of networks satisfying on average a given in-degree and out-degree sequences. It should be noted that the popular microcanonical implementation [14] of this null model is biased [19], and that its unbiased modification is computationally demanding. Moreover, in order to measure the expected value of any quantity of interest, the generation of several randomized networks, on each of which the quantity needs to be calculated, is necessary. For this reason, we adopt the unbiased canonical approach based on an analytical and fast Maximum Likelihood method [20, 21, 22]. Such method, explained in more details in the Appendix, is able to provide the exact probabilities of occurrence of the graphs of the ensemble having the same average constraints as the real network. Based on these probabilities, it is then possible to compute the expectation values of the quantities of interest; in particular, the average link probability p_{ij}^{α}, from any node i to any node j in layer α, can be easily calculated. As we said before, such probabilities are computed independently for each layer.

We can introduce a notation able to describe the pairwise interactions between layers of the multiplex. Defining $p_{ij}^{\alpha\rightarrow\beta}$ as the joint probability of observing a directed link from i to j in α and no link from j to i in β, and $p_{ij}^{\alpha\leftrightarrow\beta}$ as the joint probability of observing a link from i to j in α and a link in the opposite direction in β, the previous single-layer average link probability can be expressed as:

$$p_{ij}^{\alpha} = p_{ij}^{\alpha\rightarrow\beta} + p_{ij}^{\alpha\leftrightarrow\beta} \tag{1}$$

With the aforementioned notation, from Bayes' theorem we have, for each ordered pair of layers (α, β):

$$p_{ij}^{\alpha\leftrightarrow\beta} = r_{ij}^{\alpha\beta} \cdot p_{ji}^{\beta} \tag{2}$$

where $r_{ij}^{\alpha\beta}$ is the conditional probability of observing a direct link from i to j in layer α given that we observe a link from j to i in layer β. In the general case, the value of $r_{ij}^{\alpha\beta}$ depends both on the considered pair of nodes and on the pair of layers.
In this context, the expected degree sequences of the single layers read:

$$\langle k_i^{\alpha,in} \rangle = \sum_{j\neq i} p_{ji}^{\alpha}\,; \qquad \langle k_i^{\alpha,out} \rangle = \sum_{j\neq i} p_{ij}^{\alpha} \tag{3}$$

while the reciprocated degree sequence, for any pair of layers, is:

$$\langle k_i^{\alpha\leftrightarrow\beta} \rangle = \sum_{j\neq i} r_{ij}^{\alpha\beta} p_{ji}^{\beta} \tag{4}$$

In Section 3.4 we will show that such quantities may be used to build the minimal model able to reproduce the observed level of reciprocation between layers of the multiplex.

Analogously, in order to build suitable null models in the weighted case, the enforced constraints are chosen to be the in-strength and out-strength sequences of the real network - separately for each layer - and the key property that can be computed is the average weight $\langle w_{ij}^{\alpha} \rangle$ related to each directed link connecting i to j in layer α. Moreover, we can also define quantities taking into account interactions between pairs of layers: $\langle w_{ij}^{\alpha\leftrightarrow\beta} \rangle$ is the reciprocated component of the weights associated to the links from i to j in α and from j to i in β while $\langle w_{ij}^{\alpha\rightarrow\beta} \rangle$ is the non-reciprocated component between the same two directed weighted links. Hence, the expected weights in a given layer α are therefore given by:

$$\langle w_{ij}^{\alpha} \rangle = \langle w_{ij}^{\alpha\rightarrow\beta} \rangle + \langle w_{ij}^{\alpha\leftrightarrow\beta} \rangle \tag{5}$$

In particular, the reciprocated component can be written in terms of a joint probability, in order to keep the same structure adopted for the binary case [15]:

$$\begin{aligned} \langle w_{ij}^{\alpha\leftrightarrow\beta} \rangle &= \langle \min\{w_{ij}^{\alpha}, w_{ji}^{\beta}\} \rangle = \\ &= \sum_{w=1}^{\infty} P\left(\min\{w_{ij}^{\alpha}, w_{ji}^{\beta}\} \geq w\right) = \end{aligned}$$

$$= \sum_{w=1}^{\infty} P\left(w_{ij}^{\alpha} \geq w \cap w_{ji}^{\beta} \geq w\right) =$$
$$= \sum_{w=1}^{\infty} R_{ij}^{\alpha\beta}\left(w_{ij}^{\alpha} \geq w | w_{ji}^{\beta} \geq w\right) P\left(w_{ji}^{\beta} \geq w\right) \tag{6}$$

where $R_{ij}^{\alpha\beta}$ is now the probability of observing a weight w_{ij}^{α} in α larger than w given that a weight w_{ji}^{β} larger than w has been observed in β. Further steps in the simplification of this expression can be done based on the empirical evidence of the considered system, as we will show in Section 3.4. From the previous definitions we can directly compute the expected in- and out-strengths in each layer:

$$\langle s_i^{\alpha,in}\rangle = \sum_{j\neq i}\langle w_{ji}^{\alpha}\rangle\ ; \qquad \langle s_i^{\alpha,out}\rangle = \sum_{j\neq i}\langle w_{ij}^{\alpha}\rangle \tag{7}$$

and the reciprocated strength sequence, for any ordered pair of layers:

$$\langle s_i^{\alpha\leftrightarrow\beta}\rangle = \sum_{j\neq i}\langle w_{ij}^{\alpha\leftrightarrow\beta}\rangle \tag{8}$$

Furthermore, the previous average values of the link probability or link weight will be crucial for the computations of other quantities such as, in this work, the expected overlap between links in different layers or the expected inter-layer reciprocity.

2.2. Binary multiplexity and multireciprocity

In order to analyze the similarity and reciprocity between layers of a directed unweighted multiplex, we define the binary directed multiplexity and the multireciprocity as the extension of the multiplexity introduced in [23] to directed networks:

$$m_{bin}^{\alpha,\beta} = \frac{2\sum_{i\neq j}\min\{a_{ij}^{\alpha}, a_{ij}^{\beta}\}}{L^{\alpha}+L^{\beta}}\ ; \qquad r_{bin}^{\alpha,\beta} = \frac{2\sum_{i\neq j}\min\{a_{ij}^{\alpha}, a_{ji}^{\beta}\}}{L^{\alpha}+L^{\beta}} \tag{9}$$

with $\alpha,\beta = 1,\ldots,M$ standing for the different layers, $a_{ij}^{\alpha} = 0,1$ depending on the presence of a directed link from node i to node j in layer α and L^{α} representing the total number of directed links in layer α (analogously for layer β). The quantities defined in (9) provide information about the overlap between directed links connecting nodes in any pair of layers. Indeed, such raw quantities range in $[0,1]$ and are maximal only when layers α and β are respectively identical - hence, if there is full similarity between the considered layers - or fully multireciprocated; therefore, they evaluate the tendency of nodes to share directed connections in the various layers (possibly, in different directions). Though, the proper way to extract information from such measures consists in a comparison with some kind of expected value under a given null model. Hence, we introduce the transformed directed multiplexity and multireciprocity as:

$$\mu_{bin}^{\alpha,\beta} = \frac{m_{bin}^{\alpha,\beta} - \langle m_{bin}^{\alpha,\beta}\rangle}{1-\langle m_{bin}^{\alpha,\beta}\rangle}\ ; \qquad \rho_{bin}^{\alpha,\beta} = \frac{r_{bin}^{\alpha,\beta} - \langle r_{bin}^{\alpha,\beta}\rangle}{1-\langle r_{bin}^{\alpha,\beta}\rangle} \tag{10}$$

where $m_{bin}^{\alpha,\beta}$ and $r_{bin}^{\alpha,\beta}$ are the previously defined observed directed multiplexity and multireciprocity, while $\langle m_{bin}^{\alpha,\beta}\rangle$ and $\langle r_{bin}^{\alpha,\beta}\rangle$ are their expected values with respect to the chosen reference model. These filtered quantities are directly informative about the real correlations between layers: indeed, since $\mu^{\alpha,\beta}$ and $\rho^{\alpha,\beta}$ range in $[-1,1]$, positive values represent positive correlations (respectively, stronger inter-layer reciprocity), negative values are associated to anticorrelated (respectively, antireciprocated) pairs of layers, while pairs of uncorrelated layers are characterized by multiplexity and multireciprocity values comparable with 0.

It can be shown, with a straightforward extension of the calculations reported in [23], that such expected values, when the directed binary configuration model (DBCM) is considered as benchmark, just require the computation of the minimum between two binary variables, which is given by:

$$\langle\min\{a_{ij}^{\alpha},a_{ij}^{\beta}\}\rangle_{DBCM}=p_{ij}^{\alpha}p_{ij}^{\beta}\;;\qquad\langle\min\{a_{ij}^{\alpha},a_{ji}^{\beta}\}\rangle_{DBCM}=p_{ij}^{\alpha}p_{ji}^{\beta}\tag{11}$$

with the same notation introduced previously.

It is worth pointing out that, while the intra-layer multiplexity $m_{bin}^{\alpha,\alpha}$ leads by construction to the maximum value of similarity (that is, $m_{bin}^{\alpha,\alpha}=1$), the intra-layer multireciprocity $r_{bin}^{\alpha,\alpha}$ provides important information regarding the reciprocity observed within a given layer. Indeed, it can be shown that it is nothing but the usual expression of reciprocity introduced for unweighted monoplex networks [12].

In this context, we can therefore adopt an aggregation procedure, in order to compare the values of multireciprocity for the different pairs of layers with the global reciprocity of the aggregated network; we define the latter in the following way:

$$a_{ij}^{aggr}=\begin{cases}1 & if\ \exists\ \alpha\ \mid\ a_{ij}^{\alpha}=1\\ 0 & \text{otherwise}\end{cases}\tag{12}$$

The global raw reciprocity is therefore given by:

$$r_{bin}^{aggr}=\frac{\sum_{i\neq j}\min\{a_{ij}^{aggr},a_{ji}^{aggr}\}}{L^{aggr}}\tag{13}$$

with the same notation as before, but referred to the aggregated network; it is then possible to define the corresponding filtered quantity, in analogy with (10).

2.3. Weighted multiplexity and multireciprocity

Similarly to the unweighted analysis, we can study weighted correlations and inter-layer reciprocities in a directed multiplex by means of the weighted directed multiplexity and multireciprocity:

$$m_{w}^{\alpha,\beta}=\frac{2\sum_{i\neq j}\min\{w_{ij}^{\alpha},w_{ij}^{\beta}\}}{W^{\alpha}+W^{\beta}}\;;\qquad r_{w}^{\alpha,\beta}=\frac{2\sum_{i\neq j}\min\{w_{ij}^{\alpha},w_{ji}^{\beta}\}}{W^{\alpha}+W^{\beta}}\tag{14}$$

where w_{ij}^{α} is the weight associated to the directed link from node i to node j in layer α and W^{α} is the total weight of the links in layer α (analogue notation for layer β).

In this case, the required introduction of the rescaled directed multiplexity and multireciprocity:

$$\mu_{w}^{\alpha,\beta}=\frac{m_{w}^{\alpha,\beta}-\langle m_{w}^{\alpha,\beta}\rangle}{1-\langle m_{w}^{\alpha,\beta}\rangle}\;;\qquad\rho_{w}^{\alpha,\beta}=\frac{r_{w}^{\alpha,\beta}-\langle r_{w}^{\alpha,\beta}\rangle}{1-\langle r_{w}^{\alpha,\beta}\rangle}\tag{15}$$

are based on the following expressions for the expected value of the minimum, when the directed weighted configuration model (DWCM) is taken into account as a benchmark (calculations are straightforward modifications of [23]):

$$\langle \min\{w_{ij}^{\alpha}, w_{ij}^{\beta}\}\rangle_{DWCM} = \frac{p_{ij}^{\alpha} p_{ij}^{\beta}}{1 - p_{ij}^{\alpha} p_{ij}^{\beta}} \tag{16}$$

$$\langle \min\{w_{ij}^{\alpha}, w_{ji}^{\beta}\}\rangle_{DWCM} = \frac{p_{ij}^{\alpha} p_{ji}^{\beta}}{1 - p_{ij}^{\alpha} p_{ji}^{\beta}} \tag{17}$$

We can furthermore analyze the weighted reciprocity of the network resulting from the aggregation of the various layers by means of the usual weighted reciprocity [15]:

$$r_w^{aggr} = \frac{\sum_{i \neq j} \min\{w_{ij}^{aggr}, w_{ji}^{aggr}\}}{W^{aggr}} \tag{18}$$

together with its corresponding filtered value, similarly to (15). In this case, the aggregating procedure consists of:

$$w_{ij}^{aggr} = \sum_{\alpha=1}^{M} w_{ij}^{\alpha} \tag{19}$$

In the following Section, we will measure the above quantities on a real multi-layer system and explore many empirical patterns that allow us to model the observed levels of multiplexity and multiplexity.

3. Results

3.1. Data

In order to test the previous definitions on a real-world system, we analyze the World Trade Multiplex (WTM), namely the graph representing the import-export bilateral relations between countries in different products, as provided by the BACI-Comtrade dataset [24]. Indeed, it is possible to characterize this system as a multiplex [25], where each layer stands for a different commodity, for instance exploiting the standard 2-digits HS1996 classification [26] of traded goods. In particular, we will consider a multi-layer representation defined by 207 nodes (countries) and 96 layers (commodities), as reported in 2011.

Firstly, we will take into account only the topology of the various layers, thus disregarding the information about the volume of the disaggregated trade. Secondly, we will consider a weighted approach, therefore taking into account also the value of import and export between each pair of countries and adding a further level of complexity.

It should be noted that we can simply recover the aggregated directed trade relations between any two countries i and j (i.e. the collapsed monoplex trade network) by summing all the values of i's exports to j over all commodities, as stated in the previous Section. We will therefore compare the obtained values of multireciprocity for any pair of commodities with the usual definition of reciprocity (both binary and weighted) for the monoplex represented by the aggregated trade network.

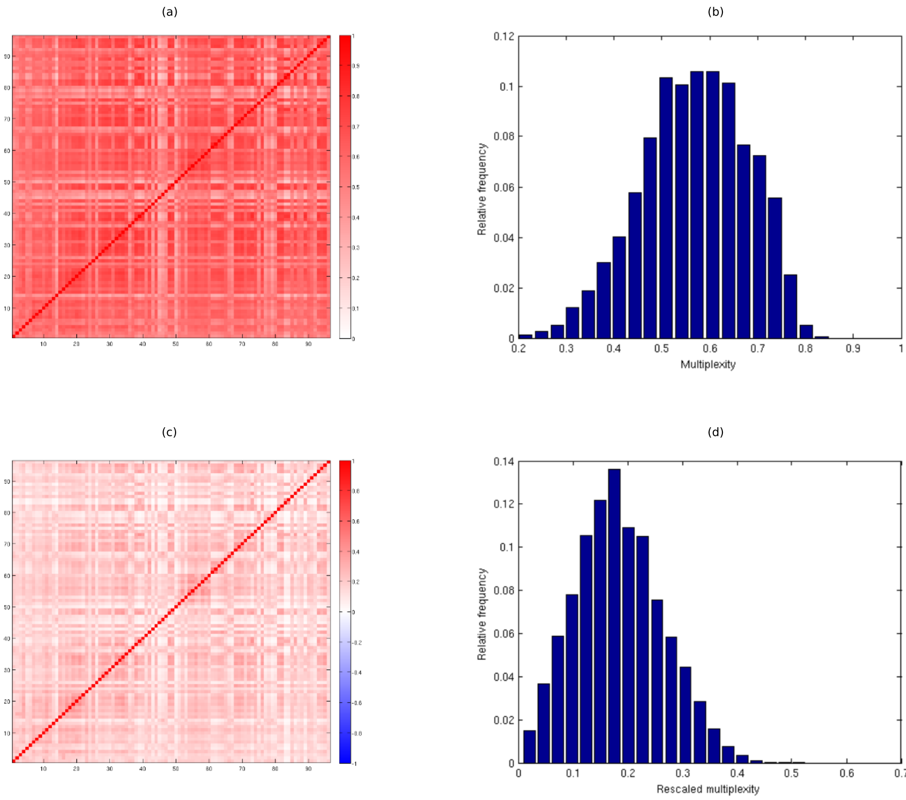

Figure 1. Top panels: binary directed raw multiplexity color-coded matrix (a) and the corresponding distribution of pairwise multiplexities $m_{bin}^{\alpha,\beta}$ (b). Bottom panels: binary directed rescaled multiplexity color-coded matrix (c) and the corresponding distribution of pairwise values $\mu_{DBCM}^{\alpha,\beta}$ (d).

3.2. Binary multiplexity and multireciprocity

In Figure 1(a) we report the color-coded matrix showing the directed overlap between pairs of layers, together with its corresponding distribution (Figure 1(b)). A high overlap can be observed for most of the pairs of commodities even when both import and export are taken into consideration separately, in agreement to what has been reported in [23]. Furthermore, Figure 1(c) shows the color-coded matrix of rescaled values of directed multiplexity $\mu_{DBCM}^{\alpha,\beta}$ after discarding the information already encoded into the corresponding binary configuration model; clearly, a significant amount of correlation is destroyed, but we can see from the corresponding distribution in Figure 1(d) that all the values are still strictly positive, thus pointing out a positive correlation between each pair of traded commodities.

It has been shown [12] that the International Trade Network exhibits a high level of reciprocity when the aggregated trade between countries is considered; however, it is interesting to see whether such property is preserved also at the single-layer level. We check this by means of the previously defined multireciprocity.

Figure 2(a) shows the color-coded matrix reporting the pairwise values of binary multireciprocity. Significant raw inter-layer reciprocities can be measured when most of the pairs of commodities are taken into account, but the multireciprocity distribution reported in Figure 2(b) provides a novel kind of information: the aggregated network exhibits a higher value of binary reciprocity with respect to any pair of commodities. As we have already mentioned, the intra-layer multireciprocity $m^{\alpha,\alpha}$, unlike the intra-layer multiplexity, is well-defined and corresponds to the usual measure of reciprocity introduced for monoplex networks. We observe that the intra-layer reciprocity values (shown in the main diagonal of the matrix) do not look different from the neighbouring inter-layer multireciprocity values, as the diagonal is indistinguishable from the rest of the matrix. This means 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 tendency to reciprocate is not specific to each individual layer, but rather to groups of layers.

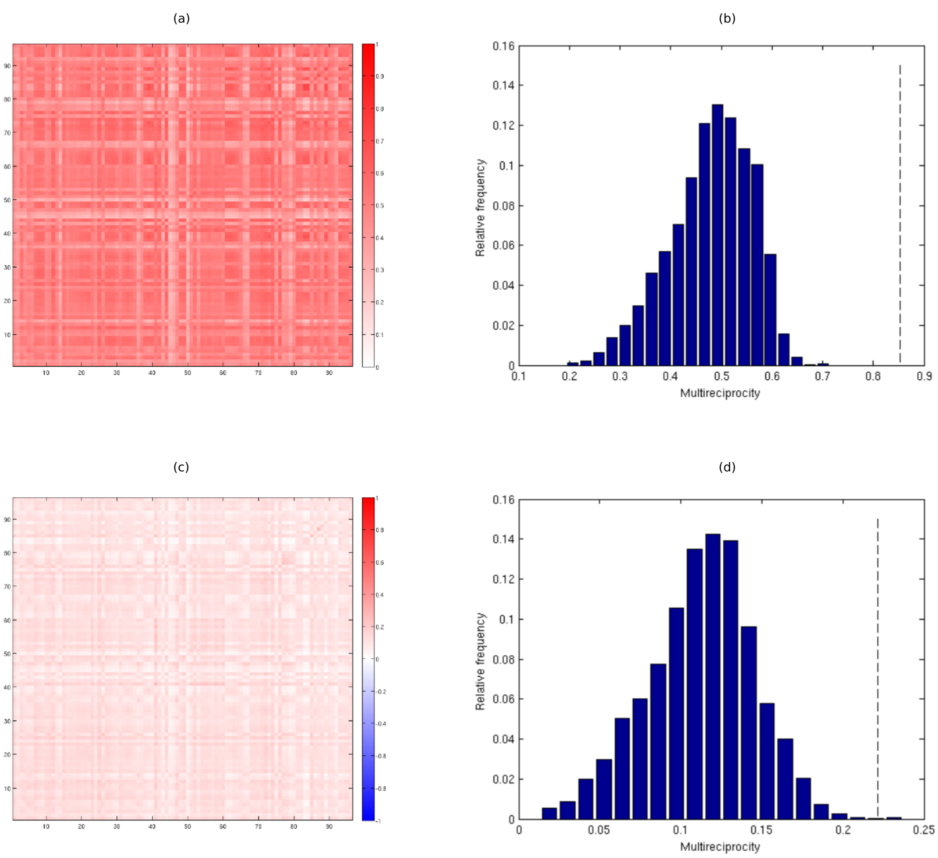

Figure 2. Top panels: binary raw multireciprocity color-coded matrix (a) and the corresponding distribution of pairwise values $r_{bin}^{\alpha,\beta}$ (b). Bottom panels: binary rescaled multireciprocity color-coded matrix (c) and the corresponding distribution of pairwise values $\rho_{DBCM}^{\alpha,\beta}$ (d). The dashed lines represent the value of reciprocity associated to the aggregated network.

In Figure 2(c) and (d) we report the color-coded matrix and the corresponding distribution of binary rescaled multireciprocity $\rho_{bin}^{\alpha,\beta}$. As already said for the multiplexity, we show that most of the correlations between links a_{ij}^{α} and a_{ji}^{β} are actually encoded in the degree sequences of the considered layers. Indeed, we observe that all the pairs of layers exhibit positive multireciprocities, but such values are significantly reduced with respect to the corresponding raw values. Moreover, we show that the aggregated reciprocity is still higher than almost all the pairs of commodities ‡.

When we look at the raw multiplexity matrices (for instance, in Figure 1(a)) and the corresponding raw multireciprocity matrices (such as in Figure 2(a)), we see the appearance of similar patterns. Such behaviour is explicitly shown in Figure 3(a), where we report the scatter plots of pairwise multreciprocity values versus the corresponding directed multiplexity values. However, such behaviour is partially lost when we look at the filtered values, as shown in Figure 3(b).

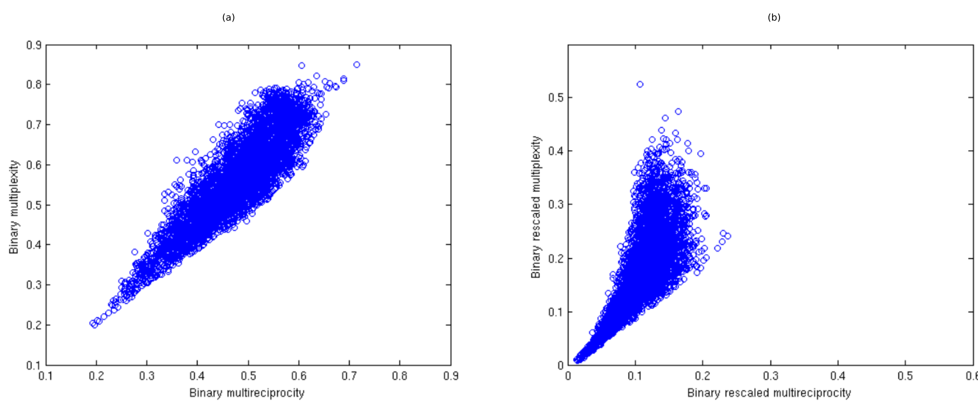

Figure 3. Scatter plots of binary multireciprocity values versus binary directed multiplexity values. Left: raw values ($r_{bin}^{\alpha,\beta}$ vs $m_{bin}^{\alpha,\beta}$); right: rescaled values ($\rho_{DBCM}^{\alpha,\beta}$ vs $\mu_{DBCM}^{\alpha,\beta}$).

3.3. Weighted multiplexity and multireciprocity

We now perform a weighted analysis of the World Trade Multiplex, that is, we also take into account the value of imports and exports observed between countries. In this context, we apply the weighted quantities introduced in the previous Section in order to study correlations between weighted layers of the multiplex.

‡ It is worth noticing that values of reciprocity reported in [12] and [15], although referred to the same real-world system, are actually calculated on a different dataset with respect to our analysis, based on the BACI–Comtrade dataset [24].

In Figure 4(a) and (b) we show the color-coded matrix referred to the weighted directed multiplexity values $m_w^{\alpha,\beta}$ and the corresponding distribution. We clearly see that, even though several pairs of commodities are still strongly overlapping, the highest fraction of pairs of layers are characterized by values of multiplexity lower than 0.2; indeed, such weighted overlap provides a more refined information with respect to the unweighted case, and this generalized reduction in the amount of correlation (w.r.t Figure 1(a)), is therefore expected. In Figure 4(c) and (d) we report the color-coded

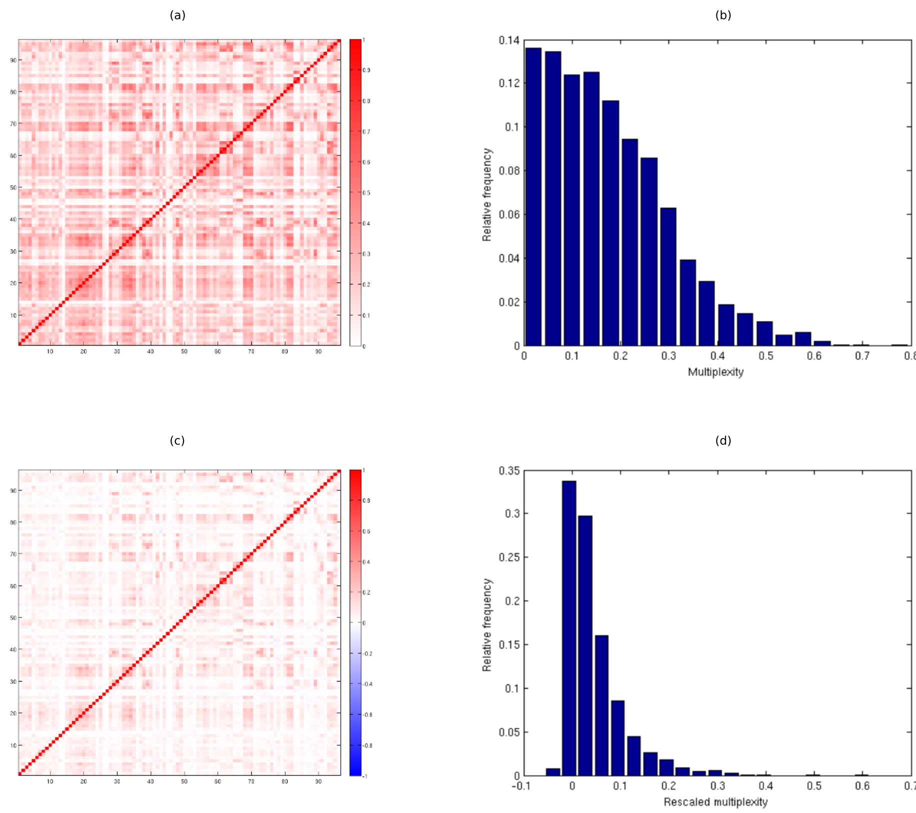

Figure 4. Top panels: weighted directed raw multiplexity color-coded matrix (a) and the corresponding distribution of pairwise multiplexities $m_w^{\alpha,\beta}$ (b). Bottom panels: weighted directed rescaled multiplexity color-coded matrix (c) and the corresponding distribution of pairwise values $\mu_{DWCM}^{\alpha,\beta}$ (d).

matrix and the corresponding distribution related to the weighted rescaled multiplexity values $\mu_{DWCM}^{\alpha,\beta}$; in this case, we show that much information about the correlations between commodities are encoded into the strength sequences of the different layers. Moreover, we see that some pairs actually exhibit negative correlations, after applying the weighted configuration model, even though the distribution is far from being symmetric.

We then analyze the weighted reciprocity of the World Trade Multiplex at the commodity level by means of the weighted multireciprocity. This property has been studied at the aggregated level [15] and it has been shown that the International Trade

Network is a clear example of strongly positively reciprocated system. In Figure 5(a) and (b) we report the pairwise multireciprocity matrix together with its distribution; we can clearly see, analogously to the binary case, that the aggregated network exhibits a reciprocity which is significantly higher than the multireciprocity associated to any pair of layers; however, several pairs of commodities are characterized by noticeable inter-layer reciprocity. Similar considerations can be done when the rescaled multireciprocity

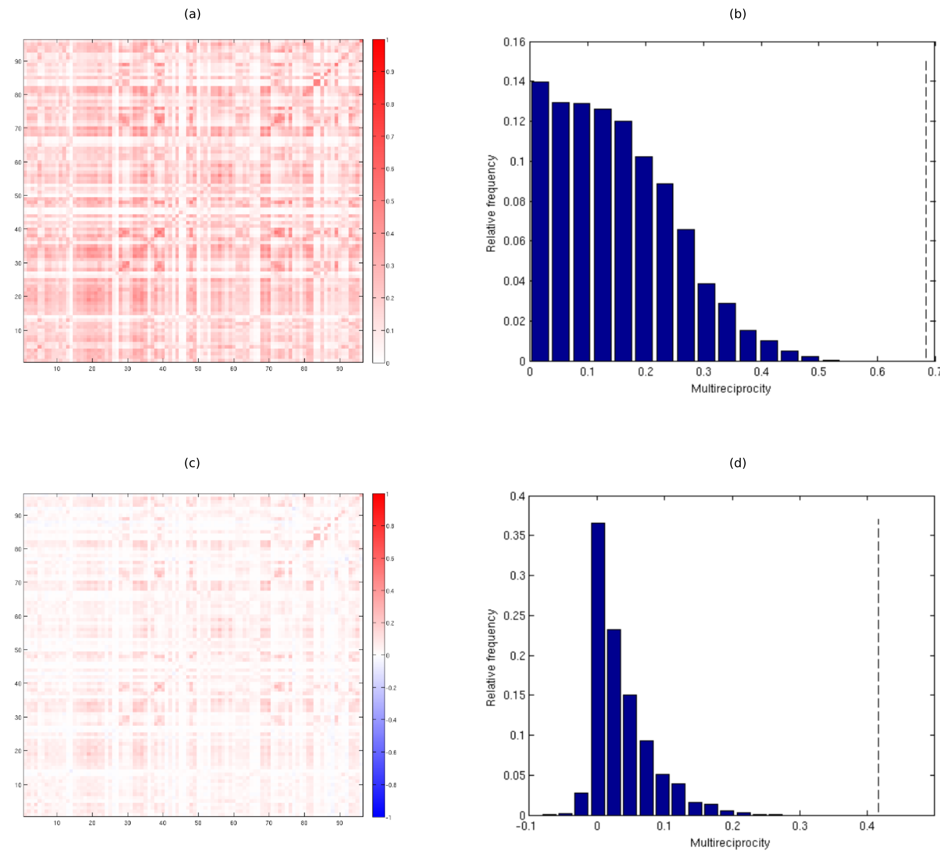

Figure 5. Top panels: weighted raw multireciprocity color-coded matrix (a) and the corresponding distribution of pairwise values $r_w^{\alpha,\beta}$ (b). Bottom panels: weighted rescaled multireciprocity color-coded matrix (c) and the corresponding distribution of pairwise values $\rho_{DWCM}^{\alpha,\beta}$ (d). The dashed lines represent the value of weighted reciprocity associated to the aggregated network.

is considered, as shown in Figure 5(c) and (d).

Furthermore, if we look at the plots reporting the relation between values of weighted multireciprocity and weighted directed multiplexity (Figure 6), we observe a clear linear trend - although more scattered than the corresponding unweighted case when we consider the raw values (a) - while the trend becomes even more clear for the filtered values, as shown in (b).

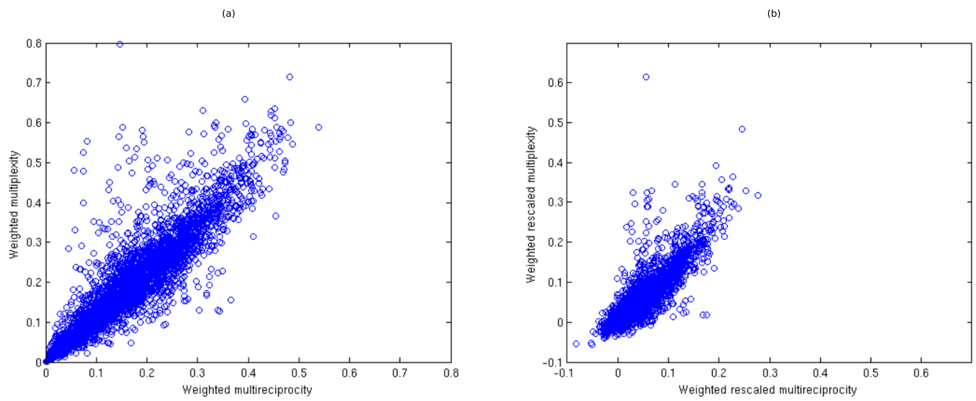

Figure 6. Scatter plots of weighted multireciprocity values versus weighted directed multiplexity values. Left: raw values ($r_w^{\alpha,\beta}$ vs $m_w^{\alpha,\beta}$); right: rescaled values ($\rho_{DWCM}^{\alpha,\beta}$ vs $\mu_{DWCM}^{\alpha,\beta}$).

3.4. Patterns of multireciprocity: reconstructing joint multi-layer connection probabilities from the multireciprocity matrix

In this Section, we identify robust empirical patterns in the WTM and show that the multireciprocity matrices allow us to reconstruct the joint multi-layer connection probabilities from the marginal single-layer ones, thus retrieving inter-layer information like $p_{ij}^{\alpha\leftrightarrow\beta}$ from single-layer properties such as p_{ji}^{α}.
In the previous Section we introduced the expressions for the expected degree sequences, both single-layer (3) and pairwise reciprocated (4); though, we can easily apply such definitions to the observed system, as this reduces to a shift from the variables p_{ji}^{α} and $p_{ij}^{\alpha\leftrightarrow\beta}$ (given by the model) to a_{ji}^{α} and $a_{ij}^{\alpha\leftrightarrow\beta}$ (directly measured on the observed network). Their application to the World Trade Multiplex results in what is shown in Figure 7; clearly, a linear trend can be inferred from the scattered plot of the pairwise reciprocated degree sequence versus the in-degree sequence in one of the layers (four pairs of commodities are shown, but similar plots can be observed in the other cases).

Hence, phenomenologically we observe that the conditional probability defined in (2) is actually independent from the considered pair of nodes:

$$r_{ij}^{\alpha\beta} = \frac{p_{ij}^{\alpha\leftrightarrow\beta}}{p_{ji}^{\beta}} = \frac{\langle a_{ij}^{\alpha} a_{ji}^{\beta}\rangle}{\langle a_{ji}^{\beta}\rangle} \sim r^{\alpha\beta} \tag{20}$$

Since the transformation $i \mapsto j$ together with $\alpha \mapsto \beta$ keeps the quantities unaffected, we have:

$$r^{\alpha\beta}\langle a_{ji}^{\beta}\rangle \sim \langle a_{ij}^{\alpha} a_{ji}^{\beta}\rangle = \langle a_{ji}^{\beta} a_{ij}^{\alpha}\rangle \sim r^{\beta\alpha}\langle a_{ij}^{\alpha}\rangle \tag{21}$$

Furthermore, summing (21) over i and j, we get:

$$r^{\alpha\beta}L^{\beta}=r^{\beta\alpha}L^{\alpha}\ ; \tag{22}$$

From (20), we immediately have:

$$r^{\alpha\beta}\langle a_{ji}^{\beta}\rangle\sim\langle a_{ij}^{\alpha}a_{ji}^{\beta}\rangle \tag{23}$$

Summing the previous (23) and inverting the obtained expression:

$$r^{\alpha\beta}=\frac{\sum_{i,j}\langle a_{ij}^{\alpha}a_{ji}^{\beta}\rangle}{\sum_{i,j}\langle a_{ji}^{\beta}\rangle}=\frac{\sum_{i,j}a_{ij}^{\alpha}a_{ji}^{\beta}}{L^{\beta}} \tag{24}$$

Therefore we have:

$$\frac{2}{r_{bin}^{\alpha\beta}}=\frac{2\left(L^{\alpha}+L^{\beta}\right)}{2\sum_{i,j}a_{ij}^{\alpha}a_{ji}^{\beta}}=\frac{1}{r^{\alpha\beta}}+\frac{1}{r^{\beta\alpha}} \tag{25}$$

where $r_{bin}^{\alpha\beta}$ is the entry of the raw multireciprocity matrix shown in Figure 2(a) and $r^{\alpha\beta}$ is derived from the empirical relationship between $k_i^{\beta,in}$ and $k_i^{\alpha\leftrightarrow\beta}$. Thus, $r_{bin}^{\alpha\beta}$ is the harmonic mean of the conditional probabilities $r^{\alpha\beta}$ and $r^{\beta\alpha}$. Applying (22) to the previous expression, we get:

$$\begin{aligned}\frac{2}{r_{bin}^{\alpha\beta}}&=\frac{1}{r^{\alpha\beta}}\left(1+\frac{r^{\alpha\beta}}{r^{\beta\alpha}}\right)=\\&=\frac{1}{r^{\alpha\beta}}\left(1+\frac{L^{\alpha}}{L^{\beta}}\right)=\\&=\frac{L^{\alpha}+L^{\beta}}{r^{\alpha\beta}L^{\beta}}\end{aligned} \tag{26}$$

Hence, the value of the angular coefficient in the plots $k_i^{\alpha\leftrightarrow\beta}$ vs $k_i^{\beta,in}$ should be:

$$r^{\alpha\beta}=\frac{L^{\alpha}+L^{\beta}}{2L^{\beta}}r_{bin}^{\alpha\beta} \tag{27}$$

Indeed, in Figure 7 we show that the actual best fit curves (in the form $y=a\cdot x$ since if $k_i^{\beta,in}=0$ then it is necessarily true that $k_i^{\alpha\leftrightarrow\beta}=0$) almost coincide with the expected ones according to (27). In this context, it turns out phenomenologically that the most minimal model one can design in order to reproduce the observed values of pairwise multireciprocity builds on the information about the total number of reciprocated links $L^{\alpha\leftrightarrow\beta}$ for any ordered pair of layers (α,β) (together with the aforementioned in- and out-degree sequences in each layer).

Similarly to the binary case, from Figure 8 phenomenologically we observe (again switching to the observed values $w_{ij}^{\alpha\leftrightarrow\beta}$ and w_{ji}^{β} instead of the expected ones) that the conditional probability defined in (6) is actually independent from the considered pair of nodes:

$$R_{ij}^{\alpha\beta}=\frac{\langle\min\{w_{ij}^{\alpha},w_{ji}^{\beta}\}\rangle}{\langle w_{ji}^{\beta}\rangle}\sim R^{\alpha\beta} \tag{28}$$

Applying the same transformations $i\mapsto j$ and $\alpha\mapsto\beta$ we get:

$$R^{\alpha\beta}\langle w_{ji}^{\beta}\rangle\sim\langle\min\{w_{ij}^{\alpha},w_{ji}^{\beta}\}\rangle=\langle\min\{w_{ji}^{\beta},w_{ij}^{\alpha}\}\rangle\sim R^{\beta\alpha}\langle w_{ij}^{\alpha}\rangle \tag{29}$$

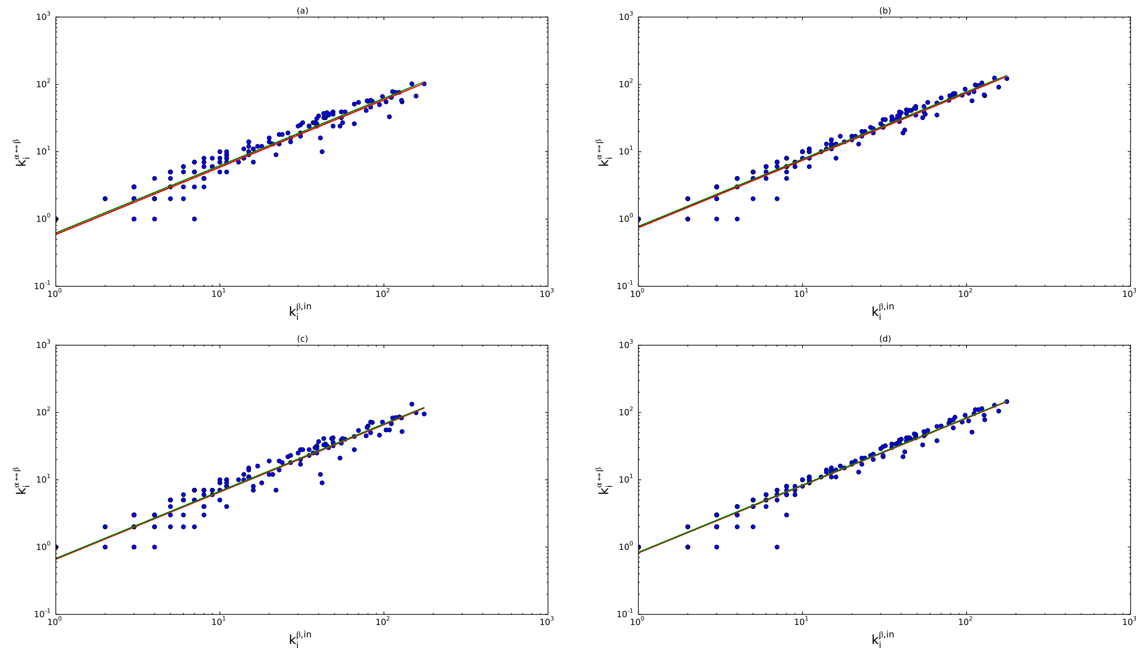

Figure 7. In-degree of layer β versus inter-layer reciprocated degree for 4 different pairs of commodities: inorganic chemicals (a), plastics (b), iron and steel (c), electric machinery (d) versus trade in cereals. Blue dots: real data; red line: best fit; green line: expected trend according to (27).

Summing (29) over i and j, we have:

$$R^{\alpha\beta}W^{\beta} = R^{\beta\alpha}W^{\alpha} \tag{30}$$

Similarly, inverting (28) we obtain:

$$R^{\alpha\beta}\langle w_{ji}^{\beta}\rangle \sim \langle \min\{w_{ij}^{\alpha}, w_{ji}^{\beta}\}\rangle \tag{31}$$

and summing the previous expression, as in the binary case:

$$R^{\alpha\beta} = \frac{\sum_{i,j}\langle \min\{w_{ij}^{\alpha}, w_{ji}^{\beta}\}\rangle}{\sum_{i,j}\langle w_{ji}^{\beta}\rangle} = \frac{\sum_{i,j} \min\{w_{ij}^{\alpha}, w_{ji}^{\beta}\}}{W^{\beta}} \tag{32}$$

Therefore we get:

$$\frac{2}{r_w^{\alpha\beta}} = \frac{2\left(W^{\alpha} + W^{\beta}\right)}{2\sum_{i,j}\min\{w_{ij}^{\alpha}, w_{ji}^{\beta}\}} = \frac{1}{R^{\alpha\beta}} + \frac{1}{R^{\beta\alpha}} \tag{33}$$

where $r_w^{\alpha\beta}$ represents the entry of the raw weighted multireciprocity matrix (Figure5(a)) and $R^{\alpha\beta}$ is derived from the empirical relationship between $s_i^{\beta,in}$ and $s_i^{\alpha\leftrightarrow\beta}$. In analogy with the binary case, $r_w^{\alpha\beta}$ is therefore the harmonic mean of the conditional probabilities $R^{\alpha\beta}$ and $R^{\beta\alpha}$, as previously defined. Applying (30) to the previous expression, we get:

$$\frac{2}{r_w^{\alpha\beta}} = \frac{1}{R^{\alpha\beta}}\left(1 + \frac{R^{\alpha\beta}}{R^{\beta\alpha}}\right) =$$

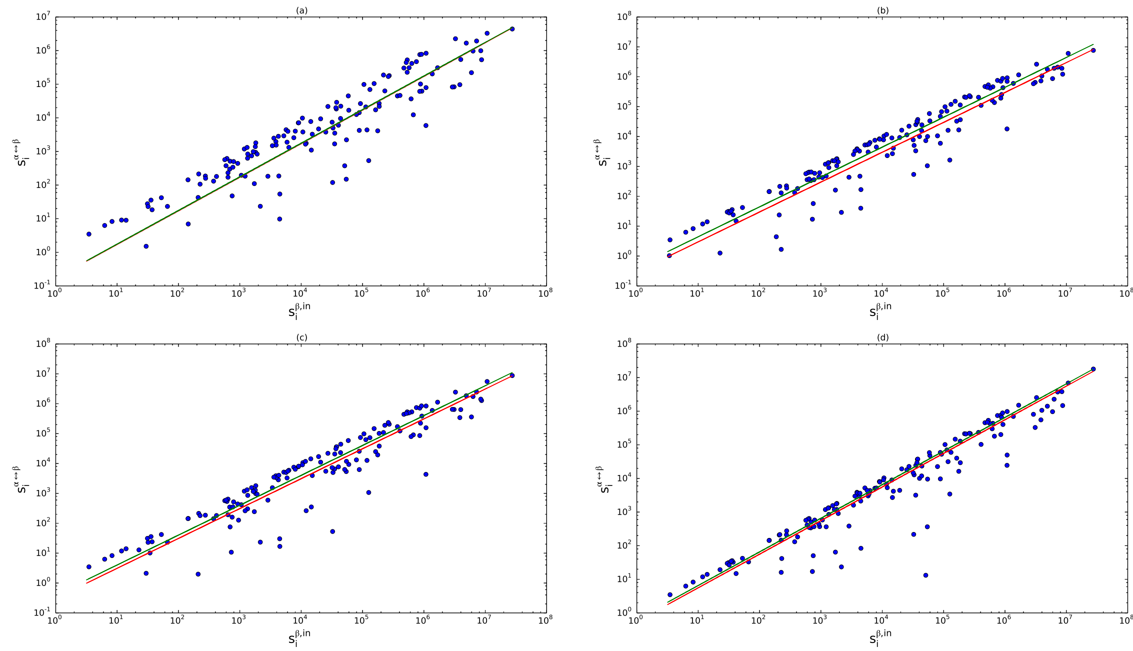

Figure 8. In-strength of layer β versus inter-layer reciprocated strength for 4 different pairs of commodities: inorganic chemicals (a), plastics (b), iron and steel (c), electric machinery (d) versus trade in cereals. Blue dots: real data; red line: best fit; green line: expected trend according to (35).

$$= \frac{1}{R^{\alpha\beta}}\left(1 + \frac{W^\alpha}{W^\beta}\right) =$$
$$= \frac{W^\alpha + W^\beta}{R^{\alpha\beta} W^\beta} \tag{34}$$

Thus, the value of the angular coefficient in the plots $s_i^{\alpha\leftrightarrow\beta}$ vs $s_i^{\beta,in}$ should be, in the weighted case:

$$R^{\alpha\beta} = \frac{W^\alpha + W^\beta}{2W^\beta} r_w^{\alpha\beta} \tag{35}$$

in perfect analogy with the unweighted case. Indeed, in Figure 8 we show the comparison between the actual fit lines (again in the form $y = a \cdot x$ since $s_i^{\beta,in} = 0$ implies $s_i^{\alpha\leftrightarrow\beta} = 0$) and the expected ones according to (35): the agreement is clear and is robust across different pairs of commodities. Therefore, in analogy to the unweighted case, here the most minimal model suitable to reproduce the observed values of pairwise weighted multireciprocity is based on the total reciprocated weight $W^{\alpha\leftrightarrow\beta}$ for any ordered pair of layers (α, β), accompanied by the in- and out-strength sequences measured in any layer.

4. Conclusions

The study of multi-layer networks has been deeply pursued in the last few years, by means of the introduction of several novel quantities characterizing such systems and the dynamical processes acting on top of them. However, a crucial property like the reciprocity has not been generalized for these multiplexes yet. In this work, we have faced this issue by defining the so-called multireciprocity, in order to detect the tendency of pairs of nodes to form mutual connections (in opposite directions) in different layers; such quantity has been introduced for unweighted and weighted systems. Furthermore, we have defined the directed multiplexity - again, for both binary and weighted networks - to characterize the similarity between layers of a directed multiplex network.

We have shown, by testing such quantities on the World Trade Multiplex, that significant correlation and inter-layer reciprocity can be observed for most of the pairs of layers in the binary case. In addition, when the weighted links are taken into account such properties are still present, even though the overlap between commodities and the multireciprocity exhibit lower values with respect to the binary case, due to the unbalance between weights in different layers (that is, between values of import and export in different commodities).

Moreover, our results show that, for the World Trade Multiplex, a large amount of correlation and inter-layer reciprocity is actually encoded in the degree and strength sequences associated to the various layers of the system, as we observe after a comparison between the raw measured values and the expected ones according to proper null models.

In order to exploit these new measures of multireciprocity to model the system, we have analyzed the behaviour of the pairwise reciprocated degree as a function of the in-degree for any node of the multiplex network. This dependence is based on the the conditional probabilities of observing a link from i to j in layer α given that a link from j to i is observed in β and is in principle peculiar of each node.

However, we have phenomenologically observed a linear trend between the pairwise reciprocated degree sequence and the in-degree sequence in one of the layers; in this context, these conditional connection probabilities do not actually depend on the considered pair of nodes, but only on the pair of layers. Moreover, the entries of the multireciprocity matrix are the harmonic mean of the aforementioned probabilities. Similar considerations can be done in the weighted case, except for a different definition of the conditional probabilities.

This evidence shows that the multireciprocity matrices allow to reconstruct the joint connection probabilities from the marginal ones, hence bridging the gap between single-layer information and truly multiplex properties.

Such results highlight some crucial properties of the WTM, such as the high reciprocity, but provide new insights into the understanding of the characteristics of this network at the disaggregated level, therefore pointing out the importance of a multiplex approach to the study of such system.

Furthermore, these considerations open new perspectives in the definition of proper null models for directed multi-layer networks, since the introduction of the notions of multireciprocity and multiplexity as constraints, in addition to the degree or strength sequences, may be pursued. We believe that our findings can be important in order to properly characterize multi-layer networks and may affect the understanding of several dynamical processes acting on such systems.

Acknowledgments

This work was supported by the EU project MULTIPLEX (contract 317532). DG also acknowledges support from the Netherlands Organization for Scientific Research (NWO/OCW) and the Dutch Econophysics Foundation (Stichting Econophysics, Leiden, the Netherlands) with funds from beneficiaries of Duyfken Trading Knowledge BV, Amsterdam, the Netherlands.

Appendix A. Maximum Likelihood Method

We consider the multiplex $\overrightarrow{G} = (G_1, G_2, \ldots, G_M)$ as the superposition of M layers G_k ($k = 1, 2, \ldots, M$), each of them represented by a graph having in common with the others a given set of N nodes with the other ones [3]. It is a usual practice to define null models for networks as randomized network ensembles, represented by the family of graphs satisfying a given set of constraints on average (canonical ensembles). In this context, such approach has been generalized to multi-layer networks: a multiplex ensemble can be introduced by assigning a probability $P(\overrightarrow{G})$ to each multiplex, therefore the entropy S of the ensemble is:

$$S = -\sum_{\overrightarrow{G}} P(\overrightarrow{G}) \ln P(\overrightarrow{G}) \tag{A.1}$$

In order to design null models for multi-layers networks, the maximization of (A.1) subject to given constraints has to be performed. In particular, as we stated in the main text, we make use of the concept of uncorrelated multiplex ensembles, thus the probability of any multiplex in the ensemble can be factorized into the probabilities of the different layers G_k building that particular multiplex (this is due to the supposed lack of correlation between the presence of links in any two layers α and β); hence, such probability is given by:

$$P(\overrightarrow{G}) = \prod_{k=1}^{M} P_k(G_k) \tag{A.2}$$

Based on such probabilities, it is then possible to compute the expected values of any quantity of interest under the chosen null model.

We can therefore focus our analysis on a given layer α, thus exploiting the framework used for single-layer networks, as the same will hold separately for all the others. It has been shown that a proper choice of constraints, when designing a null model for a real network, is represented by the degree sequence of that network. Since we deal with directed graphs, we therefore enforce as constraints both the in-degree and the out-degree sequences, defining the so-called directed (binary, as we are now considering unweighted networks) configuration model (DBCM) [14].

However, finding such probabilities (and then the average values of the quantities we are interested in) for real networks is computationally demanding, as it requires the generation of several randomized networks on top of which we can measure such quantities. Hence, we use a fast and completely analytical method, building on the maximization of the likelihood [20, 21, 22]; such approach is able to provide the exact occurrence probabilities of the randomized networks with the same average constraints as the real network and, as a consequence, other properties such as the expected link probability p_{ij}^{α}.

In the unweighted case, the Maximum Likelihood Method reduces to solving the solution of following set of $2N$ coupled nonlinear equation (as we said, independently for each layer $\alpha = 1, 2, \ldots, M$ due to the uncorrelated assumption):

$$\sum_{i \neq j} \frac{x_i^{\alpha} y_j^{\alpha}}{1 + x_i^{\alpha} y_j^{\alpha}} = k_{i,out}^{\alpha} \quad \forall i = 1, 2, \ldots, N \tag{A.3}$$

$$\sum_{i \neq j} \frac{x_j^{\alpha} y_i^{\alpha}}{1 + x_j^{\alpha} y_i^{\alpha}} = k_{i,in}^{\alpha} \quad \forall i = 1, 2, \ldots, N \tag{A.4}$$

where $k_{i,out}^{\alpha}$ is the observed out-degree of node i in layer α, $k_{i,in}^{\alpha}$ the observed in-degree and the unknown variables $\{x_i^{\alpha}\}$ and $\{y_i^{\alpha}\}$ $(i = 1, \ldots, N)$ of the equations are the so-called $2N$ hidden variables associated to layer α. Hence, the expected value of the link probability p_{ij}^{α} is given by, for any pair of nodes (i, j) in any layer α:

$$p_{ij}^{\alpha} = \frac{x_i^{\alpha} y_j^{\alpha}}{1 + x_i^{\alpha} y_j^{\alpha}} \tag{A.5}$$

where now $\{x_i^{\alpha}\}$ and $\{y_i^{\alpha}\}$ are the values solving the previous set of equations.

Similar considerations can be made for directed weighted multiplexes, except for a change in the chosen constraints: we now enforce the in-strength and out-strength sequences of the real system (again, independently for each layer, due to the uncorrelated assumption that we still take into account), designing the so-called directed weighted configuration model (DWCM) [16]. Analogously to the binary case, the Maximum Likelihood Method, when applied to weighted networks, reduces to finding the solution to a set of $2N$ coupled nonlinear equations. Indeed, for any node i in any layer α, we have:

$$\sum_{i \neq j} \frac{x_i^{\alpha} y_j^{\alpha}}{1 - x_i^{\alpha} y_j^{\alpha}} = s_{i,out}^{\alpha} \tag{A.6}$$

$$\sum_{i \neq j} \frac{x_j^{\alpha} y_i^{\alpha}}{1 - x_j^{\alpha} y_i^{\alpha}} = s_{i,in}^{\alpha} \tag{A.7}$$

where $s_{i,out}^{\alpha}$ is the observed out-strength of node i in layer α, $s_{i,in}^{\alpha}$ the observed in-strength and the unknown variables of the equation represent the $2N$ hidden variables associated to that particular layer. It is then possible to use the solutions to the previous system of equations in order to find the expected link weight w_{ij}^{α} form node i to node j, which becomes:

$$w_{ij}^{\alpha} = \frac{x_i^{\alpha} y_j^{\alpha}}{1 - x_i^{\alpha} y_j^{\alpha}} \tag{A.8}$$

Based on such expressions for p_{ij}^{α} and w_{ij}^{α}, we can compute the expected values of higher-order properties, such as the directed multiplexity and the multireciprocity reported in the main text.

References

[1] Ebel H, Mielsch L-I and Bornholdt S 2002 *Phys. Rev. E* **66** 035103(R)
[2] Newman M E J 2004 *Phys. Rev. E* **70** 056131
[3] Boccaletti S, Bianconi G, Criado R, del Genio C I, Gomez-Gardeñes J, Romance M, Sendiña-Nadal I, Wang Z and Zanin M 2014 *Phys. Rep.* **544** 1
[4] Kivelä M, Arenas A, Barthelemy M, Gleeson J P, Moreno Y and Porter M A 2014 *J. Complex Netw.* **2** 203
[5] Magnani M, Micenková B and Rossi L 2013 arXiv:1303.4986
[6] Menichetti G, Remondini D, Panzarasa P, Mondragón R J and Bianconi G 2014 *PLoS ONE* **9(6)**: e97857
[7] Meyers L A, Newman M J and Pourbohloul B 2006 *J. Theor. Biol.* **240** 400
[8] Boguñá M and Serrano M A 2005 *Phys. Rev. E* **72** 016106
[9] Schnegg M 2006 *Int. J. Mod. Phys. C* **17** 1067
[10] Garlaschelli G and Loffredo M I 2005 *Physica A* **355(1)** 138
[11] Zhou C, Zemanova L, Zamora G, Hilgetag C C and Kurths J 2006 *Phys. Rev. Lett.* **97** 238103
[12] Garlaschelli D and Loffredo M I 2004 *Phys. Rev. Lett.* **93** 268701
[13] Newman M E J, Forrest S and Balthrop J 2002 *Phys. Rev. E* **66** 035101(R)
[14] Maslov S and Sneppen K 2002 *Science* **296** 910
[15] Squartini T, Picciolo F, Ruzzenenti F and Garlaschelli D 2013 *Sci. Rep.* **3** 2729
[16] Serrano M A and Boguñá M 2005 *AIP Conf. Proc.* **776** 101
[17] Bianconi G 2013 *Phys. Rev. E* **87** 062806
[18] Park J and Newman M E J 2004 *Phys. Rev. E* **70** 066117
[19] Roberts E S and Coolen A C C 2012 *Phys. Rev. E* **85** 046103
[20] Garlaschelli D and Loffredo M I 2008 *Phys. Rev. E* **78** 015101
[21] Squartini T and Garlaschelli D 2011 *New J. Phys.* **13** 083001
[22] T. Squartini, R. Mastrandrea, and D. Garlaschelli, New J. Phys. 17, 023052 (2015).
[23] Gemmetto V and Garlaschelli D 2015 *Sci. Rep.* **5** 9120
[24] Gaulier G and Zignago S 2010 *CEPII Working Paper* **23**
[25] Barigozzi M, Fagiolo G and Garlaschelli D 2010 *Phys. Rev. E* **81** 046104
[26] http://www.wcoomd.org