Reciprocity of weighted networks
- Supplementary Information -

Tiziano Squartini[*, 1], Francesco Picciolo[2], Franco Ruzzenenti[2], Diego Garlaschelli[1]
[1] *Instituut-Lorentz for Theoretical Physics, Leiden Institute of Physics, University of Leiden, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands and*
[2] *Department of Chemistry, University of Siena, Via A. De Gasperi 2, 53100 Siena, Italy*
(Dated: July 23, 2013)

In directed networks, reciprocal links have dramatic effects on dynamical processes, network growth, and higher-order structures such as motifs and communities. While the reciprocity of binary networks has been extensively studied, that of weighted networks is still poorly understood, implying an ever-increasing gap between the availability of weighted network data and our understanding of their dyadic properties. Here we introduce a general approach to the reciprocity of weighted networks, and define quantities and null models that consistently capture empirical reciprocity patterns at different structural levels. As we show, counter-intuitively, previous reciprocity measures based on the similarity of mutual weights are uninformative. By contrast, our measures allow to consistently classify different weighted networks according to their reciprocity, track the evolution of a network's reciprocity over time, identify patterns at the level of dyads and vertices, and distinguish the effects of flux (im)balances or other (a)symmetries from a true tendency towards (anti-)reciprocation.

PACS numbers: Valid PACS appear here

RECIPROCITY OF BINARY NETWORKS

Before considering the reciprocity of weighted networks, we briefly recall the basic definitions in the binary case, that were originally introduced to describe the mutual relations taking place between vertex pairs [1, 2].

Reciprocity as the fraction of bidirectional links

For binary, directed networks the reciprocity is defined as the fraction of links having a "partner" pointing in the opposite direction:

$$r^b \equiv \frac{L^{\leftrightarrow}}{L} \tag{1}$$

where $L = \sum_{i\neq j} a_{ij}$ and $L^{\leftrightarrow} = \sum_{i\neq j} a_{ij}a_{ji}$. The above quantity, r^b, is not independent on the link density (or connectance) $c \equiv \frac{L}{N(N-1)} = \frac{\sum_{i\neq j} a_{ij}}{N(N-1)} \equiv \bar{a}$: on the contrary, it can be shown that c is the expected value of r^b under the Directed Random Graph Model (DRG in what follows) [3, 4]. In the DRG, a directed link is placed with probability p between any two vertices, i.e. $\langle a_{ij}\rangle_{DRG} = p$, $\forall\, i, j$ (with $i \neq j$). This implies

$$\langle r^b\rangle_{DRG} \equiv \frac{\langle L^{\leftrightarrow}\rangle}{\langle L\rangle} = \frac{N(N-1)p^2}{N(N-1)p} = p \equiv \frac{L}{N(N-1)} = c \tag{2}$$

showing that the expected value of r^b coincides with the fundamental parameter of this null model, and hence depends on L and N. In order to assess whether there is positive or negative reciprocity, one should compare the measured r^b with its expected value $\langle r^b\rangle_{DRG}$. This means that r^b cannot be used to consistently rank networks with different values of L and N, because they have different reference values. Also, and consequently, r^b cannot be used to track the evolution of a network that changes in time, because L and/or N will also change [3].

Reciprocity as a correlation coefficient

This is why a different definition of reciprocity was proposed [3], trying to control for the time-varying properties by means of the Pearson correlation coefficient between the transpose elements of the adjacency matrix [5]:

$$\rho^b \equiv \frac{\sum_{i\neq j}(a_{ij}-c)(a_{ji}-c)}{\sum_{i\neq j}(a_{ij}-c)^2} = \frac{r^b - c}{1-c} = \frac{r^b - \langle r^b\rangle_{DRG}}{1-\langle r^b\rangle_{DRG}}. \tag{3}$$

A symmetrical adjacency matrix (as those for binary, undirected networks) represents a network with the highest values of r^b and ρ (both equal to 1), whereas a fully asymmetrical one, with *zero* values mirroring *unit* values on opposite sides of the main diagonal (like a triangular matrix), displays the lowest value, being $r^b = 0$ and $\rho = -c/(1-c)$) [3]. This meaningful definition of reciprocity automatically discounts density effects, i.e. the expectation value of r^b (under the DRG). As a result, consistent rankings and temporal analyses become possible in terms of ρ.

RECIPROCITY OF WEIGHTED NETWORKS

In what follows we provide additional information about the possible generalization of the reciprocity to the weighted case.

From binary to weighted: the first route

By looking at eq.(3), it is not clear whether a generalization to the weighted case should start from the first term on the left (i.e. as a correlation coefficient) or from the last term on the right (i.e. as the normalized excess from a

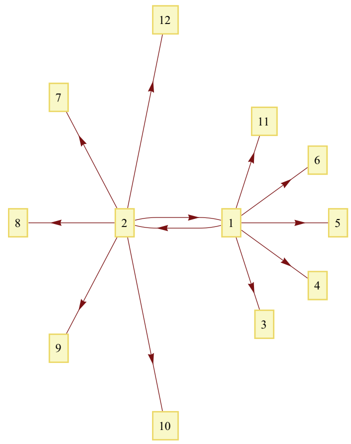

FIG. 1. A double-star network, with only one reciprocated pair of weights.

random expectation). This ambiguity comes from the fact that, for weighted networks, those two terms are no longer equivalent (as we now show). We therefore start by attempting the first route, and then consider the second one.

If we follow the binary recipe from left to right, we define the weighted reciprocity as the Pearson correlation coefficient (where, as usual, $\bar{w} = \frac{\sum_{i \neq j} w_{ij}}{N(N-1)} = \frac{W_{tot}}{N(N-1)}$). After some algebra, this implies

$$\rho \equiv \frac{\sum_{i \neq j}(w_{ij} - \bar{w})(w_{ji} - \bar{w})}{\sum_{i \neq j}(w_{ij} - \bar{w})^2} = \frac{r - c^w}{1 - c^w} \tag{4}$$

where, in order to produce a result formally equivalent to eq.(3), we have defined the weighted analogues of r and c as follows:

$$r \equiv \frac{\sum_{i \neq j} w_{ij} w_{ji}}{\sum_{i \neq j} w_{ij}^2}, \; c^w \equiv \frac{\bar{w}^2}{\sum_{i \neq j} w_{ij}^2 / N(N-1)} \tag{5}$$

Note that the equivalence $\bar{a} = c$, valid for the binary case, no longer holds: $\bar{w} \neq c^w$. The previous expressions generalize the binary ones and reduce to them when substituting the a_{ij}'s in place of the w_{ij}'s. Moreover, interestingly enough, the coefficient c^w can be expressed as a function of the weights' distribution mean, m, and standard deviation, s, or, in an equivalent way, as a function of the so-called coefficient of variation, $c_v = s/m$, as

$$c^w = \frac{m^2}{m^2 + s^2} = \frac{1}{1 + c_v^2}. \tag{6}$$

We could be tempted to interpret c^w as the weighted counterpart of the binary connectance and, r as the weighted counterpart of eq.(1). However, we can show a simple case for which the above "product-over-squares" definition above fails in measuring our intuitive notion of reciprocity. Let us consider a simple network like that in Fig. 1.

If we calculate r by choosing $w_{12} = w_{21}$, we obtain

$$r_1 = \frac{2w_{12}^2}{2w_{12}^2 + \sum_{i,j \neq (1,2),(2,1)} w_{ij}^2} \tag{7}$$

where the sum in the denominator includes all the weights different from the central ones. Now, let us imagine a second situation where $w_{21} = w_{12} + 1$; the calculations, now, would give

$$r_2 = \frac{2w_{12}(w_{12}+1)}{w_{12}^2 + (w_{12}+1)^2 + \sum_{i,j \neq (1,2),(2,1)} w_{ij}^2} \tag{8}$$

and we would intuitively require that $r_2 < r_1$, for every choice of the involved weights, because of the greater disparity between the two central flows. However, it can be shown that under certain circumstances exactly the opposite result is obtained, by simply changing the non-central weights. In fact, by choosing the latter to satisfy the condition

$$\sum_{i,j\neq(1,2),(2,1)} w_{ij}^2 > w_{12} \tag{9}$$

the very counter-intuitive result $r_2 - r_1 > 0$ is obtained. This shows that eq.(4) is not a good choice for a weighted extension of eq.(3).

Before considering the alternative route, we observe that we could also imagine to define a slightly different correlation coefficient, only between the two triangular blocks of the weighted adjacency matrix: the upper-diagonal one and the lower-diagonal one. This would be defined as

$$\rho' \equiv \frac{\sum_{i<j}(w_{ij} - \bar{w}_u)(w_{ji} - \bar{w}_l)}{\sqrt{\sum_{i<j}(w_{ij} - \bar{w}_u)^2 \sum_{i<j}(w_{ji} - \bar{w}_l)^2}} \tag{10}$$

where $\bar{w}_u \equiv \frac{\sum_{i>j} w_{ij}}{N(N-1)}$ is the upper-diagonal mean and $\bar{w}_l \equiv \frac{\sum_{i>j} w_{ji}}{N(N-1)}$ is the lower-diagonal mean. Again, this definition has an undesirable performance. This is evident if we imagine a matrix whose transposed entries are defined as w_{ij} and $w_{ji} \equiv \lambda w_{ij}$ (with $i < j$). In this case, we would have

$$\rho' = \frac{\sum_{i<j}(w_{ij} - \bar{w}_u)(\lambda w_{ij} - \lambda\bar{w}_u)}{\sqrt{\sum_{i<j}(w_{ij} - \bar{w}_u)^2 \sum_{i<j}(\lambda w_{ij} - \lambda\bar{w}_u)^2}} = 1 \tag{11}$$

independently of the value of λ! So we could arbitrarily rise or lower the value of λ, thus making the matrix more and more asymmetric, without measuring this effect at all. Note that this circumstance is impossible in the binary case, as all weights are forced to be either zero or one, and therefore the only allowed value for λ is one.

The two examples above show that correlation-based definitions of reciprocity, while having a satisfactory behaviour in the binary case, become problematic in the weighted one. Unfortunately, the few attempts that have been proposed so far in order to characterize the reciprocity of weighted networks [6–9] are all based on measures of correlation or symmetry between mutual weights. Later, we show that symmetry-based measures are also flawed. Together with our results above, this means that all the available measures fail in providing a consistent and interpretable characterizaton of the reciprocity of weighted networks.

From binary to weighted: the second route

We now consider the second route, i.e. a definition that starts from generalizing the last term in eq.(3). This means that we are now free to first generalize r in a satisfactory way, rather than as a forced effect of the correlation-based definition, and then calculate its expected value under some appropriate null model. To this end, we note that the binary nature of the variables defining r^b allows us to rewrite it in a very suggestive way:

$$r^b \equiv \frac{L^{\leftrightarrow}}{L} = \frac{\sum_{i\neq j} a_{ij}a_{ji}}{\sum_{i\neq j} a_{ij}} = \frac{\sum_{i\neq j} \min[a_{ij},\, a_{ji}]}{\sum_{i\neq j} a_{ij}}. \tag{12}$$

The previous relation is consistent with the intuitive meaning of reciprocity, as a measure of the quantity of mutually-exchanged flux between vertices. So we can extend this definition to the weighted case, to obtain

$$r \equiv \frac{W^{\leftrightarrow}}{W} = \frac{\sum_{i\neq j} \min[w_{ij},\, w_{ji}]}{\sum_{i\neq j} w_{ij}}. \tag{13}$$

where we have defined the *total reciprocated weight* as $W^{\leftrightarrow} \equiv \sum_{i\neq j} \min[w_{ij},\, w_{ji}]$. This definition does not suffer from the same limitations of the previous one. On the contrary, the more the difference between mutual links, the less the

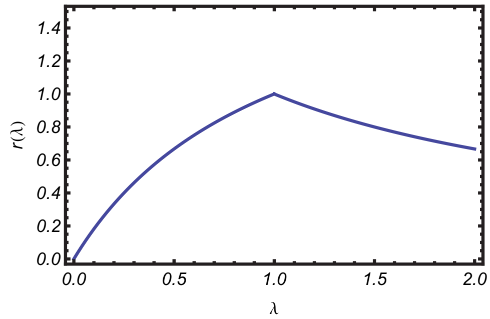

FIG. 2. The quantity r as a function of λ.

reciprocity, because the numerator would not change, while the denominator would become larger. Note that $r \leq 1$: in fact, since we are considering pairs of nodes at a time, we can rewrite it as

$$r = \frac{\sum_{i<j} (\min[w_{ij},\, w_{ji}] + \min[w_{ij},\, w_{ji}])}{\sum_{i<j} (\min[w_{ij},\, w_{ji}] + \max[w_{ij},\, w_{ji}])}. \tag{14}$$

Another advantage of this second definition is the possibility of mutuating from it the concept of *reciprocated strength* in the same way as the concept of reciprocated degree was defined:

$$k_i^{\leftrightarrow} \equiv \sum_{j(\neq i)} a_{ij} a_{ji} \quad \rightarrow \quad s_i^{\leftrightarrow} \equiv \sum_{j(\neq i)} \min[w_{ij},\, w_{ji}] \tag{15}$$

so that a very impressive definition of reciprocity can be given, as

$$r^b = \frac{\sum_i k_i^{\leftrightarrow}}{L} \quad \rightarrow \quad r = \frac{\sum_i s_i^{\leftrightarrow}}{W}. \tag{16}$$

A further feature of this quantity is its scale-invariance: if all the weights are multiplied by a scale factor, $w_{ij} \rightarrow \lambda w_{ij}$, r does not change, as shown below:

$$r_\lambda = \frac{\sum_{i\neq j} \min[\lambda w_{ij},\, \lambda w_{ji}]}{\sum_{i\neq j} \lambda w_{ij}} = \frac{\lambda \sum_{i\neq j} \min[w_{ij},\, w_{ji}]}{\lambda \sum_{i\neq j} w_{ij}} = r. \tag{17}$$

Moreover, in the case we had a matrix with transposed entries defined as w_{ij} and $w_{ji} \equiv \lambda w_{ij}$ (with $i < j$) as in the example considered before, we would find

$$r = \begin{cases} \frac{\sum_{i<j} 2w_{ij}}{\sum_{i<j} (\lambda+1) w_{ij}} = \frac{2}{(\lambda+1)}, & \text{if } \lambda > 1 \\ \frac{\sum_{i<j} 2w_{ij}}{\sum_{i<j} 2w_{ij}} = 1, & \text{if } \lambda = 1 \\ \frac{\sum_{i<j} 2\lambda w_{ij}}{\sum_{i<j} (\lambda+1) w_{ij}} = \frac{2\lambda}{(\lambda+1)}, & \text{if } \lambda < 1 \end{cases} \tag{18}$$

thus obtaining a continuous function with a global maximum in $\lambda = 1$ as it should be (see Fig. 2).

It follows that the appropriate weighted generalization of eq.(3) is

$$\rho_{NM} \equiv \frac{r - \langle r \rangle_{NM}}{1 - \langle r \rangle_{NM}} \tag{19}$$

where r is defined by eq.(13) and its expected value has to be computed according to a chosen null model (NM). Indeed, this choice also gives us the possibility to choose different null models, and compare their effects on ρ. From $r \leq 1$, it follows that $\rho \leq 1$.

NULL MODELS

In this section we describe in detail the three null models we considered in order to carry out our analysis. We adopt the formalism of Exponential Random Graphs or p^* models, which allows to obtain maximally random ensembles of networks with specified constraints. Exponential random graphs were first introduced in social network analysis [1, 2, 10, 11] and then recently rephrased within a maximum-entropy approach typical of statistical physics [12]. We adopt the latter notation, as it is more practical when, rather than approaching the problem using approximate techniques such as Markov Chain Monte Carlo or pseudo-likelihood [10, 11], one can solve the model analytically and obtain exact results as we do below.

Exponential Random Graphs are very useful when one needs to understand, as in our case, the expected effects of a given set of topological properties, $\vec{C}$ (such as the total weight, or the strength sequence) on the structure of networks. Recently, a method based on the maximum-likelihood principle was proposed [13] in order to fit exponential random graphs to a real-world graph $\mathbf{G}^*$ exactly [13]. This method provides null models which specify the effects of one or more constraints on the structure of the *particular* network $\mathbf{G}^*$, and hence allows to empirically detect patterns in the latter, identified as deviations from the model's predictions [13]. In the method, maximum-entropy exponential random graphs are generated by specifying an ensemble $\mathcal{G}$ of allowed graphs, and by looking for the probability $P(\mathbf{G}|\vec{\theta})$ of generating a single graph $\mathbf{G}$ in the ensemble in such a way that the Shannon entropy

$$S(\vec{\theta}) \equiv -\sum_{\mathbf{G}\in\mathcal{G}} P(\mathbf{G}|\vec{\theta}) \ln P(\mathbf{G}|\vec{\theta}) \tag{20}$$

is maximum, under the constraints that the probability is properly normalized, $\sum_{\mathbf{G}\in\mathcal{G}} P(\mathbf{G}|\vec{\theta}) = 1$, $\forall\vec{\theta}$, and that the expected value

$$\langle\vec{C}\rangle_{\vec{\theta}} \equiv \sum_{\mathbf{G}\in\mathcal{G}} \vec{C}(\mathbf{G})P(\mathbf{G}|\vec{\theta}) \tag{21}$$

of the set $\vec{C}$ of enforced topological properties equals the particular value $\vec{C}^* \equiv \vec{C}(\mathbf{G}^*)$ observed on the real network $\mathbf{G}^*$:

$$\langle\vec{C}\rangle_{\vec{\theta}^*} = \vec{C}^*. \tag{22}$$

In the above expressions, $\vec{\theta}$ is a vector of Langrange multipliers allowing to tune the value of $\langle\vec{C}\rangle_{\vec{\theta}}$, and $\vec{\theta}^*$ is the specific value of $\vec{\theta}$ that makes $\langle\vec{C}\rangle_{\vec{\theta}}$ coincide with $\vec{C}^*$, as dictated by the maximum-likelihood principle [14]. The solution to the above constrained maximization problem is

$$P(\mathbf{G}|\vec{\theta}^*) = \frac{e^{-H(\mathbf{G}|\vec{\theta}^*)}}{Z(\vec{\theta}^*)} \tag{23}$$

where

$$H(\mathbf{G}|\vec{\theta}^*) = \vec{\theta}^* \cdot \vec{C}(\mathbf{G}) \tag{24}$$

is sometimes called the *graph Hamiltonian* and

$$Z(\vec{\theta}^*) = \sum_{\mathbf{G}\in\mathcal{G}} e^{-H(\mathbf{G}|\vec{\theta}^*)} \tag{25}$$

is the *partition function*, ensuring that the probability is properly normalized. The above formal results translate into specific quantitative expectations when a particular choice of the constraints, $\vec{C}$, is made.

Once the numerical values of the Lagrange multipliers are found, they can be used to find the ensemble average, $\langle X\rangle^*$, of any topological property X of interest:

$$\langle X\rangle^* = \sum_{\mathbf{G}\in\mathcal{G}} X(\mathbf{G})P(\mathbf{G}|\vec{\theta}^*). \tag{26}$$

The exact computation of the expected values can be very diffcult. For this reason it is often necessary to rest on the linear approximation method even if, in what follows, the only approximation will be that of treating the expected value of a ratio, as the ratio of the expected values: $\langle n/d\rangle \simeq \langle n\rangle/\langle d\rangle$.

The next subsections will be devoted to the description of the null modes used in the main text.

The Directed Weighted Random Graph (WRG) model

We start with the simplest case, which is the most direct generalization of the binary, undirected random graph (Erdős-Rényi) model. For an ensenble of binary, undirected networks, it was shown [12] that, if the only constraint C is the total number L of links (i.e. $H(\mathbf{G},\theta) = \theta L$), then the probability $P(\mathbf{G}|\theta)$ coincides with that of the *Erdős-Rényi Random Graph Model*. In the latter, each pair of vertices is connected with the same probability p, all pairs of vertices being sampled independently of each other. In the framework of exponential random graphs, the probability p is simply a function of θ.

The random graph model has already been generalized to the undirected, weighted case [15], by considering an ensemble of networks with non-negative, integer-valued edge weights ($w_{ij} \in \mathbf{N}$, $\forall\, i,j$) and imposing, as the only constraint, the total weight, $W = \sum_{i<j} w_{ij}$. The result is the *Undirected Weighted Random Graph* model [15], where each pair of vertices is still independent as in its binary counterpart, and connected by an edge of weight w with probability $q(w) = p^w(1-p)$, where $p \equiv e^{-\theta}$.

Here we introduce the directed version of the weighted random graph. The hamiltonian of the WRG is

$$H(\mathbf{G}|\theta) = \theta W = \theta\sum_{i\neq j} w_{ij}; \tag{27}$$

thus, the partition function becomes

$$Z(\theta) = \sum_{\mathbf{G}\in\mathcal{G}} e^{-H(\mathbf{G}|\theta)} = \sum_{\mathbf{G}\in\mathcal{G}} e^{-\theta\sum_{i\neq j} w_{ij}} = \prod_{i\neq j}\sum_{w_{ij}=0}^{+\infty} e^{-\theta w_{ij}} = \prod_{i\neq j}(1-e^{-\theta})^{-1} \tag{28}$$

(provided that $e^{-\theta} < 1$), that is a product over the $N(N-1)$ independent random variables, identified with the orderd pairs of the network's N nodes. So, every (non-negative, integer-valued) weighted network in the grandcanonical ensemble has the following probability

$$P(\mathbf{G}) = \frac{\prod_{i\neq j} e^{-\theta w_{ij}}}{\prod_{i\neq j}(1-e^{-\theta})^{-1}} \equiv \prod_{i\neq j} p^{w_{ij}}(1-p) \equiv \prod_{i\neq j} q_{ij}(w_{ij}) \tag{29}$$

by defining $p \equiv e^{-\theta}$. Note that this parameter has a precise probabilistic meaning, making even more evident the above prescription, $p < 1$. In fact, $\langle a_{ij}\rangle = \sum_{w_{ij}=0}^{+\infty} a_{ij} q_{ij}(w_{ij}) = p = 1 - q_{ij}(0)$. According to the maximum-likelihood principle [13, 14], p has to be calculated in terms of the observed quantities, by maximizing the function

$$\ln\mathcal{L}(\theta) = \ln P(\mathbf{G}^*|\theta) = \sum_{i\neq j}\left[w^*_{ij}\ln(e^{-\theta}) + \ln(1-e^{-\theta})\right] \tag{30}$$

with respect to θ. The solution to this optimization problem can be found by isolating θ in the above equation

$$W(\mathbf{G}^*) = \sum_{i\neq j}\frac{e^{-\theta^*}}{1-e^{-\theta^*}} \equiv N(N-1)\frac{p^*}{1-p^*} = \langle W\rangle_{p^*} \tag{31}$$

and, then, by inverting eq. 31 (note that the condition expressed by eq. 22 is satisfied because $\langle w_{ij}\rangle = \sum_{w_{ij}=0}^{+\infty} w_{ij} q_{ij}(w_{ij}) = \frac{p}{1-p}$):

$$p^* = \frac{W(\mathbf{G}^*)}{W(\mathbf{G}^*) + N(N-1)}. \tag{32}$$

To calculate ρ we need the expected value of r. Looking at its definition, we need the expected value of the minimum between w_{ij} and w_{ji}:

$$\langle r\rangle \equiv \frac{\langle W^{\leftrightarrow}\rangle}{\langle W\rangle} = \frac{\sum_{i\neq j}\langle \min[w_{ij},\, w_{ji}]\rangle}{\langle W\rangle}. \tag{33}$$

By considering that w_{ij} and w_{ji} are independent random variables, the cumulative distribution for the minimum is relatively easy to calculate:

$$P(\min[w_{ij},\, w_{ji}] \geq w) = P(w_{ij} \geq w)P(w_{ji} \geq w) = p^w p^w; \tag{34}$$

from this, it follows that its expected value is

$$\langle \min[w_{ij},\, w_{ji}]\rangle_{WRG} = \sum_{w=1}^{+\infty} P(\min[w_{ij},\, w_{ji}] \geq w) = \frac{p^2}{1-p^2}. \tag{35}$$

Now, the expected value (that is, the ensemble average) of r, computed in correspondence of the maximum-likelihood parameters, can be found by using the result of eq. 32:

$$\langle r\rangle^*_{WRG} = \frac{\sum_{i\neq j} \frac{(p^*)^2}{1-(p^*)^2}}{\sum_{i\neq j} \frac{p^*}{1-p^*}} = \frac{p^*}{1+p^*}. \tag{36}$$

The Directed Weighted Configuration Model (WCM)

This second null model is the weighted version of the Directed Configuration Model, fully specified by the in-degree and out-degree sequences [12, 13]. The weighted counterparts of these constraints are the in-strength and out-strength sequences [16]:

$$H(\mathbf{G}|\vec{\theta}) = \sum_i (\alpha_i s_i^{out} + \beta_i s_i^{in}) = \sum_{i\neq j} (\alpha_i + \beta_j) w_{ij}; \tag{37}$$

the partition function of the WCM is

$$Z(\vec{\theta}) = \sum_{\mathbf{G}\in\mathcal{G}} e^{-H(\mathbf{G}|\vec{\theta})} = \sum_{\mathbf{G}\in\mathcal{G}} e^{\sum_{i\neq j} -(\alpha_i+\beta_j)w_{ij}} = \prod_{i\neq j} \sum_{w_{ij}=0}^{+\infty} e^{-(\alpha_i+\beta_j)w_{ij}} = \prod_{i\neq j} \left[1 - e^{-(\alpha_i+\beta_j)}\right]^{-1}$$

(where $\vec{\theta} \equiv \{\vec{\alpha},\, \vec{\beta}\}$ and provided that $e^{-(\alpha_i+\beta_j)} < 1$). Again, it is a product over $N(N-1)$ independent random variables. The reason becomes clearer when considering the WCM: when the contraints are *local*, that is expressable as linear combinations of the adjacency matrix elements, the partition function factorizes and the probability of a given configuration factorizes as well, as a product of the independent random variables probability coefficients [13]. In this case every (non-negative, integer-valued) weighted network in the grandcanonical ensemble has a probability of the following form

$$P(\mathbf{G}) = \frac{\prod_{i\neq j} e^{-(\alpha_i+\beta_j)w_{ij}}}{\prod_{i\neq j}\left[1 - e^{-(\alpha_i+\beta_j)}\right]^{-1}} \equiv \prod_{i\neq j} p_{ij}^{w_{ij}}(1-p_{ij}) \equiv \prod_{i\neq j} (x_i y_j)^{w_{ij}}(1 - x_i y_j) \equiv \prod_{i\neq j} q_{ij}(w_{ij}) \tag{38}$$

by defining $p_{ij} \equiv e^{-(\alpha_i+\beta_j)} = e^{-\alpha_i}e^{-\beta_j} \equiv x_i y_j$. Now, two parameters per vertex have to be calculated in terms of the observed quantities: the maximum-likelihood principle [13] prescribes to maximize

$$\ln \mathcal{L}(\vec{\theta}) = \ln P(\mathbf{G}^*|\vec{\theta}) = \sum_{i\neq j} \left[w_{ij}^* \ln(x_i y_j) + \ln(1 - x_i y_j) \right] \tag{39}$$

with respect to $\vec{x}$ and $\vec{y}$. The solution to the optimization problem can be found by solving the system

$$\begin{cases} s_i^{out}(\mathbf{G}^*) &= \sum_{j\neq i} \frac{x_i^* y_j^*}{1-x_i^* y_j^*} = \langle s_i^{out} \rangle_{\vec{\theta}^*}, \quad \forall i \\ s_i^{in}(\mathbf{G}^*) &= \sum_{j\neq i} \frac{x_j^* y_i^*}{1-x_j^* y_i^*} = \langle s_i^{in} \rangle_{\vec{\theta}^*}, \quad \forall i \end{cases} \tag{40}$$

(again, the condition expressed by eq. 22 is satisfied because $\langle w_{ij} \rangle = \sum_{w_{ij}=0}^{+\infty} w_{ij} q_{ij}(w_{ij}) = \frac{p_{ij}}{1-p_{ij}} \equiv \frac{x_i y_j}{1-x_i y_j}$). The expected value of the minimum between w_{ij} and w_{ji} can be easily found by generalizing eq. 35

$$\langle \min[w_{ij},\, w_{ji}] \rangle_{WCM} = \frac{p_{ij} p_{ji}}{1 - p_{ij} p_{ji}} \tag{41}$$

and the expected value of r, computed in correspondence of the maximum-likelihood parameters, can be found by using the results of eq. 40:

$$\langle r \rangle_{WCM}^* = \frac{\sum_{i\neq j} \frac{p_{ij}^* p_{ji}^*}{1-p_{ij}^* p_{ji}^*}}{\sum_{i\neq j} \frac{p_{ij}^*}{1-p_{ij}^*}}. \tag{42}$$

The Balanced Configuration Model (BCM)

In addition to the WCM, we further developed a version of it that is intended to model networks where the observed differences between s_i^{out} and s_i^{in} are interpreted as statistical fluctuations around a balanced condition, i.e. $\langle s_i^{out} \rangle = \langle s_i^{in} \rangle$. We can start from the WCM equations, to specify them in this particular case. The condition $s_i^{out} \simeq s_i^{in}$ implies that $x_i \simeq y_i \equiv z_i$ and this reduce the number of equations to solve, from $2N$ to N:

$$s_i^{out} + s_i^{in} = \sum_{j\neq i} \frac{2 z_i z_j}{1 - z_i z_j} \Longrightarrow s_i^{tot}(\mathbf{G}^*) = \sum_{j\neq i} \frac{2 z_i^* z_j^*}{1 - z_i^* z_j^*}, \quad \forall i. \tag{43}$$

This, in turn, implies that $p_{ij} = p_{ji} = z_i z_j$ and that $\langle w_{ij} \rangle = \langle w_{ji} \rangle$. So, under the BCM, the expected value of the minimum and of r become, respectively,

$$\langle \min[w_{ij},\, w_{ji}] \rangle_{BCM} = \frac{p_{ij}^2}{1 - p_{ij}^2} \tag{44}$$

and

$$\langle r \rangle_{BCM}^* = \frac{\sum_{i<j} \frac{(p_{ij}^*)^2}{1-(p_{ij}^*)^2}}{\sum_{i<j} \frac{p_{ij}^*}{1-p_{ij}^*}} \tag{45}$$

which is nothing more that a simplified version of eq. 42.

The fundamental insight given by the BCM is that, in networks with node balance, i.e. where $\langle s_i^{out} \rangle = \langle s_i^{in} \rangle$, the expected weights are symmetric:

$$\langle w_{ij} \rangle_{BCM}^* = \frac{z_i z_j}{1 - z_i z_j} = \langle w_{ji} \rangle_{BCM}^* \tag{46}$$

This means that, in networks where the observed differences between s_i^{out} and s_i^{in} are consistent with statistical fluctuations around a balanced condition, one automatically expects symmetric weights, even without introducing any tendency towards reciprocation. This shows that measures of reciprocity based on the symmetry of mutual weights necessarily receive spurious contributions from other sources of flow balance. This observation concludes our statement that previously attempted correlation- and symmetry-based measures [6–9] cannot properly separate reciprocity from other factors.

FROM NULL MODELS TO TRUE MODELS

The three previous null models are defined in terms of constraints as the total weight, the in- and out-strength sequences and the total-strength sequence. So, not being included in the list of the constraints, the reciprocity r and the index ρ were a sort of target quantities, to test the power of the considered null models in reproducing them.

Now, we can make a step forward and include some information about the reciprocity structure of the network.

The Weighted Reciprocity Model (WRM)

We start by generalizing the WCM, by adding to its hamiltonian a sort of "global reciprocity" defined over the whole network, thus fixing the total number of reciprocal links. This means to consider, as a further constraint, the quantity

$$W^{\leftrightarrow} = \sum_{i\neq j} \min[w_{ij},\, w_{ji}] = \sum_i s_i^{\leftrightarrow} \tag{47}$$

to obtain the following Hamiltonian

$$H(\mathbf{G}|\vec{\theta}) = \sum_i (\alpha_i s_i^{out} + \beta_i s_i^{in}) + \gamma W^{\leftrightarrow} \tag{48}$$

(where $\vec{\theta} \equiv \{\vec{\alpha},\, \vec{\beta},\, \gamma\}$). The resolution of this null model is considerably simplified by considering an equivalent way of rewriting it,

$$H(\mathbf{G}|\vec{\theta}) = \sum_i [\alpha_i s_i^{\rightarrow} + \beta_i s_i^{\leftarrow} + (\alpha_i + \beta_i + \gamma) s_i^{\leftrightarrow}] \tag{49}$$

having posed, to uniform the formalism, $\min[w_{ij},\, w_{ji}] \equiv w_{ij}^{\leftrightarrow}$ and having defined

$$\begin{aligned} s_i^{out} &\equiv s_i^{\rightarrow} + s_i^{\leftrightarrow} \Longrightarrow w_{ij} \equiv w_{ij}^{\rightarrow} + w_{ij}^{\leftrightarrow}, \\ s_i^{in} &\equiv s_i^{\leftarrow} + s_i^{\leftrightarrow} \Longrightarrow w_{ij} \equiv w_{ji}^{\leftarrow} + w_{ij}^{\leftrightarrow}. \end{aligned} \tag{50}$$

Now, the most challenging calculation is about the partition function. This can be done by rewriting the hamiltonian solely in terms of the variables $w_{ij}^{\rightarrow}$, $w_{ij}^{\leftarrow}$ and $w_{ij}^{\leftrightarrow}$,

$$H(\mathbf{G}|\vec{\theta}) = \sum_{i<j} [(\alpha_i + \beta_j) w_{ij}^{\rightarrow} + (\alpha_j + \beta_i) w_{ij}^{\leftarrow} + (\alpha_i + \beta_j + \alpha_j + \beta_i + 2\gamma) w_{ij}^{\leftrightarrow}] \tag{51}$$

and considering the admissible states for them:

$$(w_{ij}^{\rightarrow},\, w_{ij}^{\leftarrow},\, w_{ij}^{\leftrightarrow}) = \{(0,\, 0,\, \mathbf{N}),\, (\mathbf{N}^+,\, 0,\, \mathbf{N}),\, (0,\, \mathbf{N}^+,\, \mathbf{N})\} \tag{52}$$

where $\mathbf{N} \equiv [0 \ldots \infty)$ and $\mathbf{N}^+ \equiv [1 \ldots \infty)$. So the partition function becomes

$$Z(\vec{\theta}) = \sum_{\mathbf{G}\in\mathcal{G}} e^{-H(\mathbf{G}|\vec{\theta})} = \sum_{(w_{ij}^{\rightarrow},\, w_{ij}^{\leftarrow},\, w_{ij}^{\leftrightarrow})} e^{-H(\mathbf{G}|\vec{\theta})} = \prod_{i<j} \frac{(1-x_i x_j y_i y_j)}{(1-x_i y_j)(1-x_j y_i)(1-x_i x_j y_i y_j z^2)} \equiv \prod_{i<j} Z_{ij}^{WRM}(\vec{\theta}) \quad (53)$$

(having posed $x_i \equiv e^{-\alpha_i}$, $y_i \equiv e^{-\beta_i}$ and $z \equiv e^{-\gamma}$) and, consequently, the probability coefficient for the generic configuration $\mathbf{G}$ is

$$P(\mathbf{G}) = \prod_{i<j} \frac{(x_i y_j)^{w_{ij}} (x_j y_i)^{w_{ji}} z^{2w_{ij}^{\leftrightarrow}}}{Z_{ij}^{WRM}(\vec{\theta})}. \quad (54)$$

Now, the maximum-likelihood principle prescribes to maximize

$$\ln P(\mathbf{G}^*|\vec{\theta}) = \sum_{i<j} [w_{ij}^* \ln(x_i y_j) + w_{ji}^* \ln(x_j y_i) + (2w_{ij}^{\leftrightarrow})^* \ln z - \ln Z_{ij}^{WRM}(\vec{\theta})] \quad (55)$$

with respect to $\vec{x}$, $\vec{y}$ and z. The solution to the previous optimization problem can be found by solving the system

$$\begin{cases} s_i^{out}(\mathbf{G}^*) & = \sum_{j\neq i} \langle w_{ij} \rangle_{\vec{\theta}^*} = \langle s_i^{out} \rangle_{\vec{\theta}^*}, \quad \forall i \\ s_i^{in}(\mathbf{G}^*) & = \sum_{j\neq i} \langle w_{ji} \rangle_{\vec{\theta}^*} = \langle s_i^{in} \rangle_{\vec{\theta}^*}, \quad \forall i \\ W^{\leftrightarrow}(\mathbf{G}^*) & = \sum_{i<j} 2\langle w_{ij}^{\leftrightarrow} \rangle_{\vec{\theta}^*} = \langle W^{\leftrightarrow} \rangle_{\vec{\theta}^*} \end{cases} \quad (56)$$

where

$$\langle w_{ij} \rangle_{\vec{\theta}^*} = \frac{x_i^* y_j^* (1 - x_j^* y_i^*)}{(1-x_i^* y_j^*)(1-x_i^* x_j^* y_i^* y_j^*)} + \frac{x_i^* x_j^* y_i^* y_j^* (z^*)^2}{1 - x_i^* x_j^* y_i^* y_j^* (z^*)^2}, \quad (57)$$

$$\langle w_{ji} \rangle_{\vec{\theta}^*} = \frac{x_j^* y_i^* (1 - x_i^* y_j^*)}{(1-x_j^* y_i^*)(1-x_i^* x_j^* y_i^* y_j^*)} + \frac{x_i^* x_j^* y_i^* y_j^* (z^*)^2}{1 - x_i^* x_j^* y_i^* y_j^* (z^*)^2}, \quad (58)$$

$$\langle w_{ij}^{\leftrightarrow} \rangle_{\vec{\theta}^*} = \frac{x_i^* x_j^* y_i^* y_j^* (z^*)^2}{1 - x_i^* x_j^* y_i^* y_j^* (z^*)^2}. \quad (59)$$

Now, the expected value of the minimum between w_{ij} and w_{ji} is $\langle \min[w_{ij},\, w_{ji}] \rangle_{WRM}^* = \langle w_{ij}^{\leftrightarrow} \rangle_{\vec{\theta}^*}$. Even if it is possible to write down the analytical expression of the expected value of r, by using it, this can be avoided, by considering that

$$\langle r \rangle_{WRM}^* = \frac{\langle W^{\leftrightarrow} \rangle_{\vec{\theta}^*}}{\langle W \rangle_{\vec{\theta}^*}} = \frac{W^{\leftrightarrow}(\mathbf{G}^*)}{W(\mathbf{G}^*)} = r; \quad (60)$$

this, in turn, implies that

$$\rho_{WRM}^* = \frac{r - \langle r \rangle_{WRM}^*}{1 - \langle r \rangle_{WRM}^*} \equiv \frac{r - r}{1 - r} = 0. \quad (61)$$

So, by definition, the index ρ is trivially reproduced by the WRM.

Note also that the only difference between the predicted quantities $\langle w_{ij}^{\leftrightarrow} \rangle_{WCM}$ and $\langle w_{ij}^{\leftrightarrow} \rangle_{WRM}$ lies in the presence of the extra-parameter z in the second expression. Recalling that $z < 1$, if the hidden variables $\vec{x}$ and $\vec{y}$ are kept fixed, changing z means lowering the expected reciprocal weight with respect to the WCM prediction. This makes the WRM best suited to reproduce networks that are anti-reciprocal (i.e., less reciprocal than the WCM prediction).

The Non-Reciprocated Strength Model (NSM)

A second null model including the information about the global reciprocity structure of the network can be defined, starting by the WRM hamiltonian. This time, the imposed constraints are the in- and out-strength sequences, diminished by the reciprocal strength sequence (see eq. 50), and the total number of reciprocal links:

$$H(\mathbf{G}|\vec{\theta}) = \sum_i (\alpha_i s_i^{\not\to} + \beta_i s_i^{\not\leftarrow}) + \gamma W^{\leftrightarrow} = \sum_{i<j} [(\alpha_i + \beta_j) w_{ij}^{\not\to} + (\alpha_j + \beta_i) w_{ij}^{\not\leftarrow} + 2\gamma w_{ij}^{\leftrightarrow}].$$

Following the calculations of the WRM, the partition function is

$$Z(\vec{\theta}) = \sum_{\mathbf{G}\in\mathcal{G}} e^{-H(\mathbf{G}|\vec{\theta})} = \sum_{(w_{ij}^{\not\to},\, w_{ij}^{\not\leftarrow},\, w_{ij}^{\leftrightarrow})} e^{-H(\mathbf{G}|\vec{\theta})} = \prod_{i<j} \frac{(1 - x_i x_j y_i y_j)}{(1 - x_i y_j)(1 - x_j y_i)(1 - z^2)} \equiv \prod_{i<j} Z_{ij}^{NSM}(\vec{\theta}) \tag{62}$$

(having posed $x_i \equiv e^{-\alpha_i}$, $y_i \equiv e^{-\beta_i}$ and $z \equiv e^{-\gamma}$). The probability coefficient for a generic configuration, $\mathbf{G}$, is

$$P(\mathbf{G}) = \prod_{i<j} \frac{(x_i y_j)^{w_{ij}^{\not\to}} (x_j y_i)^{w_{ij}^{\not\leftarrow}} z^{2 w_{ij}^{\leftrightarrow}}}{Z_{ij}^{NSM}(\vec{\theta})} \tag{63}$$

and the maximum-likelihood principle prescribes to maximize

$$\ln P(\mathbf{G}^*|\vec{\theta}) = \sum_{i<j} [(w_{ij}^{\not\to})^* \ln(x_i y_j) + (w_{ij}^{\not\leftarrow})^* \ln(x_j y_i) + (2 w_{ij}^{\leftrightarrow})^* \ln z - \ln Z_{ij}^{NSM}(\vec{\theta})] \tag{64}$$

with respect to $\vec{x}$, $\vec{y}$ and z. The solution to the previous optimization problem can be found by solving the system

$$\begin{cases} s_i^{\not\to}(\mathbf{G}^*) &= \sum_{j\neq i} \langle w_{ij}^{\not\to} \rangle_{\vec{\theta}^*} = \langle s_i^{\not\to} \rangle_{\vec{\theta}^*}, \quad \forall i \\ s_i^{\not\leftarrow}(\mathbf{G}^*) &= \sum_{j\neq i} \langle w_{ij}^{\not\leftarrow} \rangle_{\vec{\theta}^*} = \langle s_i^{\not\leftarrow} \rangle_{\vec{\theta}^*}, \quad \forall i \\ W^{\leftrightarrow}(\mathbf{G}^*) &= \sum_{i<j} 2 \langle w_{ij}^{\leftrightarrow} \rangle_{\vec{\theta}^*} = \langle W^{\leftrightarrow} \rangle_{\vec{\theta}^*} \end{cases} \tag{65}$$

where

$$\langle w_{ij}^{\not\to} \rangle_{\vec{\theta}^*} = \frac{x_i^* y_j^* (1 - x_j^* y_i^*)}{(1 - x_i^* y_j^*)(1 - x_i^* x_j^* y_i^* y_j^*)} + \frac{(z^*)^2}{1 - (z^*)^2}, \tag{66}$$

$$\langle w_{ij}^{\not\leftarrow} \rangle_{\vec{\theta}^*} = \frac{x_j^* y_i^* (1 - x_i^* y_j^*)}{(1 - x_j^* y_i^*)(1 - x_i^* x_j^* y_i^* y_j^*)} + \frac{(z^*)^2}{1 - (z^*)^2}, \tag{67}$$

$$\langle w_{ij}^{\leftrightarrow} \rangle_{\vec{\theta}^*} = \frac{(z^*)^2}{1 - (z^*)^2}. \tag{68}$$

As for the WRM

$$\langle r \rangle_{NSM}^* = \frac{\langle W^{\leftrightarrow} \rangle_{\vec{\theta}^*}}{\langle W \rangle_{\vec{\theta}^*}} = \frac{W^{\leftrightarrow}(\mathbf{G}^*)}{W(\mathbf{G}^*)} = r; \tag{69}$$

this, in turn, implies that

$$\rho_{NSM}^* = \frac{r - \langle r \rangle_{NSM}^*}{1 - \langle r \rangle_{NSM}^*} \equiv \frac{r - r}{1 - r} = 0. \tag{70}$$

The Reciprocated Strength Model (RSM)

Until now, we have defined three null models with no constraints about the reciprocity (the WRG, the WCM and the BCM) and two null models with the total number of reciprocal links (thus implementing a global notion of reciprocity), as a constraint.

Now, we can define more refined null models, by considering, as constraints, the local notion of reciprocity, as defined by eq. 15. We start by considering the following hamiltonian:

$$H(\mathbf{G}|\vec{\theta}) = \alpha W + \sum_i \delta_i s_i^{\leftrightarrow} \tag{71}$$

(where $\vec{\theta} \equiv \{\alpha,\ \vec{\delta}\}$). The resolution of the null model described by this hamiltonian is, again, considerably simplified by considering the equivalent expression

$$H(\mathbf{G}|\vec{\theta}) = \sum_{i<j} [\alpha w_{ij}^{\rightarrow} + \alpha w_{ij}^{\leftarrow} + (2\alpha + \delta_i + \delta_j) w_{ij}^{\leftrightarrow}]; \tag{72}$$

by summing over the states defined in eq. 52 we find the partition function

$$Z(\vec{\theta}) = \sum_{\mathbf{G}\in\mathcal{G}} e^{-H(\mathbf{G}|\vec{\theta})} = \sum_{(w_{ij}^{\rightarrow},\, w_{ij}^{\leftarrow},\, w_{ij}^{\leftrightarrow})} e^{-H(\mathbf{G}|\vec{\theta})} = \prod_{i<j} \frac{(1+x)}{(1-x)(1-x^2 z_i z_j)} \tag{73}$$

(having posed $x \equiv e^{-\alpha}$, $z_i \equiv e^{-\delta_i}$) the probability coefficient for the generic configuration $\mathbf{G}$ is

$$\begin{aligned} P(\mathbf{G}) &= \prod_{i<j} \frac{x^{w_{ij}^{\rightarrow} + w_{ij}^{\leftarrow} + 2w_{ij}^{\leftrightarrow}} (z_i z_j)^{w_{ij}^{\leftrightarrow}} (1 - x^2 z_i z_j)(1-x)}{(1+x)} = \\ &= \prod_{i<j} \frac{x^{w_{ij} + w_{ji}} (z_i z_j)^{w_{ij}^{\leftrightarrow}} (1 - x^2 z_i z_j)(1-x)}{(1+x)} \end{aligned} \tag{74}$$

and the likelihood function is, of course, the logarithm of the previous probability coefficient. The solution to this optimization problem prescribes to solve the following system

$$\begin{cases} s_i^{\leftrightarrow}(\mathbf{G}^*) &= \sum_{j\neq i} \langle w_{ij}^{\leftrightarrow} \rangle_{\vec{\theta}^*} = \langle s_i^{\leftrightarrow} \rangle_{\vec{\theta}^*}, \qquad \forall i \\ W(\mathbf{G}^*) &= \sum_{i\neq j} \langle w_{ij} \rangle_{\vec{\theta}^*} = \langle W \rangle_{\vec{\theta}^*} \end{cases} \tag{75}$$

where

$$\langle w_{ij} \rangle_{\vec{\theta}^*} = \frac{x^*}{(1-(x^*)^2)} + \frac{(x^*)^2 z_i^* z_j^*}{1-(x^*)^2 z_i^* z_j^*}, \tag{76}$$

$$\langle w_{ij}^{\leftrightarrow} \rangle_{\vec{\theta}^*} = \frac{(x^*)^2 z_i^* z_j^*}{1-(x^*)^2 z_i^* z_j^*}. \tag{77}$$

This model allows to solve for the x value analitically. In fact, by summing eq. 76 over the ordered pairs of nodes, we find

$$W(\mathbf{G}^*) = \frac{N(N-1)x^*}{(1-(x^*)^2)} + W^{\leftrightarrow}(\mathbf{G}^*) \tag{78}$$

and by solving this second-order equation w.r.t. x, and taking the positive solution, we have the maximum-likelihood estimation of this parameter. Also this model exactly reproduces the observed reciprocity, because $W^{\leftrightarrow}(\mathbf{G}^*) = \langle W^{\leftrightarrow} \rangle_{\vec{\theta}^*}$ and $W(\mathbf{G}^*) = \langle W \rangle_{\vec{\theta}^*}$. This means that $\rho_{RSM} = \frac{r - \langle r \rangle_{RSM}}{1 - \langle r \rangle_{RSM}} \equiv \frac{r-r}{1-r} = 0$ and the local quantities as the reciprocal strength sequence are now trivially reproduced.

The Weighted Reciprocated Configuration Model (WRCM)

The last step is the definion of a very general null model, to finally include those local quantities not fixed by the NSM and the RSM. This implies a slight generalization of the formulas in the previous two paragraphs. The graph hamiltonian becomes

$$H(\mathbf{G}|\vec{\theta}) = \sum_i (\alpha_i s_i^{\rightarrow} + \beta_i s_i^{\leftarrow} + \gamma_i s_i^{\leftrightarrow}) \tag{79}$$

where, now, $\vec{\theta} \equiv \{\vec{\alpha},\ \vec{\beta},\ \vec{\gamma}\}$ and

$$s_i^{\rightarrow} \equiv \sum_{j(\neq i)} w_{ij}^{\rightarrow}\,,\ s_i^{\leftarrow} \equiv \sum_{j(\neq i)} w_{ij}^{\leftarrow},\ s_i^{\leftrightarrow} \equiv \sum_{j(\neq i)} w_{ij}^{\leftrightarrow} \tag{80}$$

with obvious meaning of the symbols (defined above). The partition function now becomes

$$Z(\vec{\theta}) = \prod_{i<j} \frac{(1 - x_i x_j y_i y_j)}{(1 - x_i y_j)(1 - x_j y_i)(1 - z_i z_j)} \equiv \prod_{i<j} Z_{ij}^{WRCM}(\vec{\theta}) \tag{81}$$

and the likelihood is

$$\ln P(\mathbf{G}^*|\vec{\theta}) = \sum_{i<j} [(w_{ij}^{\rightarrow})^* \ln(x_i y_j) + (w_{ij}^{\leftarrow})^* \ln(x_j y_i) + (w_{ij}^{\leftrightarrow})^* \ln(z_i z_j) - \ln Z_{ij}^{WRCM}(\vec{\theta})]. \tag{82}$$

The solution to this optimization problem, with respect to $\vec{x}$, $\vec{y}$ and $\vec{z}$, can be found by solving the following system:

$$\begin{cases} s_i^{\rightarrow}(\mathbf{G}^*) &= \sum_{j\neq i} \langle w_{ij}^{\rightarrow} \rangle_{\vec{\theta}^*} = \langle s_i^{\rightarrow} \rangle_{\vec{\theta}^*}, \quad \forall i \\ s_i^{\leftarrow}(\mathbf{G}^*) &= \sum_{j\neq i} \langle w_{ij}^{\leftarrow} \rangle_{\vec{\theta}^*} = \langle s_i^{\leftarrow} \rangle_{\vec{\theta}^*}, \quad \forall i \\ s_i^{\leftrightarrow}(\mathbf{G}^*) &= \sum_{j\neq i} \langle w_{ij}^{\leftrightarrow} \rangle_{\vec{\theta}^*} = \langle s_i^{\leftrightarrow} \rangle_{\vec{\theta}^*}, \quad \forall i \end{cases} \tag{83}$$

where

$$\langle w_{ij}^{\rightarrow} \rangle_{\vec{\theta}^*} = \frac{x_i^* y_j^* (1 - x_j^* y_i^*)}{(1 - x_i^* y_j^*)(1 - x_i^* x_j^* y_i^* y_j^*)}, \tag{84}$$

$$\langle w_{ij}^{\leftarrow} \rangle_{\vec{\theta}^*} = \frac{x_j^* y_i^* (1 - x_i^* y_j^*)}{(1 - x_j^* y_i^*)(1 - x_i^* x_j^* y_i^* y_j^*)}, \tag{85}$$

$$\langle w_{ij}^{\leftrightarrow} \rangle_{\vec{\theta}^*} = \frac{z_i^* z_j^*}{1 - z_i^* z_j^*}. \tag{86}$$

By the definition of the WRCM model, we not only recover the result that the global reciprocity is equal to the observed one (implying $r \equiv \langle r \rangle_{WRCM}$ and $\rho_{WRCM} \equiv 0$, also valid for the WRM): now, all the vertex-level, strength sequences are exactly reproduced, impyling that the reciprocity is reproduced at a *local* level.

The WRCM is now powerful enough to allow for the analysis of the weighted motifs (to understand which all the dyadic information has to be fixed) and for the community detection, especially for those networks where the reciprocity plays an important role in shaping its structure.

MODELS: A SUMMARY

The first six models explained in the previous sections can be recovered from the last and most general one (the WRCM) by means of simple substitutions in the graph hamiltonian (as shown by the following table).

Non-reciprocal models	
$\alpha_i = \beta_i = \gamma;\ \gamma_i = 2\gamma,\ \forall i$	WRCM $\rightarrow$ WRG
$\alpha_i = \beta_i;\ \gamma_i = \alpha_i + \beta_i,\ \forall i$	WRCM $\rightarrow$ BCM
$\alpha_i \neq \beta_i;\ \gamma_i = \alpha_i + \beta_i,\ \forall i$	WRCM $\rightarrow$ WCM
Global-reciprocity models	
$\alpha_i \neq \beta_i;\ \gamma_i = \alpha_i + \beta_i + \gamma,\ \forall i$	WRCM $\rightarrow$ WRM
$\alpha_i \neq \beta_i;\ \gamma_i = \gamma,\ \forall i$	WRCM $\rightarrow$ NSM
Local-reciprocity models	
$\alpha_i = \beta_i = \alpha;\ \gamma_i = \alpha_i + \beta_i + \delta_i,\ \forall i$	WRCM $\rightarrow$ RSM

THE JACKKNIFE METHOD

The jackknife method [17, 18] is an expedient to mimic resampling and it is usually used to estimate the variance of a given function of the population mean, $f(\langle x \rangle)$ (being x the random variable of interest). Doing this in the least biased way, would imply to have a whole collection of samples. However, we observe only a single realization. How can compensate for the lack of such observations? We can build a set of artificial samples by considering the following sets:

$$\begin{aligned} s_1^J &= \{x_2,\ x_3 \ldots x_M\}, \\ s_2^J &= \{x_1,\ x_3 \ldots x_M\}, \\ &\vdots \\ s_M^J &= \{x_1,\ x_2 \ldots x_{M-1}\}; \end{aligned} \tag{87}$$

that is a list of vectors for each of which a single observation has been removed. Then, we calculate the so called *jackknife averages*

$$\bar{s_1}^J = \frac{\sum_{i \neq 1} x_i}{M-1},\ \bar{s_2}^J = \frac{\sum_{i \neq 2} x_i}{M-1} \ldots \bar{s_M}^J = \frac{\sum_{i \neq M} x_i}{M-1}, \tag{88}$$

the estimates of the first two moments

$$\mu_1^J \equiv \frac{\sum_i f(\bar{s_i}^J)}{M};\ \mu_2^J \equiv \frac{\sum_i f(\bar{s_i}^J)^2}{M}, \tag{89}$$

from which the estimate of the jackkinfe-standard deviation follows

$$\sigma_f^J \simeq \sqrt{\mu_2^J - (\mu_1^J)^2}, \tag{90}$$

and, finally [18],

$$\sigma_{f(\langle x \rangle)} \simeq \sqrt{M-1}\, \sigma_f^J. \tag{91}$$

How can we implement all this for our weighted networks? The quantity we are interested in is ρ. It is a function of the expected value of r, taken over the whole grandcanonical ensemble: $\langle r \rangle$. By applying the jackknife method, we can build L artificial samples by removing one *weight* at a time. By rewriting the above formulas, the final estimates become

$$\rho_{NM} = \frac{r - \langle r \rangle_{NM}}{1 - \langle r \rangle_{NM}}, \tag{92}$$

$$\sigma^2_{\rho_{NM}} = \sum_i^L (\rho_{i,\,NM} - \rho_{NM})^2 = \frac{\sigma_r^2}{(1 - \langle r_{NM} \rangle)^2}, \tag{93}$$

where NM can be WRG, WCM, BCM, WRM, NSM, RSM, $WRCM$ and where the sum over the index i means that we are summing over the realizations with the i-th weight removed.

DESCRIPTION OF THE DATASET

In what follows a brief description of the analysed networks is given.

Interbank network. This is the network of the Italian interbank monetary exchanges [19], in the year 1999. We analysed the monthy transactions for May ($N = 215$, $L = 5269$), June ($N = 215$, $L = 5229$), August ($N = 215$, $L = 5071$), October ($N = 215$, $L = 4712$) and December ($N = 215$, $L = 4685$). **Food webs.** We analysed eight different food webs [20–22], from different ecosystems (lagoons, marshes, lakes, bays, estuaries, grasses, rivers), with a prevalence of aquatic habitats: *Chesapeake Bay* ($N = 34$, $L = 177$) and *Mondego Bay* ($N = 46$, $L = 400$), *Everglades Marshes* ($N = 69$, $L = 916$), *Maspalomas Lagoon* ($N = 24$, $L = 82$), *Michigan Lake* ($N = 39$, $L = 221$), *St. Marks Seagrass* ($N = 54$, $L = 536$), *Crystal River Creek* ($N = 24$, $L = 125$ and $N = 24$, $L = 100$). **Neural networks.** We analysed the neural network [23] of *C. Elegans* ($N = 297$, $L = 2345$). **Social networks.** We analysed three different social networks [24–30]: *BK-Office*, *BK-Tech* and *BK-Fraternity*. BK-Tech and BK-Fraternity are completely connected (that is, $L = N(N-1)$). Bernard and Killworth (and, later, also with the help of Sailer), collected five sets of data on human interactions in bounded groups. *BK-Office* ($N = 40$, $L = 1558$) is the network of the human interactions (conversations) frequency between the employees of a small business-office, as recorded at time intervals of fifteen minutes (during two four-days periods), by an external observer, along a fixed route through the office itself. *BK-Tech* ($N = 34$, $L = 1122$) is the network of the human interactions (conversations) frequency between collaborators in a technical research group at a West Virginia University, as recorded at time intervals of half-hour (during one five-days working week), by an external observer. *BK-Fraternity* ($N = 58$, $L = 3306$) is the network of the human interactions (conversations) frequency between the students living in a fraternity at a West Virginia College, as recorded by an external observer at time intervals of fifteen minutes (during a five-days week, twenty-one hours per day) who walked through the public areas of the building. **The World Trade Network.** We analyse the series of yearly bilateral data on exports and imports among world countries from the database in ref.[31], from 1948 to 2000 ($N \in [82,\ 186]$ and $L \in [2539,\ 19903]$).

ACKNOWLEDGMENTS

D. G. acknowledges support from the Dutch Econophysics Foundation (Stichting Econophysics, Leiden, the Netherlands) with funds from beneficiaries of Duyfken Trading Knowledge BV, Amsterdam, the Netherlands.

F. R. acknowledges support from the FESSUD project on “Financialisation, economy, society and sustainable development”, Seventh Framework Programme, EU.

ADDITIONAL INFORMATION

The authors declare no competing financial interests.

[1] Holland P. & Leinhardt S. in Sociological Methodology. ed Heise D (Jossey-Bass, San Francisco), 1-45 (1975).
[2] Wasserman S. & Faust K. Social Network Analysis. Cambridge University Press (Cambridge, New York) (1994).
[3] Garlaschelli D. & Loffredo M. I. Patterns of Link Reciprocity in Directed Networks. *Phys. Rev. Lett.* **93**, 268701 (2004).
[4] Garlaschelli D. & Loffredo M. I. Multispecies grand-canonical models for networks with reciprocity. *Phys. Rev. E* **73**, 015101 (2006).

[5] Garlaschelli D., Ruzzenenti F. & Basosi R. Complex Networks and Symmetry I: a Review. *Symmetry* **2**(3) 1683-1709 (2010).
[6] Kovanen L., Saramaki J. & Kaski K. Reciprocity of mobile phone calls. *Arxiv:1002.0763v1* (2010).
[7] Fagiolo G. Directed or Undirected? A New Index to Check for Directionality of Relations in Socio-Economic Networks. *Econ. Bull.* **3** (34), 1-12 (2006).
[8] Wang C. et al. Weighted reciprocity in human communication networks. *arXiv:1108.2822v5* (2011).
[9] Akoglu L., Vaz de Melo P. O. S. & Faloutsos C. Quantifying Reciprocity in Large Weighted Communication Networks. *Lec. Notes Comp. Science* **7302**, 85-96 (2012).
[10] Snijders T. A. B., Pattison P. E., Robins G. L. & Handcock M. S. New specifications for exponential random graph models. *Sociological Methodology* **36** (1), 99-153 (2006).
[11] Robins G., Pattison P., Kalish Y. & Lusher D. An introduction to exponential random graph (p*) models for social networks. *Social Networks* **29** (2), 173191 (2007).
[12] Park J. & Newman M. E. J. The statistical mechanics of networks. *Phys. Rev. E* **70**, 066117 (2004).
[13] Squartini T. & Garlaschelli D. Analytical maximum-likelihood method to detect patterns in real networks. *New J. Phys.* **13** 083001 (2011).
[14] Garlaschelli D. & Loffredo M. I. Maximum likelihood : Extracting unbiased information from complex networks. *Phys. Rev. E* **78**, 015101(R) (2008).
[15] Garlaschelli D. The weighted random graph model. *New. J. Phys.* **11**, 073005 (2009).
[16] Squartini T., Garlaschelli D. & Fagiolo G. Randomizing world trade. II. A weighted network analysis. *Phys. Rev. E* **84**, 046118 (2011).
[17] Efron B. Computers and the theory of statistics: thinking the unthinkable. *SIAM Rev.* **21**, 460 (1979).
[18] Newman M. E. J. Mixing patterns in networks. *Phys. Rev. E* **67**, 026126 (2003).
[19] De Masi G., Iori G. & Caldarelli G. A fitness model for the Italian Interbank Money Market. *Phys. Rev. E* **74**, 066112 (2006).
[20] Pajek datasets: `http://vlado.fmf.uni-lj.si/pub/networks/data/bio/foodweb/foodweb.htm` (accessed: March 2012).
[21] Martinez N. D. Artifacts or Attributes? Effects of Resolution on the Little Rock Lake Food Web. *Ecol. Monogr.* **61**, 367-392 (1991).
[22] Stouffer D. B., Camacho J., Jiang W. & Amaral L. A. N. Evidence for the existence of a robust pattern of prey selection in food webs. *Proc. R. Soc. B* **274**, 1931-1940 (2007).
[23] Oshio K. et al. Database of Synaptic Connectivity of *C. elegans* for Computation. *Tech Rep CCeP, Keio Future, Keio University* **3** (2003).
[24] UCINET IV Datasets: `http://vlado.fmf.uni-lj.si/pub/networks/data/ucinet/ucidata.htm` (accessed: March 2012).
[25] Killworth P. D. & Bernard H. R. Informant accuracy in social network data. *Human Organization* **35**, 269-386 (1976).
[26] Killworth P. D. & Bernard H. R. Informant accuracy in social network data II. *Human Communication Research* **4**, 3-18 (1977).
[27] Killworth P. D. & Bernard H. R. Informant accuracy in social network data III: A comparison of triadic structure in behavioral and cognitive data. *Social Networks* **2**, 19-46 (1979).
[28] Killworth P. D., Bernard H. R. & Sailer L. Informant accuracy in social network data IV: a comparison of clique-level structure in behavioral and cognitive network data. *Social Networks* **2**, 191-218 (1979).
[29] Killworth P. D., Bernard H. R. & Sailer L. Informant accuracy in social-network data V. An experimental attempt to predict actual communication from recall data. *Social Science Research* **11**, 30-66 (1982).
[30] Romney A. K. & Weller S. C. Predicting informant accuracy from patterns of recall among individuals. *Social Networks* **6**, 59-77 (1984).
[31] Gleditsch K. S. Expanded Trade and GDP Data. *J. Conflict Resol.* **46**(5) 712-724 (2002).