Patterns of link reciprocity in directed, signed networks

Anna Gallo [1,2,*] Fabio Saracco [3,4,1] Renaud Lambiotte [5] Diego Garlaschelli [1,2,6] and Tiziano Squartini [1,2]

[1] *IMT School for Advanced Studies, Piazza San Francesco 19, 55100 Lucca, Italy*

[2] *INdAM-GNAMPA Istituto Nazionale di Alta Matematica "Francesco Severi", P.le Aldo Moro 5, 00185 Rome, Italy*

[3] *"Enrico Fermi" Research Center (CREF), Via Panisperna 89A, 00184 Rome, Italy*

[4] *Institute for Applied Computing "Mauro Picone" (IAC), National Research Council, Via dei Taurini 19, 00185 Rome, Italy*

[5] *Mathematical Institute, University of Oxford, Woodstock Road, Oxford OX2 6GG, United Kingdom*

[6] *Lorentz Institute for Theoretical Physics, University of Leiden, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands*

(Received 12 July 2024; accepted 6 December 2024; published 21 February 2025)

Most of the analyses concerning signed networks have focused on balance theory, hence identifying frustration with undirected, triadic motifs having an odd number of negative edges; much less attention has been paid to their directed counterparts. To fill this gap, we focus on signed, directed connections, with the aim of exploring the notion of frustration in such a context. When dealing with signed, directed edges, frustration is a multifaceted concept, admitting different definitions at different scales: if we limit ourselves to consider cycles of length 2, frustration is related to reciprocity, i.e., the tendency of edges to admit the presence of partners pointing in the opposite direction. As the reciprocity of signed networks is still poorly understood, we adopt a principled approach for its study, defining quantities and introducing models to consistently capture empirical patterns of the kind. In order to quantify the tendency of empirical networks to form either mutualistic or antagonistic cycles of length 2, we extend the exponential random graph framework to binary, directed, signed networks with global and local constraints and then compare the empirical abundance of the aforementioned patterns with the one expected under each model. We find that the (directed extension of the) balance theory is not capable of providing a consistent explanation of the patterns characterizing the directed, signed networks considered in this work. Although part of the ambiguities can be solved by adopting a coarser definition of balance, our results call for a different theory, accounting for the directionality of edges in a coherent manner. In any case, the evidence that the empirical, signed networks can be highly reciprocated leads us to recommend to explicitly account for the role played by bidirectional dyads in determining frustration at higher levels (e.g., the triadic one).

DOI: 10.1103/PhysRevE.111.024312

I. INTRODUCTION

A. A brief history of balance

The interest in the study of networks with positive, negative, or missing edges can be traced back to the formulation of the so-called balance theory, first proposed by Heider [1] and further developed by Cartwright and Harary, who adopted *signed graphs* to model it [2]. The balance theory deals with the concept of balance: a complete, signed graph is said to be balanced if *all triads* have an even number of negative edges, i.e., either zero (in this case, the three edges are all positive) or two. The so-called structure theorem states that a complete, signed graph is balanced if and only if its set of nodes can be partitioned into $k = 2$ disjoint subsets whose intramodular edges are all positive and whose intermodular edges are all negative. Cartwright and Harary extended the definition of balance to incomplete graphs [2] by including cycles of length larger than three: a network is deemed balanced if *all cycles* have an even number of negative edges (although the points of each subset are no longer required to be connected). Taken together, the criteria above define the so-called (structural) strong balance theory. Such a framework has been further extended by Davis [3], who introduced the concept of k-balanced networks, according to which signed graphs are balanced if their set of nodes can be partitioned into $k \geqslant 2$ disjoint subsets with positive, intramodular edges and negative, intermodular edges. This generalized definition of balance has led to the formulation of the so-called (structural) weak balance theory, according to which triads whose edges are all negative are balanced as well, since each node can be thought of as a group on its own. Taken together, the strong and the weak variants of the balance theory define the so-called (structural) traditional balance theory; hence, k-balanced networks are traditionally balanced.

From a mesoscopic perspective, however, both versions of the balance theory require the presence of positive blocks along the main diagonal of the adjacency matrix ($k = 2$, according to the strong variant; $k > 2$, according to the weak variant) and of negative, off-diagonal blocks. As noticed in [4], the block structure defining the traditional balance theory

*Contact author: anna.gallo@imtlucca.it

is overly restrictive, dooming the vast majority of real-world, signed networks to be quickly dismissed as frustrated: in order to overcome what was perceived as a major limitation of the traditional balance theory, the proposal of replacing the (traditional) notion of frustration, $F(\boldsymbol{\sigma}) = L_{\bullet}^{-} + L_{\circ}^{+}$ (i.e., the total amount of misplaced connections, coinciding with the number of negative edges found within communities, $L_{\bullet}^{-}$, plus the number of positive edges found between communities, $L_{\circ}^{+}$), with its softened variant, $G(\boldsymbol{\sigma}|\alpha) = \alpha L_{\bullet}^{-} + (1-\alpha)L_{\circ}^{+}$, was advanced. Even ignoring the ambiguity due to the lack of a principled way for selecting α (the so-called α problem in [5]), the criterion embodied by the G test is still too strict [6]. In the light of such a result, the second attempt pursued by Doreian and Mrvar to overcome the perceived limitations of the traditional balance theory was more radical, as they proposed to relax it by allowing for the presence of positive, off-diagonal blocks and negative, diagonal blocks—a generalization that has gained the name of *relaxed balance theory* [4]. Such a formulation of the relaxed balance theory lacks a proper mathematization, as a score function such as $F(\boldsymbol{\sigma})$, or $G(\boldsymbol{\sigma}|\alpha)$, cannot be easily individuated. Besides, it is affected by the problem highlighted in [7]: "if the number of clusters is left unspecified *a priori,* the best partition is the singletons partition (i.e. each node in its own cluster)." A recent attempt of overcoming such a limitation is represented by the contribution from [6]: recasting the theory of balance within a statistical framework solves several problems at once, allowing an inference scheme for assessing if a signed graph is traditionally or relaxedly balanced to be defined.

The task of evaluating which variant of the traditional balance theory is best supported by empirical data has recently received considerable attention: several metrics have been proposed to evaluate the level of balance [8–14], or of the complementary notion of *frustration* [5,15–17], and the question concerning the variant of the traditional balance theory that is best supported by data has been reformulated in statistical terms, by comparing the empirical abundance of the set of patterns that characterizes each of them with the one predicted by a properly defined benchmark model [6,18–26]. Surprisingly, the answer crucially depends on a number of factors, such as (i) the nature of the data, (ii) the measure adopted to quantify balance, and (iii) the null model employed to carry out the analysis [26]. In particular, (i) null models induced by global constraints tend to favor the weak variant of the balance theory (according to which only the triangle with one negative edge should be underrepresented in real-world networks); (ii) null models induced by local constraints tend to favor the strong variant of the balance theory, according to which the triangle with three negative edges should be underrepresented as well.

Although popular, the balance theory disregards the direction of edges, a piece of information that is accounted for by the *status theory*, implicitly proposed in [27], where the meaning of the negative, directed edges was first discussed. There, the authors focused on the model called the "web of trust," according to which users establish connections on the basis of trust or rating scores: "distrust" is, thus, modeled via a negative, directed edge, a methodological innovation allowing it to be explicitly distinguished from "absence of an opinion," which instead is represented by the absence of an edge. The

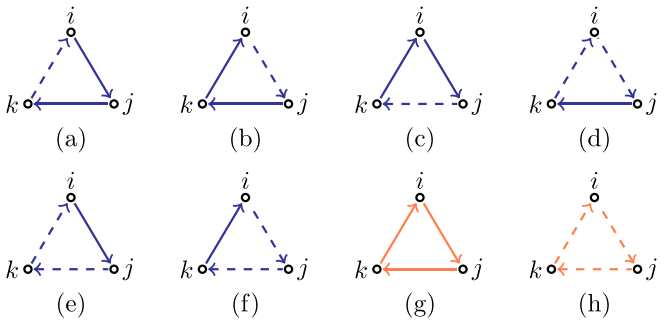

FIG. 1. Cyclic triads headed at node i: according to the status theory, [(a)–(f)] blue triangles are balanced while [(g), (h)] orange triangles are unbalanced. While solid lines represent positive edges, dashed lines represent negative edges.

status theory has been formally developed in [28]. According to it, the sign of an edge between any two nodes depends on the (perceived) difference between their status: while a positive edge directed from node i to node j indicates that the status of j is perceived by i as "higher" than its own status, a negative edge directed from i to j indicates that the status of j is perceived by i as "lower" than its own status. Figures 1 and 2 depict the (directed, triadic) patterns classified as balanced (in blue) and unbalanced (in orange) according to the status theory.

The study of directed, signed networks has been also approached with different purposes. In [29–32] the problems of edge and sign prediction are addressed. In [14,16,33] a statistical validation of the level of frustration characterizing directed, signed networks is carried out—to be noticed that the notion of frustration employed there is the one defined for undirected networks. In [34] the concept of *status factor* is formalized, i.e., an index quantifying the social status of the nodes belonging to overlapping communities. In [35] the authors address the problem of opinion formation by labeling each edge with a sign that is determined by the similarity of the two interacting agents. In [36] the level of frustration characterizing directed, signed networks is proxied by the percentage of acyclic triads not satisfying the principle of transitivity while cyclic triads are excluded from the analysis.

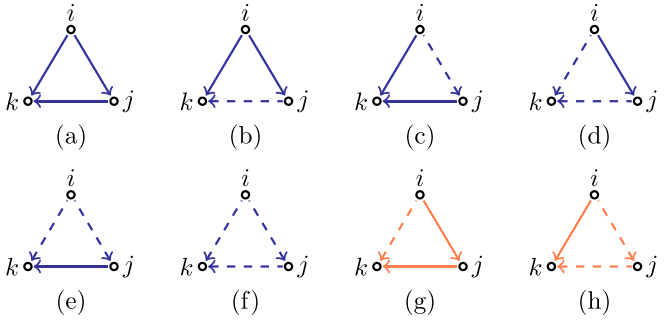

FIG. 2. Acyclic triads headed at node i: according to the status theory, [(a)–(f)] blue triangles are balanced while [(g), (h)] orange triangles are unbalanced. While solid lines represent positive edges, dashed lines represent negative edges.

B. A brief history of reciprocity

The study of reciprocity in binary, directed networks [37,38], i.e., the tendency of vertex pairs to form mutual connections, has received increasing attention over the years [39–50]. Reciprocity has been shown to play a crucial role in classifying [39] and modeling [40] directed networks, understanding the effects of a network structure on dynamical processes such as diffusion and percolation [41–43], explaining patterns of growth in out-of-equilibrium networks such as Wikipedia [44] and the World Trade Web [45,46], shaping social relationships [51,52], and studying the onset of higher-order structures such as correlations [47,48] and triadic motifs [49,50,53,54]. Reciprocity also quantifies how much information is lost when a directed network is regarded as undirected: if the reciprocity of the original network is a maximum, then the full information can be retrieved from the undirected projection; if, on the other hand, no reciprocity is observed, the uncertainty about the directionality of the original edges that have been converted into undirected ones is maximal [39].

Generally speaking, directed networks range between two extremes: being purely bidirectional, such as the Internet, where information always travels both ways along computer cables, or being purely unidirectional, such as citation networks, where recent papers can cite less recent ones while the opposite cannot occur. A traditional way of quantifying where a real network lies is measuring the ratio of the number of edges pointing in both directions, $L^{\leftrightarrow}$, to the total number of edges, L[38,55,56], i.e.,

$$r = \frac{L^{\leftrightarrow}}{L} = \frac{\sum_{i=1}^{N}\sum_{j(\neq i)} a_{ij}a_{ij}}{\sum_{i=1}^{N}\sum_{j(\neq i)} a_{ij}}. \tag{1}$$

Clearly, $r = 1$ for purely bidirectional networks while $r = 0$ for purely unidirectional ones. Social networks [38], email networks [55], the World Wide Web [55] and the World Trade Web [56] display intermediate values of r.

In order to assess if mutual edges occur more or less (or just as) often than expected by chance, the empirical value of reciprocity must be compared with the one expected in a random graph with the same number of vertices [55]. To this aim, the definition of reciprocity reading

$$\rho = \frac{r - \langle r\rangle}{1 - \langle r\rangle}, \tag{2}$$

with

$$\langle r\rangle = \frac{\langle L^{\leftrightarrow}\rangle}{\langle L\rangle} = \frac{\sum_{i=1}^{N}\sum_{j(\neq i)}\langle a_{ij}a_{ij}\rangle}{\sum_{i=1}^{N}\sum_{j(\neq i)}\langle a_{ij}\rangle} = \frac{N(N-1)p^2}{N(N-1)p} = p, \tag{3}$$

was proposed. Still, such a measure suffers from two major limitations: (i) it implements a comparison with the overall too simple benchmark known as the directed random graph model [39], according to which $p = L/N(N-1)$, and (ii) it does not admit a clear, statistical interpretation. In order to overcome these limitations, ρ has been replaced by the z-score

$$z[L^{\leftrightarrow}] = \frac{L^{\leftrightarrow} - \langle L^{\leftrightarrow}\rangle}{\sigma[L^{\leftrightarrow}]}, \tag{4}$$

where $\sigma[L^{\leftrightarrow}] = \sqrt{\mathrm{Var}[L^{\leftrightarrow}]}$ and

$$\begin{aligned}\mathrm{Var}[L^{\leftrightarrow}] &= \sum_{i=1}^{N}\sum_{j(>i)}\mathrm{Var}[2a_{ij}a_{ji}]\\ &= 4\sum_{i=1}^{N}\sum_{j(>i)} p_{ij}p_{ji}(1 - p_{ij}p_{ji}).\end{aligned} \tag{5}$$

The z-score $z[L^{\leftrightarrow}]$ quantifies the number of standard deviations by which the empirical abundance of reciprocal edges differs from the expected one [57]. After checking for the Gaussianity of $L^{\leftrightarrow}$ under the chosen benchmark—since it is a sum of dependent random variables, this is ensured by the generalization of the central limit theorem—a result $|z_m| \leqslant 2$ ($|z_m| \leqslant 3$) indicates that the empirical abundance of the pattern under study is compatible with the one expected under the chosen benchmark at the 5% (1%) level of statistical significance. On the other hand, a value $|z_m| > 2$ ($|z_m| > 3$) indicates that the empirical abundance of pattern m is not compatible with the chosen benchmark at the same significance level: in the latter case, a value $z_m > 0$ ($z_m < 0$) indicates the tendency of the pattern to be overrepresented (underrepresented) in the data with respect to the null model. Real-world directed networks have been shown [54,57] to display a value of reciprocity that neither the directed random graph model nor the directed configuration model is capable of reproducing [54,57].

The present paper is devoted to exploring the concept of signed reciprocity as well as its relationship with that of frustration. To this aim, we adopt a principled approach for its study, introducing models to quantify the tendency of empirical networks to form either mutualistic or antagonistic cycles of length 2, that are suitable to analyze graphs with "plus one," "minus one," and "zero" edges as well as those with just "plus one" and "minus one" edges. More specifically, we extend the exponential random graph framework to binary, directed, signed networks with global and local constraints. As in the undirected case [26], our approach has the clear advantage of being analytically tractable; besides, its versatility allows it to be employed either in the presence of full information—to spot the degree of self-organization of a signed network by detecting the patterns that are not explained by lower-level constraints [57–61]—or in presence of partial information—to infer the missing portion of a given signed network [62].

The rest of the paper is organized as follows. Section II introduces the formalism and the basic quantities we will consider in our analysis. Section III is devoted to the description of the maximum-entropy models we will employ to carry out our analysis. Section IV illustrates their application to a bunch of real-world networks. Section V discusses the main results of the paper and presents an outlook on future extensions of our work.

II. SETTING UP THE FORMALISM

In a signed graph, each edge can be positive, negative, or missing. In what follows, we will focus on binary, directed, signed networks: each edge will be, thus, "plus one," "minus one," or "zero." More formally, the generic entry of the adjacency matrix $\mathbf{A}$ reads $a_{ij} = -1, 0, +1$. Since the total number

of node pairs reads $N(N-1)$ and any "directed" node pair can be "positively connected," "negatively connected," or "disconnected," the cardinality of the ensemble of binary, directed, signed graphs is $|\mathbb{A}| = 3^{N(N-1)}$. Mathematical manipulations can be eased by employing Iverson's brackets to define the three functions of the entries reading $a_{ij}^{-} = [a_{ij} = -1]$, $a_{ij}^{0} = [a_{ij} = 0]$, and $a_{ij}^{+} = [a_{ij} = +1]$ These new variables, inducing the definition of the matrices $\mathbf{A}^{+}$ and $\mathbf{A}^{-}$ (see Appendix A in the Supplemental Material [63]), are mutually exclusive, i.e., $\{a_{ij}^{-}, a_{ij}^{0}, a_{ij}^{+}\} = \{(1,0,0),(0,1,0),(0,0,1)\}$, and sum to 1, i.e., $a_{ij}^{-} + a_{ij}^{0} + a_{ij}^{+} = 1$. The advantage of adopting such a formalism becomes evident when considering that each quantity is, now, computed on a matrix whose entries are, by definition, positive; as a consequence, it is positive as well.

A. Total number of positive and negative links

Our formalism allows us to define a number of quantities of interest. For instance, the number of positive and negative edges can be, respectively, defined as

$$L^{+} = \sum_{i=1}^{N} \sum_{j(\neq i)} a_{ij}^{+} \quad \text{and} \quad L^{-} = \sum_{i=1}^{N} \sum_{j(\neq i)} a_{ij}^{-}. \tag{6}$$

Naturally, $L = L^{+} + L^{-} = \sum_{i=1}^{N} \sum_{j(\neq i)} (a_{ij}^{+} + a_{ij}^{-})$.

B. Degree sequences

The positive and negative out-degrees of node i (i.e., the total number of positive and negative edges "outgoing from" node i) can be, respectively, defined as

$$k_i^{+} = \sum_{\substack{j=1 \\ (j\neq i)}}^{N} a_{ij}^{+} \quad \text{and} \quad k_i^{-} = \sum_{\substack{j=1 \\ (j\neq i)}}^{N} a_{ij}^{-}, \tag{7}$$

while the positive and negative in-degrees of node i (i.e., the total number of positive and negative edges "entering into" node i) can be, respectively, defined as

$$h_i^{+} = \sum_{\substack{j=1 \\ (j\neq i)}}^{N} a_{ji}^{+} \quad \text{and} \quad h_i^{-} = \sum_{\substack{j=1 \\ (j\neq i)}}^{N} a_{ji}^{-}. \tag{8}$$

Naturally, $L^{+} = \sum_{i=1}^{N} k_i^{+} = \sum_{i=1}^{N} h_i^{+}$ and $L^{-} = \sum_{i=1}^{N} k_i^{-} = \sum_{i=1}^{N} h_i^{-}$.

C. Reciprocity

In the case of unsigned networks, the percentage of edges having a "companion" pointing in the opposite direction is defined as in Eq. (1). In the case of signed networks, four types of reciprocal patterns can be identified [64], i.e., those leading to (i) positive reciprocity, when nodes i and j trust each other; (ii) negative reciprocity, when nodes i and j distrust each other; (iii) positive antireciprocity, when node i trusts node j but node j distrusts node i; and (iv) negative antireciprocity, when node i distrusts node j but node j trusts node i. These patterns can be compactly represented via the family of dyadic motifs reading

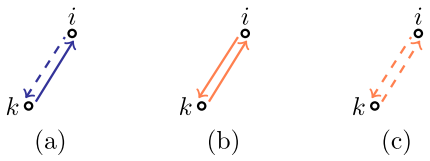

FIG. 3. Dyadic motifs for binary, directed, signed networks: (a) a pair of edges, with discordant signs, pointing in opposite directions; (b) a pair of positive edges pointing in opposite directions; and (c) a pair of negative edges pointing in opposite directions. According to the balance theory, pattern (a) is unbalanced while patterns (b) and (c) are balanced; according to the status theory, pattern (a) is balanced while patterns (b) and (c) are unbalanced.

$$L_{+}^{\leftrightarrow} = \sum_{i=1}^{N} \sum_{j(\neq i)} a_{ij}^{+} a_{ji}^{+} = 2 \sum_{i=1}^{N} \sum_{j(>i)} a_{ij}^{+} a_{ji}^{+}, \tag{9}$$

counting the pairs of positive edges pointing in opposite directions,

$$L_{-}^{\leftrightarrow} = \sum_{i=1}^{N} \sum_{j(\neq i)} a_{ij}^{-} a_{ji}^{-} = 2 \sum_{i=1}^{N} \sum_{j(>i)} a_{ij}^{-} a_{ji}^{-}, \tag{10}$$

counting the pairs of negative edges pointing in opposite directions,

$$\begin{aligned} L_{\pm}^{\leftrightarrow} &= \sum_{i=1}^{N} \sum_{j(\neq i)} (a_{ij}^{+} a_{ji}^{-} + a_{ij}^{-} a_{ji}^{+}) \\ &= 2 \sum_{i=1}^{N} \sum_{j(>i)} (a_{ij}^{+} a_{ji}^{-} + a_{ij}^{-} a_{ji}^{+}), \end{aligned} \tag{11}$$

and counting the pairs of edges, with different signs, pointing in opposite directions (see Fig. 3). The tendency to establish antireciprocal (or single) connections can be quantified as well, upon defining

$$\begin{aligned} L_{+}^{\rightarrow} &= \sum_{i=1}^{N} \sum_{j(\neq i)} a_{ij}^{+} (1 - a_{ji}^{+} - a_{ji}^{-}) \\ &= \sum_{i=1}^{N} \sum_{j(>i)} [a_{ij}^{+} (1 - a_{ji}^{+} - a_{ji}^{-}) + a_{ji}^{+} (1 - a_{ij}^{+} - a_{ji}^{-})] \end{aligned} \tag{12}$$

and

$$\begin{aligned} L_{-}^{\rightarrow} &= \sum_{i=1}^{N} \sum_{j(\neq i)} a_{ij}^{-} (1 - a_{ji}^{+} - a_{ji}^{-}) \\ &= \sum_{i=1}^{N} \sum_{j(>i)} [a_{ij}^{-} (1 - a_{ji}^{+} - a_{ji}^{-}) + a_{ji}^{-} (1 - a_{ji}^{+} - a_{ij}^{-})]. \end{aligned} \tag{13}$$

As a consequence, the following notions of reciprocity,

$$r_{+}^{\leftrightarrow} = \frac{L_{+}^{\leftrightarrow}}{L}, \quad r_{-}^{\leftrightarrow} = \frac{L_{-}^{\leftrightarrow}}{L}, \quad r_{\pm}^{\leftrightarrow} = \frac{L_{\pm}^{\leftrightarrow}}{L}, \tag{14}$$

remain naturally defined, with $r_{+}^{\leftrightarrow}$ quantifying the percentage of edges appearing in reciprocal positive pairs, $r_{-}^{\leftrightarrow}$ quantifying

the percentage of edges appearing in reciprocal negative pairs, and $r_{\pm}^{\leftrightarrow}$ quantifying the percentage of edges appearing in a reciprocal pair with discordant signs. Besides,

$$r_{+}^{\rightarrow} = \frac{L_{+}^{\rightarrow}}{L}, \quad r_{-}^{\rightarrow} = \frac{L_{-}^{\rightarrow}}{L}, \tag{15}$$

with $r_{+}^{\rightarrow}$ quantifying the percentage of edges appearing in a positive, antireciprocal pair and $r_{-}^{\rightarrow}$ quantifying the percentage of edges appearing in a negative, antireciprocal pair. Similarly,

$$s_{+}^{\leftrightarrow} = \frac{L_{+}^{\leftrightarrow}}{L^{\leftrightarrow}}, \quad s_{-}^{\leftrightarrow} = \frac{L_{-}^{\leftrightarrow}}{L^{\leftrightarrow}}, \quad s_{\pm}^{\leftrightarrow} = \frac{L_{\pm}^{\leftrightarrow}}{L^{\leftrightarrow}}, \tag{16}$$

quantify the percentage of reciprocal edges that are both positive, the percentage of reciprocal edges that are both negative, and the percentage of reciprocal edges with discordant signs.

As first noticed in [65], the concepts of reciprocity and frustration are not independent: in [14], the authors consider reciprocated dyads as generating two semicycles, whose degree of frustration is to be evaluated in the light of the balance theory; in [28,51], the authors notice that the balance theory has a larger statistical evidence on the subgraph defined by reciprocal edges. For the moment, we would simply like to stress that frustration and reciprocity are related by the meaning one attributes to (anti)reciprocal patterns: if the collected data reflect a relationship such as "friendship," then Fig. 3(b) is balanced (i sees j as a friend, hence $a_{ij} = +1$, and j sees i as a friend, hence $a_{ji} = +1$); if, on the contrary, the collected data reflect a relationship like "subordination," then Fig. 3(b) is frustrated (i sees j as leading, hence $a_{ij} = +1$, but j sees i as leading as well, hence $a_{ji} = +1$).

III. BENCHMARKS FOR BINARY, DIRECTED, SIGNED NETWORKS

Let us now generalize the exponential random graph (ERG) framework to binary, directed, signed networks. To this aim, we will follow the analytical approach introduced in [26] and carry out a constrained maximization of Shannon entropy,

$$S = -\sum_{\mathbf{A}\in\mathbb{A}} P(\mathbf{A}) \ln P(\mathbf{A}), \tag{17}$$

where the sum runs over the ensemble of $|\mathbb{A}| = 3^{N(N-1)}$ binary, directed, signed networks, a generic entry of which can assume the values $-1, 0, +1$. Let us now discuss the models induced by two different sets of constraints.

A. Signed directed random graph model

The signed directed random graph model (SDRGM) is induced by the Hamiltonian

$$H(\mathbf{A}) = \alpha L^{+}(\mathbf{A}) + \beta L^{-}(\mathbf{A}), \tag{18}$$

i.e., by the two global constraints $L^{+}(\mathbf{A})$ and $L^{-}(\mathbf{A})$.

1. Free-topology SDRGM

According to the free-topology variant of the SDRGM, each entry of a directed, signed network is a random variable whose behavior is described by the finite scheme

$$a_{ij} \sim \begin{pmatrix} -1 & 0 & +1 \\ p^{-} & p^{0} & p^{+} \end{pmatrix}, \quad \forall\, i \neq j, \tag{19}$$

with

$$p^{-} \equiv \frac{e^{-\beta}}{1 + e^{-\alpha} + e^{-\beta}} \equiv \frac{y}{1 + x + y}, \tag{20}$$

$$p^{+} \equiv \frac{e^{-\alpha}}{1 + e^{-\alpha} + e^{-\beta}} \equiv \frac{x}{1 + x + y}, \tag{21}$$

and $p^{0} \equiv 1 - p^{-} - p^{+}$. In other words, a_{ij} obeys a generalized Bernoulli distribution whose probability coefficients are determined by the (Lagrange multipliers of the) imposed constraints (see Appendix B in the Supplemental Material [63]): each positive edge appears with probability p^{+}, each negative edge appears with probability p^{-}, and each missing edge has a probability p^{0}.

In order to employ the SDRGM to study real-world networks, the parameters that define it need to be properly tuned. More specifically, one needs to ensure that $\langle L^{+}\rangle_{\text{SDRGM}} = L^{+}(\mathbf{A}^{*})$ and that $\langle L^{-}\rangle_{\text{SDRGM}} = L^{-}(\mathbf{A}^{*})$. To this aim, the log-likelihood maximization principle can be invoked [66]: it prescribes to maximize the function

$$\begin{aligned}\mathcal{L}_{\text{SDRGM}}(x, y) &= \ln P_{\text{SDRGM}}(\mathbf{A}^{*}|x, y) \\ &= \ln \left[\prod_{i=1}^{N} \prod_{j(\neq i)} (p^{-})^{a_{ij}^{-}} (p^{0})^{a_{ij}^{0}} (p^{+})^{a_{ij}^{+}} \right]\end{aligned} \tag{22}$$

with respect to the unknown parameters that define it. Such a recipe leads to the set of relationships

$$p^{+} = \frac{L^{+}(\mathbf{A}^{*})}{N(N-1)}, \quad p^{-} = \frac{L^{-}(\mathbf{A}^{*})}{N(N-1)}, \tag{23}$$

and $p^{0} \equiv 1 - p^{-} - p^{+}$.

2. Fixed-topology SDRGM

The way the SDRGM has been defined allows a network topological structure to vary along with the signs of the edges. A variant of the SDRGM that keeps the (directed) topology of the network under analysis fixed while (solely) randomizing the signs of the edges is, however, definable. The role of random variables is now played by the entries of the adjacency matrix corresponding to connected pairs of nodes. Each of them obeys the finite scheme

$$a_{ij} \sim \begin{pmatrix} -1 & +1 \\ p^{-} & p^{+} \end{pmatrix}, \quad \forall\, i \neq j \;|\; |a_{ij}| = 1, \tag{24}$$

with

$$p^{-} \equiv \frac{e^{-\beta}}{e^{-\alpha} + e^{-\beta}} \equiv \frac{y}{x + y}, \tag{25}$$

$$p^{+} \equiv \frac{e^{-\alpha}}{e^{-\alpha} + e^{-\beta}} \equiv \frac{x}{x + y}. \tag{26}$$

In other words, each entry satisfying $|a_{ij}| = 1$ obeys a Bernoulli distribution whose probability coefficients are determined by the (Lagrange multipliers of the) imposed constraints (see Appendix B in the Supplemental Material

[63]): each existing edge is assigned a "plus one" with probability p^+ and a "minus one" with probability p^-.

The maximization of the log-likelihood function

$$\mathcal{L}_{\text{SDRGM-FT}}(x, y) = \ln P_{\text{SDRGM-FT}}(\mathbf{A}^*|x, y) = \ln\left[\prod_{i=1}^{N}\prod_{j(\neq i)}(p^-)^{a_{ij}^-}(p^+)^{a_{ij}^+}\right] \tag{27}$$

with respect to the unknown parameters that define it leads to the set of relationships

$$p^+ = \frac{L^+(\mathbf{A}^*)}{L(\mathbf{A}^*)}, \quad p^- = \frac{L^-(\mathbf{A}^*)}{L(\mathbf{A}^*)}, \tag{28}$$

with $L(\mathbf{A}^*)$ representing the (empirical) number of edges characterizing the fixed topology under consideration.

B. Signed directed configuration model

The two aforementioned versions of the SDRGM are defined by global constraints; let us now consider a more refined null model, induced by local constraints. The signed directed configuration model (SDCM) is induced by the Hamiltonian

$$H(\mathbf{A}) = \sum_{i=1}^{N}[\alpha_i k_i^+(\mathbf{A}) + \beta_i k_i^-(\mathbf{A}) + \gamma_i h_i^+(\mathbf{A}) + \delta_i h_i^-(\mathbf{A})], \tag{29}$$

i.e., by the four vectors of local constraints, $\{k_i^+(\mathbf{A})\}_{i=1}^N$, $\{k_i^-(\mathbf{A})\}_{i=1}^N$, $\{h_i^+(\mathbf{A})\}_{i=1}^N$, and $\{h_i^-(\mathbf{A})\}_{i=1}^N$.

1. Free-topology SDCM

According to the free-topology variant of the SDCM, each entry of a directed, signed network is a random variable whose behavior is described by the finite scheme

$$a_{ij} \sim \begin{pmatrix} -1 & 0 & +1 \\ p_{ij}^- & p_{ij}^0 & p_{ij}^+ \end{pmatrix}, \quad \forall\, i \neq j, \tag{30}$$

with

$$p_{ij}^- \equiv \frac{e^{-(\beta_i+\delta_j)}}{1 + e^{-(\alpha_i+\gamma_j)} + e^{-(\beta_i+\delta_j)}} \equiv \frac{y_i w_j}{1 + x_i z_j + y_i w_j}, \tag{31}$$

$$p_{ij}^+ \equiv \frac{e^{-(\alpha_i+\gamma_j)}}{1 + e^{-(\alpha_i+\gamma_j)} + e^{-(\beta_i+\delta_j)}} \equiv \frac{x_i z_j}{1 + x_i z_j + y_i w_j}, \tag{32}$$

and $p_{ij}^0 \equiv 1 - p_{ij}^- - p_{ij}^+$. In other words, a_{ij} obeys a generalized Bernoulli distribution whose probability coefficients are determined by the (Lagrange multipliers of the) imposed constraints (see Appendix B in the Supplemental Material [63]): given any two nodes i and j, they are connected by a positive edge with probability p_{ij}^+, connected by a negative edge with probability p_{ij}^-, and are disconnected with probability p_{ij}^0.

In order to ensure that $\langle k_i^+\rangle_{\text{SDCM}} = k_i^+(\mathbf{A}^*)$, $\langle k_i^-\rangle_{\text{SDCM}} = k_i^-(\mathbf{A}^*)$, $\langle h_i^+\rangle_{\text{SDCM}} = h_i^+(\mathbf{A}^*)$, and $\langle h_i^-\rangle_{\text{SDCM}} = h_i^-(\mathbf{A}^*)$, $\forall\, i$, let us maximize the log-likelihood function

$$\mathcal{L}_{\text{SDCM}}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}) = \ln P_{\text{SDCM}}(\mathbf{A}^*|\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}) = \ln\left[\prod_{i=1}^{N}\prod_{j(\neq i)}(p_{ij}^-)^{a_{ij}^-}\left(p_{ij}^0\right)^{a_{ij}^0}(p_{ij}^+)^{a_{ij}^+}\right] \tag{33}$$

with respect to the unknown parameters that define it. Such a recipe leads to the set of relationships

$$k_i^+(\mathbf{A}^*) = \sum_{\substack{j=1\\(j\neq i)}}^{N} \frac{x_i z_j}{1 + x_i z_j + y_i w_j} = \langle k_i^+\rangle, \quad \forall\, i, \tag{34}$$

$$k_i^-(\mathbf{A}^*) = \sum_{\substack{j=1\\(j\neq i)}}^{N} \frac{y_i w_j}{1 + x_i z_j + y_i w_j} = \langle k_i^-\rangle, \quad \forall\, i, \tag{35}$$

$$h_i^+(\mathbf{A}^*) = \sum_{\substack{j=1\\(j\neq i)}}^{N} \frac{x_j z_i}{1 + x_j z_i + y_j w_i} = \langle h_i^+\rangle, \quad \forall\, i, \tag{36}$$

$$h_i^-(\mathbf{A}^*) = \sum_{\substack{j=1\\(j\neq i)}}^{N} \frac{y_j w_i}{1 + x_j z_i + y_j w_i} = \langle h_i^-\rangle, \quad \forall\, i. \tag{37}$$

Such a system can be solved only numerically (see Appendixes C and D in the Supplemental Material [63]).

2. Fixed-topology SDCM

As for the SDRGM, a variant of the SDCM that keeps the (directed) topology of the network under analysis fixed while (solely) randomizing the signs of the edges is definable. Again, the role of random variables is played by the entries of the adjacency matrix corresponding to connected pairs of nodes. Each of them obeys the finite scheme

$$a_{ij} \sim \begin{pmatrix} -1 & +1 \\ p_{ij}^- & p_{ij}^+ \end{pmatrix}, \quad \forall\, i \neq j \;|\; |a_{ij}| = 1, \tag{38}$$

with

$$p_{ij}^- \equiv \frac{e^{-(\beta_i+\delta_j)}}{e^{-(\alpha_i+\gamma_j)} + e^{-(\beta_i+\delta_j)}} \equiv \frac{y_i w_j}{x_i z_j + y_i w_j}, \tag{39}$$

$$p_{ij}^+ \equiv \frac{e^{-(\alpha_i+\gamma_j)}}{e^{-(\alpha_i+\gamma_j)} + e^{-(\beta_i+\delta_j)}} \equiv \frac{x_i z_j}{x_i z_j + y_i w_j}. \tag{40}$$

In other words, each entry satisfying $|a_{ij}| = 1$ obeys a Bernoulli distribution whose probability coefficients are determined by the (Lagrange multipliers of the) imposed constraints (see Appendix B in the Supplemental Material [63]): given any two connected nodes i and j, their edge is assigned a "plus one" with probability p_{ij}^+ and a "minus one" with probability p_{ij}^-.

The maximization of the log-likelihood function

$$\mathcal{L}_{\text{SDCM-FT}}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}) = \ln P_{\text{SDCM-FT}}(\mathbf{A}^*|\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}) = \ln\left[\prod_{i=1}^{N}\prod_{j(\neq i)}(p_{ij}^-)^{a_{ij}^-}(p_{ij}^+)^{a_{ij}^+}\right] \tag{41}$$

with respect to the unknown parameters that define it leads to the set of relationships

$$k_i^+(\mathbf{A}^*) = \sum_{\substack{j=1 \\ (j\neq i)}}^{N} |a_{ij}^*| \frac{x_i z_j}{x_i z_j + y_i w_j} = \langle k_i^+ \rangle, \quad \forall\, i, \tag{42}$$

$$k_i^-(\mathbf{A}^*) = \sum_{\substack{j=1 \\ (j\neq i)}}^{N} |a_{ij}^*| \frac{y_i w_j}{x_i z_j + y_i w_j} = \langle k_i^- \rangle, \quad \forall\, i, \tag{43}$$

$$h_i^+(\mathbf{A}^*) = \sum_{\substack{j=1 \\ (j\neq i)}}^{N} |a_{ji}^*| \frac{x_j z_i}{x_j z_i + y_j w_i} = \langle h_i^+ \rangle, \quad \forall\, i, \tag{44}$$

$$h_i^-(\mathbf{A}^*) = \sum_{\substack{j=1 \\ (j\neq i)}}^{N} |a_{ji}^*| \frac{y_j w_i}{x_j z_i + y_j w_i} = \langle h_i^- \rangle, \quad \forall\, i. \tag{45}$$

Such a system can be solved only numerically (see Appendixes C and D in the Supplemental Material [63]).

The meaning of our benchmarks can be understood from a behavioral perspective: according to our intuition, in fact, they can be employed to set the "level of tolerance" characterizing the agents that constitute a given social system. More specifically, null models constraining global quantities assume all agents to be animated by a comparable level of tolerance while null models constraining local quantities explicitly account for the different levels of tolerance that animate different agents [26].

C. The SIMONA matlab package

As an additional result, we release a matlab-coded package that implements all the aforementioned models: its name is SIMONA, an acronym standing for "signed models for network analysis," and is freely downloadable [67].

IV. INFERENCE UNDER BINARY, DIRECTED, SIGNED BENCHMARKS

Let us now compare the empirical abundance of the quantities defined in the previous sections with the one expected under (any of) our null models. As anticipated, a useful indicator is the so-called z-score, i.e.,

$$z_m = \frac{N_m(\mathbf{A}^*) - \langle N_m \rangle}{\sigma[N_m]}, \tag{46}$$

where $N_m(\mathbf{A}^*)$ is the empirical abundance of motif m as measured on $\mathbf{A}^*$, $\langle N_m \rangle$ is its expected value under the chosen null model, and $\sigma[N_m] = \sqrt{\langle N_m^2 \rangle - \langle N_m \rangle^2}$ is the standard deviation of N_m under the same null model. For some of the aforementioned quantities, the z-score can be explicitly computed. In particular, the expected values of the abundances of the reciprocal dyads read

$$\langle L_+^{\leftrightarrow} \rangle = 2 \sum_{i=1}^{N} \sum_{j(>i)} p_{ij}^+ p_{ji}^+, \tag{47}$$

$$\langle L_-^{\leftrightarrow} \rangle = 2 \sum_{i=1}^{N} \sum_{j(>i)} p_{ij}^- p_{ji}^-, \tag{48}$$

and

$$\langle L_\pm^{\leftrightarrow} \rangle = 2 \sum_{i=1}^{N} \sum_{j(>i)} (p_{ij}^+ p_{ji}^- + p_{ij}^- p_{ji}^+), \tag{49}$$

while the corresponding variances read

$$\begin{aligned} \mathrm{Var}[L_+^{\leftrightarrow}] &= 4 \sum_{i=1}^{N} \sum_{j(>i)} \mathrm{Var}[a_{ij}^+ a_{ji}^+] \\ &= 4 \sum_{i=1}^{N} \sum_{j(>i)} p_{ij}^+ p_{ji}^+ (1 - p_{ij}^+ p_{ji}^+), \end{aligned} \tag{50}$$

$$\begin{aligned} \mathrm{Var}[L_-^{\leftrightarrow}] &= 4 \sum_{i=1}^{N} \sum_{j(>i)} \mathrm{Var}[a_{ij}^- a_{ji}^-] \\ &= 4 \sum_{i=1}^{N} \sum_{j(>i)} p_{ij}^- p_{ji}^- (1 - p_{ij}^- p_{ji}^-), \end{aligned} \tag{51}$$

and

$$\begin{aligned} \mathrm{Var}[L_\pm^{\leftrightarrow}] &= 4 \sum_{i=1}^{N} \sum_{j(>i)} \mathrm{Var}[a_{ij}^+ a_{ji}^- + a_{ij}^- a_{ji}^+] \\ &= 4 \sum_{i=1}^{N} \sum_{j(>i)} \mathrm{Var}[a_{ij}^+ a_{ji}^-] + \mathrm{Var}[a_{ij}^- a_{ji}^+] + 2\mathrm{Cov}[a_{ij}^+ a_{ji}^-, a_{ij}^- a_{ji}^+] \\ &= 4 \sum_{i=1}^{N} \sum_{j(>i)} [p_{ij}^+ p_{ji}^- (1 - p_{ij}^+ p_{ji}^-) + p_{ij}^- p_{ji}^+ (1 - p_{ij}^- p_{ji}^+) - 2 p_{ij}^+ p_{ji}^+ p_{ij}^- p_{ji}^-] \\ &= 4 \sum_{i=1}^{N} \sum_{j(>i)} [p_{ij}^+ p_{ji}^- (1 - p_{ij}^+ p_{ji}^- - p_{ij}^- p_{ji}^+) + p_{ij}^- p_{ji}^+ (1 - p_{ij}^- p_{ji}^+ - p_{ij}^+ p_{ji}^-)], \end{aligned} \tag{52}$$

given that $\mathrm{Cov}[a_{ij}^+ a_{ji}^-, a_{ij}^- a_{ji}^+] = \langle a_{ij}^+ a_{ji}^- a_{ij}^- a_{ji}^+ \rangle - \langle a_{ij}^+ a_{ji}^- \rangle \langle a_{ij}^- a_{ji}^+ \rangle = -p_{ij}^+ p_{ij}^- p_{ji}^+ p_{ji}^-$, as the events $a_{ij}^+ a_{ji}^- = 1$ and $a_{ij}^- a_{ji}^+ = 1$ are mutually exclusive. The expected values and the variances of the abundances of the single dyads are provided in Appendix E in the Supplemental Material [63]. Let us remind that z_m returns the number of standard

TABLE I. Descriptive statistics of the largest connected component of each of the 11 snapshots of the MMOG dataset [33]. The table shows the total number of nodes, N, the total number of edges, L, the total number of positive and negative edges, L^+ and L^-, the edge density, $c = L/N(N-1)$, the percentage of reciprocated edges, $r = L^{\leftrightarrow}/L$, and the empirical value of the reciprocity measures that we have introduced. Overall, these networks are very sparse but very reciprocated: the few existing edges are paired in dyads, the vast majority of which have a $+/+$ signature.

MMOG	N	L	L^+	L^-	c	r	$r_+^{\leftrightarrow}$	$r_-^{\leftrightarrow}$	$r_\pm^{\leftrightarrow}$	$r_+^{\rightarrow}$	$r_-^{\rightarrow}$	$s_+^{\leftrightarrow}$	$s_-^{\leftrightarrow}$	$s_\pm^{\leftrightarrow}$
Day 10	1001	5 448	5402	46	0.0054	0.89	0.89	0.002	0.0011	0.10	0.006	0.99	0.003	0.0012
Day 11	1079	6 203	6118	85	0.0053	0.88	0.87	0.004	0.0010	0.11	0.010	0.99	0.004	0.0011
Day 12	1142	6 833	6651	182	0.0052	0.87	0.86	0.008	0.0015	0.11	0.018	0.99	0.009	0.0017
Day 13	1220	7 400	7152	248	0.0050	0.87	0.86	0.012	0.0016	0.11	0.021	0.98	0.013	0.0019
Day 14	1296	7 936	7617	319	0.0047	0.86	0.84	0.012	0.0025	0.11	0.027	0.98	0.014	0.0029
Day 15	1328	8 416	8055	361	0.0048	0.86	0.85	0.014	0.0024	0.11	0.028	0.98	0.016	0.0028
Day 16	1371	8 789	8352	437	0.0047	0.86	0.84	0.015	0.0036	0.11	0.033	0.98	0.017	0.0042
Day 17	1414	9 184	8646	538	0.0046	0.85	0.83	0.016	0.0046	0.11	0.040	0.98	0.019	0.0054
Day 18	1448	9 451	8847	604	0.0045	0.85	0.83	0.017	0.0042	0.11	0.045	0.98	0.020	0.0050
Day 19	1489	9 780	9104	676	0.0044	0.84	0.82	0.018	0.0037	0.11	0.049	0.97	0.021	0.0044
Day 20	1520	10 109	9384	725	0.0044	0.84	0.82	0.018	0.0042	0.11	0.052	0.97	0.021	0.0049

deviations by which the empirical and the expected abundance of motif m differ: while a result $|z_m| \leqslant 2$ ($|z_m| \leqslant 3$) would indicate that the empirical abundance of m is compatible with the chosen null model at the 5% (1%) level of statistical significance, a result $|z_m| > 2$ ($|z_m| > 3$) would indicate that the empirical abundance of m is not compatible with the chosen null model at the same significance level.

V. DATASETS DESCRIPTION

Let us now provide a brief description of the three datasets we have considered for the present analysis.

MMOG dataset. This dataset collects information about the relationships among the $\simeq 300\,000$ players of a massively multiplayer online game (MMOG). A positive edge directed from node i to node j indicates that user i perceives user j positively; conversely, a negative edge directed from node i to node j indicates that user i perceives user j negatively [33].

Honduras villages dataset. This dataset collects information about the relationships among the citizens of 11 rural villages in the Copan Province (Western Honduras). The sample population included people older than 12 years to whom questions such as "Who do you spend your free time with?," "Who is your closest friend?," "Who do you discuss personal matters with?," and "Who are the people with whom you do not get along well?" were asked (either via the Trellis software platform [68], or via tablet-based surveys, administered in a face-to-face fashion). While the answers to the first three questions induce a positive edge, the answer to the fourth question induces a negative edge [69].

Spanish high schools dataset. This dataset collects information about the relationships among the students of 13 high schools, 3 of which are located in Madrid and 10 of which are located in Andalucia. Students were asked to name the other students they had a connection with and rate each relationship with a value ranging from -2 to $+2$. Since the data describe a weighted configuration, we binarized it by considering each positive weight as a $+1$ and each negative weight as a -1. As three schools are characterized by two disconnected components and one school is characterized by four disconnected components, we end up with 19 different snapshots.

VI. PATTERNS OF RECIPROCITY IN BINARY, DIRECTED, SIGNED NETWORKS

MMOG dataset. Let us start commenting on the empirical patterns characterizing the MMOG dataset (see also Table I for a summary of descriptive statistics). Overall, these networks are very sparse but very reciprocated: more specifically, 80–90 &*percnt*; of the edges come in positive, reciprocal dyads, and the vast majority of the remaining ones coming in positive, single dyads: as $s_+^{\leftrightarrow}$ confirms, 99% of the reciprocal edges are organized in dyads with a $+/+$ signature. Interestingly, there are more negative single dyads than negative reciprocal dyads: more quantitatively, $L_-^{\leftrightarrow}/L^- \simeq 0.25$, i.e., only a quarter of the negative edges appear in negative reciprocal dyads. Lastly, it is very unlikely to observe reciprocal dyads with discordant signs.

The observations above point out the existence of a strong asymmetry between positive and negative relationships, only partly confirmed when our dataset is observed through the lenses of statistics. Figure 4 shows the evolution of the aforementioned z-scores, i.e., $z_{L_+^{\leftrightarrow}}$, $z_{L_-^{\leftrightarrow}}$, $z_{L_\pm^{\leftrightarrow}}$, $z_{L_+^{\rightarrow}}$, and $z_{L_-^{\rightarrow}}$. The SDRGM and the SDCM return the same picture, i.e., $z_{L_+^{\leftrightarrow}} > z_{L_-^{\leftrightarrow}} > z_{L_\pm^{\leftrightarrow}} \simeq z_{L_-^{\rightarrow}} \simeq 0 > z_{L_+^{\rightarrow}}$. In words, the reciprocal dyads with concordant signs are overrepresented in the data while the single positive dyads are underrepresented. Overall, these results point out the tendency of pairs of nodes to establish reciprocal relationships having the same sign—irrespective of the nature of the latter one—under free-topology benchmarks.

The SDRGM-FT and the SDCM-FT instead return a different picture: the quantities $z_{L_+^{\leftrightarrow}}$, $z_{L_-^{\leftrightarrow}}$, and $z_{L_-^{\rightarrow}}$ are, in fact, positive and of comparable magnitude while $0 > z_{L_\pm^{\leftrightarrow}} > z_{L_+^{\rightarrow}}$. In words, the reciprocal dyads with concordant signs and the single negative dyads are overrepresented in the data while the reciprocal dyads with discordant signs and the single, positive dyads are underrepresented. Overall, these results point out the tendency of pairs of nodes to establish reciprocal relationships having the same sign—irrespective of the nature of

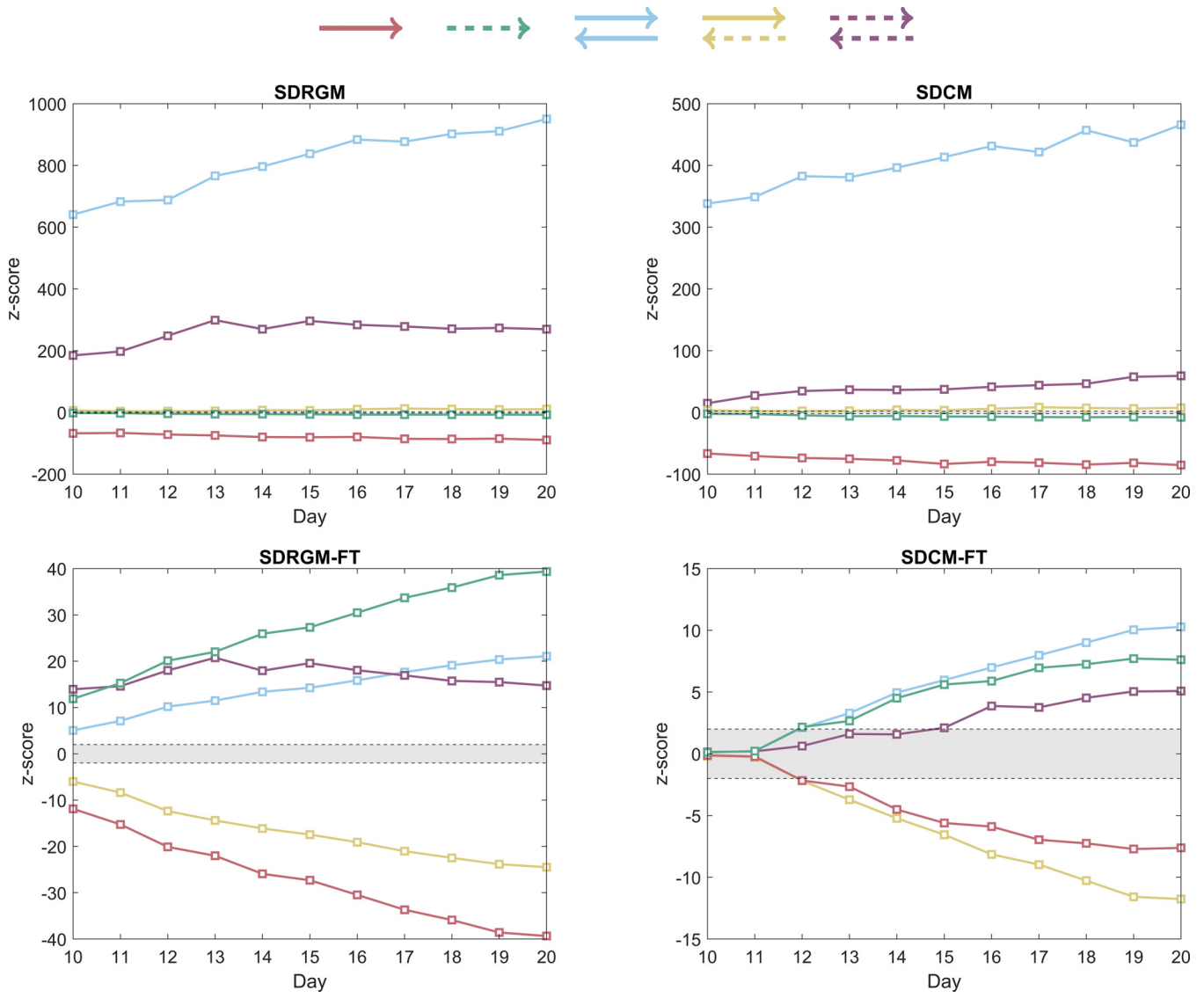

FIG. 4. Top: Evolution of the z-scores of our dyadic motifs, under free-topology benchmarks (left, SDRGM; right, SDCM), for the MMOG dataset. The z-scores induced by the SDRGM are larger, in absolute value, than the z-scores induced by the SDCM, an evidence indicating that accounting for the heterogeneity of nodes greatly enhances the explanatory power of a model. Moreover, the reciprocal dyads with concordant signs (in light blue and purple) are overrepresented in the data while the single, positive dyads (in red) are underrepresented in the data. Bottom: Evolution of the z-scores of our dyadic motifs, under fixed-topology benchmarks (i.e., the SDRGM-FT and the SDCM-FT), for the MMOG dataset. While the explanatory power of the fixed-topology benchmarks is larger than that of their free-topology counterparts, the reciprocal dyads with concordant signs (in light-blue and purple) and the single negative dyads (in green) are overrepresented in the data, while the reciprocal dyads with discordant signs (in yellow) and the single positive dyads (in red) are underrepresented in the data. While the free-topology benchmarks favor the directed version of the balance theory, the fixed-topology benchmarks lead to contradictory results. The gray area corresponds to $z \in [-2, +2]$.

the latter one—as well as single negative dyads under fixed-topology benchmarks.

Honduras villages dataset. Let us now comment the empirical patterns characterizing the Honduras villages dataset (see also Table II for a summary of descriptive statistics). Overall, these networks are very sparse and only moderately reciprocated. In fact, (only) one-third of the edges have a companion pointing in the opposite direction. More specifically, 25–35 &*percnt*; of the edges come in positive reciprocal dyads while more than half of the edges come in positive single dyads and more than 10% of the edges come in negative single dyads: as $s_+^{\leftrightarrow}$ reveals, more than 95% of the reciprocal edges are organized in dyads with a $+/+$ signature. Interestingly, there are many more negative single dyads than negative reciprocal dyads: more quantitatively, $L_-^{\leftrightarrow}/L^- \simeq 0.05$, i.e., only 5% of the negative edges appear in negative reciprocal dyads. Lastly, it is very unlikely to observe reciprocal dyads with discordant signs.

As for the MMOG dataset, a strong asymmetry between positive and negative relationships is present. Figure 5 shows the evolution of the aforementioned z-scores: again, the SDRGM and the SDCM return the same picture, pointing out

TABLE II. Descriptive statistics of the largest connected component of each of the 11 rural villages composing our dataset [69]. The table shows the total number of nodes, N, the total number of edges, L, the total number of positive and negative edges, L^+ and L^-, the edge density, $c = L/N(N-1)$, the percentage of reciprocated edges, $r = L^{\leftrightarrow}/L$, and the empirical value of the reciprocity measures that we have introduced. Overall, these networks are very sparse and only moderately reciprocated; still, the vast majority of the existing dyads have a $+/+$ signature.

Honduras villages	N	L	L^+	L^-	c	r	$r_+^{\leftrightarrow}$	$r_-^{\leftrightarrow}$	$r_\pm^{\leftrightarrow}$	$r_+^{\rightarrow}$	$r_-^{\rightarrow}$	$s_+^{\leftrightarrow}$	$s_-^{\leftrightarrow}$	$s_\pm^{\leftrightarrow}$
A	148	1423	1242	181	0.07	0.36	0.33	0.013	0.02	0.53	0.10	0.91	0.035	0.054
B	101	473	444	29	0.05	0.35	0.34	0.004	0	0.60	0.06	0.99	0.012	0
C	59	388	348	40	0.01	0.36	0.35	0.015	0	0.55	0.09	0.96	0.042	0
D	113	363	306	57	0.03	0.31	0.29	0.011	0.011	0.55	0.14	0.93	0.036	0.036
E	199	652	602	50	0.02	0.34	0.33	0	0.003	0.59	0.08	0.99	0	0.009
F	101	417	353	64	0.04	0.32	0.31	0.005	0.005	0.54	0.15	0.97	0.015	0.015
G	233	1070	924	146	0.02	0.34	0.33	0.008	0.004	0.53	0.13	0.97	0.022	0.011
H	99	582	507	75	0.06	0.38	0.37	0.003	0.007	0.50	0.12	0.97	0.009	0.018
I	98	321	285	36	0.03	0.39	0.38	0	0.006	0.50	0.11	0.98	0	0.016
J	118	715	535	180	0.05	0.27	0.24	0.017	0.014	0.50	0.23	0.89	0.062	0.052
K	206	890	708	182	0.02	0.32	0.29	0.016	0.018	0.50	0.18	0.89	0.049	0.056

the tendency of pairs of nodes to establish reciprocal relationships having the same sign—irrespective of the nature of the latter one. More quantitatively, $z_{L_+^{\leftrightarrow}} > z_{L_-^{\leftrightarrow}} > z_{L_\pm^{\leftrightarrow}} \simeq z_{L_-^{\rightarrow}} \simeq 0 > z_{L_+^{\rightarrow}}$, a result confirming that the reciprocal dyads with concordant signs are overrepresented in the data while the single positive dyads are underrepresented.

The SDRGM-FT and the SDCM-FT, instead, return a different picture, pointing out the tendency of pairs of nodes to establish reciprocal relationships with a $+/+$ signature, as well as single negative dyads. More quantitatively, $z_{L_+^{\leftrightarrow}}$ and $z_{L_-^{\leftrightarrow}}$ are, in fact, positive and of comparable magnitude, while $z_{L_-^{\rightarrow}} \simeq 0 > z_{L_\pm^{\leftrightarrow}} > z_{L_+^{\rightarrow}}$, a result confirming that the reciprocal dyads with a $+/+$ signature and the single negative dyads are overrepresented in the data while the reciprocal dyads with discordant signs and the single positive dyads are underrepresented.

Spanish schools dataset. Lastly, let us comment on the empirical patterns characterizing the Spanish schools dataset (see also Table III for a summary of descriptive statistics). Overall, the density of connections of these networks varies over one order of magnitude, ranging from 0.02 to 0.2; the reciprocity, instead, varies from 0.3 to 0.5. More specifically, 25–50 &*percnt*; of the edges come in positive reciprocal dyads while the same percentage of edges comes in positive single dyads and 10–30 &*percnt*; of the edges come in negative single dyads: as $s_+^{\leftrightarrow}$ reveals, 80–90 &*percnt*; of the reciprocal edges are organized in dyads with a $+/+$ signature. Interestingly, there are many more negative single dyads than negative reciprocal dyads: more quantitatively, $L_-^{\leftrightarrow}/L^-$ ranges from 0.1 to 0.2, i.e., only 10–20 &*percnt*; of the negative edges appear in negative reciprocal dyads. Lastly, it is quite unlikely to observe reciprocal dyads with discordant signs–although it is more likely than observing negative reciprocal dyads, for some of the snapshots.

Figure 6 shows the evolution of the aforementioned z-scores: according to both the SDRGM and the SDCM, the reciprocal dyads with concordant signs are overrepresented in the data while the single positive dyads are underrepresented. More quantitatively, $z_{L_+^{\leftrightarrow}} > z_{L_-^{\leftrightarrow}} > z_{L_\pm^{\leftrightarrow}} \simeq z_{L_-^{\rightarrow}} \simeq 0 > z_{L_+^{\rightarrow}}$.

According to both the SDRGM-FT and the SDCM-FT, instead, the reciprocal dyads with a $+/+$ signature and the single negative dyads are overrepresented in the data while the reciprocal dyads with discordant signs and the single positive dyads are underrepresented. More quantitatively, $z_{L_+^{\leftrightarrow}}$ and $z_{L_-^{\leftrightarrow}}$ are, in fact, positive and of comparable magnitude while $z_{L_-^{\rightarrow}} \simeq 0 > z_{L_\pm^{\leftrightarrow}} > z_{L_+^{\rightarrow}}$.

VII. PATTERNS OF FRUSTRATION IN BINARY, DIRECTED, SIGNED NETWORKS

The structural patterns characterizing our datasets highlight the tendency of nodes to establish reciprocal connections—preferentially, with concordant signs. Do these findings have a meaning from the point of view of balance? According to Cartwright and Harary, this question should be answered by disregarding the directionality of the connections constituting the cycles. In our opinion, such an approach is not fully convincing and some index accounting for directionality should, instead, be considered.

Luckily, the three datasets considered in the present work have been collected in a similar way, i.e., by letting users "rate" each other. As a consequence, the concept of balance better framing our results seems to be closer to the one informing the (directed extension of the) traditional balance theory than to the one informing the status theory. In other words, a larger number of positive reciprocal dyads seems to characterize more balanced configurations.

A. Strong or weak balance?

The negative reciprocal dyads, instead, play a role that is analogous to the role played by undirected triangles with all negative edges. In other words, they may be considered either as a signature of balance or not. The intuition would suggest to not consider a couple of agents "hating" each other as balanced; according to references like [28,51], instead negative reciprocal dyads are balanced, the rationale being that of identifying the notion of "balance" with the notion of "agreement." That is, any two agreeing agents

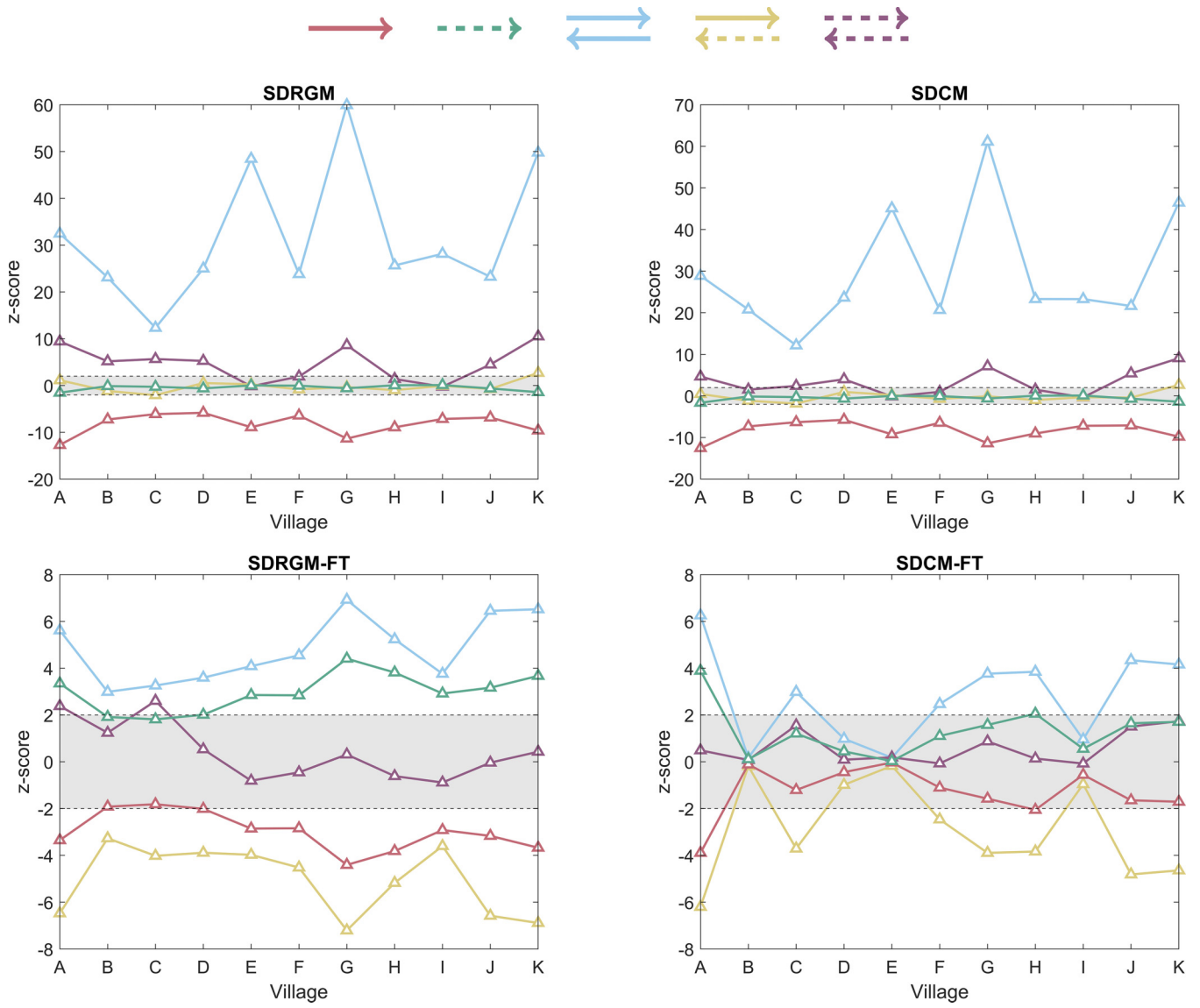

FIG. 5. Top: Evolution of the z-scores of our dyadic motifs, under free-topology benchmarks (left, SDRGM; right, SDCM), for the Honduras villages dataset. The reciprocal dyads with concordant signs (in light blue and purple) are overrepresented in the data while the single positive dyads (in red) are underrepresented in the data. Bottom: Evolution of the z-scores of our dyadic motifs, under fixed-topology benchmarks (i.e., the SDRGM-FT and the SDCM-FT), for the Honduras villages dataset. While the explanatory power of the fixed-topology benchmarks is larger than that of their free-topology counterparts, the reciprocal dyads with a $+/+$ signature (in light blue) and the single negative dyads (in green) are overrepresented in the data while the reciprocal dyads with discordant signs (in yellow) and the single positive dyads (in red) are underrepresented in the data. While the free-topology benchmarks favor the directed version of the balance theory, the fixed-topology benchmarks lead to contradictory results. The gray area corresponds to $z \in [-2, +2]$.

constitute a balanced dyad: while in the first case one may speak of *strong balance*, in the second, one may speak of *weak balance*. Interestingly, the interpretation sustaining the weak balance obeys the (directed extension of the) following definition: cycles with an even number of negative edges are balanced while cycles with an odd number of negative edges are frustrated. More quantitatively, the number of directed cycles with zero and two negative edges is quantified by $L_{+}^{\leftrightarrow}$ and $L_{-}^{\leftrightarrow}$, respectively, while the number of directed cycles with one negative edge is quantified by $L_{\pm}^{\leftrightarrow}$—and they are regarded as a signature of frustration irrespective of the variant of the (directed extension of the) balance theory that is adopted.

B. Free-topology or fixed-topology benchmarks?

Analogously to what has been noticed in [26], different benchmarks provide different answers to the question about the balancedness of a given directed, signed network. More specifically, free-topology benchmarks such as the SDRGM and the SDCM point out that the reciprocal dyads with concordant signs are overrepresented in the data: as a consequence, our three datasets are weakly balanced if free-topology benchmarks are adopted.

On the contrary, fixed-topology benchmarks such as the SDRGM-FT and the SDCM-FT do not lead to the same conclusions, as they point out that the negative single dyads are overrepresented as well. Should these dyads be considered?

TABLE III. Descriptive statistics of the 19 snapshots of the Spanish schools dataset [32]. The table shows the total number of nodes, N, the total number of edges, L, the total number of positive and negative edges, L^+ and L^-, the edge density, $c = L/N(N-1)$, the percentage of reciprocated edges, $r = L^{\leftrightarrow}/L$, and the empirical value of the reciprocity measures that we have introduced. Overall, the density of connections of these networks varies over one order of magnitude while the reciprocity ranges from 0.3 to 0.5. As for the MMOG and the Honduras villages datasets, the vast majority of the existing dyads have a $+/+$ signature.

Spanish schools	N	L	L^+	L^-	c	r	$r_+^{\leftrightarrow}$	$r_-^{\leftrightarrow}$	$r_\pm^{\leftrightarrow}$	$r_+^{\rightarrow}$	$r_-^{\rightarrow}$	$s_+^{\leftrightarrow}$	$s_-^{\leftrightarrow}$	$s_\pm^{\leftrightarrow}$
1	409	8 557	7302	1255	0.05	0.48	0.44	0.014	0.023	0.40	0.12	0.92	0.030	0.047
2	238	3 755	3397	358	0.07	0.46	0.42	0.008	0.033	0.47	0.07	0.91	0.016	0.072
3	534	12 812	11024	1788	0.05	0.51	0.47	0.016	0.029	0.38	0.11	0.91	0.030	0.057
4	291	5 152	3959	1193	0.06	0.35	0.29	0.031	0.026	0.46	0.19	0.84	0.089	0.074
5	702	11 190	8242	2948	0.02	0.33	0.28	0.026	0.021	0.45	0.23	0.86	0.080	0.064
6	97	1 805	1055	750	0.19	0.29	0.22	0.041	0.029	0.35	0.36	0.76	0.140	0.100
7	98	2 080	1243	837	0.22	0.44	0.34	0.072	0.029	0.25	0.32	0.77	0.160	0.066
8	121	1 622	1081	541	0.11	0.55	0.40	0.12	0.036	0.25	0.20	0.72	0.210	0.065
9	60	920	687	233	0.26	0.50	0.38	0.039	0.074	0.33	0.18	0.77	0.079	0.150
10	186	3 197	2470	727	0.09	0.51	0.45	0.026	0.032	0.31	0.18	0.88	0.050	0.064
11	145	2 229	1844	385	0.11	0.31	0.27	0.016	0.027	0.54	0.14	0.86	0.052	0.086
12	105	1 311	975	336	0.12	0.35	0.31	0.015	0.027	0.42	0.23	0.88	0.043	0.078
13	91	935	791	144	0.11	0.27	0.24	0.009	0.019	0.60	0.14	0.90	0.032	0.072
14	132	1 035	855	180	0.06	0.27	0.25	0.012	0.013	0.57	0.16	0.91	0.042	0.049
15	89	593	467	126	0.08	0.28	0.08	0.230	0.030	0.02	0.55	0.17	0.810	0.110
16	66	448	401	47	0.10	0.34	0.31	0.005	0.270	0.57	0.09	0.91	0.013	0.079
17	124	1 365	1097	268	0.09	0.29	0.26	0.012	0.019	0.53	0.18	0.89	0.040	0.065
18	391	4 511	3472	1039	0.03	0.29	0.24	0.017	0.028	0.51	0.20	0.84	0.060	0.098
19	458	7 028	4701	2327	0.03	0.33	0.27	0.028	0.028	0.38	0.29	0.83	0.087	0.084

If so, how? The answer to the first question should be, in our opinion, positive, for (at least) two reasons: (i) the percentage of reciprocal links can be small, hence ignoring single dyads would imply ignoring a substantial portion of a directed, signed configuration, and (ii) the magnitude of the z-scores of the negative single dyads is comparable with the magnitude of the z-scores of the reciprocal dyads with concordant signs. What about the answer to the second question? The way our data have been collected suggests that the absence of an edge should be regarded as a lack of interest in reciprocating the interaction, whether positive or negative. As a consequence, one is led to consider the single dyads as a signature of frustration. If this is the case, however, the (directed extension of the) balance theory leads to contradictory results, as *both* balanced *and* frustrated patterns are overrepresented in the data; as a consequence, no univocal conclusions can be drawn if fixed-topology benchmarks are adopted.

C. A coarser theory of balance?

As previously highlighted, while classifying reciprocal patterns is relatively straightforward, classifying single dyads is more problematic. As we will show, a possible solution is that of grouping all patterns of the same kind by defining the total number of balanced patterns as $N_B = L_+^{\leftrightarrow} + L_-^{\leftrightarrow}$ and the total number of frustrated patterns as $N_F = L_\pm^{\leftrightarrow} + L_+^{\rightarrow} + L_-^{\rightarrow}$. Upon doing so, the z-scores reading

$$z_B = \frac{N_B - \langle N_B \rangle}{\sigma[N_B]} \tag{53}$$

and

$$z_F = \frac{N_F - \langle N_F \rangle}{\sigma[N_F]} \tag{54}$$

remain naturally defined. As Fig. 7 shows, the balanced patterns become overrepresented in the data while the frustrated patterns become underrepresented in the data—consistently across all datasets considered in the present work. In other words, the two indices above make the contradictory patterns disappear, leading balance to be recovered at a "coarser" level.

VIII. DISCUSSION

Extending the notion of balance to directed networks is absolutely not trivial. Although such a topic has been already addressed in several papers, the later ones have focused on triangular motifs, disregarding the information brought by the dyadic ones. Our paper represents an attempt to fill this gap, by exploring the role played by reciprocity in shaping directed, signed networks as well as understanding its relationship with the notion of frustration. Our results point out the tendency of pairs of nodes to establish reciprocal relationships having the same sign and to avoid reciprocal relationships having different signs. Is this a signature of balance or a signature of frustration?

A nonambigouous answer can be provided only after having clarified the way data have been collected: in fact, as stressed in [27], the presence of a signed edge, directed from one node to another, can be interpreted in several ways, depending on the "intention" pushing the source node to bind with the other nodes. In other words, a positive edge from i to j can either mean "node j is a friend of mine"—and the most suitable framework to interpret the results would be the one of the (directed extension of the) balance theory—or "the status of node j is higher than mine"—and the most suitable framework to interpret the results would be the one of the status theory.

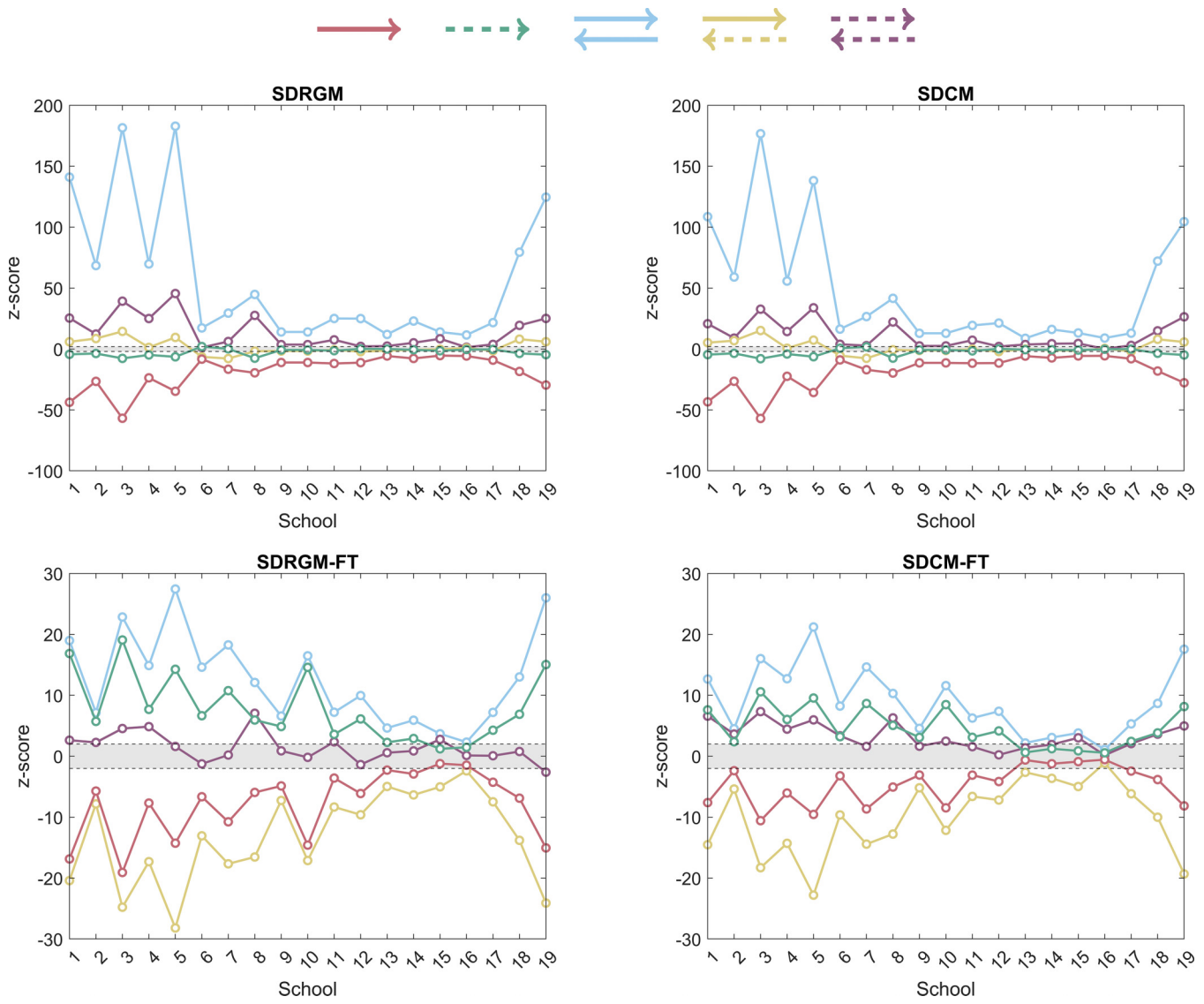

FIG. 6. Top: Evolution of the z-scores of our dyadic motifs, under free-topology benchmarks (left, SDRGM; right, SDCM), for the Spanish schools dataset. The reciprocal dyads with concordant signs (in light blue and purple) are overrepresented in the data while the single, positive dyads (in red) are underrepresented in the data. Bottom: Evolution of the z-scores of our dyadic motifs, under fixed-topology benchmarks (i.e., the SDRGM-FT and the SDCM-FT), for the Spanish schools dataset. The reciprocal dyads with a $+/+$ signature (in light blue) and the single negative dyads (in green) are overrepresented in the data while the reciprocal dyads with discordant signs (in yellow) and the single positive dyads (in red) are underrepresented in the data. As for the MMOG and the Honduras villages datasets, the free-topology benchmarks favor the directed version of the balance theory while the fixed-topology benchmarks lead to contradictory results. The gray area corresponds to $z \in [-2, +2]$.

As the way our datasets have been constructed supports the first hypothesis, we need to interpret reciprocal and single dyads in the light of the (directed extension of the) balance theory. A first route would prescribe to consider only cyclic dyads, the ones with discordant signs being a signature of frustration and the ones with concordant signs being a signature of balance. At a finer level, one may even distinguish a weak, directed extension, considering both $+/+$ and $-/-$ dyads as balanced, from a strong, directed extension, solely considering $+/+$ dyads as balanced. A second route would prescribe to consider single dyads as well. From this perspective, the literature has basically ignored the problem, just focusing on the reciprocal subgraph: while this may represent an acceptable solution for the analysis of configurations with a large number of edges having a "companion" pointing in the opposite direction,[1] it is no longer so when r is small. A choice like this may, in fact, severely bias the analysis: as also reported in [28], the statistical evidence gained by the (directed extension of the) balance theory is larger than

[1]Observing a large reciprocity usually leads to the conclusion that specifying the direction of the edges is, after all, unnecessary; see [39,70,71].

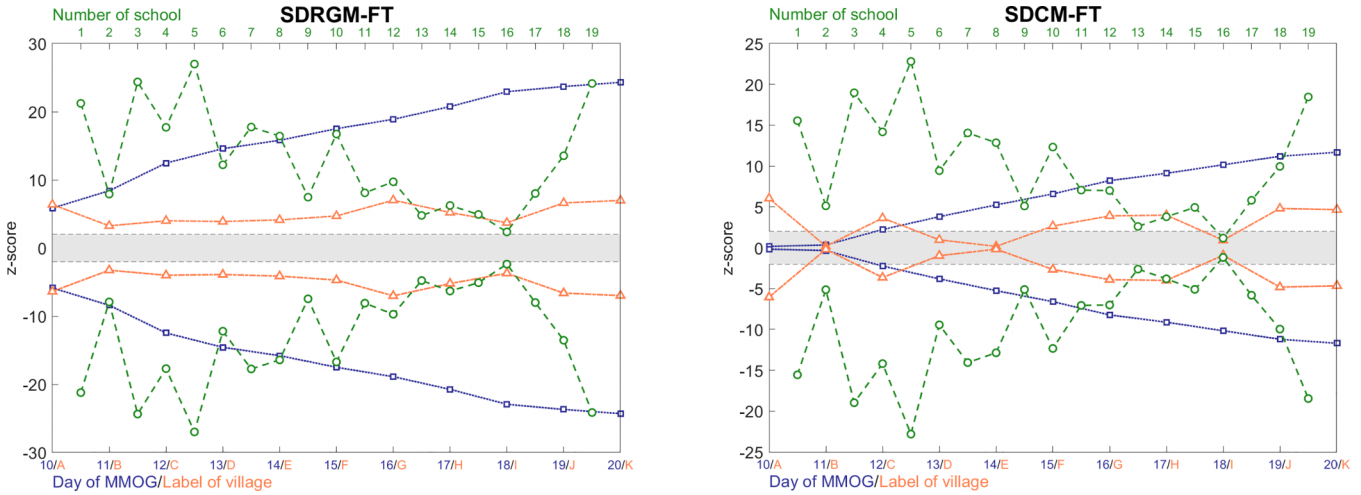

FIG. 7. Evolution of the z-scores of the coarser dyadic motifs, under fixed-topology benchmarks (i.e., the SDRGM-FT and the SDCM-FT), for the MMOG (dotted blue lines), Honduras villages (dash-dotted orange lines), and Spanish schools (dashed green lines) datasets. The reciprocal dyads with concordant signs (i.e., $N_{\mathrm{B}} = L_{+}^{\leftrightarrow} + L_{-}^{\leftrightarrow}$, in blue) are overrepresented in the data while the reciprocal dyads with discordant signs and the single dyads (i.e., $N_{\mathrm{F}} = L_{\pm}^{\leftrightarrow} + L_{+}^{\rightarrow} + L_{-}^{\rightarrow}$, in orange) are underrepresented in the data. In words, balance seems to be recovered at a "coarser" level. The gray area corresponds to $z \in [-2, +2]$.

the one gained by the status theory on the relatively small subgraph induced by reciprocal edges.

Although such a framework may be adopted to interpret the results of the analysis carried out on the MMOG dataset, its use would be much less justifiable to interpret the results of the analyses carried out on the Honduras villages and Spanish schools datasets: those results call for a finer definition of balance, accounting for the meaning of single dyads. Should they be interpreted as a signature of frustration, however, the (directed extensions of the) balance theory may lead to inconsistencies preventing us from drawing univocal conclusions.

A possible way out has been proposed in [28,29,51,52] where it has been noticed that a substantial difference exists between single and reciprocal connections and the adoption of *different* theories to explain the patterns characterizing *different* portions of the *same* directed, signed network has been proposed—more specifically, the balance theory for the subgraph induced by reciprocal edges and the status theory for the remaining portion of the network. As such an approach would assume the *simultaneous* presence of *different*—if not *contradictory*—processes behind the formation of a directed, signed configuration, we believe it not to represent the most suitable explanation for the emergence of certain empirical patterns that, instead, call for a more comprehensive analysis of directed connections in the context of social theories.

This may be especially relevant to individuate the factors driving people's interactions with other people—both offline and online—hence affecting the perception of the social world most. From this perspective, a prominent role is played by (relationships inducing) negative ties, known to be associated with a person's core-ness, facilitating information diffusion and reducing polarization [72]. Since, however, people are more likely to know who likes them than who dislikes them—a kind of interaction often kept private, even when compiling anonymous questionnaires—such information requires extra care to be collected, fluctuating more than the positive counterpart.

IX. CONCLUSIONS

As our analysis highlights, directed, signed networks differ from their undirected counterparts: while the latter ones are, overall, balanced [26], assessing if the former ones are balanced or not is much less straightforward. In our opinion, this is ultimately due to an incomplete formulation of the balance theory for directed networks that ignores the role played by single dyads. While such a result calls for a rethinking of the definition of social theories for directed, signed networks, it also paves the way for a number of future research directions, such as (i) enlarging the number of datasets over which to repeat the analysis proposed here, (ii) inspecting higher-order signatures of balance (e.g., triadic motifs) by employing a benchmark constraining the signed reciprocity, and (iii) exploring the notion of balance at the mesoscopic level by extending the approach pursued in [6].

ACKNOWLEDGMENTS

D.G. and T.S. acknowledge support from the European Union - NextGenerationEU - National Recovery and Resilience Plan (Piano Nazionale di Ripresa e Resilienza, PNRR), M4C2 I.3.1. - project 'SoBigData.it - Strengthening the Italian RI for Social Mining and Big Data Analytics' - Grant IR0000013 (n. 3264) CUP B53C22001760006. D.G. also acknowledges support from the Dutch Econophysics Foundation (Stichting Econophysics, Leiden, the Netherlands) and the Netherlands Organization for Scientific Research (NWO/OCW). This work is also supported by the project "Reconstruction, Resilience and Recovery of Socio-Economic Networks" RECON-NET EP_FAIR_005,

PE0000013 'FAIR', PNRR M4C2 Investment 1.3, financed by the European Union – NextGenerationEU. R.L. acknowledges support from the EPSRC Grants No. EP/V013068/1 and No. EP/V03474X/1. We thank M. Szell for sharing the Pardus dataset employed for the present analysis.

DATA AVAILABILITY

The Honduras villages dataset is described in [69] and can be found at [73]. The Spanish schools dataset is described in [32] and is publicly available at [74]. The MMOG dataset, described in [33], is subject to proprietary restrictions and cannot be shared.

[1] F. Heider, Attitudes and cognitive organization, J. Psychol. **21**, 107 (1946).

[2] D. Cartwright and F. Harary, Structural balance: A generalization of Heider's theory, Psychol. Rev. **63**, 277 (1956).

[3] J. A. Davis, Clustering and structural balance in graphs, Hum. Relat. **20**, 181 (1967).

[4] P. Doreian and A. Mrvar, Partitioning signed social networks, Soc. Networks **31**, 1 (2009).

[5] V. Traag, P. Doreian, and A. Mrvar, Partitioning signed networks, in *Advances in Network Clustering and Blockmodeling* (Wiley, New York, 2019), pp. 225–249.

[6] A. Gallo, D. Garlaschelli, and T. Squartini, Assessing frustration in real-world signed networks: A statistical theory of balance, Phys. Rev. Res. **6**, L042065 (2024).

[7] V. A. Traag and J. Bruggeman, Community detection in networks with positive and negative links, Phys. Rev. E **80**, 036115 (2009).

[8] R. Singh and B. Adhikari, Measuring the balance of signed networks and its application to sign prediction, J. Stat. Mech.: Theory Exp. (2017) 063302.

[9] A. Kirkley, G. T. Cantwell, and M. E. J. Newman, Balance in signed networks, Phys. Rev. E **99**, 012320 (2019).

[10] E. Estrada and M. Benzi, Walk-based measure of balance in signed networks: Detecting lack of balance in social networks, Phys. Rev. E **90**, 042802 (2014).

[11] E. Estrada, Rethinking structural balance in signed social networks, Discrete Appl. Math. **268**, 70 (2019).

[12] A. Kargaran, M. Ebrahimi, M. Riazi, A. Hosseiny, and G. R. Jafari, Quartic balance theory: Global minimum with imbalanced triangles, Phys. Rev. E **102**, 012310 (2020).

[13] M. H. H. Siboni, A. Kargaran, and G. R. Jafari, Hybrid balance theory: Heider balance under higher-order interactions, Phys. Rev. E **105**, 054105 (2022).

[14] S. Talaga, M. Stella, T. J. Swanson, and A. S. Teixeira, Polarization and multiscale structural balance in signed networks, Commun. Phys. **6**, 349 (2023).

[15] S. Aref and M. C. Wilson, Balance and frustration in signed networks, J. Complex Netw. **7**, 163 (2019).

[16] S. Aref, L. Dinh, R. Rezapour, and J. Diesner, Multilevel structural evaluation of signed directed social networks based on balance theory, Sci. Rep. **10**, 1 (2020).

[17] S. Aref, A. J. Mason, and M. C. Wilson, A modeling and computational study of the frustration index in signed networks, Networks **75**, 95 (2020).

[18] G. Facchetti, G. Iacono, and C. Altafini, Computing global structural balance in large-scale signed social networks, Proc. Natl. Acad. Sci. USA **108**, 20953 (2011).

[19] H. Saiz, J. Gómez-Gardeñes, P. Nuche, A. Girón, Y. Pueyo, and C. L. Alados, Evidence of structural balance in spatial ecological networks, Ecography **40**, 733 (2017).

[20] T. Derr, C. Aggarwal, and J. Tang, Signed network modeling based on structural balance theory, in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management* (Association for Computing Machinery, New York, 2018), pp. 557–566.

[21] G. Huitsing, M. A. Van Duijn, T. A. Snijders, P. Wang, M. Sainio, C. Salmivalli, and R. Veenstra, Univariate and multivariate models of positive and negative networks: Liking, disliking, and bully–victim relationships, Soc. Netw. **34**, 645 (2012).

[22] J. Lerner, Structural balance in signed networks: Separating the probability to interact from the tendency to fight, Soc. Netw. **45**, 66 (2016).

[23] C. Becatti, G. Caldarelli, and F. Saracco, Entropy-based randomization of rating networks, Phys. Rev. E **99**, 022306 (2019).

[24] C. Fritz, M. Mehrl, P. W. Thurner, and G. Kauermann, Exponential random graph models for dynamic signed networks: An application to international relations, arXiv:2205.13411.

[25] B. Hao and I. A. Kovács, Proper network randomization is key to assessing social balance, Sci. Adv. **10**, eadj0104 (2024).

[26] A. Gallo, D. Garlaschelli, R. Lambiotte, F. Saracco, and T. Squartini, Testing structural balance theories in heterogeneous signed networks, Commun. Phys. **7**, 154 (2024).

[27] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins, Propagation of trust and distrust, in *Proceedings of the 13th International Conference on World Wide Web* (Association for Computing Machinery, New York, NY, USA, 2004), pp. 403–412.

[28] J. Leskovec, D. Huttenlocher, and J. Kleinberg, Signed networks in social media, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Association for Computing Machinery, New York, NY, USA, 2010), pp. 1361–1370.

[29] J. Leskovec, D. Huttenlocher, and J. Kleinberg, Predicting positive and negative links in online social networks, in *Proceedings of the 19th International Conference on World Wide Web* (Association for Computing Machinery, New York, NY, USA, 2010), pp. 641–650.

[30] K.-Y. Chiang, N. Natarajan, A. Tewari, and I. S. Dhillon, Exploiting longer cycles for link prediction in signed networks, in *Proceedings of the 20th ACM International Conference on Information and Knowledge Management* (Association for Computing Machinery, New York, 2011), pp. 1157–1162.

[31] D. Song and D. A. Meyer, Link sign prediction and ranking in signed directed social networks, Soc. Netw. Anal. Min. **5**, 52 (2015).

[32] M. Ruiz-García, J. Ozaita, M. Pereda, A. Alfonso, P. Brañas-Garza, J. A. Cuesta, and A. Sánchez, Triadic influence as a

proxy for compatibility in social relationships, Proc. Natl. Acad. Sci. USA **120**, e2215041120 (2023).
[33] M. Szell, R. Lambiotte, and S. Thurner, Multirelational organization of large-scale social networks in an online world, Proc. Natl. Acad. Sci. USA **107**, 13636 (2010).
[34] N. Girdhar and K. Bharadwaj, Social status computation for nodes of overlapping communities in directed signed social networks, in *Integrated Intelligent Computing, Communication and Security* (Springer, Singapore, 2019), pp. 49–57.
[35] J. Linczuk, P. J. Górski, B. K. Szymanski, and J. A. Hołyst, Multidimensional attributes expose Heider balance dynamics to measurements, Sci. Rep. **13**, 15568 (2023).
[36] L. Dinh, R. Rezapour, L. Jiang, and J. Diesner, Enhancing structural balance theory and measurement to analyze signed digraphs of real-world social networks, Front. Hum. Dyn. **4**, 1028393 (2023).
[37] P. W. Holland and S. Leinhardt, Local structure in social networks, Sociol. Methodol. **7**, 1 (1976).
[38] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Structural Analysis in the Social Sciences (Cambridge University Press, Cambridge, UK, 1994).
[39] D. Garlaschelli and M. I. Loffredo, Patterns of link reciprocity in directed networks, Phys. Rev. Lett. **93**, 268701 (2004).
[40] D. Garlaschelli and M. I. Loffredo, Multispecies grand-canonical models for networks with reciprocity, Phys. Rev. E **73**, 015101(R) (2006).
[41] L. A. Meyers, M. Newman, and B. Pourbohloul, Predicting epidemics on directed contact networks, J. Theor. Biol. **240**, 400 (2006).
[42] M. Boguñá and M. Á. Serrano, Generalized percolation in random directed networks, Phys. Rev. E **72**, 016106 (2005).
[43] N. Perra, V. Zlatić, A. Chessa, C. Conti, D. Donato, and G. Caldarelli, Pagerank equation and localization in the WWW, Europhys. Lett. **88**, 48002 (2009).
[44] V. Zlatić and H. Štefančić, Model of Wikipedia growth based on information exchange via reciprocal arcs, Europhys. Lett. **93**, 58005 (2011).
[45] D. Garlaschelli and M. I. Loffredo, Structure and evolution of the world trade network, Physica A **355**, 138 (2005).
[46] D. Garlaschelli, F. Ruzzenenti, and R. Basosi, Complex networks and symmetry I: A review, Symmetry **2**, 1683 (2010).
[47] G. Zamora-López, V. Zlatić, C. Zhou, H. Štefančić, and J. Kurths, Reciprocity of networks with degree correlations and arbitrary degree sequences, Phys. Rev. E **77**, 016106 (2008).
[48] V. Zlatić and H. Štefančić, Influence of reciprocal edges on degree distribution and degree correlations, Phys. Rev. E **80**, 016117 (2009).
[49] D. B. Stouffer, J. Camacho, W. Jiang, and L. A. Nunes Amaral, Evidence for the existence of a robust pattern of prey selection in food webs, Proc. R. Soc. B **274**, 1931 (2007).
[50] T. Squartini and D. Garlaschelli, Analytical maximum-likelihood method to detect patterns in real networks, New J. Phys. **13**, 083001 (2011).
[51] E. Evmenova and D. Gromov, Analysis of directed signed networks: Triangles inventory, in *Fifth Networks in the Global World Conference* (Springer, Berlin, 2020), pp. 120–132.
[52] E. Evmenova and D. Gromov, Structural analysis of directed signed networks, in *International Conference Dedicated to the Memory of Professor Vladimir Zubov* (Springer, Berlin, 2020), pp. 621–628.
[53] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Network motifs: Simple building blocks of complex networks, Science **298**, 824 (2002).
[54] T. Squartini and D. Garlaschelli, Triadic motifs and dyadic self-organization in the World Trade Network, in *International Workshop on Self-Organizing Systems* (Springer, Berlin, 2012), pp. 24–35.
[55] M. E. J. Newman, S. Forrest, and J. Balthrop, Email networks and the spread of computer viruses, Phys. Rev. E **66**, 035101(R) (2002).
[56] M. A. Serrano and M. Boguná, Topology of the world trade web, Phys. Rev. E **68**, 015101(R) (2003).
[57] T. Squartini, I. Van Lelyveld, and D. Garlaschelli, Early-warning signals of topological collapse in interbank networks, Sci. Rep. **3**, 1 (2013).
[58] F. Saracco, M. J. Straka, R. Di Clemente, A. Gabrielli, G. Caldarelli, and T. Squartini, Inferring monopartite projections of bipartite networks: An entropy-based approach, New J. Phys. **19**, 053022 (2017).
[59] C. Becatti, G. Caldarelli, R. Lambiotte, and F. Saracco, Extracting significant signal of news consumption from social networks: The case of Twitter in Italian political elections, Palgrave Commun. **5**, 91 (2019).
[60] G. Caldarelli, R. De Nicola, F. Del Vigna, M. Petrocchi, and F. Saracco, The role of bot squads in the political propaganda on Twitter, Commun. Phys. **3**, 81 (2020).
[61] Z. P. Neal, R. Domagalski, and B. Sagan, Comparing alternatives to the fixed degree sequence model for extracting the backbone of bipartite projections, Sci. Rep. **11**, 23929 (2021).
[62] F. Parisi, T. Squartini, and D. Garlaschelli, A faster horse on a safer trail: Generalized inference for the efficient reconstruction of weighted networks, New J. Phys. **22**, 053053 (2020).
[63] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.111.024312, for additional data and analysis, which includes Refs. [25,75–77].
[64] M. Shariff, M. R. Mutalikdesai, K. Malekafzali, S. N. Arayilakath, V. Sudhakaran, and M. Altaf, Novel measures for reciprocal behavior and equivalence in signed networks, in *Digital Libraries: Social Media and Community Networks: 15th International Conference on Asia-Pacific Digital Libraries, ICADL 2013, Bangalore, India* (Springer, Berlin, 2013), pp. 193–194.
[65] R. B. Zajonc and E. Burnstein, Structural balance, reciprocity, and positivity as sources of cognitive bias, J. Pers. **33**, 570 (1965).
[66] D. Garlaschelli and M. I. Loffredo, Maximum likelihood: Extracting unbiased information from complex networks, Phys. Rev. E **78**, 015101(R) (2008).
[67] https://it.mathworks.com/matlabcentral/fileexchange/167426-signed-models-for-network-analysis
[68] trellis.yale.edu
[69] A. Isakov, J. H. Fowler, E. M. Airoldi, and N. A. Christakis, The structure of negative social ties in rural village networks, Sociol. Sci. **6**, 197 (2019).

[70] G. Fagiolo, Directed or undirected? A new index to check for directionality of relations in socio-economic networks, Econ. Bull. **3**, 1 (2006).

[71] G. Fagiolo, Clustering in complex directed networks, Phys. Rev. E **76**, 026107 (2006).

[72] A. Ghasemian and N. A. Christakis, The structure and function of antagonistic ties in village social networks, Proc. Natl. Acad. Sci. USA **121**, e2401257121 (2024).

[73] https://github.com/NDS-VU/signed-network-datasets?tab=readme-ov-file#network-village

[74] https://zenodo.org/records/7647000#.Y-5eDtLMJH4

[75] J. Park and M. E. J. Newman, Statistical mechanics of networks, Phys. Rev. E **70**, 066117 (2004).

[76] T. Squartini and D. Garlaschelli, *Maximum-Entropy Networks: Pattern Detection, Network Reconstruction and Graph Combinatorics* (Springer International Publishing, Berlin, 2017), p. 116.

[77] N. Vallarano, M. Bruno, E. Marchese, G. Trapani, F. Saracco, G. Cimini, M. Zanon, and T. Squartini, Fast and scalable likelihood maximization for exponential random graph models with local constraints, Sci. Rep. **11**, 15227 (2021).