Testing structural balance theories in heterogeneous signed networks

Check for updates

Anna Gallo [1,2], Diego Garlaschelli[1,2,3], Renaud Lambiotte [4], Fabio Saracco[1,5,6] & Tiziano Squartini [1,2,7] ✉

The abundance of data about social relationships allows the human behavior to be analyzed as any other natural phenomenon. Here we focus on balance theory, stating that social actors tend to avoid establishing cycles with an odd number of negative links. This statement, however, can be supported only after a comparison with a benchmark. Since the existing ones disregard actors' heterogeneity, we extend Exponential Random Graphs to signed networks with both global and local constraints and employ them to assess the significance of empirical unbalanced patterns. We find that the nature of balance crucially depends on the null model: while homogeneous benchmarks favor the weak balance theory, according to which only triangles with one negative link should be under-represented, heterogeneous benchmarks favor the strong balance theory, according to which also triangles with all negative links should be under-represented. Biological networks, instead, display strong frustration under any benchmark, confirming that structural balance inherently characterizes social networks.

Network theory has emerged as a powerful framework in many disciplines to model different kinds of real-world systems, by representing their units as nodes and the interactions between them as links. In social science, the study of networks with signed edges has recently seen its popularity revived[1–4], because the signed character of links can be used to represent the positive as well as the negative social interactions that are currently identifiable in empirical data.

From a historical perspective, the interest towards signed networks is rooted into the psychological theory named *balance theory* (BT), firstly proposed by Heider[5]. The choice of adopting signed graphs to model it has, then, led Cartwright and Harary[6] to introduce its structural version (SBT), which has found application not only in the study of human relationships, but also in that of biological, ecological and economic systems[7–10].

BT deals with the concept of *balance*: a complete, signed graph is said to be balanced if all its triads have an even number of negative edges, i.e. either zero (in this case, the three edges are all positive) or two (see Fig. 1). Informally speaking, BT formalizes the principles 'the friend of my friend is my friend' and 'the enemy of my enemy is my friend'. 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 two, disjoint subsets whose intra-modular links are all positive and whose inter-modular links are all negative.

Cartwright and Harary extended the definition of balance to incomplete graphs[6] by including cycles of length larger than three: a (connected) network is said to be balanced when all cycles are positive, i.e. they contain an even number of negative edges. Taken together, the criteria above form the so-called *structural strong balance theory* (SSBT).

The framework of SSBT has been extended by Davis[11] by introducing the concept of *k-balanced* networks, according which signed graphs are balanced if their set of nodes can be partitioned into k disjoint subsets with positive intra-modular links and negative inter-modular links. This generalized definition of balance leads to the formulation of *structural weak balance theory* (SWBT), according to which triads with all negative edges are balanced, since each of their nodes can be thought of as a group on its own if necessary (see Fig. 1).

Several metrics to decide whether signed networks are strongly or weakly balanced have been proposed. For instance, the level of balance of a signed network has been quantified as the number of edges that need be removed, or whose sign need be reversed, in order to obtain a network where each cycle has an even number of negative links[12,13]. Alternatively, it has been defined as the number of balanced, closed walks (i.e. closed walks with an even number of negative links) that are present in the network[14–17]. In[18] an incomplete, signed network is considered balanced if it is possible to fill in all

[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]Lorentz Institute for Theoretical Physics, University of Leiden, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands. [4]Mathematical Institute, University of Oxford, Woodstock Road, Oxford OX2 6GG, UK. [5]'Enrico Fermi' Research Center (CREF), Via Panisperna 89A, 00184 Rome, Italy. [6]Institute for Applied Computing 'Mauro Picone' (IAC), National Research Council, Via dei Taurini 19, 00185 Rome, Italy. [7]Institute for Advanced Study, University of Amsterdam, Oude Turfmarkt 145, 1012 GC Amsterdam, The Netherlands. ✉e-mail: tiziano.squartini@imtlucca.it
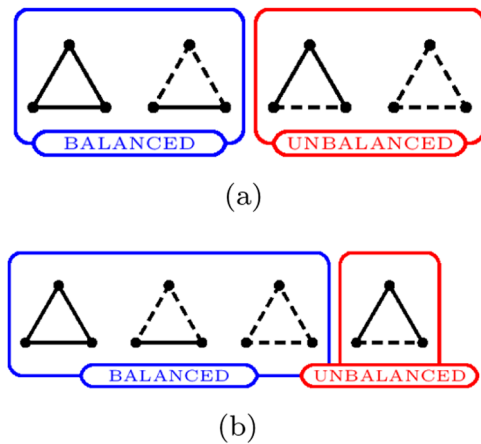
Fig. 1 | Balanced and unbalanced motifs. Fundamental triadic patterns, or motifs, considered as balanced (blue) and unbalanced (red) by the strong (**a**) and weak (**b**) versions of the balance theory.

its missing links to obtain a complete, balanced graph according to SSBT. In[19] the authors define three different levels of balance: at the micro-scale, involving triads; at the meso-scale, involving larger subgraphs; at the macro-scale, involving the entire network. Still, as firstly noticed in[6], '*it may happen that only cycles of length 3 and 4 are important for the purpose of determining balance*'; this is further stressed in[20], where it can be read that '*this intuition has been later justified empirically by demonstrating that it is easier for people to memorize the valences of ties in shorter cycles*', and confirmed in[21], where it is noticed that '*analyses based on counting simple cycles demonstrated that real networks often have a relatively low cycle length threshold after which the degree of balance measures quickly decrease*'.

Other approaches have been adopted in[22–24], where the problem is studied from a spectral perspective, and in[25], where the problem is studied by employing concepts borrowed from statistical physics (each signed triad is assigned an energy and the networks at the 'lowest temperature' have triangles without negative edges).

Other authors, instead, have focused on the complementary notion of *frustration*, trying to quantify the extent to which signed networks are far from balanced[19,26–28]. In[26], the authors define the so-called *balanced decomposition number*, i.e. the (minimum) number of balanced groups into which nodes can be partitioned, and evaluate it by counting the (minimum) number of edges whose removal increases a network balance. In[29], instead, the same index is evaluated by adopting the so-called *switching signs method* introduced in[30] and prescribing to count the (minimum) number of signs that must be reversed to balance a network. In[22], the level of (im)balance of a network is proxied by the magnitude of the smallest eigenvalue of the Laplacian matrix.

Empirical observations seem to point out that real-world, signed networks tend to be *k*-balanced, i.e. to avoid establishing the patterns that are considered as frustrated by SWBT: as an example, in[24] the authors study a pair of online, social networks induced by the relationships between users, showing that balance increases as the number of clusters into which nodes are partitioned is larger than two. In[17], the authors notice that the weak formulation of SBT allows a better performance in predicting signs to be achieved.

In the present paper, we approach the concept of balance (or frustration) from a statistical perspective, comparing the empirical value of a chosen metric with the outcome of a properly defined benchmark model, i.e. a reference model preserving some of the network properties while randomizing the rest. The most common null model for signed graphs is perhaps the one obtained by keeping the positions of edges fixed while shuffling their signs[2,17]. Reference[31] implements what we may call (for reasons that will be clear later) the canonical variant of the aforementioned exercise, assigning signs by means of a Bernoulli distribution. Reference[15]

introduces a null model for randomizing both the presence and the sign of links. In[10], the signed version of the Local Rewiring Algorithm is implemented (at each step, two edges with the same sign are selected and rewired, to preserve the total number of signed links incident to each node). The canonical variant of this model is implemented in[32], where the Balanced Signed Chung-Lu model (BSCL) is proposed (although it additionally constrains also the average number of signed triangles each edge is part of). Finally, refs. [33–36] define models constraining the structural properties of signed networks within the framework of Exponential Random Graphs (ERG).

Our contribution here focuses on binary, undirected signed networks and is motivated by two key considerations. First, real-world social networks have different levels of sparsity and we therefore aim at extending the ERG framework to include null models suitable for the analysis of signed graphs with plus (positive), minus (negative) and additionally zero (missing) edges. Second, as in the analysis of most other networks, we recognize the importance of preserving the inherent heterogeneity of different nodes and we therefore define new null models that can constrain the number of plus, minus and zero edges of each node separately. As we shall see, controlling for the different tendencies of actors of establishing friendly and unfriendly relationships can change the estimated statistical significance of balance quite dramatically. After defining a suite of such null models, we will use them to inspect the statistical significance of the most commonly studied (un)balanced patterns at both local and global levels, i.e. signed triangles and signed communities.

Results
Datasets description
We now employ the benchmarks introduced and discussed in "Materials and Methods" and summed up in Table 1 to analyze various real-world networks. Although most of them represent social relationships, we have also considered biological data as a comparison to check for specific patterns characterizing social structures.

The first dataset is the *Correlates of Wars* (CoW) dataset[37]. It provides a picture of the international political relationships over the years 1946–1997 and consists of 13 snapshots of 4 years each. A positive edge between any two countries indicates an alliance, a political agreement or the membership to the same governmental organization. Conversely, a negative edge indicates that the two countries are enemies, have a political disagreement or are part of different, governmental organizations.

The second dataset collects information about the relationships among the $\simeq$ 300.000 players of a massive multiplayer online game (MMOG)[38]. A positive edge between two players indicates a friendship, an alliance, or an economic relation. Conversely, a negative edge indicates the existence of an enmity, a conflict, or a fight. Since the network is directed, we have made it undirected by applying the following rules: if any two agents have the same opinion about the other, the undirected connection preserve the sign (i. e. $+1 \cdot +1 = +1$ and $-1 \cdot -1 = -1$); if any two agents have opposite opinions, we assume their undirected connection to have a negative sign (i.e. $+1 \cdot -1 = -1 \cdot +1 = -1$). Furthermore, in order to preserve the total number of nodes, we treat non-reciprocal connections as reciprocal, by preserving the original sign (i.e. $+1 \cdot 0 = 0 \cdot +1 = +1$ and $-1 \cdot 0 = 0 \cdot -1 = -1$).

The remaining datasets we consider are those collected in[39] and analyzed in[27]. These include three socio-political networks (SPNs): *N.G.H. Tribes*, *Senate US*, *Monastery*; two financial networks (FNs): *Bitcoin Alpha* and *Bitcoin OTC*; and three gene-regulatory networks (GRNs): *E. Coli*, *Macrophage*, *Epidermal Growth Factor Receptor*.

In the SPNs, N.G.H. Tribes collects data about New Guinean Highland Tribes (here, a positive/negative link denotes alliance/rivalry), Monastery corresponds to the last frame of Sampson's data about the relationships between novices in a monastery[40] (here, a positive/negative link indicates a positive/negative interaction), and Senate US collects data about the members of the 108th US Senate Congress (here, a positive/negative link indicates trust/distrust or similar/dissimilar political opinions).

Table 1 | Descriptive summary of signed benchmarks

Null model	Topology: free	Topology: fixed
Homogenous	SRGM: each pair of nodes is assigned a plus, a minus or a zero edge with a probability that is pair-independent; all nodes are statistically equivalent. Differently from the recipe adopted in[15,49], the parameters defining our SRGM can be unambiguously tuned to reproduce the empirical number of plus and minus edges of any (binary, undirected, signed) network.	SRGM-FT: the topology is the same as in the real network and the connected pairs of nodes are assigned either a plus one or a minus one, with a probability that is pair-independent. Differently from the recipe adopted in[2,31], the parameters defining our SRGM-FT can be unambiguously tuned to reproduce the empirical number of plus and minus edges of any (binary, undirected, signed) network. The SRGM-FT is the conditional version of the SRGM.
Heterogenous	SCM: each pair of nodes is assigned a plus, a minus or a zero edge, with a probability that is pair-dependent and determined by the different tendencies of nodes to establish positive and negative interactions. This model represents the canonical variant of the one employed in[10].	SCM-FT: the topology is the same as in the real network and the connected pairs of nodes are assigned either a plus one or a minus one, with a probability that is pair-dependent and determined by the different tendencies of nodes to establish positive and negative interactions. The SCM-FT is the conditional version of the SCM.

Table summarizing the properties of the four, signed null models introduced in this article, i.e. the Signed Random Graph Model (SRGM), the Signed Random Graph Model with Fixed Topology (SRGM-FT), the Signed Configuration Model (SCM), the Signed Configuration Model with Fixed Topology (SCM-FT).

The FNs are 'who-trust-whom' networks of Bitcoin traders on an online platform: a positive/negative link indicates trust/distrust between users[41]. The networks representing the FNs are weighted, directed ones: hence, after having binarized them by replacing each positive (negative) weight with a $+1$ (-1), we have made them undirected by applying the same rules adopted for the MMOG dataset.

Lastly, in the GRNs each node represents a gene, with positive links indicating activating connections and negative links indicating inhibiting connections. Specifically, *E. Coli* collects data about a transcriptional network of the bacterium *Escherichia Coli*; *Macrophage* collects data about a blood cell that eliminates substances such as cancer cells, cellular debris and microbes; *Epidermal Growth Factor Receptor* collects data about the protein that is responsible for cell division and survival in epidermal tissues.

The vast majority of the networks considered here is characterized by a small link density $c = 2L/N(N-1)$ but a large fraction L^+/L of positive links. The density of the CoW network decreases over time from $\simeq 0.2$ to $\simeq 0.1$ and the percentage of positive links is roughly stationary around $\simeq 88\%$; on the other hand, the link density of the MMOG network is stationary around 0.003 and the percentage of positive links decreases from $\simeq 98\%$ to $\simeq 60\%$. The SPNs have the largest values of link density among the configurations in our basket, ranging from $\simeq 0.3$ to $\simeq 0.5$, and percentages of positive links ranging from $\simeq 50\%$ to $\simeq 75\%$. *Bitcoin Alpha* has a link density of $\simeq 0.002$ and a percentage of positive links of $\simeq 90\%$, while *Bitcoin OTC* has a link density of $\simeq 0.001$ and a percentage of positive links of $\simeq 85\%$. Lastly, the GRNs have a link density ranging from $\simeq 10^{-3}$ to $\simeq 10^{-2}$ and a percentage of positive links ranging from $\simeq 58\%$ to $\simeq 66\%$. For more details on the basic descriptive statistics of the networks considered in the present work, see the Supplementary Note 4.

Assessing balance

In order to test the validity of the two formulations of SBT, at the local level, we need to compare the empirical abundance of the triadic motifs defined in the Methods section with the corresponding expected values calculated under the null models we have introduced. To this aim, a very useful indicator is represented by the z-score $z_m = [N_m(\mathbf{A}^*) - \langle N_m \rangle]/\sigma[N_m]$, where $N_m(\mathbf{A}^*)$ is the number of occurrences of pattern m in the real network $\mathbf{A}^*$, $\langle N_m \rangle$ is the expected occurrence of the same pattern 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. z_m quantifies the number of standard deviations by which the empirical abundance of pattern m differs from the expected one. For instance, after checking for the Gaussianity of N_m under the null model (since it is a sum of dependent random variables, this is ensured by the generalization of the Central Limit Theorem - see the Supplementary Note 6), a result $|z_m| \leq 2$ ($|z_m| \leq 3$) indicates that the empirical abundance of pattern m is

compatible with the one expected under the chosen null model 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 null model at those significance levels. In the latter case, a value $z_m > 0$ ($z_m < 0$) indicates the tendency of the pattern to be over- (under-)represented in the data with respect to the null model.

z-scores can be evaluated either analytically or numerically: implementing the first alternative requires employing the formulas provided in the Supplementary Note 6; implementing the second alternative requires numerically sampling the ensembles of graphs defined by our null models. Since the entries of the adjacency matrix are independent random variables, the unbiased generation of a random matrix $\mathbf{A} \in \mathbb{A}$ can be carried out by drawing a real number $u_{ij} \in U[0, 1]$ and posing: for models with varying topology, $a_{ij} = -1$ if $0 \leq u_{ij} \leq p_{ij}^-$, $a_{ij} = +1$ if $p_{ij}^- < u_{ij} < p_{ij}^- + p_{ij}^+$ and $a_{ij} = 0$ if $p_{ij}^- + p_{ij}^+ \leq u_{ij} \leq 1$, for all pairs $i < j$; for models with fixed topology, $a_{ij} = -1$ if $0 \leq u_{ij} \leq p_{ij}^-$ and $a_{ij} = +1$ if $p_{ij}^- < u_{ij} \leq 1$, for all pairs $i < j$ such that $|a_{ij}^*| = 1$ (see the Supplementary Note 5 for an estimation of the time required to sample the ensemble induced by each of our models, for each of our datasets).

Testing structural balance at the microscopic scale

We report our results starting from the network datasets that have several temporal snapshots (CoW and MMOG). Figure 2 shows the temporal trends of the z-scores for the two networks under the homogeneous null models (SRGM and SRGM-FT). Panels **(a)** − **(c)** refer to the SRGM and show that the z-scores for all triangles, irrespective of their signs, are strong and positive. This means that all triangles are over-represented in the data, with respect to a null model that completely randomizes the topology. This result is not unexpected, as it merely indicates that, given the empirical density of links, it is very unlikely to form triangles completely by chance. These results simply tell us that the SRGM is uninformative about the (im)balance in the data, as it is entirely biased by a purely topological effect. This conclusion is in line with the results in[17], which suggested that the SRGM-FT is to be preferred over the SRGM as it provides a better explanation of empirical network structures.

By contrast, the results generated under the SRGM-FT clearly support SWBT (see panels (b)–(d)). Indeed, the only significantly over-represented pattern in the data is precisely the only one that SWBT considers frustrated (the triangle with a single negative link), whereas the empirical abundance of the triangle with all negative edges (which SSBT would predict to be over-represented as well) remains largely compatible with the null model. Notice that also the empirical abundance of the balanced triangle with two negative edges is close to the one expected under the SRGM-FT, although its z-score is typically smaller than the z-score of the all-negative triangle. In any case, the abundance of the balanced triangle with three positive edges is
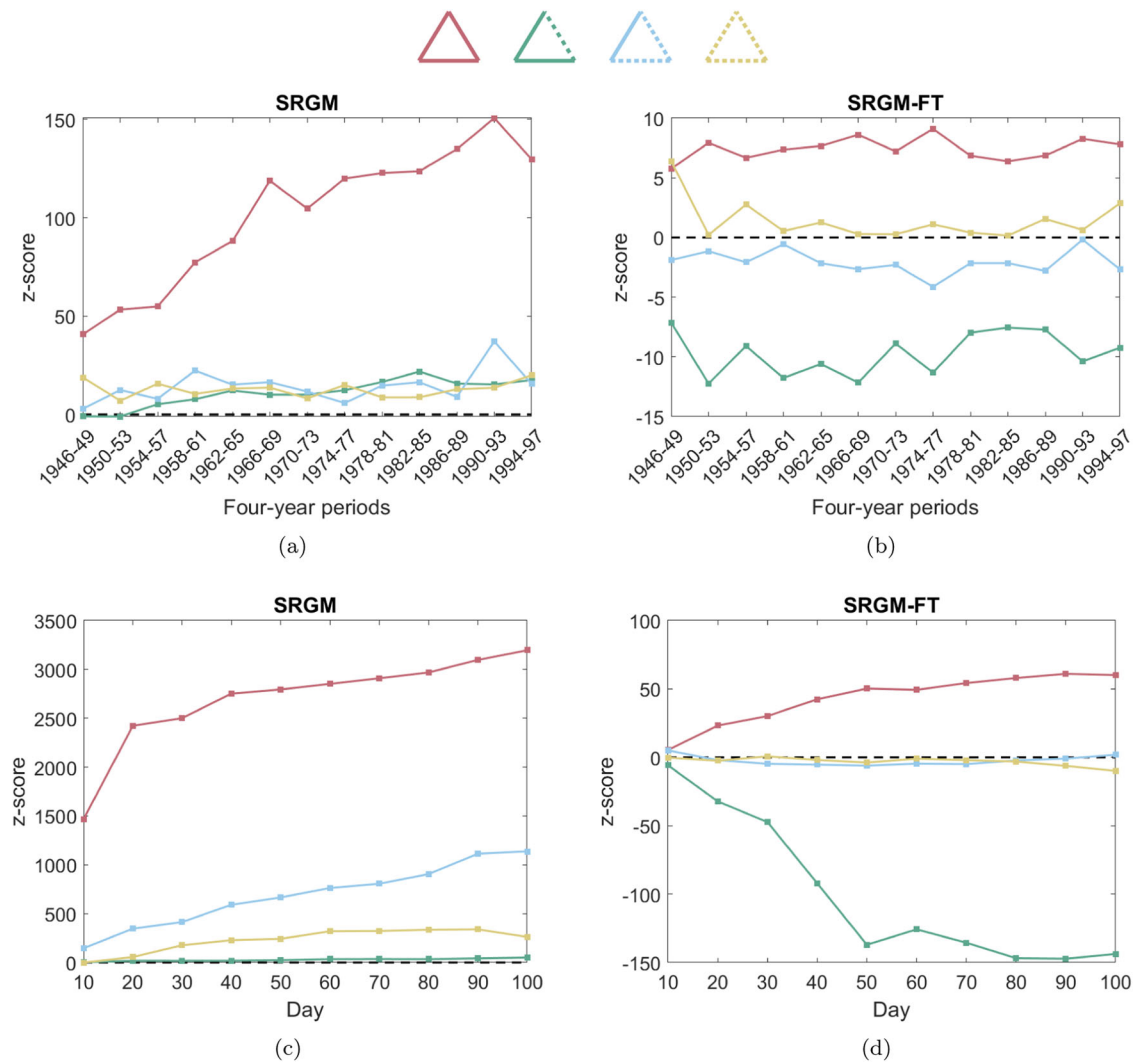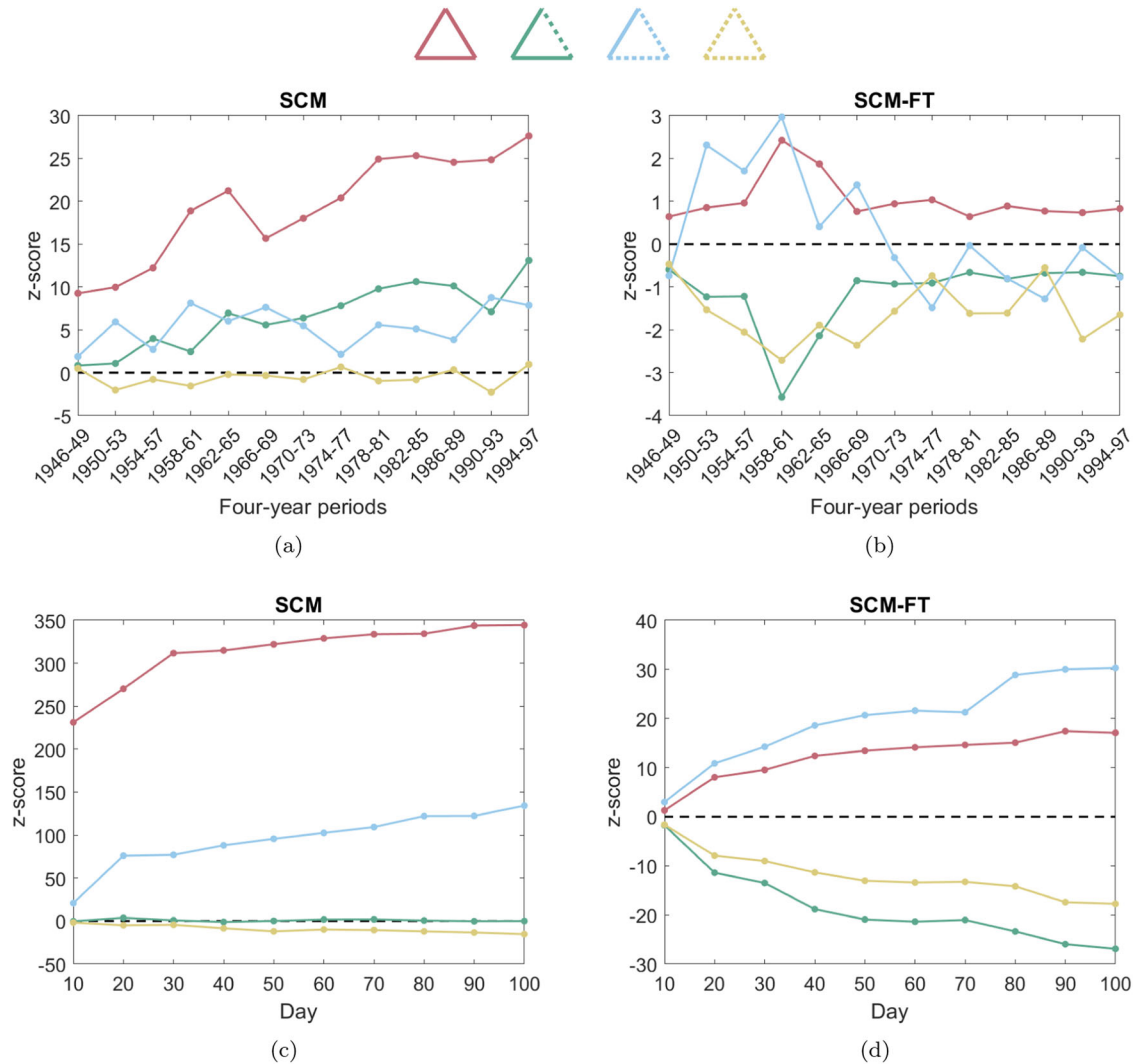
Fig. 2 | Structural (im)balance in the Correlates of Wars and MMOG datasets under homogeneous benchmarks. Structural (im)balance in the CoW and MMOG datasets: evolution of the *z*-scores of signed triangles under homogeneous benchmarks, i.e. the Signed Random Graph Model (SRGM) and the Signed Random Graph Model with Fixed Topology (SRGM-FT). **a, b** 13 snapshots (of 4 years each) of the CoW dataset, covering the period 1946–1997. **c, d** 10 snapshots of the MMOG dataset. **b–d** The SRGM-FT supports the structural weak balance theory (SWBT) because the only significantly under-represented pattern in the data is also the only one that SWBT considers frustrated (triangle with only one negative link), while the *z*-score of the triangle with all negative edges (which the structural strong balance theory would expect to be under-represented as well) is very low. In any case, the hypothesis that nodes tend to establish balanced triangles with all positive links is supported on both datasets. Results of this type constitute the backbone of the narrative according to which the weak version of the structural balance theory (SBT) is the one that is better supported by data. **a–c** Note that the SRGM has all *z*-scores positive, thereby not supporting any version of SBT, a result due to the complete randomization of the topology along with the edge signs: the over-representation of all patterns in the data is merely due to the fact that triangles form with small probability at a purely topological level, given the low link density, irrespective of their signs.

significantly over-represented on both datasets. This type of results constitute the backbone of the narrative according to which the weak version of SBT is the one that is better supported by data[2,17].

.However, since both the SRGM and the SRGM-FT do not constrain the local (node-specific) signed properties (i.e. the signed degrees of nodes), they cannot disentangle the effects of node heterogeneity from the revealed overall structural (im)balance. For this reason, In Fig. 3 we repeat the analysis of the CoW and MMOG datasets using the SCM and SCM-FT null models. As expected, the resulting *z*-scores are much smaller in absolute value, showing that node heterogeneity in the real networks is in general strong and is responsible for a significant part of the overall measured (im) balance. Therefore, controlling for the local signed degrees is a way to filter out the effects of node heterogeneity in the statistical analysis of structural balance. In general, we see that the triangle will all negative links has now negative *z*-scores in both datasets, under both null models. Similarly, the all-positive triangle remains with positive *z*-scores in all cases. The level of statistical significance (i.e. the absolute value of the *z*-score) is however quite different in the various cases: in general we see an overwhelming over-representation of the two balanced triangles (the all-positive one and the one with two negative links) in the MMOG data under both null models, while for the CoW data the only clearly significant patter is the over-representation of the all-positive triangle in the SCM. Nicely, the SCM-FT gives always negative *z*-scores to both the frustrated triangles (the all-negative one and the one with only one negative link), and most of the time positive *z*-scores to the two balanced triangles. Although the statistical evidence is much stronger for the MMOG data, this result indicates that, if any, the version of SBT supported by the data is SSBT, rather than SWBT. Therefore, as soon as the heterogeneity of the signed degrees of nodes is accounted for, SWBT loses its statistical support, and SSBT is favored by the data.

Fig. 3 | Structural (im)balance in the Correlates of Wars and MMOG datasets under heterogeneous benchmarks. Structural (im)balance in the CoW and MMOG datasets: evolution of the z-scores of signed triangles under heterogeneous benchmarks, i.e. the Signed Configuration Model (SCM) and the Signed Configuration Model with Fixed Topology (SCM-FT). **a, b** 13 snapshots (of 4 years each) of the CoW dataset, covering the period 1946–1997. **c, d** 10 snapshots of the MMOG dataset. The z-scores produced by the SCM (**a**)–(**c**) and the SCM-FT (**b**)–(**d**) are much smaller, in absolute value, than the corresponding ones produced by the Signed Random Graph Model and the Signed Random Graph Model with Fixed Topology (see Fig. 2), showing that node heterogeneity contributes significantly to the overall abundance of signed triangles. The all-positive (balanced) triangle is still strongly over-represented in all cases, but additionally the all-negative (frustrated) triangle is now always under-represented. Under the SCM-FT, the other frustrated triangle (the one with a single negative link) is also systematically under-represented, and these combined results provide support for the structural strong balance theory (SSBT) (particularly evidently for the MMOG data). By contrast, the structural weak balance theory (SWBT) (according to which one would expect the under-representation of only the triangle with a single negative link) is no longer supported. These results provide an alternative narrative w.r.t. the usual one: when the heterogeneity of the signed degrees of nodes is accounted for, statistical evidence supports SSBT rather than SWBT.

We now move to the results obtained on datasets which include other social networks as well as various biological networks, providing a different real-world benchmark where structural balance theory is not expected to apply. From Fig. 4 we confirm that the SRGM is completely uninformative about structural (im)balance, as it produces z-scores that are typically positive and very large for all triangles (balanced and unbalanced) and all networks (social and biological). This result simply means that the formation of any triangle, irrespective of its defining signs, is highly unlikely if the topology is completely randomized. By contrast, under the SRGM-FT the only pattern that is under-represented in social network data is the frustrated triangle with a single negative link, a result that largely supports SWBT (on biological data, this pattern is instead either not significant or over-represented). Heterogeneous null models (SCM and SCM-FT), instead, assign positive z-scores to the two balanced triangles (all-positive and with two negative links) and negative z-scores to the two frustrated triangles (all-negative and with one negative link), thereby systematically supporting SSBT. When used on biological networks, they instead highlight a strong tendency towards imbalance, as they tend to assign opposite signs (w.r.t. social networks) to most z-scores. These results confirm and extend what discussed above for the CoW and MMOG datasets, and additionally show that biological networks behave very differently from social networks, somehow favouring frustration. This is an indication that structural balance is indeed an inherent property of social networks.

As further evidence supporting the above conclusion, in Fig. 5 we show, for all networks and under all null models, the z-scores of the frustration indices SDoF and WDoF defined in Eqs. (9) and (10)

(a)

(b)

(c)

(d)
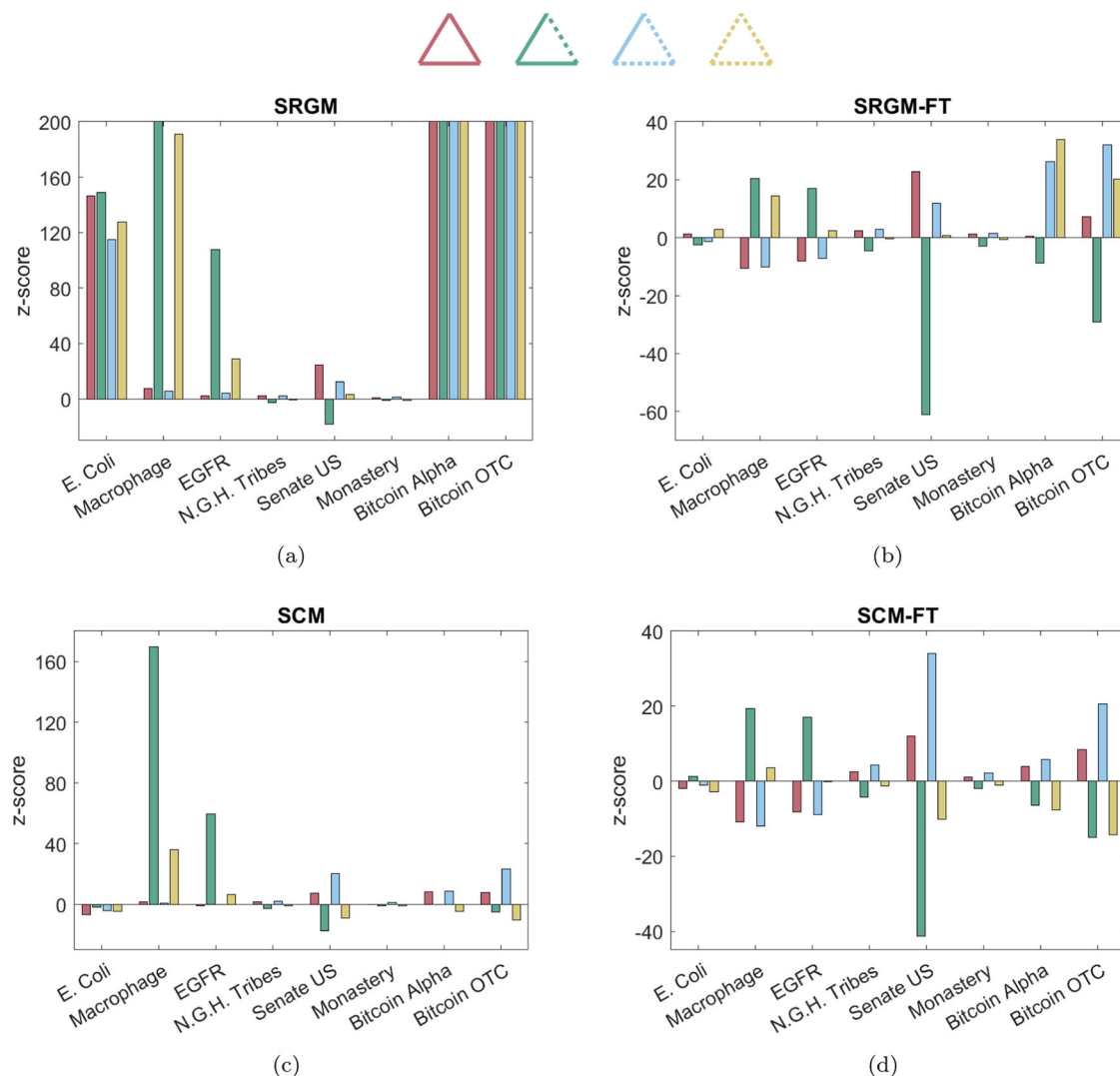
Fig. 4 | Structural (im)balance in social and biological networks under homogeneous and heterogeneous benchmarks. Structural (im)balance in social and biological networks under homogeneous (**a**, **b**) and heterogeneous (**c**, **d**) null models: z-scores of signed triangles for three, socio-political networks (N.G.H. Tribes, Senate US, Monastery), two, financial networks (Bitcoin Alpha, Bitcoin OTC) and, as a comparison, three, biological networks (E. Coli, Macrophage, EGFR). The Signed Random Graph Model (SRGM) produces z-scores that are almost always positive and very large for all triangles (balanced and unbalanced) and all signed networks (social and biological), a result confirming that this null model is completely uninformative about structural (im)balance, as it merely highlights that the formation of any triangle, irrespective of its signs, is highly unlikely if the topology is randomized completely. By contrast, the Signed Random Graph Model with Fixed Topology (SRGM-FT) largely supports the structural weak balance theory on social networks, as the only pattern under-represented in the data is the frustrated triangle with a single negative link. Heterogeneous null models, i.e. the Signed Configuration Model (SCM) and the Signed Configuration Model with Fixed Topology (SCM-FT), instead, systematically support the structural strong balance theory because they assign positive z-scores to the two balanced triangles (all-positive and with two negative links) and negative z-scores to the two frustrated triangles (all-negative and with one negative link). Additionally, in biological networks they tend to assign opposite signs (w.r.t. social networks) to most z-scores, highlight a strong tendency towards imbalance. These results are fully in line with what already observed with the Correlates of Wars and MMOG datasets in Figs. 2, 3.

respectively. Note that, while the raw values of SDoF and WDoF would not discount the effects of the imposed structural constraints on the raw values of frustration, the z-scores measure the level of statistical significance of the 'residual' frustration, after the structural constraints are accounted for. We see that, under all null models, the z-scores (when significant) are always negative for the social networks (signaling under-representation of the frustration indices in the data) and always positive for the biological networks (signaling over-representation of frustration in the data). Moreover, for the models with fixed topology, the z-scores for the heterogeneous null model (SCM-FT) are systematically smaller (in absolute value) than the ones for the corresponding homogeneous model (SRG-FT), indicating that, compared with the latter, the former model 'explains more' of the level of empirical frustration observed in the data. The same relation does not apply systematically between the models with varying topology (SCM and SRG), suggesting that models with fixed topology lead to more robust conclusions, as already observed in terms of their support for SWBT or SSBT.

Testing structural balance at the mesoscopic scale

Motivated by the last observation, we now use the null models with fixed topology to probe the patterns of structural (im)balance at a larger, mesoscopic level, i.e. as portrayed by the community structure deriving from optimally partitioning the nodes into communities with
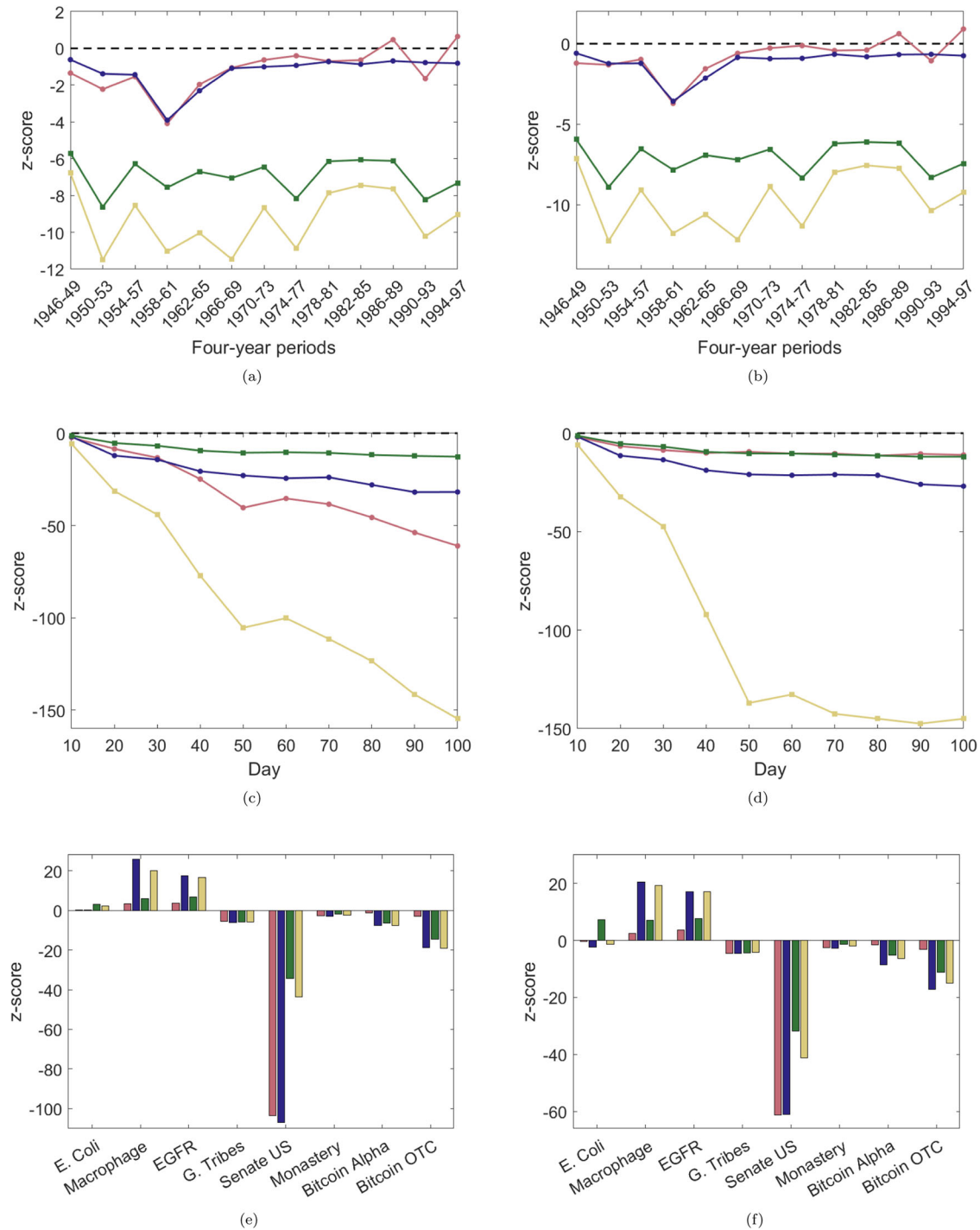
Fig. 5 | Analysis of the z-scores of the degree of frustration indices. Analysis of the z-scores of the strong degree of frustration defined in Eq. (9) and the weak degree of frustration defined in Eq. (10). **a, b** 13 snapshots of 4 years each of the Correlates of Wars dataset (covering the period 1946–1997). **c, d** 10 snapshots of the MMOG dataset. **e, f** Set of social and biological networks. z-scores are computed under the Signed Random Graph Model (SRGM) (•), the Signed Random Graph Model with Fixed Topology (SRGM-FT) (•), the Signed Configuration Model (SCM) (•) and the Signed Configuration Model with Fixed Topology (SCM-FT) (•). We see that, with respect to all null models, frustration is under-represented in all social network data and over-represented in all biological data.

positive internal links and negative external ones. As anticipated, SSBT predicts that the overall level of intra-community frustration, as measured by the FI defined in Eq. (11), should be observed after optimally partitioning the nodes into two communities, dominated by positive signs internally and negative signs across. By contrast, SWBT allows for potentially any number of communities, because it bases the

idea of balance precisely at the level of communities, so that all-negative triangles (and in principle all-negative cycles of any length) can be explained by placing the constituent nodes across distinct communities. To extract information about the signed community structure from our data, given a null model $\langle a_{ij} \rangle$ for the signed adjacency matrix entry a_{ij}, we look for the partition that maximizes the
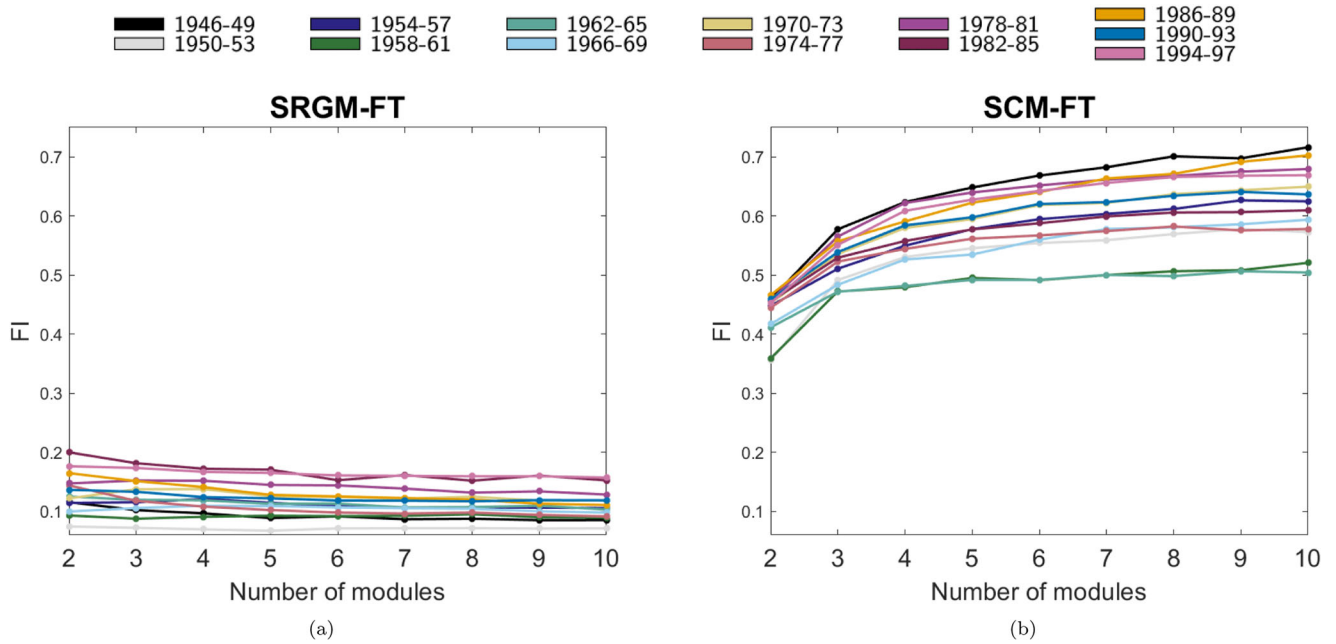
Fig. 6 | Values of the frustration index on several, optimal partitions of the Correlates of Wars dataset. Value of the frustration index (FI) on several, optimal partitions of the 13 snapshots (of 4 years each) of the CoW dataset, each obtained by maximizing the modularity $Q = -L \cdot (\text{FI} - \langle \text{FI} \rangle)$ for a given number K of modules (communities), using as null models the Signed Random Graph Model with Fixed Topology (SRGM-FT) (**a**) and the Signed Configuration Model with Fixed Topology (SCM-FT) (**b**). While the SRGM-FT reveals a rather flat profile of FI as a function of K, with the minimum obtained for a number of groups which is larger than two, the SCM-FT reveals that FI is always clearly minimized for a number of groups $K = 2$. Taken together, these results extend our findings at the mesoscale level.

signed modularity, as defined in[42]:

$$
\begin{aligned}
Q &= \sum_{i=1}^{N} \sum_{j>i} \left[a_{ij}^* - \langle a_{ij} \rangle \right] \delta_{c_i c_j} \\
&= \sum_{i=1}^{N} \sum_{j>i} \left[\left(a_{ij}^+ \right)^* - \left(a_{ij}^- \right)^* - \left(p_{ij}^+ - p_{ij}^- \right) \right] \delta_{c_i c_j} \\
&= L_\bullet^+ - L_\bullet^- - \langle L_\bullet^+ - L_\bullet^- \rangle \\
&= - \left[\left(L_o^+ + L_\bullet^- \right) - \langle L_o^+ + L_\bullet^- \rangle \right],
\end{aligned}
\tag{1}
$$

where $\bullet$ indicates quantities inside and $\circ$ outside the communities (note that $L_\bullet^+ = L^+ - L_o^+$ and that the total number of positive links is preserved under any null model considered here). For null models with fixed topology, a stronger result holds true, i.e. $Q = -L \cdot (\text{FI} - \langle \text{FI} \rangle)$ so that, since $L > 0$, maximizing Q becomes equivalent to minimizing the difference between FI and its expected value (see the Supplementary Note 7). The minimization of FI is another popular approach to finding the optimal partition[43] which, however, neglects the information embodied in a null model. Here, we consider a varying number $K = 2 \dots 10$ of communities and, for each value of K, look for the partition maximizing Q, using as null model both the SRGM-FT and the SCM-FT. We then compute the value of FI as a function of K, as plotted in Fig. 6 for the CoW dataset. We find that the trends produced under the SRGM-FT are quite flat, and in no case the minimum of FI is achieved by $K = 2$. This result is in line with SWBT, under whose assumptions there is no specific characteristic number of communities that would characterize real networks. By contrast, the SCM-FT produces clearly increasing trends, all starting from a minimum of FI at $K = 2$. This result strongly supports SSBT, according to which structural balance can be achieved by placing negative links between two communities, and positive links inside them. Taken together, these results extend our finding that SWBT (SSBT) is supported by homogeneous (heterogeneous) null models.

Discussion

Motivated by the widespread observation that actors in real social networks are characterized by a strong heterogeneity (typically signaled by broad distributions of node-specific topological properties), we have introduced a class of null models for signed networks characterized by either global or local constraints and with either fixed or varying topology. Our formalism provides the equivalent of various important ERGs to the domain of signed graphs. We have used our null models to address the problem of structural balance in real social networks. Our results show that the nature (weak or strong) and statistical strength of evidence of structural balance strongly depends on the null model adopted. In particular, we have shown that the occurrences of signed triangles favor SWBT when a homogeneous, global null model is considered. By contrast, SSBT is favoured by heterogeneous models with local constraints.

Generally speaking, adopting fixed-topology benchmarks seems to enhance the detection of frustration with the corresponding, homogeneous (heterogeneous) variant favouring SWBT (SSBT). As a possible behavioral explanation, we may advance the following one. Social agents are characterized by a certain level of tolerance. Such a level can be set by choosing the proper benchmark: null models constraining global quantities assume agents to be characterized on average by the same expected level of tolerance; by contrast, null models constraining local quantities account for the different levels of tolerance characterizing different agents. Let us imagine that relationships were established according to a random mechanism that preserves the total number of friends and enemies: should this be the case, our results indicate that equally tolerant agents would establish many more $(+,+,-)$ motifs than observed; instead, real-world agents are found to avoid engaging in relationships that lead to the formation of the $(+,+,-)$ pattern. Let us, now, refine the aforementioned picture and imagine that relationships were established according to a random mechanism that preserves the local number of friends and enemies: in this case, diversely tolerant agents would establish many more $(+,+,-)$ and $(-,-,-)$ motifs than observed; instead, real-world agents are found to avoid engaging in relationships that lead to the formation of both the $(+,+,-)$ and the

$(-, -, -)$ patterns. Overall, then, agents that cannot choose with whom to interact, but only how, adopt a behavior strongly avoiding engagement in frustrated relationships.

The same results have been extended to the mesoscale structural level, by finding that the optimal number of communities minimizing the overall level of frustration is $K = 2$ with respect to a heterogeneous null model (strongly supporting SSBT), while there is no characteristic optimal number with respect to a homogeneous null model (in line with SWBT). Importantly, we have considered a set of biological networks as a benchmark of real-world systems for which structural balance theory is not expected to apply. We have found a strong level of frustration in biological systems, indicating that structural balance (in either strong or weak form) indeed characterizes social networks.

Future directions along which the present analysis could be extended concern the possibility of defining ERGs for directed, as well as weighted, signed networks - the main technical difficulty lying in the proper definition of (binary, directed; weighted, both undirected and directed) constraints. The most natural application of such a formalism would be represented by the statistical validation of the so-called *status theory*, describing social interactions when hierarchies play a role[2].

Methods
Formalism and basic quantities

A *signed* graph is a graph where each edge can be *positive*, *negative* or *missing*. In what follows, we will focus on binary, undirected, signed networks: hence, each edge will be 'plus one', 'minus one' or 'zero'. More formally, for any two nodes i and j, the corresponding entry of the signed adjacency matrix $\mathbf{A}$ will be assumed to be $a_{ij} = -1, 0, +1$ (with $a_{ij} = a_{ji}, \forall\ i < j$). Since the total number of node pairs is $\frac{N(N-1)}{2} = \binom{N}{2}$ and any node pair can be positively connected, negatively connected or disconnected, the total number of possible graph configurations is $|\mathbb{A}| = 3^{\binom{N}{2}}$. To ease mathematical manipulations, let us define the following three quantities:

$$a_{ij}^- = [a_{ij} = -1], \quad a_{ij}^0 = [a_{ij} = 0], \quad a_{ij}^+ = [a_{ij} = +1] \quad (2)$$

where we have employed Iverson's brackets notation (see the Supplementary Note 1). These new variables are mutually exclusive, i.e. $\{a_{ij}^-, a_{ij}^0, a_{ij}^+\} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$, sum to 1, i.e. $a_{ij}^- + a_{ij}^0 + a_{ij}^+ = 1$,

and induce two non-negative matrices $\mathbf{A}^+, \mathbf{A}^-$ such that $\mathbf{A} = \mathbf{A}^+ - \mathbf{A}^-$ and $|\mathbf{A}| = \mathbf{A}^+ + \mathbf{A}^-$.

The numbers of positive and negative links are defined as

$$L^+ = \sum_{i=1}^N \sum_{j>i} a_{ij}^+ \quad \text{and} \quad L^- = \sum_{i=1}^N \sum_{j>i} a_{ij}^-. \quad (3)$$

Analogously, the positive and negative degrees of node i are

$$k_i^+ = \sum_{j \neq i} a_{ij}^+ \quad \text{and} \quad k_i^- = \sum_{j \neq i} a_{ij}^- \quad (4)$$

(naturally, $2L^+ = \sum_{i=1}^N k_i^+$ and $2L^- = \sum_{i=1}^N k_i^-$). The advantage of adopting Iverson's brackets is that each quantity is now computed from a matrix with positive entries, so that all quantities of interest are positive as well.

Let us now follow[21], according to which '*local measures attain efficiency by focusing only on cycles of particular, usually short, length, such as 3-cycles (triads)*', and consider the signed triads depicted in Fig. 7. As mentioned above, according to BT social systems tend to arrange themselves into configurations satisfying the principles 'the friend of my friend is my friend', 'the friend of my enemy is my enemy', 'the enemy of my friend is my enemy', 'the enemy of my enemy is my friend'[5]. SSBT formalizes this concept by stating that the overall network balance increases with the fraction of triangles having an even number of negative edges (said to be balanced or 'positive' since the product of the edge sings is a 'plus') and decreases with the fraction of triangles having an odd number of negative edges (said to be unbalanced or 'negative' since the product of the edge sings is a 'minus'). SWBT, on the other hand, considers the triangle with all negative edges balanced as well.

Notice that the product of an arbitrary number of matrices of type $\mathbf{A}^+$ and $\mathbf{A}^-$ allows us to count the abundance of closed walks whose signature matches the sequence of signs of the matrices. For example, the expression $[\mathbf{A}^+ \mathbf{A}^- \mathbf{A}^+]_{ii}$ counts the number of closed walks, starting from and ending at i, of length 3 and signature $(+ - +)$. Similarly, the expression $[\mathbf{A}^+ \mathbf{A}^+ \mathbf{A}^- \mathbf{A}^+]_{ii} = [(\mathbf{A}^+)^2 \mathbf{A}^- \mathbf{A}^+]_{ii}$ counts the number of closed walks, starting from and ending at i, of length 4 and signature $(+ + - +)$. Therefore, the level of balance of a network can be quantified by the abundance of (non-degenerate) triangles with an even number of negative
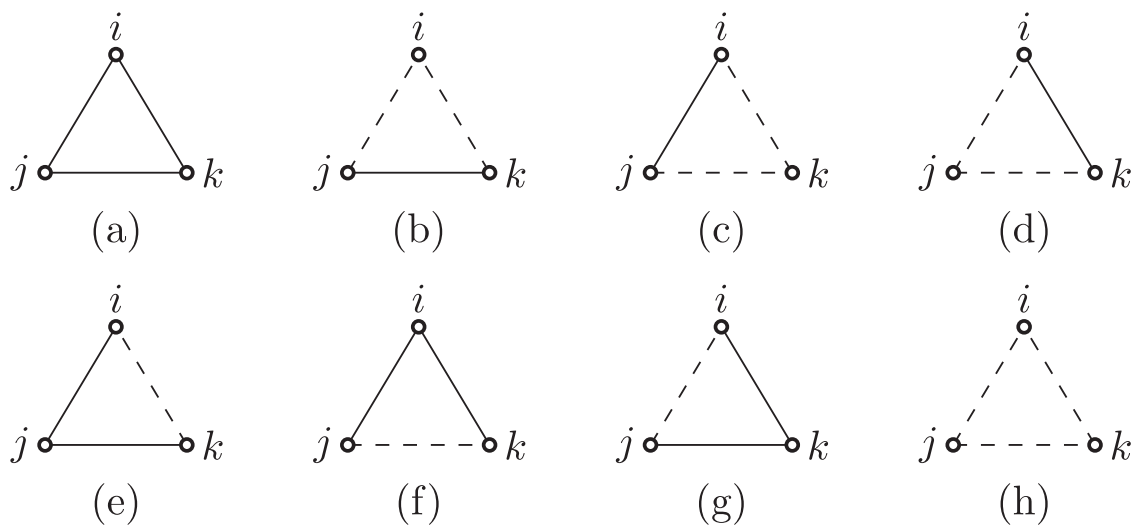
Fig. 7 | Signed triadic motifs. Signed triangles involving three representative nodes i, j, k. Solid lines denote positive edges while dashed lines denote negative edges. According to the strong version of balance theory, triangles (**a**)–(**d**) are balanced while triangles (**e**)–(**h**) are unbalanced. According to the weak version of balance theory, triangles (**a**)–(**d**) and (**h**) are balanced while triangles (**e**)–(**g**) are unbalanced.

links, i.e.

$$T^{(+++)} = \frac{1}{3}\sum_{i=1}^{N} T_i^{(+++)} = \frac{\text{Tr}\left[\left(\mathbf{A}^+\right)^3\right]}{6}, \tag{5}$$

$$T^{(+--)} = \frac{1}{2}\sum_{i=1}^{N} T_i^{(+--)} = \frac{\text{Tr}\left[\mathbf{A}^+\left(\mathbf{A}^-\right)^2\right]}{2}. \tag{6}$$

Similarly, the level of frustration of a network can be quantified by the abundance of (non-degenerate) triangles with an odd number of negative links, i.e.

$$T^{(---)} = \frac{1}{3}\sum_{i=1}^{N} T_i^{(---)} = \frac{\text{Tr}\left[\left(\mathbf{A}^-\right)^3\right]}{6}, \tag{7}$$

$$T^{(++-)} = \frac{1}{2}\sum_{i=1}^{N} T_i^{(++-)} = \frac{\text{Tr}\left[\left(\mathbf{A}^+\right)^2 \mathbf{A}^-\right]}{2} \tag{8}$$

(see the Supplementary Note 2 for more details).

The above expressions form the basis for the definition of several indices quantifying the level of balance of a network. For instance, the total number of balanced patterns according to SSBT is $\#_\triangle^{sb} = T^{(+++)} + T^{(+--)}$, while the total number of unbalanced patterns is $\#_\triangle^{su} = T^{(---)} + T^{(++-)}$. Hence, we may naturally define a 'strong degree of balance' index (SDoB) and a corresponding 'strong degree of frustration' index (SDoF) as

$$\text{SDoB} = \frac{\#_\triangle^{sb}}{\#_\triangle^{sb} + \#_\triangle^{su}}, \quad \text{SDoF} = 1 - \text{SDoB}. \tag{9}$$

On the other hand, the total number of balanced patterns according to SWBT is $\#_\triangle^{wb} = T^{(+++)} + T^{(+--)} + T^{(---)}$, while the total number of unbalanced patterns is $\#_\triangle^{wu} = T^{(++-)}$. Hence, we can introduce a 'weak degree of balance' index (WDoB) and a corresponding 'weak degree of frustration' index (WDoF) as

$$\text{WDoB} = \frac{\#_\triangle^{wb}}{\#_\triangle^{wb} + \#_\triangle^{wu}}, \quad \text{WDoF} = 1 - \text{WDoB}. \tag{10}$$

The indices defined above quantify imbalance by counting the abundance of locally frustrated, short cycles. Other indices of frustration account for the effect of structural (im)balance at larger scales. In particular, at the meso-scopic level, the effect of structural balance would result in a signed network being partitioned into communities of nodes, where intra-community links would be preferentially positive and inter-community links would be preferentially negative. Correspondingly, one can define the *frustration index*

$$\text{FI} = \frac{L_\circ^+ + L_\bullet^-}{L} \tag{11}$$

measuring the percentage of misplaced links, i.e. the total number $L_\circ^+$ of positive links between communities, plus the total number $L_\bullet^-$ of negative links within communities, divided by the total number L of links (the formalism is adapted from the one in[44]). According to SSBT, the node partition minimizing frustration (and, correspondingly, the FI) should be the one corresponding to only two communities, because such bipartition can be realized without creating all-negative triangles. By contrast, SWBT allows for a larger number of communities, because the theory justifies the presence of all-negative triangles precisely by assuming that the three participating nodes are all placed in different communities.

Null models of binary, undirected, signed graphs
Here we generalize the ERG framework to account for models of binary, undirected, signed graphs. We will follow the analytical approach introduced in[45], and further developed in[46], aimed at identifying the functional form of the maximum-entropy probability distribution (over all graphs of a chosen type) that preserves a desired set of empirical constraints on average. Specifically, this approach looks for the graph probability $P(\mathbf{A})$ that maximizes Shannon entropy

$$S = -\sum_{\mathbf{A}\in\mathbb{A}} P(\mathbf{A})\ln P(\mathbf{A}) \tag{12}$$

(where the sum runs over the set $\mathbb{A}$, of cardinality $|\mathbb{A}| = 3^{\binom{N}{2}}$, of all binary, undirected, signed graphs) under a set of constraints enforcing the expected value of a chosen set of properties. The formal solution to this problem is the exponential probability $P(\mathbf{A}) = e^{-H(\mathbf{A})}/Z$ where $H(\mathbf{A})$ (the *Hamiltonian*) is a linear combination of the constrained properties, each multiplied by a corresponding Lagrange multiplier, and $Z = \sum_{\mathbf{A}\in\mathbb{A}} e^{-H(\mathbf{A})}$ is the normalizing constant (or *partition function*).

In what follows, we will consider two classes of models, i.e. those keeping the network topology fixed and those letting the topology vary along with the edge signs. The first class is better suited for studying systems where actors cannot choose 'with whom' to interact, but only 'how' (e.g. because workers necessarily interact with colleagues at the same workplace or because countries necessarily interact with each other). On the other hand, the second class is better suited for studying systems where actors can choose their neighbors as well[17]. Whatever the situation, comparing the two types of models for the same network is in any case instructive, as it allows the role played by signed constraints to be disentangled from the one played by non-signed (purely topological) constraints.

Signed random graph model. As the simplest example, the Signed Random Graph Model (SRGM) is defined by two, global constraints: $L^+(\mathbf{A})$ and $L^-(\mathbf{A})$. The Hamiltonian

$$H(\mathbf{A}) = \alpha L^+(\mathbf{A}) + \beta L^-(\mathbf{A}) \tag{13}$$

leads to a graph probability $P_{\text{SRGM}}(\mathbf{A})$ that factorizes over the individual entries of the matrix $\mathbf{A}$, which are i.i.d. random variables described by the finite scheme

$$a_{ij} \sim \begin{pmatrix} -1 & 0 & +1 \\ p^- & p^0 & p^+ \end{pmatrix} \quad \forall i < j \tag{14}$$

with $p^0 \equiv 1 - p^- - p^+$ and

$$p^- \equiv \frac{e^{-\beta}}{1 + e^{-\alpha} + e^{-\beta}} \equiv \frac{y}{1 + x + y}, \tag{15}$$

$$p^+ \equiv \frac{e^{-\alpha}}{1 + e^{-\alpha} + e^{-\beta}} \equiv \frac{x}{1 + x + y}, \tag{16}$$

where $x \equiv e^{-\alpha}$ and $y \equiv e^{-\beta}$ are transformed Lagrange multipliers (see the Supplementary Note 3 for more details). In other words, positive, negative and missing links appear with probability p^+, p^- and p^0 respectively. The parameters (x, y) determining these probabilities are tuned by maximizing the log-likelihood function $\mathcal{L}_{\text{SRGM}}(x, y) \equiv \ln P_{\text{SRGM}}(\mathbf{A}^*|x, y)$ where $\mathbf{A}^*$ denotes the specific, empirical network under analysis. This maximization, according to a general result[47], leads to an equality between the expected and the empirical values of the constraints, i.e. $\langle L^+\rangle_{\text{SRGM}} = L^+(\mathbf{A}^*)$ and $\langle L^-\rangle_{\text{SRGM}} = L^-(\mathbf{A}^*)$. This leads to $p^0 \equiv 1 - p^- - p^+$ and

$$p^+ = \frac{2L^+(\mathbf{A}^*)}{N(N-1)}, \quad p^- = \frac{2L^-(\mathbf{A}^*)}{N(N-1)}. \tag{17}$$

Signed random graph model with fixed topology. We can also consider a variant of the SRGM that keeps the topology of the network under analysis fixed while (solely) randomizing the edge signs. The Hamiltonian is again $H(\mathbf{A}) = \alpha L^+(\mathbf{A}) + \beta L^-(\mathbf{A})$, but the random variables are now only the entries of the adjacency matrix corresponding to the connected pairs of nodes in the original network $\mathbf{A}^*$, i.e. the ones for which $|a_{ij}^*| = 1$. These entries obey the finite scheme

$$a_{ij} \sim \begin{pmatrix} -1 & +1 \\ p^- & p^+ \end{pmatrix} \quad \forall\, i < j \,\big|\, |a_{ij}^*| = 1 \tag{18}$$

with

$$p^- \equiv \frac{e^{-\beta}}{e^{-\alpha} + e^{-\beta}} \equiv \frac{y}{x+y}, \tag{19}$$

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

In other words, each entry for which $|a_{ij}^*| = 1$ obeys a Bernoulli distribution with probabilities determined by the (Lagrange multipliers of the) imposed constraints (see the Supplementary Note 3 for more details). The maximization of the likelihood function $\mathcal{L}_{\mathrm{SRGM-FT}}(x, y) \equiv \ln P_{\mathrm{SRGM-FT}}(\mathbf{A}^*|x, y)$ (where FT stands for 'fixed topology') leads to

$$p^+ = \frac{L^+(\mathbf{A}^*)}{L(\mathbf{A}^*)}, \quad p^- = \frac{L^-(\mathbf{A}^*)}{L(\mathbf{A}^*)} \tag{21}$$

with $L(\mathbf{A}^*)$ representing the (empirical) number of links.

The SRGM and the SRGM-FT are related via the simple expression

$$P_{\mathrm{SRGM}}(\mathbf{A}) = P_{\mathrm{RGM}}(\mathbf{A}) \cdot P_{\mathrm{SRGM-FT}}(\mathbf{A}) \tag{22}$$

involving the probability of the usual 'unsigned' (Erdős-Rényi) Random Graph Model (RGM) and stating that the probability of connecting any two nodes with, say, a positive link can be rewritten as the probability of connecting them with an unsigned link times the probability of assigning the latter a 'plus one': in formulas, $p^+_{\mathrm{SRGM}}/p^+_{\mathrm{RGM}} = p^+_{\mathrm{SRGM-FT}}$ (see the Supplementary Note 3 for more details). Notice that if the network under analysis is completely connected, the SRGM and the SRGM-FT coincide.

Although the recipes implemented in[15] and[2,31] are similar in spirit to the SRGM and the SRGM-FT, we provide the rigorous derivation of both models, together with the proof that the latter is nothing but the conditional version of the former.

Signed configuration model. The two aforementioned versions of the SRGM are defined by constraints which are global in nature. However, real social networks are characterized by an inherent heterogeneity of actors, which results in broad distributions of the number of connections of actors. To avoid statistical conclusions about structural balance that are biased by the application of homogeneous null models to intrinsically heterogeneous networks, it is therefore important to introduce models with local (node-specific) constraints.

We, therefore, introduce the Signed Configuration Model (SCM) via the Hamiltonian

$$H(\mathbf{A}) = \sum_{i=1}^{N} \left[\alpha_i k_i^+(\mathbf{A}) + \beta_i k_i^-(\mathbf{A}) \right] \tag{23}$$

which constraints the expected value of the signed degrees $\{k_i^+(\mathbf{A})\}_{i=1}^N$ and $\{k_i^-(\mathbf{A})\}_{i=1}^N$ of all nodes. The resulting graph probability $P_{\mathrm{SCM}}(\mathbf{A})$ is still factorized over independent entries of the matrix $\mathbf{A}$, however these entries are no longer identically distributed. Rather, they obey the finite

scheme

$$a_{ij} \sim \begin{pmatrix} -1 & 0 & +1 \\ p_{ij}^- & p_{ij}^0 & p_{ij}^+ \end{pmatrix} \quad \forall\, i < j \tag{24}$$

with

$$p_{ij}^- \equiv \frac{e^{-(\beta_i+\beta_j)}}{1 + e^{-(\alpha_i+\alpha_j)} + e^{-(\beta_i+\beta_j)}} \equiv \frac{y_i y_j}{1 + x_i x_j + y_i y_j}, \tag{25}$$

$$p_{ij}^+ \equiv \frac{e^{-(\alpha_i+\alpha_j)}}{1 + e^{-(\alpha_i+\alpha_j)} + e^{-(\beta_i+\beta_j)}} \equiv \frac{x_i x_j}{1 + x_i x_j + y_i y_j}, \tag{26}$$

and $p_{ij}^0 \equiv 1 - p_{ij}^- - p_{ij}^+$ (see the Supplementary Note 3 for more details). In other words, the two nodes i and j are connected by a positive, negative or missing link with probability p_{ij}^+, p_{ij}^- or p_{ij}^0 respectively. The parameters of the SCM are found by maximizing the log-likelihood $\mathcal{L}_{\mathrm{SCM}}(\{x_i\}_{i=1}^N, \{y_i\}_{i=1}^N) \equiv \ln P_{\mathrm{SCM}}(\mathbf{A}^*|\{x_i\}_{i=1}^N, \{y_i\}_{i=1}^N)$, and the result ensures that $\langle k_i^+ \rangle_{\mathrm{SCM}} = k_i^+(\mathbf{A}^*)$ and $\langle k_i^- \rangle_{\mathrm{SCM}} = k_i^-(\mathbf{A}^*), \forall\, i$. Explicitly,

$$k_i^+(\mathbf{A}^*) = \sum_{j\neq i} \frac{x_i x_j}{1 + x_i x_j + y_i y_j} = \langle k_i^+ \rangle \quad \forall\, i, \tag{27}$$

$$k_i^-(\mathbf{A}^*) = \sum_{j\neq i} \frac{y_i y_j}{1 + x_i x_j + y_i y_j} = \langle k_i^- \rangle \quad \forall\, i, \tag{28}$$

which is a system of $2N$ coupled non-linear equations that have a unique solution to be found numerically, e.g. following the guidelines provided in[48] (see the Supplementary Note 4). If $x_i \ll 1$ and $y_i \ll 1 \,\forall\, i$, a 'sparse' approximation of the SCM holds true and one can factorize the probabilities as $p_{ij}^+ \simeq x_i x_j$ and $p_{ij}^- \simeq y_i y_j, \forall\, i < j$. Such a manipulation leads us to

$$p_{ij}^+ \simeq \frac{k_i^+(\mathbf{A}^*) k_j^+(\mathbf{A}^*)}{2L^+(\mathbf{A}^*)}, \quad p_{ij}^- \simeq \frac{k_i^-(\mathbf{A}^*) k_j^-(\mathbf{A}^*)}{2L^-(\mathbf{A}^*)}, \tag{29}$$

a result that we may call the Signed Chung-Lu Model (SCLM).

To the best of our knowledge, the canonical SCM described here has no precedents in the literature: Ref. 10 provides a microcanonical version of the model, while the variant considered in[32] is just an approximation of the full canonical model derived here. Notice that the bipartite version of the SCM can be recovered as a special case of the Bipartite Score Configuration Model, proposed in[35].

Signed configuration model with fixed topology. As for the SRGM, a variant of the SCM that keeps the topology of the network under analysis fixed while (solely) randomizing the signs of the edges can be defined. Again, the Hamiltonian reads $H(\mathbf{A}) = \sum_{i=1}^{N} [\alpha_i k_i^+(\mathbf{A}) + \beta_i k_i^-(\mathbf{A})]$ but the only random variables are those corresponding to the connected pairs of nodes in the empirical graph, i.e. the ones for which $|a_{ij}^*| = 1$. Each of them obeys the finite scheme

$$a_{ij} \sim \begin{pmatrix} -1 & +1 \\ p_{ij}^- & p_{ij}^+ \end{pmatrix} \quad \forall\, i < j \,\big|\, |a_{ij}^*| = 1 \tag{30}$$

with

$$p_{ij}^- \equiv \frac{e^{-(\beta_i+\beta_j)}}{e^{-(\alpha_i+\alpha_j)} + e^{-(\beta_i+\beta_j)}} \equiv \frac{y_i y_j}{x_i x_j + y_i y_j}, \tag{31}$$

$$p_{ij}^+ \equiv \frac{e^{-(\alpha_i+\alpha_j)}}{e^{-(\alpha_i+\alpha_j)} + e^{-(\beta_i+\beta_j)}} \equiv \frac{x_i x_j}{x_i x_j + y_i y_j}. \tag{32}$$

Maximizing the log-likelihood $\mathcal{L}_{\text{SCM-FT}}(\{x_i\}_{i=1}^N, \{y_i\}_{i=1}^N) \equiv \ln P_{\text{SCM-FT}}(\mathbf{A}^*|\{x_i\}_{i=1}^N, \{y_i\}_{i=1}^N)$ leads to the equations

$$k_i^+(\mathbf{A}^*) = \sum_{j \neq i} |a_{ij}^*| \frac{x_i x_j}{x_i x_j + y_i y_j} = \langle k_i^+ \rangle \quad \forall i, \tag{33}$$

$$k_i^-(\mathbf{A}^*) = \sum_{j \neq i} |a_{ij}^*| \frac{y_i y_j}{x_i x_j + y_i y_j} = \langle k_i^- \rangle \quad \forall i, \tag{34}$$

which can be solved numerically - again, along the guidelines provided in[48] (see the Supplementary Note 4 for more details).

Similarly to what has been observed for the SRGM and the SRGM-FT, the SCM and the SCM-FT are related via

$$P_{\text{SCM}}(\mathbf{A}) = P_{\text{ICM}}(\mathbf{A}) \cdot P_{\text{SCM-FT}}(\mathbf{A}), \tag{35}$$

an expression involving the probability of an ordinary (unsigned) 'induced' Configuration Model (ICM) with probabilities such that $(p_{ij}^+)_{\text{SCM}}/(p_{ij}^+)_{\text{ICM}} = (p_{ij}^+)_{\text{SCM}}/[(p_{ij}^+)_{\text{SCM}} + (p_{ij}^-)_{\text{SCM}}] = (p_{ij}^+)_{\text{SCM-FT}}$, for any pair of nodes (see the Supplementary Note 3). Notice that, if the network under consideration is completely connected, then the SCM and the SCM-FT coincide.

Data availability
Data concerning CoW are described in[37] and can be found at the address http://mrvar.fdv.uni-lj.si/pajek/SVG/CoW/. Data concerning E. coli, Macrophage, EGFR, N.G.H. Tribes and Monastery are are described in[27] and can be found at the address https://figshare.com/articles/dataset/Signed_networks_from_sociology_and_political_science_biology_international_relations_finance_and_computational_chemistry/5700832. Data concerning Bitcoin Alpha and Bitcoin OTC are described in[19] and can be found at the address https://figshare.com/articles/dataset/Dataset_of_directed_signed_networks_from_social_domain/12152628. Data concerning MMOG, described in[38] are subject to proprietary restrictions and cannot be shared.

Code availability
The codes implementing the null models employed for the present analysis are available upon request.

References
1. Antal, T., Krapivsky, P. L. & Redner, S. Social balance on networks: The dynamics of friendship and enmity. *Phys. D: Nonlinear Phenom.* **224**, 130–136 (2006).
2. Leskovec, J., Huttenlocher, D. & Kleinberg, J. Signed networks in social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1361–1370 https://doi.org/10.1145/1753326.1753532 (2010).
3. Zaslavsky, T. A Mathematical bibliography of signed and gain graphs and allied areas. *Electron. J. Combinatorics* **DS8**, 1–524 (2012).
4. Tang, J., Chang, Y., Aggarwal, C. & Liu, H. A survey of signed network mining in social media. *ACM Comput. Surv.* **49**, 1–37 (2016).
5. Heider, F. Attitudes and cognitive organization. *J. Psychol.* **21**, 107–112 (1946).
6. Cartwright, D. & Harary, F. Structural balance: A generalization of Heider's theory. *Psychol. Rev.* **63**, 277 (1956).
7. Harary, F., Lim, M.-H. & Wunsch, D. C. Signed graphs for portfolio analysis in risk management. *IMA J. Manag. Math.* **13**, 201–210 (2002).
8. Ou-Yang, L., Dai, D.-Q. & Zhang, X.-F. Detecting protein complexes from signed protein-protein interaction networks. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* **12**, 1333–1344 (2015).
9. Iorio, F. et al. Efficient randomization of biological networks while preserving functional characterization of individual nodes. *BMC Bioinforma.* **17**, 1–14 (2016).
10. Saiz, H. et al. Evidence of structural balance in spatial ecological networks. *Ecography* **40**, 733–741 (2017).
11. Davis, J. A. Clustering and structural balance in graphs. *Hum. Relat.* **20**, 181–187 (1967).
12. Akiyama, J., Avis, D., Chvátal, V. & Era, H. Balancing signed graphs. *Discret. Appl. Math.* **3**, 227–233 (1981).
13. Harary, F. On the measurement of structural balance. *Behav. Sci.* **4**, 316–323 (1959).
14. Estrada, E. & Benzi, M. Walk-based measure of balance in signed networks: Detecting lack of balance in social networks. *Phys. Rev. E* **90**, 042802 (2014).
15. Singh, R. & Adhikari, B. Measuring the balance of signed networks and its application to sign prediction. *J. Stat. Mech. Theory Exp.* **2017**, 063302 (2017).
16. Estrada, E. Rethinking structural balance in signed social networks. *Discret. Appl. Math.* **268**, 70–90 (2019).
17. Kirkley, A., Cantwell, G. T. & Newman, M. E. Balance in signed networks. *Phys. Rev. E* **99**, 012320 (2019).
18. Easley, D. et al. Networks, crowds, and markets. *Cambridge Books* https://doi.org/10.1017/CBO9780511761942 (2012).
19. Aref, S., Dinh, L., Rezapour, R. & Diesner, J. Multilevel structural evaluation of signed directed social networks based on balance theory. *Sci. Rep.* **10**, 1–12 (2020).
20. Talaga, S., Stella, M., Swanson, T. J. & Teixeira, A. S. Polarization and multiscale structural balance in signed networks. *Commun. Phys.* **6**, 349 (2023).
21. Giscard, P.-L., Rochet, P. & Wilson, R. C. Evaluating balance on social networks from their simple cycles. *J. Complex Netw.* **5**, 750–775 (2017).
22. Kunegis, J. et al. Spectral analysis of signed graphs for clustering, prediction and visualization. In *Proceedings of the 2010 SIAM International Conference on Data Mining*, 559–570 (SIAM, 2010). https://doi.org/10.1137/1.9781611972801.49.
23. Terzi, E. & Winkler, M. A spectral algorithm for computing social balance. In *International Workshop on Algorithms and Models for the Web-Graph*, 1–13 (Springer, 2011). https://doi.org/10.1007/978-3-642-21286-4_1.
24. Anchuri, P. & Magdon-Ismail, M. Communities and balance in signed networks: A Spectral Approach. In *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 235–242 (IEEE, 2012). https://doi.org/10.1109/ASONAM.2012.48.
25. Belaza, A. M. et al. Statistical physics of balance theory. *PLoS One* **12**, e0183696 (2017).
26. Zaslavsky, T. Balanced decompositions of a signed graph. *J. Combinatorial Theory, Ser. B* **43**, 1–13 (1987).
27. Aref, S. & Wilson, M. C. Balance and frustration in signed networks. *J. Complex Netw.* **7**, 163–189 (2019).
28. Aref, S., Mason, A. J. & Wilson, M. C. A modeling and computational study of the frustration index in signed networks. *Networks* **75**, 95–110 (2020).
29. Traag, V., Doreian, P. & Mrvar, A. Partitioning signed networks. *Advances in Network Clustering and Blockmodeling* 225–249 https://doi.org/10.1002/9781119483298.ch8 (2019).
30. Abelson, R. P. & Rosenberg, M. J. Symbolic psycho-logic: A model of attitudinal cognition. *Behav. Sci.* **3**, 1–13 (1958).

31. Facchetti, G., Iacono, G. & Altafini, C. Computing global structural balance in large-scale signed social networks. *Proc. Natl. Acad. Sci.* **108**, 20953–20958 (2011).

32. Derr, T., Aggarwal, C. & Tang, J. Signed network modeling based on structural balance theory. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, 557–566 https://doi.org/10.1145/3269206.3271746 (2018).

33. Huitsing, G. et al. Univariate and multivariate models of positive and negative networks: Liking, disliking, and bully–victim relationships. *Soc. Netw.* **34**, 645–657 (2012).

34. Lerner, J. Structural balance in signed networks: Separating the probability to interact from the tendency to fight. *Soc. Netw.* **45**, 66–77 (2016).

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

36. Fritz, C. Mehrl, M., Thurner, P. W. & Kauermann, G. Exponential random graph models for dynamic signed networks: an application to international relations. *Polit. Anal.* (in the press).

37. Doreian, P. & Mrvar, A. Structural balance and signed international relations. *J. Soc. Struct.* **16**, 1 (2015).

38. Szell, M., Lambiotte, R. & Thurner, S. Multirelational organization of large-scale social networks in an online world. *Proc. Natl. Acad. Sci.* **107**, 13636–13641 (2010).

39. Signed networks from sociology and political science, systems biology, international relations, finance, and computational chemistry https://figshare.com/articles/dataset/Signed_networks_from_sociology_and_political_science_biology_international_relations_finance_and_computational_chemistry/5700832 (2018).

40. Sampson, S. F. *A novitiate in a period of change: An experimental and case study of social relationships* (Cornell University, 1968). https://doi.org/10.1016/0378-8733(95)00259-6

41. Kumar, S., Spezzano, F., Subrahmanian, V. & Faloutsos, C. Edge weight prediction in weighted signed networks. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*, 221–230 (IEEE, 2016). https://ieeexplore.ieee.org/document/7837846.

42. Gómez, S., Jensen, P. & Arenas, A. Analysis of community structure in networks of correlated data. *Phys. Rev. E* **80**, 016114 (2009).

43. Doreian, P. & Mrvar, A. A partitioning approach to structural balance. *Soc. Netw.* **18**, 149–168 (1996).

44. Marchese, E., Caldarelli, G. & Squartini, T. Detecting mesoscale structures by surprise. *Commun. Phys.* **5**, 1–16 (2022).

45. Park, J. & Newman, M. E. J. Statistical mechanics of networks. *Phys. Rev. E* **70**, 66117 (2004).

46. Squartini, T. & Garlaschelli, D. *Maximum-Entropy Networks. Pattern Detection, Network Reconstruction and Graph Combinatorics* (Springer International Publishing, 2017). https://doi.org/10.1007/978-3-319-69438-2.

47. Garlaschelli, D. & Loffredo, M. I. Maximum likelihood: Extracting unbiased information from complex networks. *Phys. Rev. E* **78**, 015101 (2008).

48. Vallarano, N. et al. Fast and scalable likelihood maximization for Exponential Random Graph Models with local constraints. *Sci. Rep.* **11**, 15227 (2021).

49. El Maftouhi, A., Manoussakis, Y. & Megalakaki, O. Balance in random signed graphs. *Internet Math.* **8**, 364–380 (2012).

Author contributions
Study conception and design: D.G., R.L., F.S., T.S. Data collection: A.G. Analysis and interpretation of results: A.G., D.G., R.L., F.S., T.S. Draft manuscript preparation: D.G., T.S.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42005-024-01640-7.

Correspondence and requests for materials should be addressed to Tiziano Squartini.

Peer review information *Communications Physics* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.