

Meso-scale structures in signed networks

Wei Zhang^{1,2*}, Olga Boichak³, Tristram J. Alexander^{2,4}, Tiago P. Peixoto⁵,
Eduardo G. Altmann^{1,2*}

¹School of Mathematics and Statistics, University of Sydney, Sydney, 2006 NSW, Australia.

²Centre for Complex Systems, University of Sydney, Sydney, 2006 NSW, Australia.

³School of Art, Communication and English, University of Sydney, Sydney, 2006 NSW, Australia.

⁴School of Physics, University of Sydney, Sydney, 2006 NSW, Australia.

⁵Inverse Complexity Lab, IT:U Interdisciplinary Transformation University, Linz 4040, Austria.

*Corresponding author. Email: wei.zhang4@sydney.edu.au; eduardo.altmann@sydney.edu.au

Abstract

Meso-scale structures in signed networks have been studied under the limiting assumption of the validity of social balance theory, which predicts positive connections within groups and negative connections between groups. Here, we propose and apply a methodology that overcomes this limitation and is able to find and characterize also the different possible unbalanced structures in signed networks. Applying our methodology to 24 empirical networks, from social-political, financial, and biological domains, we find that unbalanced meso-scale structures are prevalent in real-world networks, including cases with substantial balance at the micro-scale of triangles. In particular, we find that assortativity often prevails regardless of the interaction sign and that core-periphery structures are typical in online social networks. Our findings highlight the complexity of meso-scale relational structures, the importance of using computational methods that are a priori agnostic to specific patterns, and the importance of independently evaluating micro- and meso-scale predictions of social balance theory.

1 Introduction

The recent renewed interest in signed networks [1–8] reflects the need for network representations of complex phenomena to distinguish between opposing types of interactions. For instance, online interactions among users (vertices) do not always reflect friendship or agreement (positive edges) and are thus better described through networks that account for enmity or disagreement (negative edges). Representing these relationships as signed networks allows researchers not only to capture the presence and strength of connections among vertices, but also differentiate the effects of positive and negative interactions on the structure of the network. Large signed networks typically show a rich combination not only of micro-scale (e.g., triadic interactions) but also of meso-scale structures (e.g., communities), that are known to be essential to understand the underlying complex system (e.g., the partition of users in opposing groups) [9–12].

Signed networks have been primarily studied within *structural balance theory* [13], which originated in studies of social and economic networks [14]. A signed network is considered balanced (anti-balanced) if and only if all cycles have an even number of negative (positive) edges [13, 15, 16]. In the simplest case of 3-cycles, the signed network with all triads having zero or two negative edges is balanced. While this micro-scale perspective of structural balance theory is the most popular, structural balance can also be equivalently viewed from a meso-scale perspective: balanced (anti-balanced) signed networks can be partitioned into 2 communities whose intra-links are all positive (negative) and inter-links are all negative (positive) [15, 17]. In a weaker version, k -balance theory states that networks are balanced if they can be partitioned into $k \geq 2$ communities with all intra-community links being positive and all inter-community links being negative [7, 8, 18]. These balance conditions are rarely satisfied strictly in signed networks obtained from empirical datasets. Different approaches at different scales [4, 19] have been proposed to quantify the extent into which a signed network is balanced: from a micro-scale perspective, the *degree of balance* (DoB) of signed networks is usually quantified as the fraction of balanced triads [4, 13, 20–22]; from a meso-scale perspective, the network is first partitioned into k -groups and then the fractions of links *frustrating* the balance condition is usually considered as a measure of distance from balance [19, 23].

A critical step in the meso-scale approach mentioned above is the partitioning of the network into groups. Different community detection methods, originally proposed for simple networks, have been recently extended to signed networks: modularity maximization [10, 11, 24–27], spectral clustering [28–31], and spin-glass [32, 33]. The goal of these methods is to find a partition that is close to a k -balanced partition, i.e., to divide the network into separate communities so that nodes within the same community have as many positive connections as possible, while nodes in different communities have as many negative connections as possible. While such descriptive methods invariably succeed in finding balanced partitions [4, 10, 19], the fact that they search exclusively for balanced partitions and do not properly account for statistical significance makes them unsuitable to quantify their importance in describing the data or to detect the existence of unbalanced meso-scale structures. As we show below, they find balanced partitions even in random networks where they are absent. Since meso-scale structures in signed networks have been studied almost exclusively within structural balance theory and using these community-detection methods, several fundamental questions remain open: To what extent are the meso-scale structures of signed networks aligned with structural balance theory? Are there different types of unbalanced structures? What are they? What do they reveal about the organization of large signed networks?

In this work, we propose and apply a framework to overcome the limitations of previous approaches and to answer these questions. Our goal is to find, characterize, and quantify the strength of meso-scale structures of signed networks beyond the balance assumption. We achieved this by using inferential methods for community detection [34], which have already highlighted that diverse mixing patterns may simultaneously coexist in a network [35–38]. In particular, we use stochastic block models (SBMs) [12, 34, 39–41] that leverage statistical evidence and do not a priori adhere to specific types of balanced community structures, as well as a generalization to signed networks of recently proposed characterizations of mesoscale structures between communities [42, 43]. We propose a methodology to characterize the different possible meso-scale structures in signed networks and we apply it to 24 different signed networks. We find that non-balanced partitions are present in almost all networks and are dominant in a significant portion of them. In particular, we find that social networks are often best partitioned in groups in which positive and negative links happen within the same group, possibly reflecting the interaction opportunities available to individuals. We also find that relationships akin to core-periphery structures are widespread in (the negative edges of) online social networks. Finally, we show how the micro- and meso-scale characterization of structural-balance theory often leads to complementary pictures of the processes shaping complex signed networks.

2 Methodology

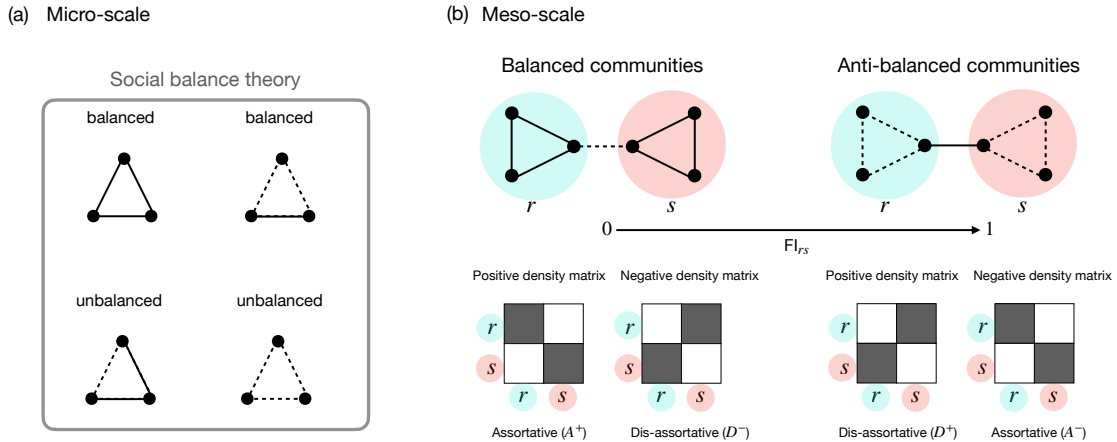


Figure 1: **Structural balance theory from (a) micro-scale and (b) meso-scale perspectives.** (a) shows fundamental triads or motifs considered as balanced and unbalanced by social balance theory. (b) Balanced (left) and anti-balanced (right) communities. The top panel shows simple signed networks, with positive links as solid lines and negative links as dashed lines. The two groups r and s are marked with circles. The bottom panels show the corresponding edge density matrices (1).

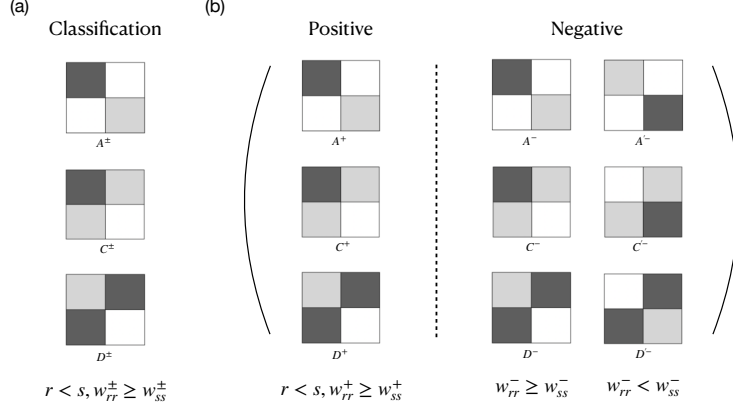


Figure 2: **Classification of community types.** (a) The three potential communities are represented by density matrices. The darker blocks indicate a higher density of edges. Assortative A^\pm (disassortative D^\pm) cases have the two largest entries in the diagonal w_{rr}^\pm, w_{ss}^\pm (antidiagonal w_{rs}^\pm, w_{sr}^\pm). The core-periphery type of relationship C^\pm has one sparser community (periphery) linked strongly to the denser community (core), which is in turn more interconnected within itself. (b) The three possible community types for positive links and six for negative links, thus leading to a total of 18 possible combination types. Note, blocks are arranged such that the left column has the highest density of positive intra-block edges. This freedom is unavailable for the negative edges, so there are twice as many types.

Table 1: **List of methods used for the partition of signed networks.**

Method	Core Idea	Balance Assumption?
Layered SBM (LSBM) [12]	Inference method. It treats the signed network as a two-layer multiplex network. The same partition applies to both layers.	No
Weighted SBM (WSBM) [40]	Inference method. Assigns edge signs using a block-pair-specific distribution (an approach used also for describing edge weights).	No
Louvain [26, 27]	Modularity-based method. Maximizes the signed modularity by iteratively merging nodes into groups until no further modularity improvement is possible.	Yes
Spectral Clustering [31]	Linear algebra-based method. Maximizes positive within-group and negative between-group edges.	Yes
Spin-glass [10]	Statistical-physics method. Uses an Ising spin-glass model, with nodes (spins) aligning/anti-aligning depending on edge sign.	Yes

We first consider how meso-scale partitions of signed networks are related to structural balance theory at its usual micro-scale representation. Figure 1 shows a simple case connecting the two scales, with balanced and anti-balanced meso-scale communities fulfilling the micro-scale definitions based on triads. A balanced network shows diagonal blocks in the positive edge density matrix and off-diagonal blocks in the negative edge density matrix (bottom of Fig. 1b). This indicates that balanced communities exhibit assortative relationships in positive interactions and disassortative relationships in negative interactions [42, 43]. Conversely, anti-balanced communities show disassortative relationships in positive interactions and assortative relationships in negative interactions. This motivates us to use the edge density matrix to measure and classify diverse community structures, including unbalanced ones.

Consider a simple undirected signed graph G with adjacency matrix $(A_{ij}) \in \mathbb{R}^{N \times N}$, where each entry A_{ij} can be written as the sum of two disjoint components $A_{ij} = A_{ij}^+ + A_{ij}^-$ with $A_{ij}^+ > 0$ represents positive edges and $A_{ij}^- < 0$ represents negative edges. The graph is partitioned into $r = 1, 2, 3, \dots, B$ disjoint subgroups, each containing N_r nodes. For a pair of communities r and s , the density of positive interactions and negative interactions is given by the ratio of existing edges and possible edges as

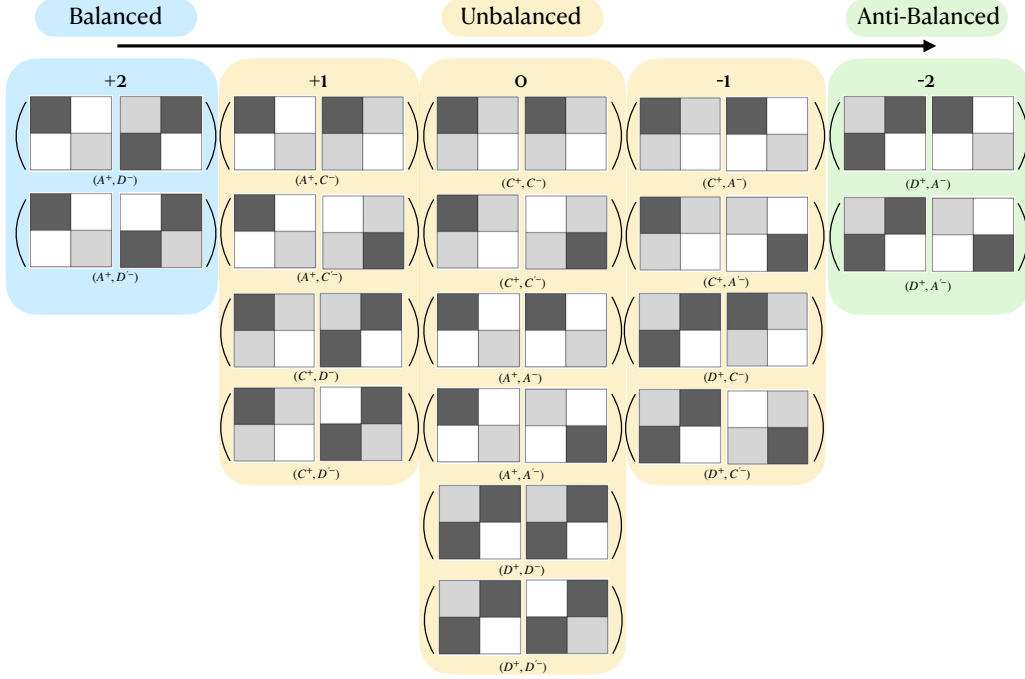


Figure 3: **Classification of meso-scale structures in signed networks.** ‘ A^\pm ’, ‘ C^\pm ’ and ‘ D^\pm ’ stand for ‘assortative’, ‘core-periphery’ and ‘disassortative’ in positive (+) or negative (−) interaction, respectively, and are defined by Eq. (2). For example, the pair (A^+, D^-) represents the positive edges are assortative and the negative edges are disassortative, and the mark ‘ $'$ labeled in D'^- indicates that the negative diagonal entries $w_{rr}^- < w_{ss}^-$. From left to right, the 18 different possible configurations are divided according to their level of balance. A relationship is balanced when positive edges are assortative (dense within communities) and negative edges are disassortative (dense between communities). In contrast, a structure is anti-balanced when positive edges are disassortative and negative edges are assortative. These balanced (anti-balanced) structures are considered to have weak structural balance (anti-balance) since they allow the limited presence of other entries in the density matrix. The remaining unbalanced configurations are scored as follows: assortative is +1, core-periphery is 0, and disassortative is −1. For negative edges, the scoring is: disassortative is +1, core-periphery is 0, and assortative is −1.

$$w_{rs}^\pm = \begin{cases} \sum_{i \in r, j \in s} \frac{|A_{ij}^\pm|}{N_r N_s}, & \text{if } r \neq s, \\ \sum_{i \in r, j \in s} \frac{|A_{ij}^\pm|}{N_r (N_s - 1)}, & \text{if } r = s. \end{cases} \quad (1)$$

We now classify the relationship between a pair of communities r and s based on w_{rs}^\pm . In simple graphs, as shown in Fig. 2a, this relationship can be classified as either assortative, core-periphery, or disassortative [42–44]. In signed networks, these relationships apply to both positive and negative edges:

$$\begin{cases} \text{Assortative}(A^\pm), & \text{if } w_{rs}^\pm < \min(w_{rr}^\pm, w_{ss}^\pm), \\ \text{Core-periphery}(C^\pm), & \text{if } \min(w_{rr}^\pm, w_{ss}^\pm) < \max(w_{rr}^\pm, w_{ss}^\pm), \\ \text{Disassortative}(D^\pm), & \text{if } w_{rs}^\pm > \max(w_{rr}^\pm, w_{ss}^\pm). \end{cases} \quad (2)$$

In the classification above, we adopt the community label convention that $w_{rr}^\pm \geq w_{ss}^\pm$ for $r < s$. By applying this convention to w^+ , this reduces the number of ranked relationships for the positive edges to three (shown in Fig. 2a), while six options remain for negative matrix w^- (shown in Fig. 2b), because the rankings of w_{rr}^- and w_{ss}^- may be oppositely ordered. To differentiate the two options (i.e., $w_{rr}^- > w_{ss}^-$ and $w_{rr}^- < w_{ss}^-$), we label this latter configuration by a prime symbol ‘ $'$ ’. For instance, the pair (C^+, C'^-) signifies that the two communities are in a core-periphery relationship in both positive and negative edges, but that the community that acts as core is *different* in each of the edge types (the notation (C^+, C^-) represents the relationship in which both cores are the same community). When no distinction between these cases is intended, we denote

the relationship as (C^+, C'^{-}) . Thus, there are a total of $3 \times 6 = 18$ possible relationships between two communities r and s for signed networks. These relationships can be grouped into balanced, unbalanced, and anti-balanced groups, as shown in Fig. 3.

In order to apply the pairwise classification described above to a network, we need to (i) identify the groups of nodes (i.e., partition the network); and (ii) consider how to apply our classification method to partitions with more than two communities ($B > 2$). Different methods for network partition can produce varying $B \times B$ density matrices, and our methodology can be applied to any of these partition methods. Below we show results for the five methods listed in Table 1, which include a variety of approaches including traditional methods that focus on balanced partitions – Louvain [26, 27], spectral clustering [31], and spin-glass [10]– and inferential approaches that are agnostic about it – Layered SBM (LSBM) [12] and Weighted SBM (WSBM) [40]. We then analyze each of the $\binom{B}{2}$ pairs, we classify their relationship type into one of the 18 possible combinations, and we focus on whether a specific relationship appears in that network (contributing to the overall prevalence of the type) and which type is the most common (dominant).

3 Results

We start by testing the methodology described in Sec. 2 in empirical networks. We collected 24 empirical signed networks, which include 18 social-political networks, 2 financial networks, and 4 gene-regulatory networks (see Methods 5.1 and Supplementary Materials for a detailed description of the data set). These networks vary substantially in size, between 16 and 10,884 nodes and between 40 and 251,406 edges (see Table S1 for a list).

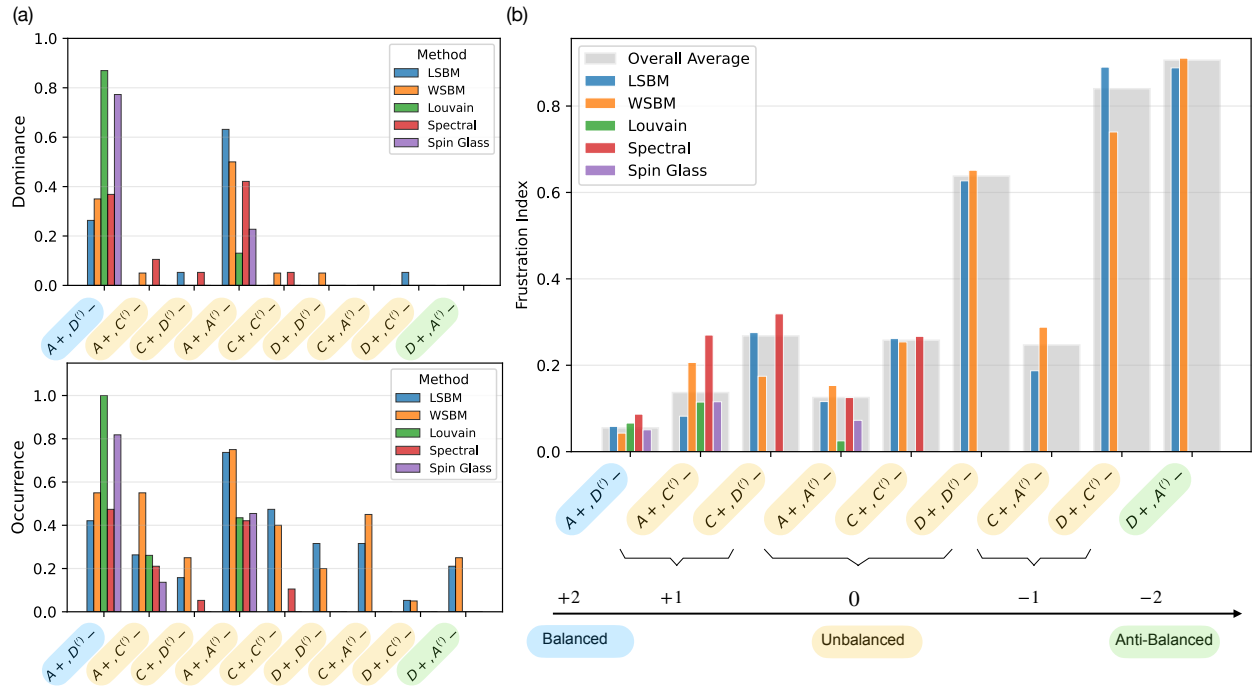


Figure 4: **Survey on 24 empirical networks.** (a) Community types in empirical networks. Top: fraction of networks in which each community type is dominant (computed as the proportion of networks in which the type is the most prevalent). Bottom: fraction of networks in which each community type occurs (computed as the proportion of networks in which the type appears at least once). (b) Frustration index in empirical networks. The x-axis represents balance categories ranging from balanced, unbalanced, to anti-balanced configurations, while the y-axis shows the frustration index. Bars correspond to different methods, with the overall average shown in gray. The calculation of the pairwise frustration index refers to Methods 5.2.

3.1 Unbalanced structures are common

The results obtained across all networks are summarized in Fig. 4. Unsurprisingly, the dominant relationship type in most networks for the three balanced-based methods (i.e., Louvain, spectral clustering, and spin-glass) is the balanced structure ($A^+, D^{(+)-}$), with few unbalanced structures detected except for ($A^+, A^{(+)-}$). This reflects the fact that these methods aim to find balanced partitions, thus providing little information about the type of structure present in the data. More interestingly, the results obtained for the two inferential methods (i.e., LSM and WSBM), which do not focus exclusively on balanced partitions and rely instead on statistical evidence, find more unbalanced structures with the dominant type of relationship over all networks being the unbalanced structure ($A^+, A^{(+)-}$) and with different core-periphery relationships observed as well. Similar results are observed when we restrict the analysis to social-political networks, financial networks, and gene-regulatory networks (Figs. S1-S3). This finding indicates that unbalanced structures among communities are not just a theoretical possibility; they are actually common in empirical networks, provided suitable partition methods are used.

Below we explore the significance and interpretation of these unbalanced relationships, focusing on the partitions obtained using SBMs because of their ability to detect such patterns. Before that, we validate our pairwise approach by comparing the consistency of the relationships we found with traditional methods from structural balance theory. We compare our classification of meso-scale structure types to the average pairwise frustration index, which measures the fraction of misplaced links, positive interlinks and negative intralinks, between communities (Methods 5.2 for details). A frustration index $FI = 0$ indicates perfect balance, while $FI = 1$ indicates complete anti-balance. The results in Fig. 4b reveal a clear trend: balanced structures correspond to low frustration, unbalanced structures exhibit intermediate levels, and anti-balanced structures show the highest frustration. This progression is consistent across different methods, and the overall averages align closely with theoretical expectations, confirming the alignment of our classification scheme to structural balance theory.

3.2 Assortativity prevails regardless of interaction sign

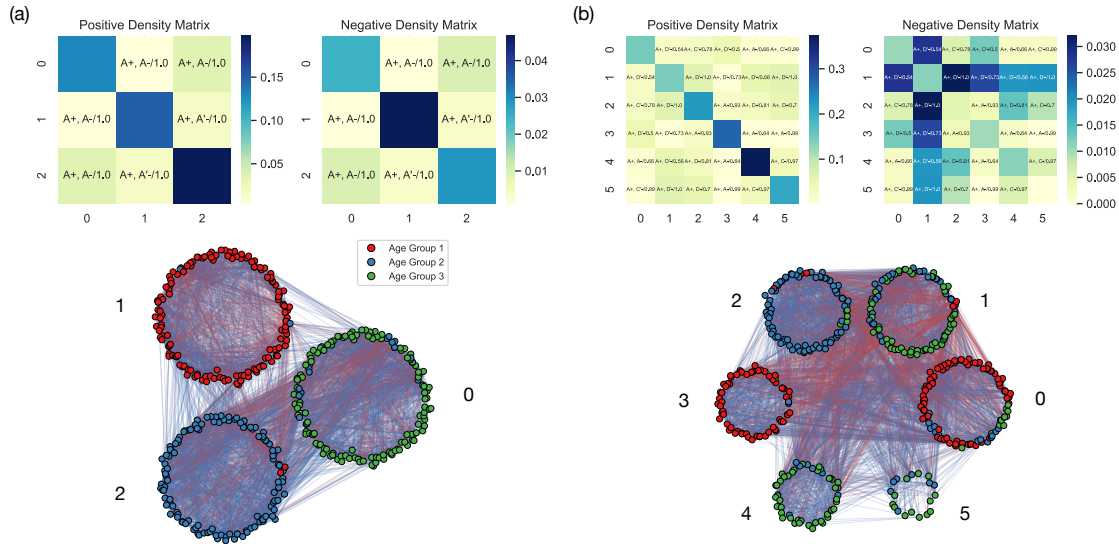


Figure 5: **Partition of Spanish High School network with (a) WSBM and (b) Louvain.** Top: The blocks are ranked according to community size. The labels in each entry of the matrix correspond to the structure type and the robustness score (see Methods 5.4). Bottom: The communities are grouped closely together, and the numbers denote the labels of each group. The colors of the nodes represent different age group memberships.

In most empirical networks investigated in Sec. 3.1, the dominant type found using SBMs is the unbalanced structure ($A^+, A^{(+)-}$), while for other methods, the dominant type is the balanced structure ($A^+, D^{(+)-}$). To understand this difference and the origin of ($A^+, A^{(+)-}$), we now focus on the example of the relationship

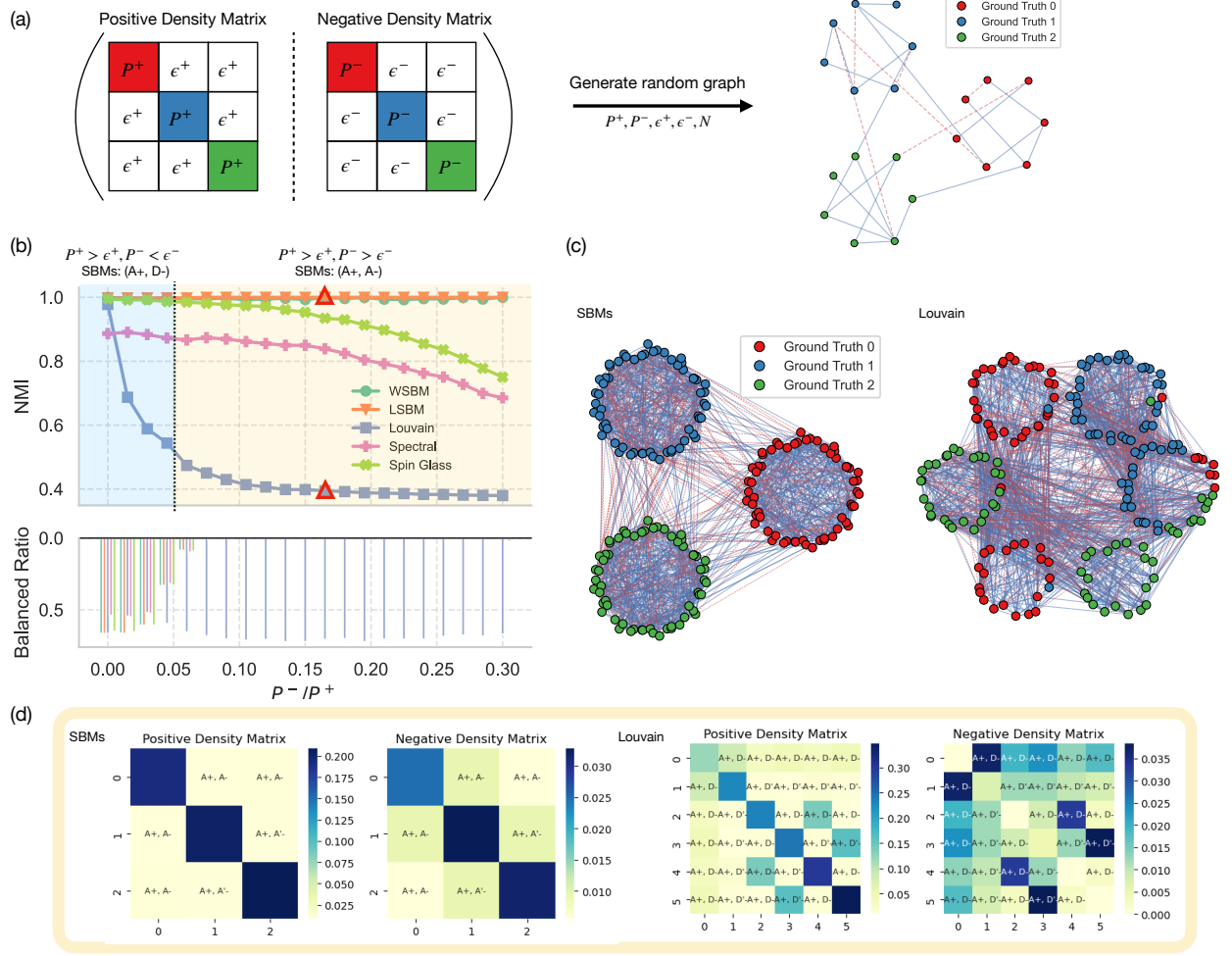


Figure 6: Comparison of partitions in synthetic networks. (a) We generate random networks that model the high-school interactions with three equal-sized age groups, each containing $N/3$ nodes. Students interact positively within their age groups with probability P^+ and negatively with probability P^- . Interactions between different age groups occur positively with probability ϵ^+ and negatively with probability ϵ^- . Given $P^+, P^-, \epsilon^+, \epsilon^-$, and N , we can generate the random graphs. Nodes are color-coded by age group: red, blue, and green, with solid blue lines indicating positive links and red dashed lines indicating negative links. (b) The similarity between detected partitions and the ground truth, measured by normalized mutual information (NMI, details can be found in Methods 5.3), as well as balanced ratio ($A^+, D^{(-)}$), is analyzed as a function of the ratio of P^-/P^+ for different partitioning methods. The dashed grey line represents the boundary where $P^- = \epsilon^-$. To the left of this line, the generated network is balanced, with all components in $(A^+, D^{(-)})$, indicated in blue. To the right of the line, the generated network becomes unbalanced, with all components in $(A^+, D^{(-)})$, highlighted in yellow. In these settings, the SBMs can successfully detect both $(A^+, D^{(-)})$ and $(A^+, A^{(-)})$, while other methods could not. Each maker is the average of 100 independent runs. Other parameters: $P^+ = 0.2, \epsilon^+ = \epsilon^- = 0.01, N = 180$. (c)(d) are the partitioned networks and blocks marked in red triangles in (b).

between students across the three years of high-school in a Spanish school. The results in Fig. 5 show that, without using any age information, the partitions obtained from SBMs align with the three distinct age groups with resulting assortative relationships between any two groups, in both positive and negative interactions ($A^+, A^{(-)}$). This suggests a picture in which links are driven by the interaction of students within their age groups, regardless of the signs of the ties, while interactions between different age groups are relatively scarce. In comparison, the groups detected by Louvain involve a mixture of age groups and are mostly in balanced relationships ($A^+, D^{(-)}$) with each other. Similar results reported in Fig. 6, obtained in synthetic networks that mimic the Spanish high school dataset, confirm that the balanced structures found by

Louvain are a consequence of the limitations of the method. Altogether, these results show that assortativity (regardless of interaction sign) is a prevailing pattern in signed networks. This pattern is masked by the use of heuristic descriptive methods (such as Louvain), which find balanced network partitions ($A^+, D^{(-)}$) that can be wrongly interpreted as a confirmation of the predictions of social balance theory.

3.3 Core-periphery is typical in social networks

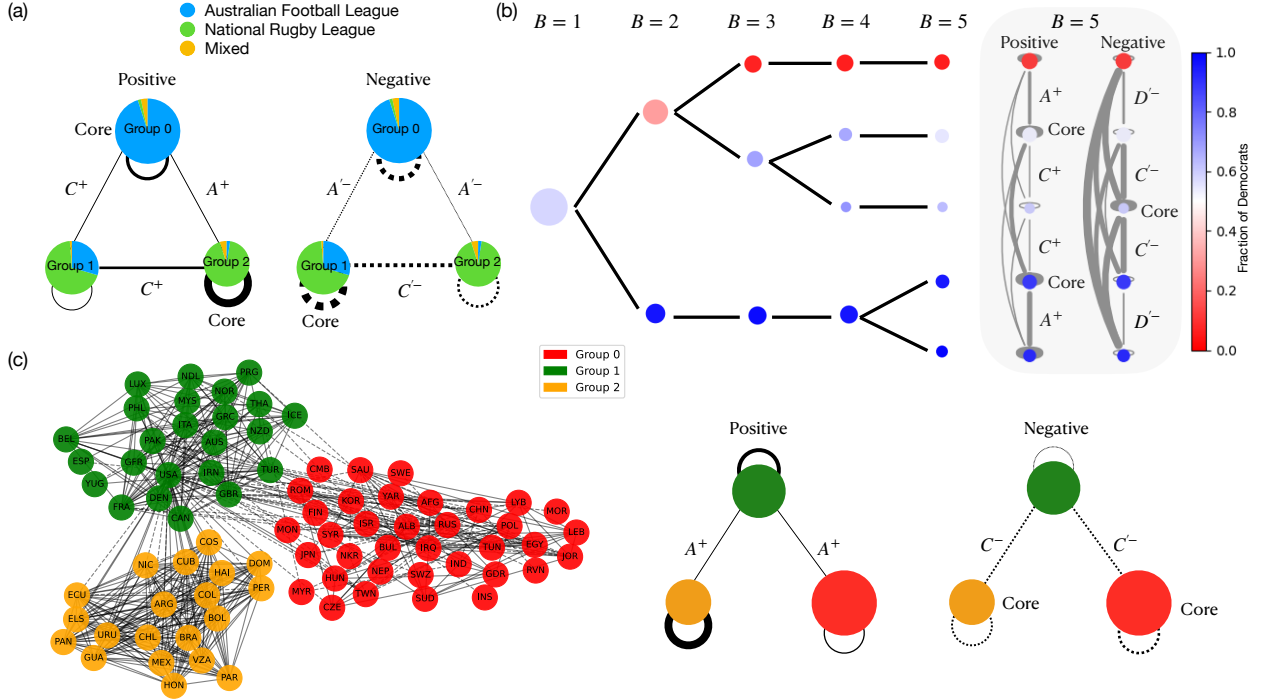


Figure 7: Examples of core-periphery structures in signed networks. (a) Twitter signed network [3]. The graphical representation of the edge density matrix corresponding to the three top-level groups identified by WSBM. Each circle (node) represents a community, and the colored sectors within each node illustrate the topic composition: Australian Football League (AFL), National Rugby League (NRL), and mixed-interest discussions. (b) USA Congress H100 [45]. The partition into groups was obtained by the LSBM method with an increasing number of communities $B \in [0, 5]$. The plot shows schematically how communities divide and the type of pairwise interaction they share. The diagram on the right displays the positive and negative density matrix for $B = 5$. Colors indicate party affiliation of group members. (c) Correlates of Wars for the years 1954-57 [46]. It shows the top-level network partitioned by WSBM (left) and the graphical representation of the edge density matrix for the communities (median and right). The label on each node is the name of the region. In all plots, the size of each node reflects the size of the group, solid lines represent positive edges, and dashed lines represent negative edges.

Another frequent unbalanced structure in our survey of empirical networks in Sec. 3.1 is the core-periphery structure. In such structures, the core represents a dense cluster of nodes with strong internal connectivity, while the periphery consists of loosely connected nodes that rely on the core for interaction [42–44]. For positive interactions in social-media networks, the core can be thought of as consisting of influencers or highly active users who act as information creators, whereas the periphery primarily consumes information. In contrast, for negative interactions, the core could be interpreted as the users who experience more conflict and mixed opinions, as its members are deeply engaged in debates. In Fig. 7, we explore the appearance of core-periphery patterns in three empirical examples:

- (a) A social-media network of 3,883 Twitter users discussing sports in Australia, with positive edges representing “like” and negative edges representing opposing replies [3]. Our analysis of the network partition (Fig. 7a) reveals two core-periphery structures in positive interactions and one core-periphery structure in negative interactions. Analyzing the users within each community, we observe that users

in the positive core communities tend to engage in more “like” actions, whereas users in the negative core are more active in the debate. Looking at the topical interest of the users, groups 0 and 2 are each dominated by a different sport (Australian Football League and National Rugby League, respectively), explaining their mutual assortative relationship. More interestingly, Group 1 is a mixed-topic group that acts as a periphery in positive interactions and a core in negative interactions, highlighting how interaction type shapes structural roles.

- (b) A well-known political network of the Congress of the United States of America (USA), constructed based on the co-sponsorship patterns of bills in the 100th House of Representatives [45]. Nodes are members of the House of Representatives, and edges represent significant positive or negative tendencies to co-sponsor bills. Our analysis of the network partition with different numbers of groups (Fig. 7b), obtained without using party affiliation, finds schematically how groups divide and uncovers several core-periphery patterns: for positive interactions, groups composed of members of the same party often form the core, exerting influence over peripheral groups with members of both parties. In contrast, for negative interactions, groups with mixed party membership frequently serve as cores, shaping debates and influencing sparser peripheral groups. This pattern reflects how bipartisan clusters can function as structural cores in legislative networks.
- (c) The well-studied Correlates of War network for the years 1954-57 [46]. Nodes are regions worldwide, positive edges denote alliances and memberships, and negative edges indicate conflicts. The partition (Fig. 7c) reveals three main groups in assortative relationships in the positive interactions, but in core-periphery relationships in the negative interactions. Notably, a large group of regions involving Europe, Africa, and Asia (Group 0) serves as the largest core in the negative interactions with the group, which includes the USA and most of Western Europe (Group 1).

These examples reveal a common pattern: core-periphery relationships often appear in the negative interactions with negative cores often composed by diverse groups which accommodate internal conflicts (becoming the main arena for conflict), sit at the center of disputes, and connect different groups. In other words, they act as mixed hubs that concentrate antagonistic ties and mediate interactions across opposing factions. This structural role suggests that unbalanced structures, such as mixed cores in negative interactions, are critical for understanding polarization, negotiation, and the dynamics of disagreement in social, political, and interstate networks [47].

3.4 Micro and meso scales exhibit different balance

Structural balance is a scale-dependent property [22]. Although perfectly balanced or anti-balanced states are equivalent across micro and meso scales under strong balance theory (Fig. 1), measuring imperfect balance reveals important differences. We quantify balance at two levels: local triadic balance (degree of balance, DoB) at the micro-scale and overall frustration at the meso-scale. Contrary to the expectation that local balance implies global balance, Fig. 8a shows only a weak correlation between them (i.e., a weak negative correlation between the micro-scale degree of balance and the meso-scale frustration). This result indicates that micro-scale and meso-scale balance are not interchangeable: local structural balance does not reliably predict global frustration. To isolate the effect of local structure from meso-scale organization, we introduce null models sampled from SBMs as a baseline for comparison. These null models preserve the meso-scale block structures and edge sign densities, allowing us to compute the z-score for observed DoB. As shown in Fig. 8b, most empirical networks have a significantly higher micro-scale balance than expected under the fixed meso-scale structures. Overall, these results confirm that the micro-level and meso-level balance are distinct dimensions of signed network structure, with networks with similar triad balance exhibiting vastly different levels of meso-scale balance (overall frustration).

4 Discussion

Social balance theory has been a cornerstone of social network analysis for over half a century [13, 14], providing a foundational framework for the study of signed networks, where positive ties denote friendship or alliance and negative ties represent enmity or conflict. Traditionally, quantitative studies of signed networks based on this theory focused on the triad level, aligning with the popular adage, “the friend of my friend is my friend” and “the enemy of my enemy is my friend”. While this micro-level analysis has provided

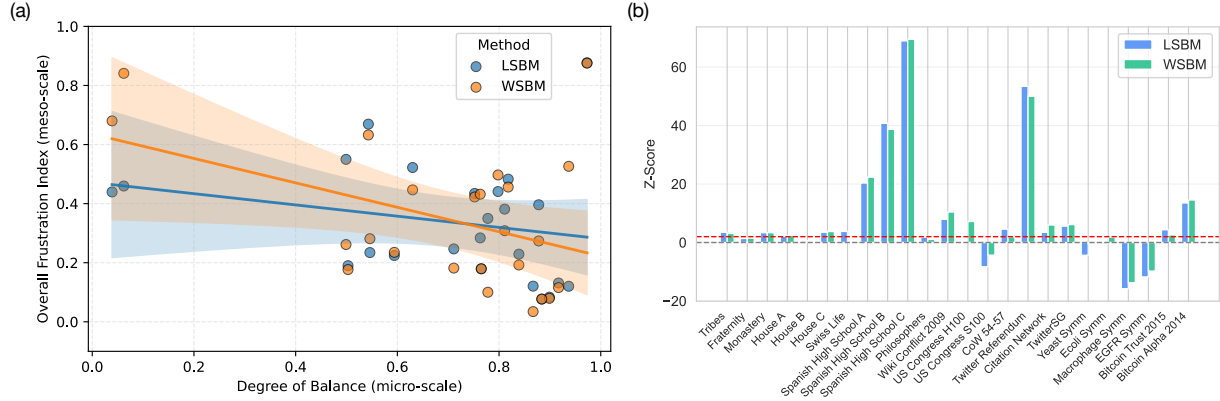


Figure 8: **Micro-scale balance vs. meso-scale frustration.** (a) Negative relationship between DoB (fraction of balanced triads) and overall frustration index. Scatter plots show data points of each network with linear regression fits and 95% confidence intervals (shaded regions). (WSBM: $R^2 = 0.180, p = 0.039$; LSBM: $R^2 = 0.054, p = 0.273$). (b) Z-scores quantify the deviation of observed DoB from null models preserving meso-scale block structure. The red dashed line is when z -score = 2. Even when meso-scale structure is fixed, micro-scale balance is significantly overrepresented in real networks, indicating that micro-level balance and meso-level balance are not interchangeable. Other relationships between meso-scale (im)balance and micro-scale balance also support this result (Fig. S4). The computation of DoB and z -score refers to Methods 5.5.

valuable insights, it has an emphasis on local configurations and overlooks the complexity of the connectivity present in empirical networks, which includes different global and meso-scale structures. To address this gap, we proposed and applied a methodology to characterize meso-scale structures in signed networks that does not a priori assume that structures are balanced, being thus suitable to evaluate the extent into which the observations are explained by structural balance theory and to reveal new (unbalanced) structures shaping the network.

The results we obtained applying our methodology to 24 empirical networks, including social, political, citation, financial, and biological networks, revealed that micro- and meso-scale structures in signed networks can show a diverse picture on the extent into which they align with social-balance predictions. Most importantly, we find that unbalanced structures are common in signed networks. In many empirical networks, particularly social and political ones, the dominant structure is not balanced but instead it is assortative in both positive and negative interactions. This suggests that individuals often interact within their social communities, regardless of the nature of their connections. In other words, the signs of interaction emerge from communities that frequently interact. Additionally, we find core-periphery structures to be common in signed social networks. This includes the expected cases of well-connected influencers who create content consumed by periphery, but also a new pattern in which a conflict (adversarial) network of mixed-group members constitutes the core of a core-periphery relationship.

Future research should question assumptions of balance and consider the presence of unbalanced meso-scale structures in network representations. As our findings suggest, we can reasonably expect that studies with signed networks in other areas will be similarly successful in identifying diverse structures and providing new insights not only on the role of specific nodes but also on the organization and function of meso-scale communities and the network as a whole. In particular, the application of our results to online social-media debates has the potential to reveal meso-scale structures beyond opinion polarization between two groups. These analyses may require a generalization of our methods to directed signed networks [43, 48] (i.e., where influence may not be reciprocal) and moving beyond a simple pairwise classification of community types to capture more multifunctional structures.

5 Methods

5.1 Data preprocessing

We collected three types of datasets: 18 social-political networks, 4 gene-regulatory networks, and 2 financial networks. A basic description of these networks can be found in the Supplementary Materials. Some networks are directed (e.g., *Spanish High School*, *TwitterSG*), and we need to convert them to simple undirected networks for our analysis. We follow the rules proposed in Ref. [49]: if two nodes share the same sign in both directions, the undirected connection preserves that sign; if the connection is non-reciprocal, we symmetrize by adopting the existing sign; if the signs are opposite, the undirected edge is negative. For graphs with a weighted configuration (e.g., *Spanish High School*, *Monastery*), we binarized the weights by assigning +1 to each positive edge and -1 to each negative edge.

Formally, given a directed signed adjacency matrix $\mathbf{S} = (S_{ij}) \in \mathbb{R}^{N \times N}$ where $S_{ij} > 0$ indicating a positive edge $i \mapsto j$, $S_{ij} < 0$ a negative edge, and $S_{ij} = 0$ no edge, we construct a simple undirected signed adjacency matrix $\mathbf{A} = (A_{ij}) \in \mathbb{R}^{N \times N}$, using the symmetrization rule:

$$A_{ij} = \begin{cases} 0, & S_{ij} = 0 \text{ and } S_{ji} = 0, \\ -1, & S_{ij}S_{ji} < 0, \\ \text{sgn}(S_{ij} + S_{ji}), & \text{otherwise,} \end{cases} \quad (3)$$

where $\text{sgn}(x) = +1$ if $x > 0$, and $\text{sgn}(x) = -1$ if $x < 0$.

After this pre-processing, we calculated the descriptive statistics for the largest connected component of each network. The results are summarized in Table S1.

5.2 Frustration index

To measure the structural (im)balance on the mesoscopic scale, we use the frustration index [6, 23] calculated based on the fraction of misplaced edges according to classical balance theory, i.e., the positive edges between communities, and the negative edges within communities. We use two frustration indices in the main text. The first calculates the pairwise imbalance between communities, and the second calculates the overall imbalance of the partition. Given a pair of communities i and j , the pairwise frustration index is computed as

$$\text{FI}_{ij} = \frac{L_{ij}^+ + L_{ii}^- + L_{jj}^-}{L_{ii} + L_{jj} + L_{ij}}, \quad (4)$$

where L_{ij}^+ represents the number of positive edges between community i and j , L_{ii}^- represents the number of negative edges within community i , L_{ij} represents the total number of edges between community i and community j , and L_{ii} represents the total number of edges within the community i .

Given a partition with B communities, the overall frustration index is defined as [6, 23]:

$$\text{FI} = \frac{\sum_{i=1}^B L_{ii}^- + \sum_{i=1}^B \sum_{j(>i)} L_{ij}^+}{L}, \quad (5)$$

where $\sum_{i=1}^B L_{ii}^-$ represents the total number of negative edges within each community, $\sum_{i=1}^B \sum_{j(>i)} L_{ij}^+$ the total number of positive edges between communities, and L the total number of edges overall. $\text{FI} \in [0, 1]$, with $\text{FI} = 1$ indicating complete anti-balance and $\text{FI} = 0$ perfect balance.

5.3 NMI

The normalized mutual information (NMI) is a widely used metric to measure the similarity of detected communities [50, 51]. Given a network with N nodes, the partition $p = \{p_1, \dots, p_B\}$ with B communities, and the ground truth $u = \{u_1, \dots, u_A\}$ with A groups, we calculate the NMI as:

$$\text{NMI} = \frac{2 \sum_{i=1}^A \sum_{j=1}^B \frac{|p_i \cap u_j|}{N} \log \left(\frac{\frac{|p_i \cap u_j|}{N}}{\frac{|p_i|}{N} \cdot \frac{|u_j|}{N}} \right)}{- \sum_{i=1}^A \frac{|p_i|}{N} \log \frac{|p_i|}{N} - \sum_{j=1}^B \frac{|u_j|}{N} \log \frac{|u_j|}{N}}, \quad (6)$$

which corresponds to the mutual information between p and u , normalized by the combined entropy. $\text{NMI} \in [0, 1]$, with $\text{NMI} = 0$ indicating perfect alignment and $\text{NMI} = 0$ no alignment (zero mutual information).

5.4 Robustness

Our community structure classification is computed for a given partition of the network. To have a better understanding of the statistical robustness of our classification, we adopt a bootstrapping method to estimate the robustness of the classification in the presence of small perturbations of the underlying network [43].

Given a signed graph G with the number of positive edges L^+ and negative edges L^- , and the partition p with its corresponding pairwise community classification matrix M , we generate a new graph G' by randomly choosing L^+ and L^- edges with replacement from the existing positive and negative edges (respectively) in G . We compute the positive and negative density matrix, w'^{+} and w'^{-} , based on G' and p . And we obtain the new interaction classification matrix M' . We repeat the process k times and compute the certainty P_{rs} for each pair of communities r and s , as the fraction of all k interaction classifications that are equal to the original type

$$P_{rs} = \frac{1}{k} \sum_1^k \delta(M_{rs}, M'_{rs}), \quad (7)$$

where $\delta(a, b) = 1$ if $a = b$ and $\delta(a, b) = 0$ if $a \neq b$.

5.5 Balanced triads

In order to measure and compare the degree of balance (DoB) [4, 13, 20–22] of the empirical networks here, we calculate the fraction of (weakly) balanced triads, which includes triads with all positive edges T_{+++} , all negative edges T_{---} , and two negative edges T_{+--} . The fraction of triads is calculated by adjacency matrix $\mathbf{A} = \mathbf{A}^+ + \mathbf{A}^-$ ($\mathbf{A}^+ \equiv \{A_{ij}^+\}^N$, $\mathbf{A}^- \equiv \{A_{ij}^-\}^N$):

$$\begin{aligned} T_{+++} &= \frac{\text{Tr}[(\mathbf{A}^+)^3]}{\text{Tr}[(\mathbf{A}^+ - \mathbf{A}^-)^3]}, \\ T_{---} &= \frac{\text{Tr}[(-\mathbf{A}^-)^3]}{\text{Tr}[(\mathbf{A}^+ - \mathbf{A}^-)^3]}, \\ T_{+--} &= \frac{3\text{Tr}[\mathbf{A}^+(-\mathbf{A}^-)^2]}{\text{Tr}[(\mathbf{A}^+ - \mathbf{A}^-)^3]}, \\ T_{++-} &= \frac{3\text{Tr}[(\mathbf{A}^+)^2(-\mathbf{A}^-)]}{\text{Tr}[(\mathbf{A}^+ - \mathbf{A}^-)^3]}. \end{aligned} \quad (8)$$

The degree of balance (DoB) is then defined as the fraction of balanced triads $\text{DoB} = T_{+++} + T_{+--}$, and the weak degree of balance (WDoB) is then $\text{WDoB} = T_{+++} + T_{+--} + T_{---}$.

However, most networks have more positive edges, so that a conclusion based only on this fraction is not sufficient to quantify balance. We thus compare the number of standard deviations that our observed balance deviates from what we would expect in a random network that acts as a null model [5, 52]. This model is created by fixing the edges and shuffling their signs [22]. We then adopt the z-score [49] to measure the deviations,

$$z(m) = \frac{[N(m) - \langle N(m) \rangle]}{\sigma(m)}, \quad (9)$$

where $N(m)$ is the number of pattern m (different kinds of triads) appears in the empirical network, $\langle N(m) \rangle$ is the expected number of the same pattern under the null model, and $\sigma(m) = \sqrt{\langle N^2(m) \rangle - \langle N(m) \rangle^2}$ is the standard deviation of $N(m)$ under the null model. $z(m) > 2$ indicates that the pattern m is overrepresented, meaning it occurs significantly more than in the null model. $z(m) < -2$ indicates that the pattern m is underrepresented, meaning it occurs significantly less than in the null model.

Acknowledgments

We thank Fakhteh Ghanbarnejad for useful discussions.

Competing interests: There are no competing interests to declare.

Funding: This project was funded by the Australian Research Council, project DP240100872.

Data and materials availability: The empirical network data that support the findings of this study were derived from citation links in the Supplementary Materials. The code to reproduce our results and apply our methodology to other networks will be made publicly available upon formal acceptance of the manuscript.

References

- [1] Bruno Ordozgoiti, Antonis Matakos, and Aristides Gionis. Finding large balanced subgraphs in signed networks. In *Proceedings of The Web Conference 2020*, pages 1378–1388, 2020.
- [2] Miguel Ruiz-Garcia, Juan Ozaita, María Pereda, Antonio Alfonso, Pablo Brañas-Garza, José A Cuesta, and Angel Sánchez. Triadic influence as a proxy for compatibility in social relationships. *Proceedings of the National Academy of Sciences*, 120(13):e2215041120, 2023.
- [3] John Pougué-Biyong, Akshay Gupta, Aria Haghighi, and Ahmed El-Kishky. Learning stance embeddings from signed social graphs. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, pages 177–185, 2023.
- [4] Szymon Talaga, Massimo Stella, Trevor James Swanson, and Andreia Sofia Teixeira. Polarization and multiscale structural balance in signed networks. *Communications Physics*, 6(1):349, 2023.
- [5] Bingjie Hao and István A Kovács. Proper network randomization is key to assessing social balance. *Science Advances*, 10(18):eadj0104, 2024.
- [6] Anna Gallo, Diego Garlaschelli, and Tiziano Squartini. Assessing frustration in real-world signed networks: A statistical theory of balance. *Physical Review Research*, 6(4):L042065, 2024.
- [7] Anna Gallo, Fabio Saracco, Renaud Lambiotte, Diego Garlaschelli, and Tiziano Squartini. Patterns of link reciprocity in directed, signed networks. *Physical Review E*, 111(2):024312, 2025.
- [8] Fernando Diaz-Diaz, Elena Candellone, Miguel A. González-Casado, Emma Fraxanet, Antoine Vendeville, Irene Ferri, and Andreia Sofia Teixeira. Signed networks: theory, methods, and applications. *arXiv preprint arXiv:2511.17247*, 2025.
- [9] Santo Fortunato. Community detection in graphs. *Physics Reports*, 486(3-5):75–174, 2010.
- [10] Vincent A Traag and Jeroen Bruggeman. Community detection in networks with positive and negative links. *Physical Review E*, 80(3):036115, 2009.
- [11] Peter J Mucha, Thomas Richardson, Kevin Macon, Mason A Porter, and Jukka-Pekka Onnela. Community structure in time-dependent, multiscale, and multiplex networks. *Science*, 328(5980):876–878, 2010.
- [12] Tiago P Peixoto. Inferring the mesoscale structure of layered, edge-valued, and time-varying networks. *Physical Review E*, 92(4):042807, 2015.
- [13] Dorwin Cartwright and Frank Harary. Structural balance: a generalization of heider’s theory. *Psychological Review*, 63(5):277, 1956.
- [14] Fritz Heider. Attitudes and cognitive organization. *The Journal of Psychology*, 21(1):107–112, 1946.
- [15] Frank Harary. Structural duality. *Behavioral Science*, 2(4):255–265, 1957.
- [16] Yu Tian and Renaud Lambiotte. Spreading and structural balance on signed networks. *SIAM Journal on Applied Dynamical Systems*, 23(1):50–80, 2024.
- [17] Frank Harary. On the notion of balance of a signed graph. *Michigan Mathematical Journal*, 2(2):143–146, 1953.

- [18] Patrick Doreian and Andrej Mrvar. A partitioning approach to structural balance. *Social Networks*, 18(2):149–168, 1996.
- [19] Samin Aref, Ly Dinh, Rezvaneh Rezapour, and Jana Diesner. Multilevel structural evaluation of signed directed social networks based on balance theory. *Scientific Reports*, 10(1):15228, 2020.
- [20] Ernesto Estrada and Michele Benzi. Walk-based measure of balance in signed networks: Detecting lack of balance in social networks. *Physical Review E*, 90(4):042802, 2014.
- [21] Ranveer Singh and Bibhas Adhikari. Measuring the balance of signed networks and its application to sign prediction. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(6):063302, 2017.
- [22] Alec Kirkley, George T Cantwell, and Mark EJ Newman. Balance in signed networks. *Physical Review E*, 99(1):012320, 2019.
- [23] Samin Aref and Mark C Wilson. Balance and frustration in signed networks. *Journal of Complex Networks*, 7(2):163–189, 2019.
- [24] Mark EJ Newman and Michelle Girvan. Finding and evaluating community structure in networks. *Physical Review E*, 69(2):026113, 2004.
- [25] Alessia Amelio and Clara Pizzuti. Community mining in signed networks: a multiobjective approach. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 95–99, 2013.
- [26] Vincent A Traag, Rodrigo Aldecoa, and J-C Delvenne. Detecting communities using asymptotical surprise. *Physical Review E*, 92(2):022816, 2015.
- [27] John N Pougué-Biyong and Renaud Lambiotte. Signedlouvain: Louvain for signed networks. *arXiv preprint arXiv:2407.19288*, 2024.
- [28] Jérôme Kunegis, Stephan Schmidt, Andreas Lommatzsch, Jürgen Lerner, Ernesto W De Luca, and Sahin Albayrak. Spectral analysis of signed graphs for clustering, prediction and visualization. In *Proceedings of the 2010 SIAM International Conference on Data Mining*, pages 559–570. SIAM, 2010.
- [29] Kai-Yang Chiang, Joyce Jiyoun Whang, and Inderjit S Dhillon. Scalable clustering of signed networks using balance normalized cut. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, pages 615–624, 2012.
- [30] Jérôme Kunegis. Applications of structural balance in signed social networks. *arXiv preprint arXiv:1402.6865*, 2014.
- [31] Mihai Cucuringu, Peter Davies, Aldo Glielmo, and Hemant Tyagi. Sponge: A generalized eigenproblem for clustering signed networks. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 1088–1098. PMLR, 2019.
- [32] Jörg Reichardt and Stefan Bornholdt. Statistical mechanics of community detection. *Physical Review E*, 74(1):016110, 2006.
- [33] Giuseppe Facchetti, Giovanni Iacono, and Claudio Altafini. Exploring the low-energy landscape of large-scale signed social networks. *Physical Review E*, 86(3):036116, 2012.
- [34] Tiago P Peixoto. *Descriptive vs. inferential community detection in networks: Pitfalls, myths and half-truths*. Cambridge University Press, 2023.
- [35] Jonathan Q Jiang. Stochastic block model and exploratory analysis in signed networks. *Physical Review E*, 91(6):062805, 2015.
- [36] Roger Guimerà. One model to rule them all in network science? *Proceedings of the National Academy of Sciences*, 117(41):25195–25197, 2020.
- [37] Leto Peel, Tiago P Peixoto, and Manlio De Domenico. Statistical inference links data and theory in network science. *Nature Communications*, 13(1):6794, 2022.

- [38] Tiago P. Peixoto and Alec Kirkley. Implicit models, latent compression, intrinsic biases, and cheap lunches in community detection. *Physical Review E*, 108:024309, 2023.
- [39] Brian Karrer and Mark EJ Newman. Stochastic blockmodels and community structure in networks. *Physical Review E*, 83(1):016107, 2011.
- [40] Tiago P Peixoto. Nonparametric weighted stochastic block models. *Physical Review E*, 97(1):012306, 2018.
- [41] Tiago P Peixoto. Bayesian stochastic blockmodeling. *Advances in network clustering and blockmodeling*, pages 289–332, 2019.
- [42] Richard F. Betzel, Maxwell A. Bertolero, and Danielle S. Bassett. Non-assortative community structure in resting and task-evoked functional brain networks. *bioRxiv*, 355016, 2018.
- [43] Cathy Xuanchi Liu, Tristram J Alexander, and Eduardo G Altmann. Nonassortative relationships between groups of nodes are typical in complex networks. *PNAS nexus*, 2(11):pgad364, 2023.
- [44] Puck Rombach, Mason A. Porter, James H. Fowler, and Peter J. Mucha. Core-periphery structure in networks (revisited). *SIAM Review*, 59(3):619–646, 2017.
- [45] Zachary Neal. The backbone of bipartite projections: Inferring relationships from co-authorship, co-sponsorship, co-attendance and other co-behaviors. *Social Networks*, 39:84–97, 2014.
- [46] Patrick Doreian and Andrej Mrvar. Structural balance and signed international relations. *Journal of Social Structure*, 16:1, 2015.
- [47] Omid Askarisichani, Jacqueline Ng Lane, Francesco Bullo, Noah E Friedkin, Ambuj K Singh, and Brian Uzzi. Structural balance emerges and explains performance in risky decision-making. *Nature Communications*, 10(1):2648, 2019.
- [48] Bingjie Hao, Elijah Platnick, and István A Kovács. Social balance in directed networks. *Communications Physics*, 8(1):299, 2025.
- [49] Anna Gallo, Diego Garlaschelli, Renaud Lambiotte, Fabio Saracco, and Tiziano Squartini. Testing structural balance theories in heterogeneous signed networks. *Communications Physics*, 7(1):154, 2024.
- [50] LN Fred Ana and Anil K Jain. Robust data clustering. In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, volume 2, pages II–II. IEEE, 2003.
- [51] Helcio Felipe, Federico Battiston, and Alec Kirkley. Network mutual information measures for graph similarity. *Communications Physics*, 7(1):335, 2024.
- [52] Michael Szell, Renaud Lambiotte, and Stefan Thurner. Multirelational organization of large-scale social networks in an online world. *Proceedings of the National Academy of Sciences*, 107(31):13636–13641, 2010.

Supplementary Materials

Data description

We considered three kinds of datasets: social-political networks, financial networks, and gene-regulatory networks. Although some networks originate from similar contexts, they vary substantially in structural characteristics such as size, edge density, and the ratio of negative to positive edges. These differences allow us to capture heterogeneity within the same domain and test the robustness of our method across diverse network configurations. The statistics and basic description are summed up in Table S1. The majority of empirical networks here is characterized by a small density $c = 2L/N(N - 1)$ but a small fraction of negative links $L - /L+$.

There are 18 social-political networks. *New Guinea Tribes* [1] reflected political alliances and oppositions among 16 tribes, developed through fieldwork involving observations and conversations with leaders and members in the Central Highlands. *Fraternity* [2] was created from a 15-week survey of 17 fraternity brothers at the University of Michigan in 1956, representing relationships with positive or negative edge values. *Monastery* [3] was collected via a questionnaire asking each monk to list “3 brothers who you like the most and the least” among 18 monks at an American monastery. *College House A, B, C* [4] come from a small group survey on evaluations of other group members (sorority sisters) based on a range of behavioral characteristics. *Swiss Student* [5] collected information about the relationships among Swiss students who shared most of their classes and often did coursework in groups. *Spanish High School A, B, C* [6] gathers data on student relationships by having students name their connections and rate them from -2 to $+2$. *Philosophers* [7] compiled master-pupil and acquaintanceship ties among philosophers from 800 B.C.E to 1935 C.E., based on historical texts in Randall Collins’ seminal book. *Wiki Conflict* [8] represents conflicts between English Wikipedia users, where nodes represent users and edges indicate positive or negative interactions. A negative interaction occurs when one user reverts another’s edit. *U.S. Congress* [9] illustrates bill co-sponsorship tendencies in the House and Senate. This signed network is derived from a bipartite network using the Stochastic Degree Sequence Model [9]. Edges indicate significant tendencies to co-sponsor or not, with positive and negative relationships inferred using the Stochastic Degree Sequence Model. *Correlates of Wars* [10] offers data on international political relationships from 1946 to 1997, updated every four years. Positive edges indicate alliances or political agreements, while negative edges signify enmities or disagreements. *Twitter Referendum* [11] is an undirected signed network capturing Twitter user sentiments on the 2016 Italian Referendum, where edges are weighted -1 for opposing stances and $+1$ for the same stance. *Signed Citation network* [12] consists of 8,711 papers from the field of Computational Linguistics. The positive (endorsement), negative (criticism), and neutral citations are classified by keyword-based citation sentiment. Following [13], we combine neutral and positive citations as positive, since authors do not express clear opposing views in neutral citations. *TwitterSG* [14] is a signed network drawn from Twitter interactions between 20th May and 8th August 2021. Here, we select the popular topics of the Australian Football League and the National Rugby League. A positive connection exists from user A to user B if user A likes a tweet posted by user B. Conversely, a negative connection exists from user A to user B if user A expresses opposition to user B’s tweet in a reply.

In gene-regulatory networks [15], *Yeast* and *Escherichia coli* are classified as transcriptional networks, while the networks involving the Epidermal Growth Factor Receptor (*EGFR*) and *Macrophage* are signaling networks. The *Yeast* network relates to *Saccharomyces cerevisiae*, and the *Escherichia coli* network pertains to *Escherichia coli*. *Macrophage* are white blood cells that eliminate harmful substances like cancer cells, while *EGFR* is linked to a protein that promotes cell division and survival in epidermal tissues.

The financial networks consist of Bitcoin traders, with *Bitcoin OTC* and *Bitcoin Alpha* [16] built on user ratings of trust on trading platforms, ranging from -10 for fraudsters to $+10$ for complete trust.

Our dataset consists primarily of social networks, reflecting the availability of signed network data in this domain. To assess whether this overrepresentation affects our findings, we examined within each domain separately (Figs. S1-S3) and found that the result, that unbalanced structures are common, was qualitatively similar across domains. This suggests that the observed patterns are not driven solely by social networks but are consistent across different contexts.

Table S1 Descriptive statistics of the largest component of empirical networks.

The table shows the total number of nodes N , the total number of edges L , the ratio of negative and positive edges L_-/L_+ , the average degree $\langle k \rangle$, the network density $c = 2L/N(N-1)$, and T_{+++} , T_{++-} , T_{+--} and T_{---} are the fractions of 3-cycles (triads) with all positive edges, one negative edge, two negative edges, and all negative edges, respectively. Their corresponding z-scores are shown in parentheses (see Methods 5.5). DoB is the fraction of balanced triads, that is, the fraction of triads with even negative edges. Most empirical networks exhibit strong negative z-scores for triads with two positive edges. This indicates that unbalanced triads are underrepresented. These findings provide strong support for classical balance theory, as this “forbidden” configuration is notably avoided. For social-political networks, the results consistently show positive z-scores for most positive triads, but mixed z-scores for the “enemy’s enemy” triad (+ − −) and the “all enemies” triad (− − −). This suggests that some networks strongly adhere to the balance principle, while others do not. For example, the *Spanish High School A, B, C* shows a significant overrepresentation of the − − − configuration, which supports the weak balance theory. Gene-regulatory networks exhibit distinct properties that favor “enemy’s enemy” configurations while avoiding clusters of positive relationships, in contrast to social networks. Financial networks show weaker balance effects overall, suggesting different underlying mechanisms than pure social balance theory.

Name	N	L	$\frac{L_-}{L_+}$	$\langle k \rangle$	c	T_{+++} (z-score)	T_{++-} (z-score)	T_{+--} (z-score)	T_{---} (z-score)	DoB
New Guinea Tribes [1, 17]	16	58	1	7.25	0.48	0.279 (4.48)	0.029 (-6.28)	0.588 (3.70)	0.103 (-0.42)	0.867
Fraternity [2]	17	40	0.74	4.71	0.29	0.556 (4.51)	0 (-3.82)	0.222 (-0.89)	0.222 (2.73)	0.778
Monastery [3]	18	78	0.95	8.67	0.51	0.182 (1.89)	0.122 (-5.88)	0.583 (4.87)	0.113 (0.025)	0.765
College House A [4]	21	77	1.08	7.33	0.37	0.229 (4.08)	0.133 (-4.54)	0.482 (1.77)	0.157 (0.60)	0.711
College House B [4]	17	64	1.37	7.53	0.47	0.065 (-0.20)	0.221 (-1.85)	0.481 (0.91)	0.234 (1.23)	0.546
College House C [4]	20	65	1.03	6.50	0.34	0.211 (2.0)	0.038 (-4.93)	0.673 (4.63)	0.077 (-1.27)	0.884
Swiss Studentlife [5]	181	546	0.28	6.03	0.034	0.689 (6.73)	0.15 (-9.15)	0.15 (2.22)	0.011 (0.26)	0.839
Spanish High School A [6]	409	6,509	0.22	31.83	0.078	0.629 (16.36)	0.182 (-51.20)	0.169 (37.41)	0.020 (29.89)	0.798
Spanish High School B [6]	238	2,889	0.14	24.28	0.102	0.729 (5.84)	0.183 (-15.15)	0.082 (18.75)	0.005 (9.71)	0.811
Spanish High School C [6]	534	9,527	0.22	35.68	0.033	0.614 (14.67)	0.231 (-46.91)	0.138 (36.54)	0.017 (34.37)	0.752
Philosophers [7]	218	259	0.126	2.38	0.011	0.75 (0.38)	0.083 (-1.42)	0.167 (2.49)	0 (-0.14)	0.917
Wiki Conflict, 2009 [8, 17]	359	2,535	1.33	14.12	0.039	0.088 (0.36)	0.109 (-5.46)	0.675 (11.53)	0.128 (-1.41)	0.763
U.S. Congress, H100 [9, 17]	446	50,688	2.64	227	0.51	0.102 (460.54)	0.018 (-210.88)	0.492 (141.77)	0.388 (5.82)	0.594
U.S. Congress, S100 [9, 17]	101	2,143	2.52	42.44	0.42	0.116 (68.99)	0.02 (-34.48)	0.387 (-13.05)	0.478 (17.19)	0.503
Correlates of Wars 1954-57 [10]	80	492	0.178	12.3	0.16	0.869 (14.06)	0.093 (-16.91)	0.03 (-3.84)	0.008 (3.81)	0.899
Twitter Referendum [11]	10,864	251,396	0.054	46.28	0.004	0.938 (48.21)	0.027 (-71.41)	0.035 (150.56)	0 (-15.91)	0.973
Citations [12]	8,711	25,073	0.094	5.76	0.001	0.779 (1.32)	0.179 (-3.54)	0.039 (6.67)	0.003 (3.41)	0.818
TwitterSG [14]	3,883	5,480	0.952	2.82	0.001	0.495 (14.37)	0.292 (-2.90)	0.135 (-7.39)	0.078 (-1.63)	0.63
Yeast [15]	664	1,064	0.26	3.20	0.005	0.486 (-0.13)	0.371 (-0.33)	0.057 (-0.95)	0.086 (6.91)	0.543
Escherichia coli [15]	1,376	3,150	0.705	4.58	0.003	0.251 (1.23)	0.374 (-2.49)	0.248 (-1.43)	0.127 (2.76)	0.499
EGFR [15]	313	755	0.51	4.82	0.015	0.028 (-10.42)	0.873 (17.09)	0.034 (-8.96)	0.064 (2.60)	0.062

Table S1 continued from previous page

Name	N	L	$\frac{L_-}{L_+}$	$\langle k \rangle$	c	T_{+++} (z-score)	T_{++-} (z-score)	T_{+--} (z-score)	T_{---} (z-score)	DoB
Macrophage [15]	660	1,397	0.50	4.23	0.006	0.024 (-14.63)	0.806 (21.29)	0.015 (-12.62)	0.155 (16.55)	0.039
Bitcoin OTC, 2015 [16, 17]	348	665	0.12	3.82	0.011	0.908 (4.69)	0.063 (-5.38)	0.029 (-0.19)	0 (-0.56)	0.937
Bitcoin Alpha, 2014 [16, 17]	687	1,853	0.26	5.39	0.008	0.747 (9.61)	0.12 (-17.36)	0.131 (2.31)	0.002 (-2.71)	0.878

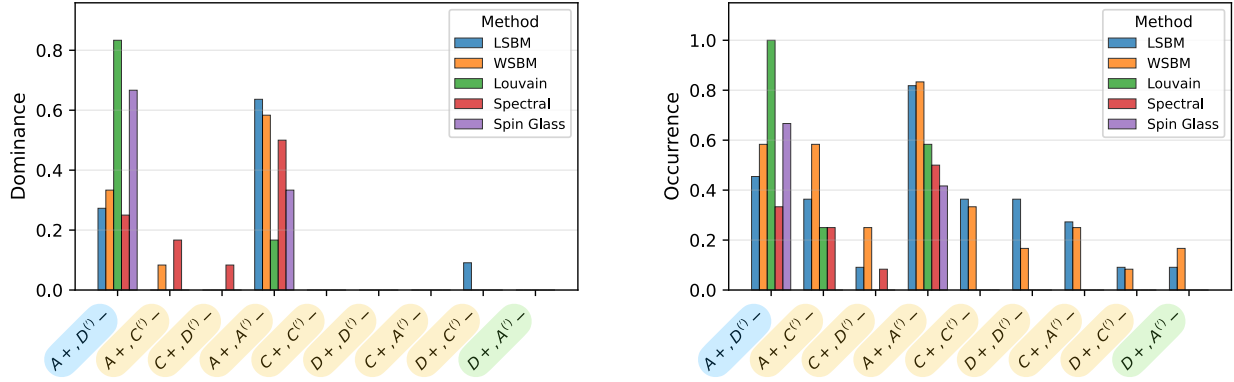


Figure S1: **Community types in social-political networks.** Left: fraction of networks in which each community type is dominant. Right: fraction of networks in which each community type occurs.

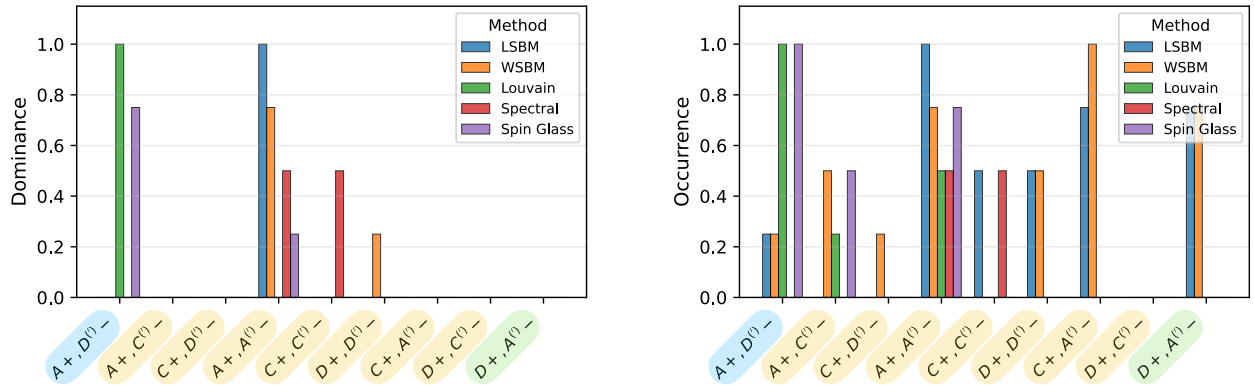


Figure S2: **Community types in gene-regulatory networks.** Left: fraction of networks in which each community type is dominant. Right: fraction of networks in which each community type occurs.

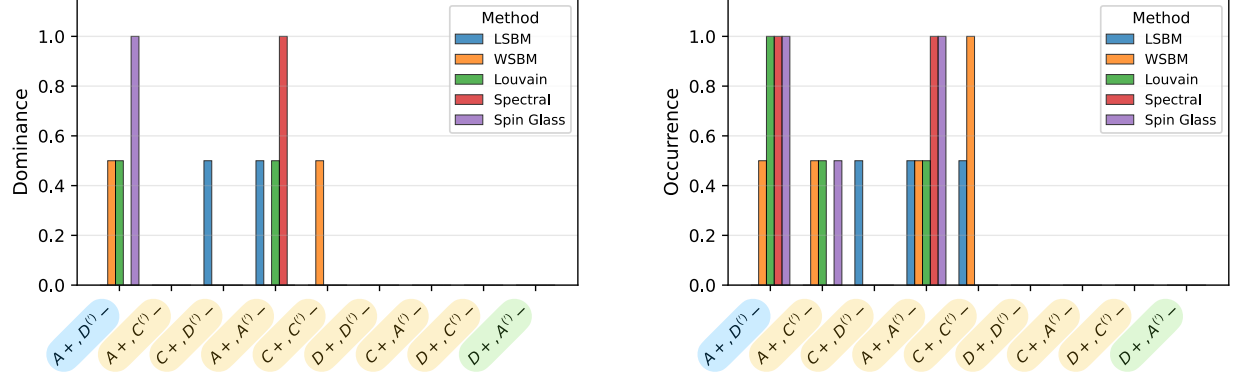


Figure S3: **Community types in financial networks.** Left: fraction of networks in which each community type is dominant. Right: fraction of networks in which each community type occurs.

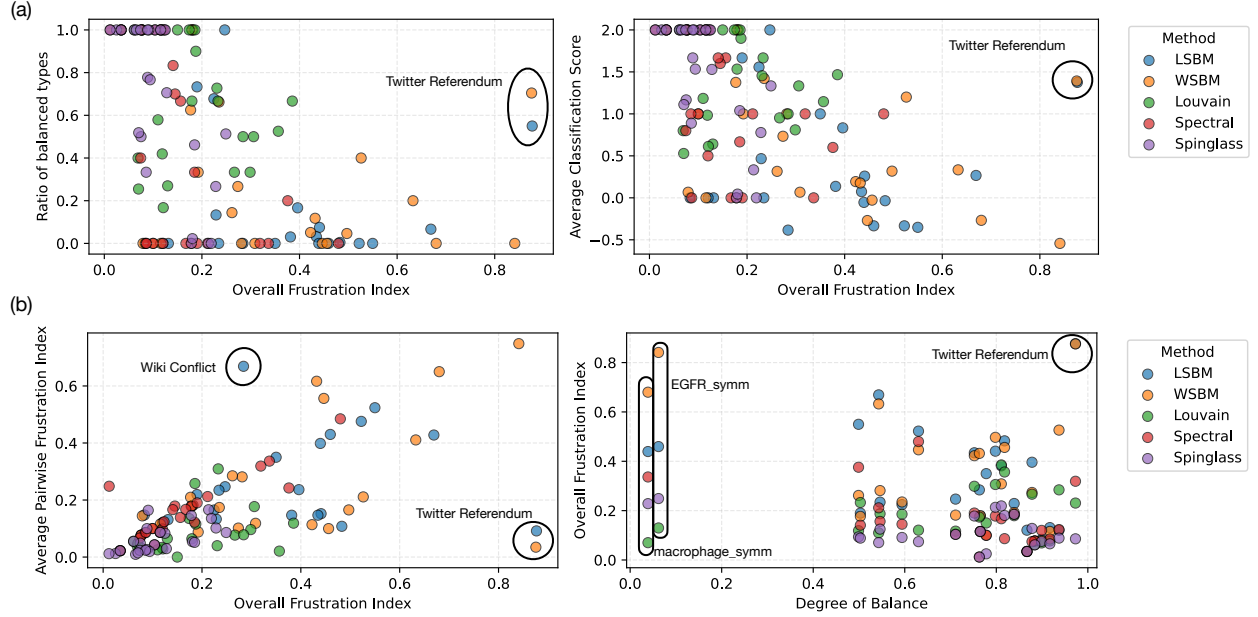


Figure S4: **Other relationships between meso-scale (im)balance and micro-scale balance.** (a) Relationship between structure classifications and the overall FI given the partition of the network. Classification of balanced structure types (left panel), and average balance score (right panel), with the prediction from traditional balance theory, qualified by the overall frustration index. The results show that our classification closely matches theoretical structural balance expectations. (b) Relationship between average pairwise FI, overall FI, and the ratio of balanced triads. The left panel confirms that the average pairwise frustration index could reflect the overall balance degree of the partition. In addition, the SBMs could find more unbalanced communities compared with other methods. The right panel shows that the ratio of balanced triads from a microscopic perspective could not reflect the balance degree of a given partition at a mesoscopic level.

Supplementary References

- [1] Kenneth E Read. Cultures of the central highlands, new guinea. *Southwestern Journal of Anthropology*, 10(1):1–43, 1954.
- [2] Theodore M. Newcomb. The acquaintance process as a prototype of human interaction. In *The acquaintance process*, pages 259–261. Holt, Rinehart & Winston, New York, NY, 1961.
- [3] Samuel Franklin Sampson. *A novitiate in a period of change: An experimental and case study of social*

relationships. Cornell University, 1968.

- [4] Thomas B Lemann and Richard L Solomon. Group characteristics as revealed in sociometric patterns and personality ratings. *Sociometry*, 15(1/2):7–90, 1952.
- [5] András Vörös, Zsófia Boda, Timon Elmer, Marion Hoffman, Kieran Mepham, Isabel J Raabe, and Christoph Stadtfeld. The swiss studentlife study: Investigating the emergence of an undergraduate community through dynamic, multidimensional social network data. *Social Networks*, 65:71–84, 2021.
- [6] Miguel Ruiz-Garcia, Juan Ozaita, María Pereda, Antonio Alfonso, Pablo Brañas-Garza, José A Cuesta, and Angel Sánchez. Triadic influence as a proxy for compatibility in social relationships. *Proceedings of the National Academy of Sciences*, 120(13):e2215041120, 2023.
- [7] Randall Collins. *The sociology of philosophies: A global theory of intellectual change*. Harvard University Press, 2009.
- [8] Ulrik Brandes, Patrick Kenis, Jürgen Lerner, and Denise Van Raaij. Network analysis of collaboration structure in wikipedia. In *Proceedings of the 18th International Conference on World Wide Web*, pages 731–740, 2009.
- [9] Zachary Neal. The backbone of bipartite projections: Inferring relationships from co-authorship, co-sponsorship, co-attendance and other co-behaviors. *Social Networks*, 39:84–97, 2014.
- [10] Patrick Doreian and Andrej Mrvar. Structural balance and signed international relations. *Journal of Social Structure*, 16:1, 2015.
- [11] Bruno Ordozgoiti, Antonis Matakos, and Aristides Gionis. Finding large balanced subgraphs in signed networks. In *Proceedings of The Web Conference 2020*, pages 1378–1388, 2020.
- [12] Srijan Kumar. Structure and dynamics of signed citation networks. In *Proceedings of the 25th International Conference Companion on World Wide Web*, pages 63–64, 2016.
- [13] Duoqi Song, Wenpei Wang, Ying Fan, Yanmeng Xing, and An Zeng. Quantifying the structural and temporal characteristics of negative links in signed citation networks. *Information Processing & Management*, 59(4):102996, 2022.
- [14] John Pougué-Biyong, Akshay Gupta, Aria Haghighi, and Ahmed El-Kishky. Learning stance embeddings from signed social graphs. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, pages 177–185, 2023.
- [15] Giovanni Iacono, Fahimeh Ramezani, Nicola Soranzo, and Claudio Altafini. Determining the distance to monotonicity of a biological network: a graph-theoretical approach. *IET Systems Biology*, 4(3):223–235, 2010.
- [16] Srijan Kumar, Francesca Spezzano, VS Subrahmanian, and Christos Faloutsos. Edge weight prediction in weighted signed networks. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*, pages 221–230. IEEE, 2016.
- [17] Tiago P Peixoto. Netzschleuder network catalogue and repository. <https://networks.skewed.de/>.