SCALING FRUSTRATION INDEX AND CORRESPONDING BALANCED STATE DISCOVERY FOR REAL SIGNED GRAPHS

Muhieddine Shebaro and Jelena Tešić

January 31, 2025

ABSTRACT

Structural balance modeling for signed graph networks presents how to model the sources of conflicts. The state-of-the-art focuses on computing the frustration index of a signed graph, a critical step toward solving problems in social and sensor networks and scientific modeling. The proposed approaches do not scale to large signed networks of tens of millions of vertices and edges. This paper proposes two efficient algorithms, a tree-based *graphBpp* and a gradient descent-based *graphL*. We show that both algorithms outperform state-of-art in terms of efficiency and effectiveness for discovering the balanced state for *any* network size. We introduce the first comparison for large graphs for the exact, tree-based, and gradient descent-based methods. The speedup of the methods is around *300+ times faster* than the state-of-the-art for large signed graphs. We find that the exact method excels at optimally finding the frustration for small graphs only. *graphBpp* scales this approximation to large signed graphs at the cost of accuracy. *graphL* produces a state with a lower frustration at the cost of selecting a proper variable initialization and hyperparameter tuning.

1 Introduction

Unstructured data requires a rich graph representation. The signed networks can model complex relationships with negative and positive edges and lack of an edge. Social dynamics and stability concerning friendship and enmity in more depth [Antal et al., 2006, Leskovec et al., 2010] and brain behavior [Saberi et al., 2021b] were modeled using signed network analysis. The challenge right now is the size of the signed graph benchmarks [Leskovec and Krevl, 2014, Kunegis, 2013] and the complexity of the existing methods [Tomasso et al., 2022a]: the proof of concepts for narrow-band tasks in finance [Aref et al., 2016], polypharmacy [Liu et al., 2021a], bioinformatics [Li et al., 2021], and sensor data analysis [Casas et al., 2020, Liu et al., 2021b] are simply too small to be deployed for modern networks and datasets and make assumptions that are not applicable in real signed networks [Cucuringu et al., 2021, Tomasso et al., 2022a]. A salient metric in signed graphs is the frustration index, and finding it is NP-hard [Aref et al., 2016]. The frustration index can also be represented as the number of edges that need to change a sign so that no cycle in the graph contains an odd number of negative edges.

In this paper, we focus on scaling the computation of the frustration index and the associated balanced state for large signed networks. We propose a novel and efficient tree-based method, *graphBpp*, and a loss optimization method, *graphL*. We demonstrate the proof-of-concept on large (millions of vertices and edges) signed graphs derived from the actual data. Balance theory represents a theory of changes in attitudes [Abelson and Rosenberg, 1958]: people's attitudes evolve in networks so that friends of a friend will likely become friends, and so will enemies of an enemy [Abelson and Rosenberg, 1958]. Heider established the foundation for social balance theory [Heider, 1946], and Harary established the mathematical foundation for signed graphs and introduced the k-way balance[Cartwright and Harary, 1956, Harary and Cartwright, 1968].

The balanced-theory-based algorithms helped solve the tasks of predicting edge sentiment, recommending content and products, and identifying unusual trends [Derr et al., 2020, Garimella et al., 2021, Interian et al., 2022, Amelkin and Singh, 2019]. The frustration index is one measure of network property in many scientific disciplines, that is, in chemistry [Seif and Ahmadi, 2014], biology [Iacono and Altafini, 2010], brain studies [Saberi et al., 2021a], physical chemistry [Zhou et al., 2022] and control [Fontan and Altafini, 2018]. Finding the maximum cut of the graph in a particular case of all opposing edges is equivalent to calculating the frustration index [Hüffner et al., 2010]. The authors showed that the process is NP-hard [Hüffner et al., 2010]. State-of-the-art methods address the computation of the frustration index for signed graphs with up to 100,000 vertices [Aref and Neal, 2021], and the approach does

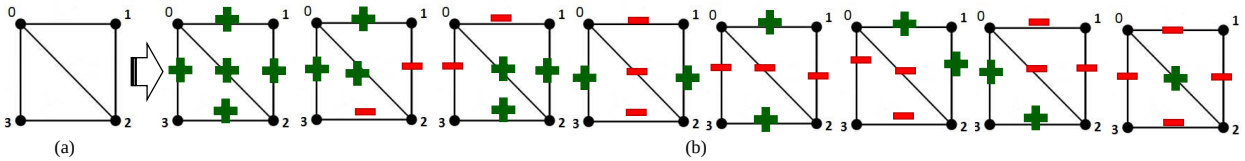

Figure 1: (a) Unsigned graph G with 4 vertices and 5 edges; (b) Eight possible edge sign combinations for G (signed graphs Σ).

not scale to modern large signed networks with tens of millions of vertices and edges. Signed networks can have multiple nearest-balanced states, and *graphB* algorithm [Rusnak and Tešić, 2021] implements the first approach to scale Algorithm 3 in the appendix. The contributions are:

• We propose *graphBpp*, an extension to graphB+, to iteratively find the frustration index for any real-world signed network of any size or density.

• We extend the frustration cloud from a set in [Rusnak and Tešić, 2021] to a *(key, value)* tuple collection. We store the nearest balanced states with their associated frequency and edge switches as a tuple in the memory-bound frustrated cloud.

• We propose *graphL*, a gradient descent algorithm that produces a more optimal balanced state with a lower index than *graphBpp* in linear time.

• We test seven spanning tree-sampling methods for *graphBpp* to find the ideal sampler that best minimizes the frustration when approximated iteratively. To the best of our knowledge, this marks the first instance where exact, tree-based, and gradient descent-based methods are evaluated and directly compared in terms of their effectiveness in estimating the frustration index.

2 Definitions and Corollaries

Nearest balanced states S are a subset of all possible balanced states of a signed graph in which *graphB* produces these states by a minimal number of edge sign changes using a tree-sampling method. In other words, the algorithm always produces this subset of balanced states by avoiding the tedious calculations of finding all balanced states, some of which are only present by passing through another balanced state [Rusnak and Tešić, 2021]. We designate $S(i)$ to indicate the i^{th} nearest balanced state produced by *graphB* in the i^{th} iteration. In the for loop in line 1, the algorithm loops over k sampled spanning trees instead of all trees (Algorithm 1 line 1). Next, the *graphB+* algorithm [Alabandi et al., 2021] scaled the computation of fundamental cycles for the spanning tree T. If T is a spanning tree of Σ and e is an edge of Σ that does not belong to T, then *fundamental cycle* C_e, defined by e, is the cycle consisting of e together with the straightforward path in T connecting the endpoints of e. If $|V|$ denotes the number of vertices and $|E|$ the number of edges in Σ, there are precise $|E| - |V| + 1$ fundamental cycles, one for each edge that does not belong to T. Each C_e is linearly independent of the remaining cycles because it includes an edge e not present in any other fundamental process. Figure 1 shows eight balanced states for an unsigned graph with four vertices and five edges. We can observe that every cycle in each state contains an even number of negative edges. We define the memory-bound frustration cloud as a container with a collection of nearest balanced states for a signed graph, restricted in size based on the computer's random access memory.

2.1 Fundamental Cycle Basis

Definition 2.1 ***Path*** *is a sequence of distinct edges m that connect a sequence of distinct vertices n in a graph.* ***Connected graph*** *has a path that joins any two vertices.* ***Cycle*** *is a path that begins and ends at the same node.* ***Simple Cycle*** *is a route that begins and concludes at an identical vertex and doesn't pass through any other vertex more than once.* ***Cycle Basis*** *is a set of simple cycles that forms a basis of the cycle space.*

Definition 2.2 *For the underlying graph G, let T be the spanning tree of G, and let an edge m be an edge in G between vertices x and y that is* NOT *in the spanning tree T. Since the spanning tree spans all vertices, a unique path in T exists between vertices x and y, which does not include m. A* ***Fundamental Cycle*** *is a cycle that combines a path in the tree T and an edge m from the graph G. The cycles, denoted as c_i, are considered fundamental if they include precisely one edge that is not part of the tree. They are a collection of cycles capable of generating all possible cycles in a graph through a linear combination of its members, which is determined based on a spanning tree. T is just one potential*

spanning tree of many spanning trees (unless the underlying graph is a tree). For instance, the cycles 0-1-2 and 0-3-2 in Figure 1 are fundamental and can generate a larger cycle 0-1-2-3, which is not a fundamental cycle.

Corollary 2.1 *A fundamental cycle basis can be derived from a spanning tree or spanning forest of the given graph by selecting the cycles formed by combining a path in the tree and a single edge outside the tree. For the graph G with a set of vertices V and a set of edges E, there are precisely $|E| - |V| + 1$ fundamental cycles for each connected component.*

2.2 Balanced Graphs and Frustration

Definition 2.3 ***Signed graph*** *$\Sigma = (G, \sigma, V, E)$ consists of underlying unsigned graph G and an edge signing function $\sigma : E \to \{+1, -1\}$. The edge $m \in E$ can be positive m^+ or negative m^-.* ***Fully Signed Graph*** *is a signed graph with vertex signs (assigned +1 or -1) [Du et al., 2016].* ***Sign*** *of a sub-graph is the* product *of the edges signs.* ***Balanced signed graph*** *is a signed graph where every cycle is positive.* ***Frustration*** *of a signed graph (Fr) is defined as the number of candidate edges whose sign needs to be switched for the graph to reach a balanced state.* ***Frustration Cloud*** *contains a collection of nearest balanced states for a particular signed graph.*

Definition 2.4 *A balanced state is* ***optimal*** *if and only if it requires a minimum number of edge sign switches in the original graph to reach a balanced state.*

Theorem 2.2 ([Cartwright and Harary, 1956]) *If a signed subgraph Σ' is balanced, the following are equivalent:*

1. *Σ' is balanced. (All cycles are positive.)*
2. *For every vertex pair (n_i, n_j) in Σ', all (n_i, n_j)-paths have the same sign.*
3. *$Fr(\Sigma') = 0$.*
4. *There exists a bipartition of the vertex set into sets U and W such that an edge is negative if, and only if, it has one vertex in U and one in W. The bipartition (U,W) is called the* Harary-bipartition.

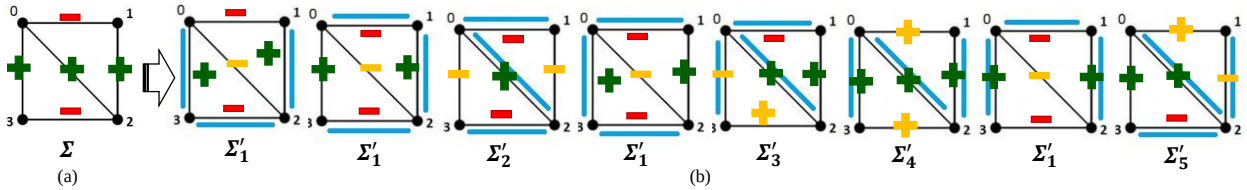

Figure 2: (a) Signed graph Σ (b) Near-balanced states of Σ, $\Sigma'_i : i \in [1, 5]$ where blue lines illustrate the spanning tree and yellow signs note the edge sign change in Algorithm 3 in the appendix. If a fundamental cycle contains an odd number of negative edges, sign switching occurs on *non-tree* edges (non-blue edges) to balance the signed network.

In this paper, Section 3 summarizes related work and state-of-the-art in the field. We introduce a *balancing* algorithmic improvements for approximating the frustration index as the *graphBpp* algorithm or *graphB++* in Section 4. We introduce the gradient descent-based heuristic for finding the frustration index and the novel loss function in the *graphL* algorithm in Section 5. We assess the effectiveness and efficiency of *graphBpp* in Section 6, and of *graphL* in Section 7, and compare and contrast them with state-of-art using real-world signed graph benchmarks [Kunegis, 2013, He and McAuley, 2016]. In Section 8, we summarize our findings.

3 Related Work

Frustration index computation has various applications in bioinformatics, engineering, and science, and the only existing open-source code for calculating the frustration index is the Binary Linear Programming (BLP) [Aref et al., 2016]. More applications for the frustration index are found in the appendix in Section B.

Computing the Frustration Index: Researchers have focused on calculating the exact frustration index. Calculating the frustration index is an NP-hard problem equivalent to calculating the ground state of the spin glass model on unstructured graphs [Schaub et al., 2016]. The frustration index for small fullerene graphs can be calculated in polynomial time [Došlić and Vukičević, 2007], and the finding was used to estimate the genetic algorithm of the frustration index in [Seif and Ahmadi, 2014]. Bansal et al. introduced the correlation clustering problem, which is a problem in computing

the minimum number of frustrated edges for several subsets [Bansal et al., 2004]. Aref et al. provided an exact algorithm to calculate the partial balance and frustration index with $O((2^b)|E|^2)$ complexity where b is a fixed parameter, and $|E|$ is the number of edges [Aref et al., 2016]. Recent improvements in the algorithm include binary programming models and the use of multiple powerful mathematical solvers by Gurobi [Gurobi, 2008], and the algorithm can handle up to $|E| = 100,000$ edges and compute the frustration index of the network in 10 hours [Aref and Neal, 2021]. The integer and binary programming models are known to be slow, computationally expensive, and have a huge search space for large problems. The use of a parallel genetic algorithm for solving large integer programming models [Fallah et al., 2020] does not scale as the authors postulate that as these models grow, the efficiency decreases greatly, making it impossible to have any output, and we demonstrate this in experiments.

Gradient Descent in Signed Networks: Tang et al. [Tang and Zhu, 2023] propose a statistically principled latent space approach for modeling signed networks and accommodating the well-known balance theory. They build a balanced inner-product model that has three different kinds of latent variables to optimize: vertex degree heterogeneity α vector of size n, z vector of size n, which encodes for the latent position, and the latent polar variable vector q of size n in which it encodes the placement of the vertices in one of the two Harary subsets. They model the distribution of signs through their product, which satisfies the balance. An edge between two vertices will likely have a positive sign when their latent variables q_i and q_j have the same sign and a negative sign otherwise. Finally, they propose a loss function to minimize and find the optimal polar variable values using the *projected* gradient descent. They present the error rates for these estimates using simulation studies.

4 graphBpp: scaling the graph balancing

This section proposes the improved graph balancing algorithm, the *graphBpp* algorithm. The *graphBpp* extends the fundamental cycle algorithm *graphB+* proposed in [Alabandi et al., 2021] to *approximate* the frustration index Fr for a signed graph Σ as Fr_Σ using a particular tree-sampling technique.

$$Fr_\Sigma = min_i(\mathcal{S}(i)) \tag{1}$$

The objective function for approximating the frustration index is outlined in Eq. 1, and S is a container that stores the number of edge sign switches for a given i^{th} nearest balanced state.

The *graphB+* is an efficient algorithm alternative for computing the fundamental cycles[Alabandi et al., 2021]. The *graphBpp* algorithm builds on the *graphB*[Rusnak and Tešić, 2021] and *graphB+*[Alabandi et al., 2021] as it combines the efficiency of *graphB+* with the functionality of *graphB*. The *graphBpp* integrates different tree-sampling approaches, as outlined in Algorithm 1 for the frustration index computation. Next, the *graphBpp* algorithm scales the calculation of the frustration index and associated optimal balanced state by iteratively keeping in memory only the subset of nearest balanced states with the smallest number of edge negations, as outlined in Algorithm 5 in the appendix. The *graphBpp* finds the approximate frustration index and the nearest balanced state associated with the index for *any* large signed graph.

Algorithm 1 Tree-Based Graph Balancing and Frustration Index

```
Input signed graph Σ and spanning trees sampling method M
Generate set T_{M^k} of k trees of Σ using M
Empty F_Σ (frustration cloud of nearest balanced states)
B = φ
for spanning trees T, T ∈ T_k where T_k is a set of k spanning trees of Σ do
    Find balanced state Σ'_i using Algorithm 3 in the appendix
    s = edge signs difference count from Σ to Σ'_i
    Transform Σ'_i balanced state to string B
    if B ∉ B then
        Add key B to B
        S(B) = s
        C(B) = 1
    else
        C(B)++
    end if
end for
Return frustration index Fr(Σ) = min_i(S) and frustration cloud F_Σ = B:(C, S)
```

We extend the definition of frustration cloud $\mathcal{F}_\Sigma$ from a set to a *(key,value)* tuple collection $\mathcal{F}_\Sigma = \mathcal{B}{:}(\mathcal{C}, \mathcal{S})$. The key is the unique balanced state $\mathcal{B}(i)$, and the value is the count of balanced states occurring in iteration $\mathcal{C}(i)$, and the

edge count switches to the balanced state $\mathcal{S}(i)$. In each balancing iteration, we examine the resulting balance state (Algorithm 1) Σ'_T about $\mathcal{B}$. We represent the balanced state Σ'_T as a string B to make the process more efficient. The balanced state Σ'_T represents the three edge vectors (src, tgt, sign). If an edge i is defined by two vertices (u, v) and a sign s, the algorithm balances the graph and stores the edges as src(i)=u, tgt(i)=v, sign(i)=s. The number of edge sign switches for each iteration of the *graphBpp* algorithm is counted by comparing Σ with the produced balanced state Σ'_i in Algorithm 3 in the appendix, The value is stored in $\mathcal{S}(i))$ for the i^{th} iteration. Thus, by choosing the nearest balanced state with the lowest number of edge sign switches, we can *approximate* the frustration index with the lowest value available from the tree sampling. Next, we introduce the update to the *frustration cloud* [Rusnak and Tešić, 2021] to be memory-bound, and we define the new frustration cloud $\mathcal{F}_\Sigma$ in Eq. 2.

$$\mathcal{F}_\Sigma = (\mathcal{B}(i), \mathcal{C}(i), \mathcal{S}(i)), i \leq \mathcal{F}_{max} \tag{2}$$

In Eq. 2, the $\mathcal{B}(i)$ is a container for storing the i^{th} balanced state, $\mathcal{C}(i)$ is a container for saving the number of i^{th} balanced state produced, and $\mathcal{S}(i)$ is the number of the edge switches to achieve the i^{th} balanced state from Σ. The $\mathcal{F}_{max}$ represents the number of balanced states where a memory limit is reached during the frustration cloud creation. Figure 2 all the nearest balanced states produced by *graphBpp*.

For *graphBpp* implementation, we propose an efficient transform ($O(|E|)$) of the balanced state output Σ' to the string hash key B for comparison with other balanced states, as outlined in Algorithm 1 line 5. The triple edge vector (src(i),tgt(i), sign(i)) is inserted into a set of tuple data structures to organize the edges and prepare for string conversion automatically. Then, it is transformed to a string format "src(i)$->$tgt(i): sign(i)", and then all edge strings are concatenated in order, separated by the delimiter "|" and stored as the B key in $\mathcal{B}$. If B is in $\mathcal{B}$, we increase the corresponding $C(B)$ value count, where B is the existing balanced state Σ'_T. If Σ'_T is not in $\mathcal{B}$, we add (Σ'_i,(1,number of switched edge signs) pair to the collection. If the state was previously unseen, we add the new balanced state to the hashmap as a string key as illustrated in Algorithm 1. Then, we add 1 to the end of the count stack $\mathcal{C}$ and add the number of edge switches in the graph for this balanced state to the frustration cloud frequency stack $\mathcal{S}$. These two values (count stack and frequency stack) are stored as a pair, and the value of the hashmap of the balanced state string as a key is that pair. If the balanced state exists in $\mathcal{B}$, we increase the count at the same string key in $\mathcal{C}$ only (the first element of the pair is modified), as illustrated in Algorithm 1. The minimum number of edge switches in all balanced states approximates the frustration index, and we approximate the frustration index as Fr(Σ) = min($\mathcal{S}$). We expand on the data structure in C++ used for storing the states, and we present the scalable version of *graphBpp* in Algorithm 1 for large signed graphs in the appendix in Section E.

4.1 Sampling Spanning Trees

To maximize the chances of discovering the optimal nearest balanced state in Algorithm 1, we propose to utilize randomization and hybridization of the standard tree sampling techniques. We use Depth-First Search (DFS), Breadth-first search (BFS), Randomized Depth First Search (RDFS), Aldous-Broder algorithm, Kruskal's algorithm, Prim's algorithm, and the RDFS-BFS sampler. The RDFS-BFS sampler aims to minimize the frustration index and maximize the number of unique stable states to increase algorithmic chances of finding the optimal state among all the nearest balanced states. We present the full definitions of the samplers used in the appendix in Section D.

5 graphL: optimization formulation

We use gradient descent to approximate the frustration index and balance the signed graph linearly. We adopt the equation from Du et al. [Du et al., 2016] that calculates the imbalance of a fully signed network. The definition of structural balance in these networks is different. According to the theory of homophily, a fully signed network is balanced if every edge is positive and the corresponding vertices have the same sign. If there is a negative edge, the vertices should have different signs. Suppose that the fully signed network is balanced based on the homophily theory. In that case, the underlying signed network (ignoring vertex signs) is also balanced because the fully signed network is a generalization of the signed network [Du et al., 2016]. The equation for computing imbalance in the fully signed network is outlined in Eq. 3

$$L = \sum_{\forall (i,j) \in \Sigma} \frac{1 - e_{ij}\theta_i\theta_j}{2} \tag{3}$$

where e_{ij} is the sign of the edge connecting vertex i to vertex j. θ_i and θ_j are the vertex signs (1 or -1) for vertices i and j respectively. The equation is a d differentiable loss function that we will attempt to minimize by treating the θ signs of the vertices as latent variables. Initially, the θ variables are relaxed to continuous random variables in the range

between -1 to 1. We denote these continuous variables as Γ, a vector with a size equal to the number of vertices in the signed graphs. The equation for optimization is then:

$$L = \sum_{\forall (i,j) \in \Sigma} \frac{1 - e_{ij}\Gamma_i\Gamma_j}{2} \tag{4}$$

The loss function used is outlined in Eq. 4. Next, for each gradient update iteration or round, the *graphL* algorithm computes the loss using Eq. 3. Note that algorithm in lines 5-6 sets θ_i and θ_j to -1 if Γ_i and Γ_j are negative respectively and to 1 if Γ_i and Γ_j are positive respectively. The algorithm computes the gradients concerning each latent variable Γ_i in Γ vector in Eq. 4, and the gradients are:

$$\Gamma_i : \frac{\partial L}{\partial \Gamma_i} = -\frac{1}{2} \sum \Gamma_j e_{ij} \tag{5}$$

Algorithm 2 Gradient Descent-Based Graph Balancing and Frustration Index

1: Input signed Network Σ, learning rate α, number of gradient updates λ
2: x=0
3: Initialize random float vector Γ of size equal to the number of nodes.
4: **while** $x < \lambda$ **do**
5: Compute loss function (also frustration index) using $L = \sum \frac{1-e_{ij}\theta_i\theta_j}{2}$ where θ_i and θ_j is 1 if Γ_i and Γ_j is greater than 0 respectively, otherwise -1
6: Induce relaxation and allow continuous values for θ vector by substituting it with Γ: $L = \sum \frac{1-e_{ij}\Gamma_i\Gamma_j}{2}$
7: Compute the gradient with respect to each Γ_i: $\frac{\partial L}{\partial v_i} = -\frac{1}{2}\sum \Gamma_j e_{ij}$ where Γ_j is the neighbor of Γ_i
8: Update Γ: $v \leftarrow \Gamma - \alpha \frac{\partial L}{\partial \Gamma}$
9: x=x+1
10: **end while**
11: Initialize frustration=0
12: Initialize set $visit = \phi$
13: Assign $\Sigma' = \Sigma$
14: **while** all edges have not been visited **do**
15: Fetch unvisited edge e_{ij} between vertices i and j
16: **if** $\Gamma_i >= 0$ **then**
17: θ_i=1
18: **end if**
19: **if** $\Gamma_j >= 0$ **then**
20: θ_j=1
21: **end if**
22: **if** $\Gamma_i < 0$ **then**
23: θ_i=-1
24: **end if**
25: **if** $\Gamma_j < 0$ **then**
26: θ_j=-1
27: **end if**
28: frustration+=$\frac{1-e_{ij}\theta_i\theta_j}{2}$
29: **if** $\frac{1-e_{ij}\theta_i\theta_j}{2}$=1 **then**
30: Flip the sign of e_{ij} in Σ'
31: **end if**
32: Add e_{ij} to $visit$
33: **end while**
34: Return Σ' and frustration

where Γ_j is the neighbor of Γ_i, and we update the values of the elements of Γ using Γ: $\Gamma \leftarrow \Gamma - \alpha \frac{\partial L}{\partial \Gamma}$. In this way, we are *directly* minimizing the loss that represents the level of imbalance in the signed network. We repeat the process until we reach a predefined number of gradient updates λ. These latent variables in Γ should converge to be either above zero or below 0. Finally, we loop over every edge in the network and discretize the values of Γ_i and Γ_j along e_{ij} to be integers 1 or -1 to be assigned back in θ vector. If Γ_i and Γ_j along edge (i, j) have values above 0, we set n_i and n_j to 1. Otherwise, we set them to -1. We use Eq. 3 to approximate the frustration and increase the frustration counter by

$\frac{1-e_{ij}\theta_i\theta_j}{2}$ for each edge. In addition, if it is causing an imbalance for each edge, we flip its sign and finally return the balanced state. Algorithm 2 summarizes the steps of the complete algorithm.

6 The *graphBpp* Proof of Concept

The setup, implementation, and data details are presented in the appendix in Section C. We compare the proposed method to BLP baseline [Aref et al., 2016] on SNAP [Leskovec and Krevl, 2014], Konect [Kunegis, 2013], and Amazon [He and McAuley, 2016] open-source benchmarks. For the Amazon signed graphs, we ran the scalable version of *graphBpp* intended for large graphs, which is Algorithm 5 in the appendix. For other smaller graphs, we ran Algorithm 1. Running Algorithm 1 for the Amazon signed graphs would not work because the 1000 balanced states for these graphs would not fit in memory and would crash the program. Thus, we cannot use both algorithm versions on all signed graphs.

	rain	**S&P**	**wikiE**	**wikiR**	**epin**	**slash**
BFS	**10,217**	**134,515**	24,827	43,971	**100,450**	117,587
RDFS	20,047	326,957	51,197	78,215	276,584	205,236
DFS	22,879	351,135	50,617	78,151	275,211	205,089
Hybr	**10,217**	134,863	**23,970**	**42,573**	118,588	**115,932**
Krusk	18,576	318,733	38,255	71,990	200,264	188,495
AB	19,403	323,657	47,252	74,438	246,941	198,785
Prim	21,312	355,819	51,732	79,482	N/A	N/A
BLP	10,150	176,965	29,257	26,778	N/A	77,283

Table 1: SNAP frustration for 1000 iterations of graphBpp with SEVEN tree samplers introduced in subsection 4.1, and the baseline. BFS and Hybrid consistently perform the best. We rename the signed graphs as follows: rainFall (rain), S&P 1500 (S&P), wikiElec (wikiE), wikiR (wikiRfa), epinions (epin), and slashdot (slash), and we rename the samplers: Kruskal (Krusk), Aldous-Broder (AB), and Hybrid (Hybr).

6.1 Selecting the Spanning Tree Method

We compare the timing and frustration computation of *graphBpp* implementation of Algorithm 1 of **SEVEN** different tree sampling methods and look for the most effective and efficient sampling method for the frustration index computation. Note that *graphBpp* runs are non-deterministic, and we run the methods multiple times. The frustration committed and completion time is always the same for smaller graphs and within 0.1% for larger graphs. We compare the findings to the BLP baseline implementation for SNAP datasets. The results are summarized in Table 1 regarding the approximated frustration index. Table 2 summarizes the resulting frustration index per method. BFS-spanning trees produce balanced states of minimum edge switches, and DFS-spanning trees make trees with maximum edge switches. RDFS/Kruskal/Aldous Broder's frustration scores are slightly better due to the randomization step. BFS discovers the optimal trees for the frustration computation, but they are repetitive without inducing the shuffling of the adjacency list when exploring a node's neighborhood.

6.2 Impact of Number of Iterations on Frustration and Timing

Here, we evaluate the efficiency of the proposed algorithm by comparing *graphBpp* timing if the number of iterations increases. Figure 3 shows the change in performance for the two best tree sampling methods when the number of iterations grows. More iterations will not impact BFS sampling in smaller graphs. The frustration shows a slight improvement for the larger graphs for both methods when the number of iterations increases in Figure 3. The iteration timing for each implemented tree-sampling technique for *graphBpp* is presented in the appendix in Section H.

6.3 graphBpp vs. SOTA

In this experiment, we compare the baseline BLP [Aref and Neal, 2021] with graphBpp implementation with breadth-first search (BFS) spanning tree sampling in 1000 iterations in terms of the frustration index and the time it takes to approximate the frustration index for 13 benchmark graphs in Table 2. The space complexity of BLP is $O(|V|^2)$, where $|V|$ is the number of vertices on the graph. Aref et al. state that the signed graphs with up to $100,000$ edges will be solved in 10 hours [Aref and Neal, 2021]. Since our computer can store all 1000 nearest balanced states for each of these 13 signed graphs in memory, we ran the *graphBpp* implementation of Algorithm 1, which is the non-scalable version of the algorithm for small and medium-sized graphs. This algorithm does not handle the case when the memory

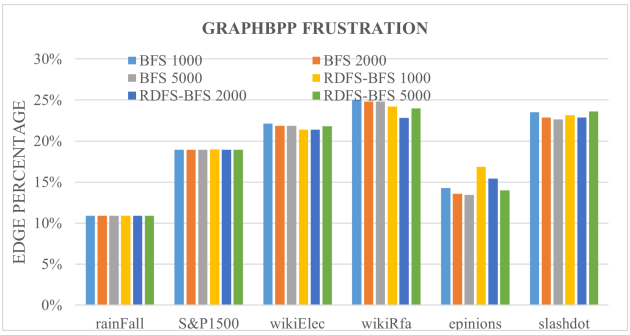

Figure 3: *graphBpp* frustration for six benchmark datasets, two spanning tree sampling approaches (BFS and RDFS-BFS), and three different iteration counts.

is complete (when storing the balanced states), and it is slightly faster than Algorithm 5 in the appendix, where it employs the nearest balanced state replacement after it reaches the memory limit. All external processes are closed to prevent interference with time measurements. The measurement for both methods includes the time it takes to input the file, process it, and output the results.

SNAP	**Cycles**	**BLP**		**graphBpp**	
[Leskovec and Krevl, 2014]		index	time	index	time
test10 [Alabandi et al., 2021]	4	**2**	0.053s	**2**	0.08s
highland [Read, 1954]	43	**7**	0.037s	**7**	0.13s
sampson18 [Sampson, 1968]	95	**39**	0.08s	**39**	0.27s
rainFall [Cucuringu et al., 2021]	93,331	**10,150**	7.26hrs	10,271	83.4s
S&P1500 [Cucuringu et al., 2021]	709,836	176,965*	N/A	**134,515**	1478s
wikiElec [Leskovec and Krevl, 2014]	104,520	29,360*	30hrs	**24,827**	184s
wikiRfa [Leskovec and Krevl, 2014]	168,154	29,971*	30 hrs	43,971	281s
epinions [Leskovec and Krevl, 2014]	585,138	N/A	N/A	**100,450**	1360s
slashdot [Leskovec and Krevl, 2014]	418,342	**77,306***	30 hrs	117,587	937s

Table 2: SNAP Signed graph baseline performance as a function of the number of fundamental cycles for the BLP [Aref and Neal, 2021] and *graphBpp* algorithm with BFS tree sampling for frustration index computation. (* indicates that BLP never completes and gives a heuristic approximation).

The last four columns of Table 2 summarize our findings on the SNAP benchmark. BLP and graphBpp computation for small graphs was fast and yielded equal indices such as highland and sampson18. Both methods retrieve correct frustration indices for the three datasets. The BLP code fails for the sparse epinions (over 100,000 vertices) and produces no results (adjacency matrix cannot fit in memory and crashes the Jupyter Notebook), where *graphBpp* finds 1000 nearest balanced states of the graph, the most optimal one with frustration 100,450 in under 23 minutes. BLP code on the S&P1500, wikiElec, WikiRfa, and slashdot signed graphs produces heuristic frustrations. (it crashes without outputting the time for S&P1500). On the other hand, *graphBpp* finds a lower frustration for S&P1500 and wikiElec. In addition, to overcome the memory restrictions of extracting and saving balanced states with their associated frequencies and frustration in the frustration cloud for *graphBpp*, we implement Algorithm 5 in the appendix and set the CAP to 75% of the total RAM size. We apply it to Amazon data in Table 3. The BLP model only worked and converged for the smallest 3 Amazon signed networks, the Core5 reviews. All Amazon ratings and graphs have several vertices $|V|$ higher than 300,000, and BLP outputs a memory error before initializing the model. The algorithm attempts to construct an adjacency matrix that does not fit into memory for any graph with more than 100,000 vertices. The serialized process takes about an extra hour for each 1 million edges, and the processing time, for a fixed CAP, is way less than BLP across the board with an increasing number of edges and vertices, see Figure 4. The BLP algorithm converges within 30 hours for smaller signed graphs and finds the optimal frustration index. *graphBpp* recovers the balanced state and associated frustration index for small graphs and in minutes for under 2 million edges. The most extensive graph we have processed is Amazon books with close to 10 million vertices and over 22 million edges, and it took 19 hours to find the nearest balanced state with frustration 3,146,316.

Amazon Ratings	**BLP** index	time	**graphBpp** index	time
Books	N/A	N/A	3,146,316	19hrs
Electronics	N/A	N/A	1,025,401	7.8 hrs
Jewelry	N/A	N/A	613,129	6hrs
TV	N/A	N/A	636,568	5.2hrs
Vinyl	N/A	N/A	412,859	4.4hrs
Outdoors	N/A	N/A	264,497	4hrs
Android App	N/A	N/A	386,947	3.6hrs
Games	N/A	N/A	173,063	2.2hrs
Automotive	N/A	N/A	85,859	50min
Garden	N/A	N/A	70,690	32.2min
Baby	N/A	N/A	106,092	30.1min
Digital Music	N/A	N/A	34,019	23min
Instant Video	N/A	N/A	32,001	20.7min
Musical Inst.	N/A	N/A	24,959	14.7min
Amazon Reviews	**BLP** index	time	**graphBpp** index	time
Digital Music	10,482*	30 hrs	19,926	101s
Instant Video	6,001*	30hrs	10,833	101s
Musical Instr	1,162	136.15s	2,311	17.3s

Table 3: The results of the Amazon runs of BLP exact method and *graphBpp*. (*) indicates that Gurobi never converged.

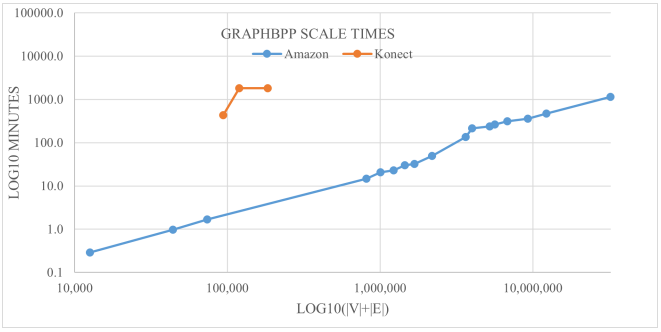

Figure 4: *graphBpp* scales better with increasing graph size $|V| + |E|$ for a fixed number of iterations (1000) for the benchmark Konect and Amazon signed graphs.

7 The graphL Proof of Concept

We pit *graphL* against *graphBpp* to approximate the frustration index and obtain better stable states. We run both heuristics on the Konect signed graphs. We use Breadth-First Search (BFS) as the tree-sampling technique for *graphBpp* using Algorithm 5 in the appendix since we proved BFS yields minimal edge sign switches previously. Since we might obtain different results for each run for the gradient descent-based method due to the random initialization of the Γ vector, we run the heuristic five times and choose the minimal edge sign switches produced out of them. Table 4 summarizes the results. First, we can observe that the gradient descent-based method is much more efficient across the board because the heuristic runs in linear time, and it does not have to extract and save multiple nearest balanced states in memory because only *graphBpp* is capable of forming the frustration cloud. Second, the gradient descent-based approach generally produces more optimal balanced states for every signed graph than that of *graphBpp* *except* Sampson, Congress, and TwitterRef. for the same number of iterations/gradient updates. However, the former comes with a downside: tuning the learning rate hyperparameter and finding the proper initialization. Hence, *graphBpp* is advantageous because it does not need any hyperparameter tuning and is a non-trainable algorithm. Unlike *graphBpp*, *graphL* produces *only* one balanced state and cannot generate multiple nearest balanced states, which are essential for computing the consensus features proposed in [Rusnak and Tešić, 2021]. These features are used in clustering and signed network analysis [Tomasso et al., 2022b].

Konect [Kunegis, 2013]	**graphL**		**graphBpp**	
Graphs	**index**	**time**	**index**	**time**
Sampson	37	0.030s	**35**	0.304s
ProLeague	**13**	0.003s	**13**	0.27s
DutchCollege	**2**	0.010s	**2**	0.80s
Congress	38	0.008s	**21**	1.021s
BitcoinAlpha	**900**	0.18s	1,105	24.38s
BitcoinOTC	**1,426**	0.27s	1,827	40.55s
Chess	**14,684**	0.70s	20,991	85.35s
TwitterRef.	19,500	2.72s	**16,183**	456.80s
SlashdotZoo	**80,787**	6.31s	109,930	837.76s
Epinions	**57,874**	9.66s	100,646	1,367.74s
WikiElec	**15,389**	1.11s	22,289	168.59s
WikiConflict	**167,003**	25.29s	252,400	6503.24s
WikiPolitics	**58,438**	9.32s	86,833	1582.30s

Table 4: Comparison of the frustration index approximation and execution time using graphBpp and graph with 1000 iterations for both on the Konect data (except TwitterRef, which is not a Konect graph).

8 Conclusion

This paper presents two novel algorithms, *graphBpp* and *graphL*, and demonstrates their scalability and superiority on large signed graphs with millions of edges and vertices. The detailed conclusion points are at the end of the appendix.

Appendix:

A The *graphB+* Algorithm [Rusnak and Tešić, 2021]

Algorithm 3 Tree-Based Signed Graph Balancing

1: Input signed graph Σ and spanning tree T of Σ
2: **for** Edges e, $e \in \Sigma \setminus T$ **do**
3: **if** fundamental cycle $T \cup e$ is negative **then**
4: Flip edge sign for edge e: $e^- \rightarrow e^+$; $e^+ \rightarrow e^-$
5: **end if**
6: **end for**
7: Return balanced graph Σ'_T

B Frustration Applications

In chemistry, the stability of fullerenes, like C60 and other spherical carbon structures, can be analyzed through the concept of the frustration index [Seif and Ahmadi, 2014]. The frustration index measures how an incoherent system responds to perturbations in large-scale signed biological networks [Iacono and Altafini, 2010]. The frustration of the network has been determining the strength of agent commitment to make a decision and win the disorder in adversarial multi-agent networks [Fontan and Altafini, 2018]. In these networks, the strength is determined by measuring the social commitment of agents, mainly when there are disorder or adversarial actions. Winning the disorder is overcoming the chaos caused by frustration in the network. The frustration arises from the mix of collaborative and antagonistic interactions, leading to an unbalanced signed graph. Agents need high social commitment to overcome this disorder and make significant decisions. The frustration in the field of neuroplasticity assesses the development of brain networks, as studies have shown that a person's cognitive performance and the frustration of the brain network have a negative correlation [Saberi et al., 2021a]. Physical chemists predict the protein-protein interaction using the frustration index of the protein signed network [Zhou et al., 2022]. Saberi et al. investigated the pattern for the formation of frustrating connections in different brain regions during multiple life stages [Saberi et al., 2022].

Amazon Ratings	$\lvert V\rvert$	$\lvert E\rvert$
Books	9,973,735	22,268,630
Electronics	4,523,296	7,734,582
Jewelry	3,796,967	5,484,633
TV	2,236,744	4,573,784
Vinyl	1,959,693	3,684,143
Outdoors	2,147,848	3,075,419
Android App	1,373,018	2,631,009
Games	1,489,764	2,142,593
Automotive	950,831	1,239,450
Garden	735,815	939,679
Baby	559,040	892,231
Digital Music	525,522	702,584
Instant Video	433,702	572,834
Musical Inst.	355,507	457,140
Amazon Reviews	$\lvert V\rvert$	$\lvert E\rvert$
Digital Music	9,109	64,706
Instant Video	6,815	37,126
Musical Instr	2,329	10,261

Table 5: Amazon ratings and reviews graph sizes [He and McAuley, 2016].

C Setup, Implementation, and Data

C.1 Signed Graph Benchmarks

In the experiments, we use SNAP [Leskovec and Krevl, 2014], Konect [Kunegis, 2013], and Amazon ratings [He and McAuley, 2016]. Table 6 summarizes SNAP signed graph benchmarks. **Konect Signed Graphs** Konect [Kunegis, 2013] signed graphs and their characteristics are described in Table 7. *Highland* is the signed social network of tribes of the GahukuGama alliance structure of the Eastern Central Highlands of New Guinea, from Kenneth Read. *CrisisInCloister* is a directed network that contains ratings between monks related to a crisis in a cloister (or monastery) in New England (USA) which led to the departure of several of the monks. *ProLeague* are results of football games in Belgium from the Pro League in 2016/2017 in the form of a directed signed graph. Vertices are teams; each directed edge from A to B denotes that team A played at home against team B. The edge weights are the goal difference, and thus positive if the home team wins, negative if the away team wins, and zero for a draw. *DutchCollege* is a directed network that contains friendship ratings between 32 first-year university students who mostly did not know each other before starting university. Each student rated the other at seven different time points. A node represents a student, and an edge between two students shows that the left rated the right. The edge weights show how good their friendship is in the eye of the left node. The weight ranges from -1 for the risk of conflict to +3 for best friend. *Congress* is a signed network where vertices are politicians speaking in the United States Congress, and a directed edge denotes that a speaker mentions another speaker. In the *Chess* network, each node is a chess player, and a directed edge represents a game with the white player having an outgoing edge and the black player having an ingoing edge. The weight of the edge represents the outcome.
BitcoinAlpha/BitcoinOTC are the user-user trust/distrust network of trading users from the Bitcoin Alpha/OTC platforms. *TwitterReferendum* captures data from Twitter concerning the 2016 Italian Referendum. Different stances between users signify a negative tie, while the same stances indicate a positive link [Lai et al., 2018]. *WikiElec* is the network of users from the English Wikipedia that voted for and against each other in admin elections. *Slashdot* is the reply network of the technology website Slashdot. The vertices are users, and the edges are replies. The edges of *WikiConflict* represent positive and negative conflicts between users of the English Wikipedia. *WikiPolitics* is an undirected signed network that contains interactions between the users of the English Wikipedia that have edited pages about politics. Each interaction, such as text editing and votes, is valued positively or negatively. *Epinions* is the trust and distrust network of Epinions, an online product rating site. It incorporates individual users connected by directed trust and distrust links.

Amazon Signed Graphs The Amazon ratings and reviews data [He and McAuley, 2016] provides rating information between 0 (low) and 5 (high) of the Amazon users on different products. We have transformed the graphs into 18 signed bipartite graphs. The raw Amazon data was originally in .json form and maintained the rating, the item I.D., and the user I.D. Here, the user I.D.s and the item I.D.s are the nodes, and the edges are constructed based on the rating value. If the rating is 5 and 4, it implies a positive edge; if the rating is 3 and 2, it gives no edge; and if the rating is 0 and 1, it provides a negative edge. Table 5 summarizes the large signed graphs stemming from the process.

SNAP	**vertices**	**edges**		
[Leskovec and Krevl, 2014]	$\|V\|$	$\|E\|$	cycles	% positive
test10 [Alabandi et al., 2021]	10	13	4	53.85
highland [Read, 1954]	16	58	43	50
sampson18 [Sampson, 1968]	18	112	95	54.4
rainFall [Cucuringu et al., 2021]	306	93,636	93,331	68.78
S&P1500 [Cucuringu et al., 2021]	1,193	711,028	709,836	75.13
wikiElec [Leskovec and Krevl, 2014]	7,539	112,058	104,520	73.33
wikiRfa [Leskovec and Krevl, 2014]	7,634	175,787	168,154	77.91
epinions [Leskovec and Krevl, 2014]	119,130	704,267	585,138	83.23
slashdot [Leskovec and Krevl, 2014]	82,140	500,481	418,342	77.03

Table 6: SNAP Signed graph Largest connected component (LCC) attributes. $|V|$ is the number of vertices, and $|E|$ is the number of edges in the largest connected component LCC; The label *% positive* is the number of positive edges divided by e .

	LCC			
Konect [Kunegis, 2013]	vertices $\|V\|$	edges $\|E\|$	cycles $\|E\| - \|V\| + 1$	% positive
Sampson	18	126	145	51.32
ProLeague	16	120	105	49.79
DutchCollege	32	422	391	31.51
Congress	219	521	303	80.44
BitcoinAlpha	3,775	14,120	10,346	93.64
BitcoinOTC	5,875	21,489	15,615	89.98
Chess	7,115	55,779	48,665	32.53
TwitterRef.	10,864	251,396	240,533	94.91
SlashdotZoo	79,116	467,731	388,616	76.092
Epinions	119,130	704,267	585,138	85.29
WikiElec	7,066	100,667	93,602	78.77
WikiConflict	113,123	2,025,910	1,912,788	43.31
WikiPolitics	137,740	715,334	577,595	87.88

Table 7: Konect Largest Connected Component (LCC) graph attributes [Kunegis, 2013] (except TwitterRef, which is not a Konect graph). $|E|$ is the number of edges, and $|V|$ is the number of vertices in the LCC of the graph. The *% positive* label marks the percentage of positive edges in the LCC.

C.2 Implementation

The baseline implementation relies on the published Binary Linear Programming (BLP) code [Aref, 2021]. The binary linear model runs on a Jupyter Notebook in Python [Aref, 2021] and is based on a Gurobi mathematical solver and has several parameters like the termination parameter where we can set a time limit on how long the optimization process should last in Gurobi [Gurobi, 2008]. The binary terms in the objective function depend on the single AND constraints and two standard XOR constraints per edge, respectively. The *graphBpp* implementation extends the open-source implementation [Alabandi et al., 2021] to include and test proposed tree sampling strategies while keeping the original speedup optimization for finding fundamental cycles intact. The *graphBpp* algorithm works with different tree-sampling strategies. It is achieved using C++ and involves minimizing the number of loops used, incorporating OpenMP directives for parallel processing, and promptly freeing up memory resources when no longer needed. In the implementation of Algorithm 5, the code checks the total RAM size of the Linux system during runtime and the amount of memory currently used by the frustration cloud. These two values are compared in each iteration to decide how many balanced states can be stored. Two replacements in the ABS model's objective function linearized two absolute value terms [Aref and Neal, 2021]. The code [Aref, 2021] was run with the following modifications: (1) -1 for the method parameter that indicates that the optimization method is automatic, and the setting will typically choose the non-deterministic concurrent method in the Gurobi's documentation [Gurobi, 2008] for this linear programming problem; (2) the lazy parameter is set to 1 with enabled speedup; (3) thread parameter is set to *multiprocessing.cpu_count(*, and (4) the time limit for the model run is set to up to 30 hours. Note that the value of the lazy attribute influences how aggressively the model is constrained. A value of 1 allows the constraint to cut off a feasible solution. The code provided [Aref, 2021] generates random graphs based on the specified number of nodes, edges, and probability of negative edges. Our improvements to the code allow for the code to (1) accept the same input format as *graphBpp* and to (2) detect and eliminate duplicates, inconsistencies, self-loops, and invalid signs in the input graph. The code for graphBpp is available on GitHub, and the data it uses is publicly accessible. The references for these resources are [Anonimoys, 2023] for the code and [Leskovec and Krevl, 2014, Kunegis, 2013, He and McAuley, 2016] for the data. On the other hand, for the gradient-based heuristic, we set λ to 1000 and the learning rate α to 0.001.

C.3 Setup

The operating system used for all experimental evaluations is Linux Ubuntu 20.04.3, running on the 11th Gen Intel(R) Core(TM) i9-11900K @ 3.50GHz with 16 physical cores. It has one socket, two threads per core, and eight cores per socket. The architecture is X86_x64. The GPU is Nvidia G Force and has 8GB of memory. Its driver version is 495.29.05, and the CUDA version is 11.5. The cache configuration is L1d : 384 KiB, L1i : 256 KiB, L2 : 4 MiB, L3 : 16 MiB. The CPU op is 32-bit and 64-bit.

D Tree-Sampling techniques for *graphBpp*

Algorithm 4 Hybridized RDFS-BFS Sampling

1: Input signed graph Σ and a root vertex n get uniformly distributed random number 0 or 1, z
2: **if** z is 0 **then**
3: Run BFS algorithm [Burtscher, 2021]
4: **else**
5: Run RDFS algorithm
6: **end if**
7: Return spanning tree T of Σ

The **Depth-First Search (DFS)** algorithm [Cormen, 2009] is a graph traversal method characterized by its time complexity of O($|V| + |E|$). The traversal commences at a root vertex and continues to explore as deeply as possible along each pathway (branch) before backtracking. This process repeats until it reaches the vertex where all adjacent vertices have been visited. The **Breadth first search (BFS)** algorithm [Cormen, 2009] with time complexity O($|V|+|E|$) is a graph traversal approach in which the algorithm first passes through all vertices on the same level before moving on to the next level. A graph traversal technique is employed to visit all the vertices of a graph. A level is a group of equidistant vertices from the root vertex. We propose to use the randomized algorithms as follows: in each iteration, we shuffle and randomize a node's neighborhood using a uniformly distributed random seed number before applying a static algorithm. The idea is that a vertex establishes a link to the first unvisited vertex based on the network's randomized order of the adjacency list. **Randomized Depth First Search (RDFS)** algorithm transforms DFS into a non-deterministic algorithm by eliminating the static ordering of the adjacency lists. The time complexity of the DFS is known to be O($|V| + |E|$), where $|V|$ is the number of vertices and $|E|$ is the number of edges in the signed network. The algorithm also runs in linear time O(n), where n is the number of vertices adjacent to a specific vertex in the network, so the total time complexity is O($|V| + |E|$). **Aldous-Broder algorithm** with complexity O($|V|$) produces a random uniform spanning tree by performing a random walk on a finite graph with any initial vertex and stops after all vertices have been visited [Hu et al., 2021]. For the popular **Kruskal's algorithm** [Wu et al., 2023] that has a time complexity O($|E|log|V|$) or O($|E|log|E|$), we intend to generate random spanning trees by assigning random weights to every edge in each iteration before running the algorithm. The method finds the minimum spanning tree of a connected and weighted graph. Randomizing the weights of **Prim's algorithm** [Gass and Fu, 2013] [Wu et al., 2023] with complexity O($|V|^2$) can also generate random spanning trees.

E Scaling *graphBpp* for Large Signed Graphs

The *graphB+* algorithm [Alabandi et al., 2021] efficiently discovers fundamental cycles for the spanning tree T and computes vertex and edge labels with linear time complexity. The algorithm requires a linear amount of memory, and the running time for balancing a cycle is linear in the length of the cycle times the vertex degrees but *independent* of the size of the graph. The labels here are specific values assigned to the vertices and edges after sampling a spanning tree as a preliminary step before finding and traversing the fundamental cycles and balancing them. These labels aid in finding these cycles and ensure efficient traversal. The algorithm speeds up the balancing to more than 14 million identified, traversed, and balanced fundamental cycles. Next, the algorithm traverses or visits each cycle in a specific order and notes which edge signs switched within each cycle to obtain an even number of negative edges along that cycle. The *graphBpp*, which builds on graphB+, scales the computation for large graphs as in Algorithm 5. First, we compute the number of balanced states that we can keep in the memory as $\mathcal{F}_{max}$ (Algorithm 5, line 3). Next, we keep only the $\mathcal{F}_{max}$ best-balanced states in $\mathcal{B}$, their count through all k iterations in $\mathcal{C}$, and the number of switched edges for each of them in $\mathcal{S}(i)$. Note that there can be different balanced states of Σ with the same number of switched edges. Finally, we compute the frustration index as a minimum of $\mathcal{S}$). The proposed approach scales well with the size of the signed graph, as we limit the number of the nearest balanced states that we keep in the memory based on their frustration index and capacity of the frustration cloud map in-memory storage, as illustrated in Algorithm 5. The comparison of balanced states now has up to k iterations times $\mathcal{F}_{max}$ closest balanced states. We determine $\mathcal{F}_{max}$ so that the size of the frustration cloud in memory is smaller than CAP. In experiments, we define CAP as 75% of the total RAM size, assuming that some vital system processes are running in the background. Note that $\mathcal{F}_{max}$ can be determined in the implementation by checking the current memory usage after every balanced state insertion and verifying that this insertion did not reach the limit. The benefit of saving balanced states in memory is that it allows us to record and study each state's frequency and use them for other tasks, such as community detection and finding the largest balanced sub-graph.

The size of the frustration cloud grows with the number of iterations as the probability of the previously unseen nearest balanced state grows. The size of the frustration cloud can become an issue for graphs with millions of vertices and

Algorithm 5 Tree-Based Graph Balancing and Frustration Index (Scalable Version for Large-sized Graphs)

Input Σ signed graph and M sampling method
Generate set $\mathcal{T}_k$ of k spanning trees of Σ
Determine $\mathcal{F}_{max}$, it is the maximum number of nearest balanced states the cloud can store before it reaches the memory limit, memory consumption of $\mathcal{F}_{Sigma} < CAP$ where $\mathcal{F}_\Sigma$ is the frustration cloud of nearest balanced states
Matrix $\mathcal{B}$, count vector $\mathcal{C}$, and edge switch count vector $\mathcal{S}$, $i = 0$, frInd=0
for T spanning tree, $T \in \mathcal{T}_k$ where $\mathcal{T}_k$ is a set of k spanning trees of Σ **do**
    Find balanced state Σ'_i
    frInd = number of edge sign switches
    if $\mathcal{S}$ is empty **then**
        $\mathcal{S}(i)$ = frInd
        $\mathcal{B}(i)$ = Σ'_i
        $\mathcal{C}(i) = 1$
    end if
    if frInd $< max_i(\mathcal{S}(i))$ **then**
        if $\Sigma'_T \notin \mathcal{B}$ **then**
            if $i < \mathcal{F}_{max}$ **then**
                i++
            else
                $i \leftarrow \arg\min_i S(i)$
            end if
            $\mathcal{S}(i)$ = frInd
            $\mathcal{B}(i)$ = Σ'_i
            $\mathcal{C}(i) = 1$
        else
            $\mathcal{C}_i$++
        end if
    end if
end for
Return frustration index $Fr_\Sigma = min_i(\mathcal{S}(i))$ and frustration cloud $\mathcal{F}_\Sigma = (B(i), C(i), S(i))_{i=1}^{F_{max}}$

vertices as the frustration cloud is too big for the main memory. The underlying data structure for the implementation in C++ is $std :: map < std :: string, std :: pair < int, int >>$. The key, as discussed above, is of type string that represents a balanced state, and the pair stores two integers, one for the number of switches and the other for the frequency of the corresponding stable state. When comparing keys, the string keys are slower than integer keys in the map data structure. However, in our case, storing the balanced states using strings or integers should have a similar comparison performance, and the integer key approach would be much more complicated. First, if we use $std :: map < int, std :: pair < int, int >>$, then it would be difficult and more complex to represent the signed network using one integer key, whereas the string can intuitively concatenate all the edges with their signs to represent that graph. There is no direct way of representing this collection of edges forming the signed graphs using solely the integer key. If we were to use an integer as a key for the map, we would still have to loop through the edges in that key for the comparisons, and the cost would be $O(|E|)$ equivalent to using the string key.

F Comparing *graphL* and Tang et al.'s Methodology

Our approach is different than the work by Tang et al. [Tang and Zhu, 2023] in several ways. First, we estimate the different types of latent variables, and we only use a random float vector Γ of size equal to the number of vertices to determine the optimal membership of each vertex in the Harary subsets. Second, we are not modeling any signed networks. Our focus is to flip the sign of the edges after estimating the latent variables of each vertex and approximate the frustration index. Tang et al. did not explicitly and directly modify the edge signs of the graph. Third, we propose using a loss function to compute the frustration index and the gradients directly. Fourth, we use the vanilla gradient descent approach instead of the projected gradient descent to minimize our loss function. We simply threshold the latent variables after optimization (ex, if an element in the latent vector is 25 after the gradient-descent step, which is above 0, we assign that element to be 1). Fifth, our loss function is adopted from Du et al. [Du et al., 2016] in which they propose to measure imbalance for fully signed networks. The loss function used by Tang et al. is

$L = \sum_{i<j}^{n} |A_{ij}| \frac{1+A_{ij}}{2} \eta_{ij} + |A_{ij}| \log(1 - \Sigma(\eta_{ij})))$ where $\eta_{ij} = v_i v_j$, A is the adjacency matrix, σ is the sigmoid function, and n is the number of nodes.

G Complexity Analysis

G.1 The graphBpp Complexity Analysis

Concerning the time complexity of *graphBpp*, it remains $O(|E|log(|V|)d_a)$ where d_a is the average degree of a vertex, $|V|$ is the number of vertices, and $|E|$ is the number of edges [Alabandi et al., 2021]. GraphB+ with (BFS) implementation has a complexity of O($|E| * log(|V| * d)$) time, where $|E|$ is the number of edges, $|V|$ is the number of vertices, and d is the average spanning-tree degree of the vertices on each cycle. The code for scaling the processing and saving of balanced states in the memory-bound frustration cloud and approximating the frustration index, which builds upon graphB+, adds O($|E|$). O($|E| * log(|V| * d)$) is still the dominant term for one iteration (generating one spanning tree and nearest balanced state). When generating $|T_k|$ spanning trees (iterations), the complexity then becomes O($|T_k| * |E| * log(|V| * d)$). On the other hand, for adapting *graphBpp* to utilize other tree-sampling techniques, the reimplemented vertex relabeling step takes O($|V| + |E|$) because DFS is used to perform the pre-order traversal on the random spanning tree generated. The edge relabeling has been implemented, resulting in a complexity of O($|V| * |E| * \alpha$) where α is the average depth from a certain vertex of an edge to the deepest relabeled vertex where the assignment of the end range of the edge takes place. The efficient fundamental cycle balancing method [Alabandi et al., 2021] has a complexity of O($|E| * log(|V| * d)$). The adapted version of graphB+ for index computation has a complexity of O($|E|$). Hence, the total time complexity for the adapted version of *graphBpp* is O($|V| * |E| * \alpha$) unless the complexity of the selected custom sampler is high enough to exceed this complexity. For $|T_k|$ iterations of the algorithm, the complexity is O($|T_k| * |V| * |E| * \alpha$).

G.2 The graphL Complexity Analysis

For every gradient update and computation, it is sufficient to loop over every edge in the signed graph to update the elements in the Γ vector. The λ is the number of gradient updates, and the total time complexity becomes O($\lambda * |E|$). We utilize the compressed sparse row (CSR) format to model the signed graph instead of the adjacency matrix because CSR scales better memory-wise. The construction of the adjacency matrix has a space complexity of O(V^2), which isn't computationally feasible when dealing with large signed graphs. Computing the imbTheccurs in constant time in each gradient update iteration does not affect the total time complexity.

H More Results on the Comparison between BLP and *graphBpp*

Konect [Kunegis, 2013]	**Computation Time for Spanning Tree Method**					
Sampling	**BFS**	**RDFS**	**DFS**	**Hybrid**	**Kruskal**	**AB**
Sampson	**0.0003s**	0.00084s	0.00053s	0.00106s	0.00055s	0.00057s
ProLeague	**0.00027s**	0.0008s	0.0035s	0.00088s	0.0005s	0.0005s
DutchCollege	**0.0008s**	0.00216s	0.00179s	0.00288s	0.00158s	N/A
Congress	**0.00102s**	0.00505s	0.00301s	0.00629s	0.00317s	0.00340s
BitcoinAlpha	**0.024s**	0.126s	0.103s	0.143s	0.098s	0.102s
BitcoinOTC	**0.041s**	0.265s	0.229s	0.268s	0.214s	0.231s
Chess	**0.085s**	0.388s	0.283s	0.375s	0.250s	0.282s
TwitterRef.	**0.457s**	2.826s	2.277s	2.566s	2.111s	2.155s
SlashdotZoo	**0.838s**	19.131s	13.803s	15.102s	9.849s	11.880s
Epinions	**1.368s**	20.794s	12.795s	16.715s	11.902s	13.373s
WikiElec	**0.16859s**	0.63977s	0.487s	0.559s	0.462s	0.502s
WikiConflict	**6.503s**	94.102s	65.282s	77.045s	40.720s	41.293s
WikiPolitics	**1.582s**	21.585s	17.374s	19.613s	15.925s	17.010s

Table 8: Average Frustration Index computation time per iteration using spanning tree sampling methods for 1000 iterations for Konect data in Table 7 (except TwitterRef, which is not a Konect graph). **BFS** sampling method is the fastest. The algorithm is highly parallelizable.

The execution time of BLP and *graphBpp* with all seven tree-sampling techniques is reported on the log 10 scale in seconds in Figure 5 (bottom) as BLP takes 30 hours for larger datasets (far right navy bar in Fig. 5 (bottom)).

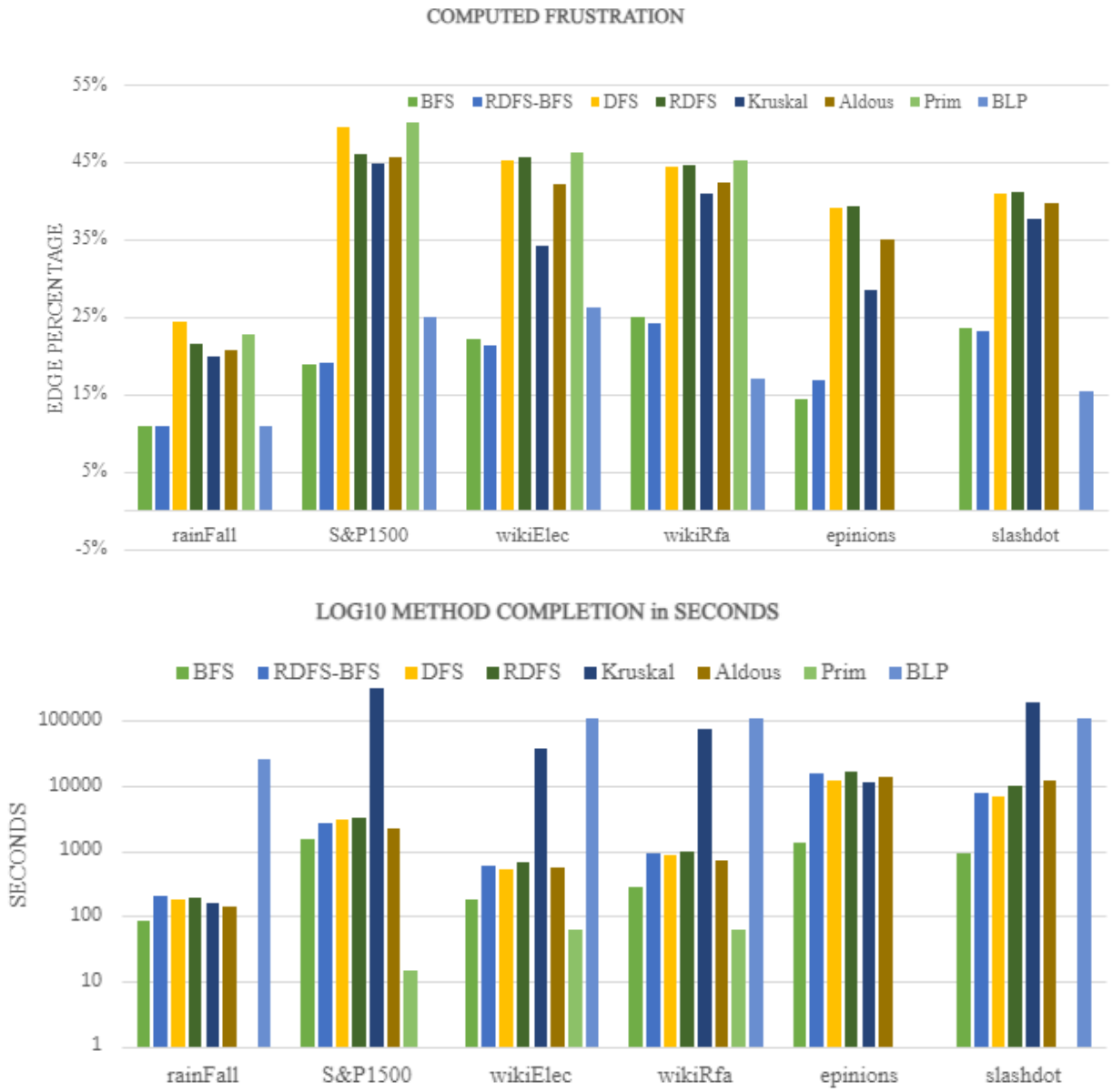

Figure 5: Frustration index (top) and timing (bottom) comparison computed using Binary Linear Programming (BLP) [Aref and Neal, 2021] and *graphBpp* 1000 iterations for different tree sampling methods over different real large signed graphs except for Prim (1 iteration). BLP never finished computing the frustration index for epinions and sp1500 within the 30 hours allocated.

RDFS-BFS is competitive with BFS in terms of frustration index as a percentage of the total number of edges in the graph (green and blue bars in Figure 5 (top)) with the small timing overhead for large graphs (Fig. 5 (bottom)): BFS produces 117,587 frustrations while BFS-RDFS produces 115,932 frustrations in 1000 iterations for the slashdot dataset. The Prim approach is too slow for large datasets, and the baseline BLP takes too long, or it does not complete. We tabulated the timing per iteration for each tree-sampling technique in Table 8. Since the time complexity of Prim is O(V^2), the number of iterations is set to 1, and it was very inefficient and slow for large graphs such as WikiConflict. Moreover, Aldous-Broder (AB) did not terminate for DutchCollege because, in uncommon scenarios, AB would get stuck looping when performing a random walk after all the current vertex's neighbors have been visited.

I Concluding Remarks

There is more than one way to achieve balance in the network. The frustration index characterizes the optimal nearest balanced state where the minimum edge switches are required to achieve balance in the network. The tree-spanning approach to graph balancing produces the nearest balanced states, e.g., there can be no other balanced state nearest balanced state derived from [Rusnak and Tešić, 2021]. In this paper, we extend our findings and propose a novel algorithm for discovering the nearest balanced states for any graph size in a fraction of the time. Our approach converges to the global optimum for the small graphs that the state-of-the-art binary linear programming (BLP) model computes. BLP does not work for graphs larger than 100,000 vertices while *graphBpp* seamlessly scales with the graph size to discover one or more nearest balanced states for the network. The state might not be optimal for a minimal number of edge switches, but it is close to optimal, and the algorithm produces a list of edges to switch to achieve the balanced state. We report the result on one computer for 1000, 2000, or 5000 iterations. In addition, we propose the use of gradient descent as a way to approximate the frustration index in linear time. We compared both *graphBpp* and *graphL*, deducing that the latter is efficient and generally yields more optimal balanced states as tested on Konect signed graphs.

References

[Abelson and Rosenberg, 1958] Abelson, R. P. and Rosenberg, M. J. (1958). Symbolic psycho-logic: A model of attitudinal cognition. *Behavioral Science*, 3(1):1–13.

[Alabandi et al., 2021] Alabandi, G., Tešić, J., Rusnak, L., and Burtscher, M. (2021). Discovering and balancing fundamental cycles in large signed graphs. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, SC '21, New York, NY, USA. Association for Computing Machinery.

[Amelkin and Singh, 2019] Amelkin, V. and Singh, A. K. (2019). Fighting opinion control in social networks via link recommendation. In *Proceedings of the 25th ACM SIGKDD Intl. Conf. Knowledge Discovery & Data Mining*, pages 677–685.

[Anonimoys, 2023] Anonymous (2023). Codebase for scaling frustration index and corresponding balanced state discovery for real signed graphs paper. `https://anonymous.4open.science/r/graphBplusplus-547A/README.md`.

[Antal et al., 2006] Antal, T., Krapivsky, P. L., and Redner, S. (2006). Social balance on networks: The dynamics of friendship and enmity. *Physica D: Nonlinear Phenomena*, 224(1):130–136.

[Aref, 2021] Aref, S. (2021). frustration-index-xor. `https://github.com/saref/frustration-index-XOR`.

[Aref et al., 2016] Aref, S., Mason, A. J., and Wilson, M. C. (2016). An exact method for computing the frustration index in signed networks using binary programming. *CoRR*, abs/1611.09030.

[Aref and Neal, 2021] Aref, S. and Neal, Z. P. (2021). Identifying hidden coalitions in the US House of Representatives by optimally partitioning signed networks based on generalized balance. *Scientific reports*, 11(1):1–9.

[Bansal et al., 2004] Bansal, N., Blum, A., and Chawla, S. (2004). Correlation clustering. *Machine Learning*, 56(1-3):89 – 113.

[Burtscher, 2021] Burtscher, M. (2021). graphB+: Balancing algorithm for large graphs. `https://userweb.cs.txstate.edu/~burtscher/research/graphB/`.

[Cartwright and Harary, 1956] Cartwright, D. and Harary, F. (1956). Structural balance: a generalization of Heider's theory. *Psychological Rev.*, 63:277–293.

[Casas et al., 2020] Casas, S., Gulino, C., Liao, R., and Urtasun, R. (2020). Spagnn: Spatially-aware graph neural networks for relational behavior forecasting from sensor data. In *2020 IEEE Intl. Conf. Robotics and Automation (ICRA)*, pages 9491–9497. IEEE.

[Cormen, 2009] Cormen, T. H. (2009). *Introduction to algorithms.* MIT Press.

[Cucuringu et al., 2021] Cucuringu, M., Singh, A. V., Sulem, D., and Tyagi, H. (2021). Regularized spectral methods for clustering signed networks. *Journal of Machine Learning Research*, 22(264):1–79.

[Derr et al., 2020] Derr, T., Wang, Z., Dacon, J., and Tang, J. (2020). Link and interaction polarity predictions in signed networks. *Social Network Analysis and Mining*, 10(1):1–14.

[Došlić and Vukičević, 2007] Došlić, T. and Vukičević, D. (2007). Computing the bipartite edge frustration of fullerene graphs. *Discrete Applied Mathematics*, 155(10):1294–1301.

[Du et al., 2016] Du, H., He, X., and Feldman, M. W. (2016). Structural balance in fully signed networks. *Complexity*, 21(S1):497–511.

[Fallah et al., 2020] Fallah, M. K., Mirhosseini, M., Fazlali, M., and Daneshtalab, M. (2020). Scalable parallel genetic algorithm for solving large integer linear programming models derived from behavioral synthesis. In *2020 28th Euromicro Intl. Conf. Parallel, Distributed and Network-Based Processing (PDP)*, pages 390–394.

[Fontan and Altafini, 2018] Fontan, A. and Altafini, C. (2018). Achieving a decision in antagonistic multi-agent networks: Frustration determines commitment strength. *2018 IEEE Conference on Decision and Control (CDC)*, pages 109 – 114.

[Garimella et al., 2021] Garimella, K., Smith, T., Weiss, R., and West, R. (2021). Political polarization in online news consumption. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 15, pages 152–162.

[Gass and Fu, 2013] Gass, S. I. and Fu, M. C., editors (2013). *Prim's Algorithm*, pages 1160–1160. Springer US, Boston, MA.

[Gurobi, 2008] Gurobi Team (2008). Gurobi optimization. `https://www.gurobi.com/`.

[Harary and Cartwright, 1968] Harary, F. and Cartwright, D. (1968). On the coloring of signed graphs. *Elemente der Mathematik*, 23:85–89.

[He and McAuley, 2016] He, R. and McAuley, J. (2016). Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering. In *Proceedings of the 25th Intl. Conf. World Wide Web*, WWW '16, pages 507–517. ACM.

[Heider, 1946] Heider, F. (1946). Attitudes and cognitive organization. *J. Psychology*, 21:107–112.

[Hu et al., 2021] Hu, Y., Lyons, R., and Tang, P. (2021). A reverse aldous-broder algorithm. *ANNALES DE L INSTITUT HENRI POINCARE-PROBABILITIES ET STATISTIQUES*, 57(2):890 – 900.

[Hüffner et al., 2010] Hüffner, F., Betzler, N., and Niedermeier, R. (2010). new block Separator-based data reduction for signed graph balancing. *Journal of combinatorial optimization*, 20(4):335–360.

[Iacono and Altafini, 2010] Iacono, G. and Altafini, C. (2010). Average frustration and phase transition in large-scale biological networks: a statistical physics approach. *IFAC Proceedings Volumes*, 43(14):320–325. 8th IFAC Symposium on Nonlinear Control Systems.

[Interian et al., 2022] Interian, R., Marzo, R. G., Mendoza, I., and Ribeiro, C. C. (2022). Network polarization, filter bubbles, and echo chambers: An annotated review of measures, models, and case studies. *arXiv preprint arXiv:2207.13799*.

[Kunegis, 2013] Kunegis, J. (2013). KONECT – The Koblenz Network Collection. In *Proceedings of the 22nd Intl. Conf. World Wide Web*, WWW '13, pages 1343–1350. ACM.

[Lai et al., 2018] Lai, M., Patti, V., Ruffo, G., and Rosso, P. (2018). Stance evolution and Twitter interactions in an Italian political debate. In Silberztein, M., Atigui, F., Kornyshova, E., Métais, E., and Meziane, F., editors, *Natural Language Processing and Information Systems*, pages 15–27, Cham. Springer International Publishing.

[Leskovec et al., 2010] Leskovec, J., Huttenlocher, D., and Kleinberg, J. (2010). Predicting positive and negative links in online social networks. In *Proceedings of the 19th Intl. Conf. World Wide Web*, WWW '10, pages 641–650. ACM.

[Leskovec and Krevl, 2014] Leskovec, J. and Krevl, A. (2014). SNAP Datasets: Stanford large network dataset collection. `http://snap.stanford.edu/data`.

[Li et al., 2021] Li, R., Yuan, X., Radfar, M., Marendy, P., Ni, W., O'Brien, T. J., and Casillas-Espinosa, P. M. (2021). Graph signal processing, graph neural network and graph learning on biological data: A systematic review. *IEEE Reviews in Biomedical Engineering*, pages 1–1.

[Liu et al., 2021a] Liu, T., Cui, J., Zhuang, H., and Wang, H. (2021a). Modeling polypharmacy effects with heterogeneous signed graph convolutional networks. *Applied Intelligence*, 51:8316–8333.

[Liu et al., 2021b] Liu, T., Sheng, A., Qi, G., and Li, Y. (2021b). Admissible bipartite consensus in networks of singular agents over signed graphs. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 68(8):2880–2884.

[Read, 1954] Read, K. (1954). Cultures of the central highlands, new Guinea. *Southwestern Journal of Anthropology*, 10(1):1–43.

[Rusnak and Tešić, 2021] Rusnak, L. and Tešić, J. (2021). Characterizing attitudinal network graphs through frustration cloud. *Data Mining and Knowledge Discovery*, 6.

[Saberi et al., 2021a] Saberi, M., Khosrowabadi, R., Khatibi, A., Misic, B., and Jafari, G. (2021a). Requirement to change of functional brain network across the lifespan. *PLoS ONE*, 16(11):1 – 19.

[Saberi et al., 2022] Saberi, M., Khosrowabadi, R., Khatibi, A., Misic, B., and Jafari, G. (2022). Pattern of frustration formation in the functional brain network. *NETWORK NEUROSCIENCE*, 6(4):1334 – 1356.

[Saberi et al., 2021b] Saberi, M., Khosrowabadi, R., Khatibi, A., Mišić, B., and Jafari, G. (2021b). Topological impact of negative links on the stability of resting-state brain network. *Scientific Reports*, 11(1):2176.

[Sampson, 1968] Sampson, S. (1968). A novitiate in a period of change: An experimental and case study of relationships. Ph.D. thesis, Cornell University.

[Schaub et al., 2016] Schaub, M. T., O'Clery, N., Billeh, Y. N., Delvenne, J.-C., Lambiotte, R., and Barahona, M. (2016). Graph partitions and cluster synchronization in networks of oscillators. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 26(9):094821.

[Seif and Ahmadi, 2014] Seif, Z. and Ahmadi, M. (2014). Computing frustration index using genetic algorithm. *Communications in Mathematical and Computer Chemistry*, 71:437–443.

[Tang and Zhu, 2023] Tang, W. and Zhu, J. (2023). Population-level balance in signed networks.

[Tomasso et al., 2022a] Tomasso, M., Rusnak, L., and Tešić, J. (2022a). Advances in scaling community discovery methods for signed graph networks. *Journal of Complex Networks*, 10(3).

[Tomasso et al., 2022b] Tomasso, M., Rusnak, L., and Tešić, J. (2022b). Cluster boosting and data discovery in social networks. In *Proceedings of the 37th ACM/SIGAPP Symposium On Applied Computing (SAC)*.

[Wu et al., 2023] Wu, Y., Meng, D., and Wu, Z.-G. geeksforgeeks. `https://www.geeksforgeeks.org/`. Accessed: 2023-06-01.

[Zhou et al., 2022] Zhou, X., Song, H., and Li, J. (2022). Residue frustration-based prediction of protein-protein interactions using machine learning. *Journal of physical chemistry*, 126(8):1719–727.