Structural Group Unfairness: Measurement and Mitigation by means of the Effective Resistance

Adrian Arnaiz-Rodriguez
adrian@ellisalicante.org
ELLIS Alicante
Spain

Georgina Curto
gcurtore@nd.edu
University of Notre Dame
Notre Dame, USA

Nuria Oliver
nuria@ellisalicante.org
ELLIS Alicante
Spain

ABSTRACT

Social networks contribute to the distribution of social capital, defined as the relationships, norms of trust and reciprocity within a community or society that facilitate cooperation and collective action. Social capital exists in the relations among individuals, such that better positioned members in a social network benefit from faster access to diverse information and higher influence on information dissemination. A variety of methods have been proposed in the literature to measure social capital at an individual level. However, there is a lack of methods to quantify social capital at a group level, which is particularly important when the groups are defined on the grounds of protected attributes. Furthermore, state-of-the-art approaches fail to model the role of long-range interactions between nodes in the network and their contributions to social capital. To fill this gap, we propose to measure the social capital of a group of nodes by means of their information flow and emphasize the importance of considering the whole network topology. Grounded in spectral graph theory, we introduce three effective resistance-based measures of group social capital, namely *group isolation*, *group diameter* and *group control*, where the groups are defined according to the value of a protected attribute. We denote the social capital disparity among different groups in a network as *structural group unfairness*, and propose to mitigate it by means of a budgeted edge augmentation heuristic that systematically increases the social capital of the most disadvantaged group. In experiments on real-world networks, we uncover significant levels of structural group unfairness when using gender as the protected attribute, with females being the most disadvantaged group in comparison to males. We also illustrate how our proposed edge augmentation approach is able to not only effectively mitigate the structural group unfairness but also increase the social capital of all groups in the network.

ACM Reference Format:
Adrian Arnaiz-Rodriguez, Georgina Curto, and Nuria Oliver. 2024. Structural Group Unfairness: Measurement and Mitigation by means of the Effective Resistance. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Online social networks play an important role in defining and sustaining the social fabric of human communities. With billions of users worldwide, they allow individuals to connect, interact and share information with one another over the internet. They have opened up new opportunities for personal and professional networking, entertainment, learning and activism. However, the formation of social networks —whether through organic growth or recommendations— can create imbalances in network positions which condition the access to resources and information [54]. These network inequalities have an impact on the social capital of its members, which exists in the relations among individuals [21]. Better positioned network members benefit from faster access to diverse information, higher influence on information dissemination and more control of the information flow [5, 14, 15, 37, 38]. In practical terms, this means that individuals with a strategic position in the network will have more influence over others, and better access to information and opportunities regarding jobs, health, education or finance.

Furthermore, link recommendation algorithms that pervade social media platforms tend to connect similar users, contributing to the homophily and clustering of the network [69, 86]. These *filter bubbles* limit the access to diverse individuals [37], exacerbate the isolation and polarization of groups [25, 35, 69], reduce the opportunities of innovation [58] and aggravate the perpetuation of societal stereotypes [40]. In sum, the topology of the network can lead to a vicious cycle where those who are disadvantaged accumulate fewer opportunities to improve their social capital [30].

A variety of graph intervention methods have been proposed in the literature to mitigate disparities in social capital at an individual level [5, 39]. However, there is a lack of methods that consider such disparities at a group level, which is particularly relevant when the groups correspond to socially vulnerable groups, *i.e.*, those defined on the grounds of sex, race, color, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other [22, 78]. Evaluating disparities at the group level provides a broader, systemic perspective that allows for the identification of overarching structural barriers or inequalities that may not be apparent when examining individual experiences. Focusing on the group level also supports the development of inclusive solutions at scale that benefit entire communities, promoting equity, diversity and the inclusion of disadvantaged groups. We denote the disparity in social capital among different groups in the network as *structural group unfairness*.

Similar to Bashardoust et al. [5], we consider a setting where each node in the network is a source of unique information and, therefore, access to all nodes is equally important. In this context, information flow is an integral component of the social capital and a distance metric that quantifies total information flow in the graph, considering high-order relations that expand beyond the immediate neighbors, is of utmost importance. We propose using the effective resistance to measure the overall information flow between pairs of nodes, since it is a theoretically grounded continuous graph diffusion metric that considers both *local* and *global* properties of the network's topology [18]. In Section 3.1.2, we introduce three measures of group social capital —*group isolation*, *group* *diameter* and *group control*— based on the effective resistance where the groups are defined according to the value of their protected attribute of interest. Based on these measures of social capital, we define three measures of structural group unfairness in Section 3.2, and frame the challenge of mitigating structural group unfairness as a budgeted edge augmentation task in Section 3.3. This section also presents the Effective Resistance Group Link (`ERG-Link`) algorithm, a greedy edge augmentation algorithm that iteratively adds the edges to the graph to increase the social capital of the most disadvantaged group. In experiments on real-world networks, described in Section 4, we uncover significant levels of structural group unfairness when using gender[1] as the protected attribute, with females being the most disadvantaged group in comparison to males. We also illustrate how our approach is able to not only mitigate structural group unfairness, but also increase the social capital of all the groups in the network.

In sum, the main contributions of our work are:
(1) We propose three effective resistance-based measures of group social capital in social networks, namely *group isolation*, *group diameter* and *group control*;
(2) We define *structural group unfairness* as a disparity in the values of such measures by different groups in the graph, where the groups are defined according to the values of a protected attribute. This approach is particularly relevant from a social perspective when the disadvantaged group in the network corresponds to a vulnerable social group;
(3) We propose `ERG-Link`, an effective resistance-based greedy edge augmentation algorithm that iteratively adds edges to the network to maximize the social capital of the most disadvantaged *group*;
(4) In experiments on real-world networks, we uncover significant levels of group structural unfairness when using gender as protected attribute, with females being the most disadvantaged group in comparison to males. We also illustrate how our approach is the most effective in reducing structural group unfairness when compared to the baselines.

2 RELATED WORK

In this section, we provide an overview of the most relevant work in the literature related to social capital, its measurement by means of computational methods, fairness in graphs from the perspective of social capital, and network interventions to mitigate unfairness.

Social capital. Social capital is as a multidimensional construct that has been extensively studied in sociology, political science, economics, and more recently, computational social science [57]. It is defined as the networks, relationships, and norms of trust and reciprocity within a community or society that facilitate cooperation and collective action [21]. In simple terms, social capital is the value derived from connections between people. It can be measured and analyzed both at an individual and collective levels [8] and it has been characterized according to different criteria. Some authors propose three main dimensions of social capital, namely: structural, emphasizing the relationships among individuals, organizations and communities; cognitive, focusing on the shared values, norms and beliefs that bind members of a group or community; and relational, highlighting the intensity and quality of relationships, including reciprocity, trust and obligations among individuals [50]. Others have proposed the distinction between bonding, bridging, and linking social capital [71]. Bonding social capital captures the aspects of "inward looking" communities that reinforce exclusive identities and homogeneous groups [21]; bridging social capital refers to "outward looking" networks across different groups that do not necessarily share similar identities [15, 37]; and linking social capital characterizes the trusting relationships and norms of respect across power or authority gradients [70, 81]. The three forms are important for the well-being of individuals and communities: bonding social capital contributes to social cohesion and support; bridging social capital to mutual understanding, solidarity and respect; and linking social capital to mobilize political resources and power.

Computational models of social capital. Network analysis offers a robust computational framework to examine and quantify social capital [15, 21, 80]. We consider a setting where all the nodes in the network may be sources of relevant information. As a consequence, access to all nodes —not just the sources or seeds of information— is equally important. In this context, *information flow* is an integral component of the social capital, and a variety of methods have been proposed to characterize it, mainly through two concepts: centrality and criticality [7].

Centrality measures the relative importance or prominence of a node in the network, quantifying its ability to reach the rest of nodes. Different approaches have been proposed in the literature to measure centrality, including the degree centrality, closeness, local clustering, the assortativity coefficient [9, 51], Katz centrality [42] and PageRank [55]. Criticality reflects the node's level of influence or vulnerability within the network [34, 74]. Nodes with high criticality are essential, such that their failure or disruption can have significant consequences, cascading effects or system-wide impact. Measures of criticality include effective size [15], redundancy [8] and shortest path betweenness [41].

However, previously proposed methods are insufficient to accurately quantify the overall information flow in the network for several reasons. First, they model the distance between nodes as the shortest path distance (geodesic distance [51]) which overlooks alternative routes and indirect connections that may exist between distant nodes, thereby underestimating the potential pathways for information diffusion, influence propagation [10, 68] or resource exchange. This myopic view can lead to oversimplified representations of network dynamics, ignoring the interplay between weak ties, bridge nodes and overlapping communities that facilitate connectivity and communication across disparate components in the network. Second, most of the proposed approaches only consider first-order —direct and local— relationships between nodes, relying on small neighborhoods of the graph. As a result, they ignore higher-order structural information [1], such as the global properties of the network topology and long-range interactions between nodes, which can lead to inaccurate insights on how information flows globally [60]. Third, popular approaches to model information flow in a network, such as the Independent Cascade model [43] assume homogeneous, deterministic and instantaneous interactions

[1]Note that we follow the same nomenclature for gender as that used in the analyzed datasets, which is a binary variable with two values: male and female.

between neighboring nodes which might lead to inaccurate predictions, biased estimations and misrepresentations of actual diffusion patterns observed in complex networks [72].

Conversely, graph diffusion metrics, such as the *effective resistance* [45, 68], offer a principled approach to quantifying distances and interactions between nodes within a network, addressing the above limitations. The effective resistance accurately captures not only short-range but also long-range relationships between nodes [18] because it considers alternative pathways, including the network dynamics, and quantifies connectivity between distant nodes. Therefore, it constitutes a natural information distance metric between nodes in a graph [10, 68]. Previous work has theoretically formulated measures of node centrality and criticality based on effective resistances [10, 11, 52, 75], yet we are not aware of any work that has modeled the social capital of a group of nodes by means of effective resistance. From a practical perspective, the concept of effective resistance has been used to measure polarization in social networks [39] and to rank user-items relations in recommender systems [32]. In this paper, we propose quantifying the social capital of a group of nodes in the network by means of three measures derived from the effective resistance: the group isolation, group diameter and group control, explained in Section 3.1.2.

Fairness in graphs. Social status plays a role in defining the structure of a network [3, 16] and a node's position in a network is a form of social capital [14, 21]. Thus, there are structural advantages in information flow depending on the position that a node occupies in the network. In the field of social networks, prior work has stud fairness from the perspective of disparities in access to information by differently positioned nodes in the graph, particularly in the case of influence maximization, *i.e.*, when a single piece of information is spread in the network [65, 79]. However, there is a scarcity of studies that model fairness considering that all the nodes are sources of information and access to all the nodes is equally important [5]. In this context and to the best of our knowledge, no previous research has considered fairness in graphs from a group perspective, when the groups are defined according to protected attributes —such as gender, ethnicity, religion or socio-economic status. In this paper, we fill this gap by defining, measuring and mitigating *structural group unfairness*, understood as disparities in social capital between different groups and where social capital is measured by information flow.

Network interventions to mitigate unfairness in graphs. Network interventions draw upon social network theory and structural analysis to understand and address the underlying mechanisms of unfairness within social networks. Interventions to mitigate structural unfairness in a network may entail redesigning network structures [39, 64] or altering (adding and/or removing) edges [5, 76] to eliminate discriminatory barriers, reduce homophily, and foster diversity within the networks [15, 37]. These interventions aim to enhance connectivity, promote inclusivity, and facilitate equitable access to resources, opportunities, and support networks [7].

When aiming to improve the social capital in a network, edge augmentation (*i.e.*, adding edges) constitutes the natural intervention to mitigate disparities [5]. Several edge augmentation strategies have been proposed in the literature, such as connecting similar nodes to improve bonding social capital [85], linking nodes with the highest product of eigenvector centralities [76] or creating edges between the most disadvantaged nodes and the central node [5]. However, these strategies are defined for individual notions of social capital and they do not consider long-range interactions between nodes. In our work, we focus on reducing the social capital disparities between groups in the graph and we propose to measure the social capital by means of the effective resistance, which is able to quantify both short- and long-range interactions between nodes. We approach this objective by adding weak ties [37] —i.e. ties that correspond to relationships that are not within one node's close-knit group (strong ties)— between the disadvantaged group and the rest of the graph.

3 MEASURING STRUCTURAL GROUP UNFAIRNESS IN A NETWORK

In this section, we first present the distance metric that quantifies the information flow between pairs of nodes in a social network and provides the theoretical basis for the proposed measures of group social capital. Next, we introduce three measures to quantify the social capital of a group of nodes in a graph and present the concept of structural group unfairness as the disparity in the measures by different groups in the graph. Finally, we describe a greedy graph intervention (edge augmentation) algorithm to mitigate structural group unfairness.

3.1 Preliminaries

3.1.1 Effective resistance and social capital. We focus on the structural dimension of social capital, which emphasizes the relationships among individuals, organizations and communities [50], and propose to measure it as the information flow of a node in the network. Such a measure is captured by the *effective resistance* [28, 45, 73] of the node. Given nodes u and v in graph $G = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V}$ is the set of nodes, $\mathcal{E} = \{(u, v) \in \mathcal{V} \times \mathcal{V} : A_{uv} = 1\}$ is the set of edges and $\mathbf{A}$ is the graph's adjacency matrix, the effective resistance R_{uv} between nodes u and v is a distance metric given by:

$$R_{uv} = (\mathbf{e}_u - \mathbf{e}_v)\mathbf{L}^{\dagger}(\mathbf{e}_u - \mathbf{e}_v)^{\top}, \tag{1}$$

where $\mathbf{e}_u$ is the unit vector with a unit value at u-th index and zero elsewhere; $\mathbf{L}^{\dagger} = \sum_{i>1} \frac{1}{\lambda_i}\phi\phi^{\top}$ is the pseudo-inverse of the graph's Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{A} = \mathbf{\Phi}\mathbf{\Lambda}\mathbf{\Phi}^{\top}$, with $\mathbf{D}$ the graph's degree matrix, $D_{u,u} = \sum_{j \in V} A_{u,v}$ and 0 elsewhere; and λ_i the i-th smallest eigenvalue of $\mathbf{L}$ corresponding to the ϕ_i eigenvector. The complete matrix of all pairwise effective resistances in a graph, $\mathbf{R}$ is given by $\mathbf{R} = \mathbf{1}\,\mathrm{diag}(\mathbf{L}^{\dagger})^{\top} + \mathrm{diag}(\mathbf{L}^{\dagger})\mathbf{1}^{\top} - 2\mathbf{L}^{\dagger}$.

The effective resistance is a distance *metric* since it satisfies the symmetry, non-negativity and triangle inequality conditions [29]. In addition, R_{uv} is proportional to the commute times between u and v, *i.e.*, the expected number of steps in a random walk starting at v to reach node u and come back: $R_{uv} \propto \mathbb{E}_u[v] + \mathbb{E}_v[u]$, where $\mathbb{E}_u[v]$, $\mathbb{E}_v[u]$ are the expected number of steps that a random walker takes to go from u to v and from v to u, respectively [17, 31, 32, 58, 73]. A high value of R_{uv} means that u and v generally struggle to visit each other in a random walk, *i.e.*, nodes with high effective resistance between them are unlikely to exchange information. $R_{u,v}$ can be expressed as

$$R_{u,v} = \sum_{i=0}^{\infty} \left(d_u^{-1} (\mathbf{A}^i)_{uu} + d_v^{-1} (\mathbf{A}^i)_{vv} - (d_u d_v)^{-1/2} 2 (\mathbf{A}^i)_{uu} \right),$$

being $\mathbf{A}^k$ the matrix that defines the number paths of length k between u and v [6]. Hence, it is able to capture both short- and long-range interactions between nodes in the graph.

The effective resistance has been characterized as the *information distance* in a network [10, 45, 68, 73] as it quantifies the amount of effort (distance) required to transmit information between the nodes. The total effective resistance R_{tot} of a graph [29] – defined as the sum of all R_{uv} ($\mathsf{R}_{\text{tot}} = \mathbf{1R1}^\top$) – is therefore inversely proportional to the expected ease of information flow in the graph.

The *total effective resistance of node u*, $\mathsf{R}_{\text{tot}}(u)$, is given by $\mathsf{R}_{\text{tot}}(u) = \sum_{v \in \mathcal{V}} R_{uv}$, *i.e.*, the sum of all the effective resistances between node u and the rest of nodes in the network. Given that the effective resistance is an information distance, the smaller the total effective resistance of a node, the larger its information flow. In other words, the effective resistance allows to identify which nodes in a graph have limited information flow (*i.e.*, high effective resistance) and thus low social capital [29, 36]. Equivalent terms to denote the effective resistance in the literature include the current-flow closeness centrality [11, 51] and the information centrality [68].

From a computational perspective, calculating R_{uv} does not require hyper-parameter tuning and can be efficiently calculated, mitigating two significant drawbacks of other diffusion or learnable graph distances [56]. An overview of the theoretical properties of R_{uv} are provided in Appendix B.

3.1.2 Effective-resistance-based measures of group social capital. Based on the definition of effective resistance above, we propose three metrics that characterize the social capital of a group of nodes in a graph. The metrics are computed based on pairwise effective resistances between the nodes in the group and the rest of nodes the graph, which are then aggregated to all members of the group to quantify its social capital. Each of the proposed metrics is based on theoretical principles derived from effective resistance, as explained in the previous section and in Appendix B.3. In the following, we refer to a group of nodes S_i in graph $G = \{\mathcal{V}, \mathcal{E}\}$ as a subset of $\mathcal{V}$, *i.e.*, $S_i \subseteq \mathcal{V}$ with $|S_i|$ nodes.

1. Group Isolation. The isolation of a group S_i, $\mathsf{R}_{\text{tot}}(S_i)$, is given by the average of the total effective resistances of all the nodes in group. $\mathsf{R}_{\text{tot}}(S_i)$ is proportional to the expected information distance when sampling one node from group S_i and another node at random. It can be interpreted as a proxy for the marginalization of a group from the perspective of information flow, such that the lower the $\mathsf{R}_{\text{tot}}(S_i)$, the less isolated the group S_i is in the network. Therefore, reducing this measure for group S_i would yield an increase in its social capital. It is given by:

$$\mathsf{R}_{\text{tot}}(S_i) = \mathbb{E}_{u \sim S_i}[R_{tot}(u)] = \frac{1}{|S_i|} \sum_{u \in S_i} \mathsf{R}_{\text{tot}}(u) = |\mathcal{V}| \mathbb{E}_{u \sim S_i, v \sim \mathcal{V}} [R_{uv}] \tag{2}$$

where $\mathsf{R}_{\text{tot}}(u) = \sum_{v \in \mathcal{V}} R_{uv}$ is the total effective resistance of node u, $\mathcal{V}$ is the set of all nodes in the graph, and S_i is group of interest. The normalization factor enables comparing groups of different sizes. Note that adding links between nodes with the highest R_{uv} —irrespective of which group they belong to— has been found to reduce the total effective resistance of a graph [6, 36] and hence the isolation of all the graph's nodes.

2. Group Diameter. The group diameter, $\mathcal{R}_{\text{diam}}(S_i)$, measures the average of the maximum distance between any node in group S_i and any node in the graph. A larger group diameter suggests that the nodes in group S_i are distant from the rest of the graph, indicating potential challenges in information exchange with the nodes outside of S_i, and hence it can be interpreted as another measure of social capital.

This measure is based on $\mathcal{R}_{\text{diam}}(G)$, which is the maximum effective resistance of the graph [58].

$$\mathcal{R}_{\text{diam}}(S_i) = \mathbb{E}_{u \sim S_i}[\mathcal{R}_{\text{diam}}(u)] = \mathbb{E}_{u \sim S_i}[\max_{v \in \mathcal{V}} R_{uv}] \tag{3}$$

where $\mathcal{R}_{\text{diam}}(u) = \max_{v \in \mathcal{V}} R_{uv}$ is the diameter of node u, *i.e.*, the maximum R_{uv} from u to any other node in the graph. $\mathcal{R}_{\text{diam}}(S_i)$ gives an indication of the information flow gap between the group S_i and the rest of the network [30]. Therefore, the larger the $\mathcal{R}_{\text{diam}}(S_i)$, the lower the social capital of group S_i.

3. Group Control. The aforementioned concepts measure the amount of information flow in a group of nodes in the graph. Another relevant variable to assess is the criticality of a node for the diffusion of information in the graph, which in the literature has been measured as betweenness [34], redundancy [8] or effective size [15]. Nodes with high levels of control serve as important connectors in the network, facilitating the flow of information and enabling communication between otherwise disconnected groups of nodes [15].

The control of a node can be computed by restricting the summation of a node's total effective resistance to the nodes that are directly linked to it (*i.e.*, its direct neighbors). Thus, it is expressed as $\mathsf{B}_{\mathsf{R}}(u) = \sum_{v \in \mathcal{N}(u)} R_{uv}$, where $\mathcal{N}(u) = \{v : (u,v) \in \mathcal{E}\}$ are the neighbors of u. The larger the $\mathsf{B}_{\mathsf{R}}(u)$, the more control a node has in the network's information flow and hence the larger its social capital. The node control of a node is bounded by $1 \leq \mathsf{B}_{\mathsf{R}}(u) \leq d_u$, being d_u the number of neighbors of node u (see Theorem C.1). $\mathsf{B}_{\mathsf{R}}(u)$ is theoretically related to the current-flow betweenness [11, 52, 74, 75], the node's information bottleneck [2], and the curvature of the node [26, 77].

We define the group control or group betweenness $\mathsf{B}_{\mathsf{R}}(S_i)$ as the average of the controls of all the nodes in S_i, *i.e.*:

$$\mathsf{B}_{\mathsf{R}}(S_i) = \mathbb{E}_{u \sim S_i}[\mathsf{B}_{\mathsf{R}}(u)] = \frac{1}{|S_i|} \sum_{u \in S_i} \mathsf{B}_{\mathsf{R}}(u), \tag{4}$$

and is bounded by $1 \leq \mathsf{B}_{\mathsf{R}}(S_i) \leq \text{vol}(S_i)/|S_i|$, being $\text{vol}(S_i)/|S_i|$ the average degree of the group S_i (see Theorem C.2). Note that the sum of all R_{uv} for all nodes in a graph is constant at $|\mathcal{V}| - 1$ and it is independent of the number of edges [45]. If an edge is added, removed or modified in the graph, all R_{uv} are updated accordingly such that their sum remains constant. The sum and the average control of all nodes in the graph are also constant with values $\sum_{u \in \mathcal{V}} \mathsf{B}_{\mathsf{R}}(u) = 2|V| - 2$ and $\mathbb{E}_{u \sim \mathcal{V}}[\mathsf{B}_{\mathsf{R}}(u)] = 2 - \frac{2}{|\mathcal{V}|}$, respectively, independently of the number of edges (see Appendix C.1.2 for more details). Consequently, the control of a node or group of nodes is distributed among the nodes in the network and cannot be optimized for every node/group in the graph by adding more edges: if a node or group of nodes increase their control over the information flow in the graph, they must do so at the cost of reducing the control of other nodes.

3.2 Structural group unfairness

To study disparities in the distribution of social capital in the network, we define the groups of nodes S_i according to their values of a sensitive attribute $i \in SA = \{sa_1, sa_2, \ldots, sa_{|SA|}\}$, which is a categorical variable with $|SA|$ possible values that refer to a socially relevant concept, such as sex, age, gender, religion or race. We denote the value of the sensitive attribute of a node v as $SA(v)$. For instance, if SA is sex with three possible values, SA={male, female, non-binary}, the groups S_{male}, S_{female} and $S_{\text{non-binary}}$ are the set of nodes whose sex is labeled as male, female and non-binary, respectively.

We define the *structural group unfairness* in a network as the disparity in information flow between the nodes belonging to different groups in the network. Since we have defined the groups in terms of protected attributes, structural group unfairness becomes particularly relevant because it informs about potential disparities in information flow (and hence social capital) between a vulnerable group –e.g. females– and the rest of the network. We present here three metrics to characterize the structural group unfairness, namely *isolation disparity*, *diameter disparity* and *control disparity*.

1. Group Isolation Disparity. Ideally, every group in the network should have the same levels of information flow and hence the same —and low— levels of group isolation, namely:

$$\mathsf{R}_{\text{tot}}(S_i) = \mathsf{R}_{\text{tot}}(S_j),\ \forall\ i, j \in SA. \tag{5}$$

Deviations from equality lead to isolation disparity $\Delta\,\mathsf{R}_{\text{tot}}$, which is defined as the maximum over all groups in the graph of the differences in group isolation: $\Delta\,\mathsf{R}_{\text{tot}} = \max_{i,j\in SA}(\mathsf{R}_{\text{tot}}(S_i) - \mathsf{R}_{\text{tot}}(S_j))$.

Reducing the isolation disparity contributes to increasing the social capital of the most disadvantaged group and equalizes the information flow between the groups in the network.

2. Group Diameter Disparity. Ideally, every social group in the network should have the same –and low– group diameter:

$$\mathcal{R}_{\text{diam}}(S_i) = \mathcal{R}_{\text{diam}}(S_j),\ \forall\ i, j \in SA. \tag{6}$$

Any deviations from equality lead to diameter disparity, $\Delta\,\mathcal{R}_{\text{diam}}$, defined as the maximum over all groups in the graph of the differences in group diameter: $\Delta\,\mathcal{R}_{\text{diam}} = \max_{i,j\in SA}(\mathcal{R}_{\text{diam}}(S_i) - \mathcal{R}_{\text{diam}}(S_j))$.

Achieving equal diameter entails equalizing the worst-case scenario in information flowing to the entire network from the perspective of any group of nodes in the graph. By promoting equal group diameter, we generate a fairer information-sharing environment.

3. Group Control Disparity. By striving for equalized control in all groups in the network, no particular group would dominate or be marginalized from the perspective of their control of information in the network:

$$\mathsf{B}_{\mathsf{R}}(S_i) = \mathsf{B}_{\mathsf{R}}(S_j) = 2 - \frac{2}{|\mathcal{V}|},\ \forall\ i, j \in SA. \tag{7}$$

Any deviations from equality lead to control disparity, $\Delta\,\mathsf{B}_{\mathsf{R}}$, defined as the maximum over all groups in the graph of the differences in group control: $\Delta\,\mathsf{B}_{\mathsf{R}} = \max_{i,j\in SA}(\mathsf{B}_{\mathsf{R}}(S_i) - \mathsf{B}_{\mathsf{R}}(S_j))$.

As previously explained, control is a bounded resource to be distributed among the groups of nodes in the graph with an expected value of $2 - \frac{2}{|\mathcal{V}|}$. Hence, reducing the control disparity entails a redistribution of the control in all the groups in the graph converging to $\mathsf{B}_{\mathsf{R}}(S_i) = 2 - \frac{2}{|\mathcal{V}|},\ \forall\ i \in SA$, leading to a more equitable allocation of the control that different groups play regarding the information flow in the network.

3.3 Edge augmentation to mitigate structural group unfairness

Edge augmentation. Edge augmentation is the natural intervention to mitigate information flow disparities in a network where all nodes are sources of unique pieces of information [5, 15].

Regarding which structural group unfairness measure we should aim to optimize, we argue that we should primarily focus on improving the isolation disparity of the most isolated group in the graph. Note that mitigating isolation will also yield an improvement in the diameter and control disparities, as illustrated in our experiments. The reduction in isolation entails creating edges between distant nodes, *i.e.*, fostering the creation of weak ties. Granovetter [37]'s work provides evidence that information spreads more effectively through weak ties than through strong ties: weak ties give peripheral nodes more visibility in the network, which leads to a decrease in group isolation and diameter. Adding weak ties reduces discontinuities in the information flow, increases redundancies in the paths between nodes and improves the control of peripheral nodes while reducing the control of dominant ones [15].

Previous work has suggested connecting peripheral isolated nodes (with low centrality and control) to salient nodes (with high centrality and control) [5, 41, 76]. However, these solutions lead to a *rich-get-richer* phenomenon that benefits the best connected nodes and potentially increases the disparities in information access [13]. Hence, we advocate for the creation of edges between the most distant nodes in the network (weak ties) without necessarily connecting them with a central node.

Problem definition. We consider a budgeted edge augmentation intervention: given a maximum number B of allowed new connections to be created in the graph, we aim to identify the B new edges $\mathcal{E}'$ to be added to the graph G that would maximally reduce the group isolation disparity of the most disadvantaged group in the graph. This leads to a new graph G' with lower levels of structural group unfairness:

$$G' = \min_{G'=(\mathcal{V},\mathcal{E}')} \quad \mathbb{E}_{u,v\sim V\times V}\,[R_{uv}] \quad \text{s.t.} \quad |\mathcal{E}' \setminus \mathcal{E}| = B \quad \mathcal{E} \subset \mathcal{E}' \tag{8}$$

Algorithm. To tackle the problem above, we introduce `ERG-Link`, a greedy algorithm that adds edges between the nodes with the largest effective resistance between them, where at least one of the nodes belongs to the most isolated group as per Section 3.1.2, and groups in the graph are defined in the grounds of a protected attribute. Note that this strategy also reduces the isolation (total effective resistance) of the entire graph [36].

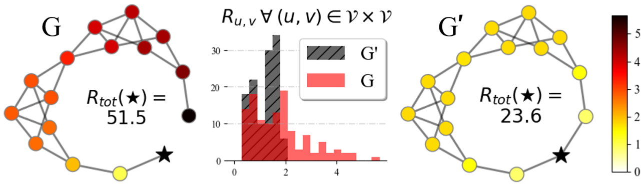

Figure 1: Illustration of the impact of adding one edge on the information flow of G. The color denotes information distances (effective resistances) w.r.t. to the star node. The histogram shows the change of all $R'_{u,v}s$ in G. Note how all $R'_{u,v}s$ between the star node and the rest of nodes in the network ($R_{\bigstar,v}$) decrease in G' even if there is no change in the geodesic distance between them.

Algorithm 1 outlines the main steps of ERG-Link. Given a graph $G = (V, E)$, a protected attribute S and a total budget B of new edges to add, the group isolation $R_{tot}(S)$ is computed for each group according to S. The most disadvantaged group S_d is identified as the group with largest $R_{tot}(S)$. Then, the $R_{uv}\ \forall(u,v) \in \mathcal{V} \times \mathcal{V}$ is computed, and a ranking of all potential new edges in the graph is created from the highest to the lowest values of effective resistance. In each iteration, ERG-Link adds the new edge to the graph that yields the largest improvement in the information flow of S_d, *i.e.*, the edge that connects the two nodes with largest effective resistance between them where at least one of the nodes belongs to S_d. See Black et al. [6], Ghosh et al. [36] for a proof that such an edge is the one that maximally improves the information flow in the graph.

Algorithm 1: ERG-Link

Data: Graph $G = (\mathcal{V}, \mathcal{E})$, a protected attribute SA, budget B of total number of edges to add
Result: New Graph $G' = (\mathcal{V}', \mathcal{E}')$ with B new edges

1 $\mathbf{L} = \mathbf{D} - \mathbf{A}$;
2 $S_d = \text{argmax}_{S_i \forall i \in SA} \text{R}_{\text{tot}}(S_i)$; `// Identify the most disadvantaged group`
3 **Repeat**
4 $\mathbf{L}^{\dagger} = \sum_{i>0} \frac{1}{\lambda_i} \phi_i \phi_i^{\top} = \left(\mathbf{L} + \frac{\mathbf{1}\mathbf{1}^{\top}}{n}\right)^{-1} - \frac{\mathbf{1}\mathbf{1}^{\top}}{n}$;
5 $\mathbf{R} = \mathbf{1}\,\text{diag}(\mathbf{L}^{\dagger})^{\top} + \text{diag}(\mathbf{L}^{\dagger})\mathbf{1}^{\top} - 2\mathbf{L}^{\dagger}$; `// Compute effective resistance`
6 $C = \{(u,v) \mid u \in S_d \text{ or } v \in S_d, (u,v) \notin \mathcal{E}'\}$; `// Select edge candidates`
7 $\mathcal{E}' = \mathcal{E}' \cup \arg\max_{(u,v) \in C} R_{uv}$; `// Add edge with maximum effective resistance from C`
8 $\mathbf{L} = \mathbf{L} + (\mathbf{e}_u - \mathbf{e}_v)(\mathbf{e}_u - \mathbf{e}_v)^{\top}$; `// Fast update of L`
9 **Until** $|\mathcal{E}' \setminus \mathcal{E}| = B$;
10 **return** G';

ERG-Link leverages *Rayleigh's monotonicity principle* [28, 29], according to which the total effective resistance of a graph can only decrease when new edges are added to it, as illustrated in Fig. 1. The new edges that most highly reduce both the graph's total and maximum effective resistances are those that connect the two nodes with the largest effective resistance between them [2, 6, 36, 58]. Therefore, creating an edge between nodes with maximum R_{uv} not only improves the information flow between the two nodes (increasing their social capital) but it also improves the information flow of the entire graph.

Note that the addition of each new edge changes all the pairwise information distances between nodes in the graph, requiring the re-computation of all distances (effective resistances) in each iteration. Therefore, this type of edge augmentation is not feasible by means of Independent Cascade distance estimation [5, 43], random-walk embeddings [56] or graph neural networks [44, 82]. These methods require training expensive neural networks or running complex simulations for the estimation of the distances in each iteration. In addition, they only capture short-range interactions between nodes. Conversely, the effective resistance captures both short and long-range interactions between nodes in the graph and it is efficient to update. While it requires the computation of the Laplacian pseudo-inverse, Woodbury's formula [6] can be used to avoid recomputing $\mathbf{L}^{\dagger}$ in line 3 of Algorithm 1, as reflected in Algorithm 2.

4 EXPERIMENTS

4.1 Datasets and set-up

Table 1: Group social capital in the original graphs. Group with the largest social capital is highlighted in bold.

G	$R_{tot} \downarrow$	$\mathcal{R}_{diam} \downarrow$	$\text{B}_\text{R} \uparrow$
Facebook (female)	221.4	2.29	1.93
Facebook (male)	**179.8**	**2.25**	**2.03**
UNC28 (female)	608.6	2.11	1.99
4051			
UNC28 (male)	**586.3**	2.11	**2.00**
	5.14	3767	
Google+ (female)	564.1	1.31	1.81
Google+ (male)	**287.7**	**1.24**	**2.32**

To empirically evaluate ERG-Link, we tackle the challenge of mitigating structural group unfairness in three real-world networks (school and online social networks), where the nodes are users and the edges correspond to connections between them, *i.e.*, friendships. The three datasets are commonly used in the graph fairness literature, namely:
(1) The Facebook dataset [47], a dense graph of 1,034 Facebook users ($|\mathcal{V}|$) and 26,749 edges ($|\mathcal{E}|$). It corresponds to a large ego-network where nodes are connected if they are friends in the social network;
(2) The UNC28 dataset [62], consisting of a 2005 snapshot from the Facebook network of the university of North Carolina ($|\mathcal{V}|$=3985, $|\mathcal{E}|$=65287);
(3) The Google+ dataset [47], an ego-network of G+, the social network developed by Google, with 3,508 nodes ($|\mathcal{V}|$) and 25,3930 edges ($|\mathcal{E}|$).

Table 2: Structural group unfairness before and after the graph interventions.

(a) Facebook (B=50)

	Δ R$_{tot}$	$\Delta \mathcal{R}_{diam}$	Δ B$_R$
G (original)	41.62	0.042	0.107
Random	38.7	0.039	0.108
DW	36.3	0.031	0.104
Cos	28.7	0.029	0.120
ERG	**10.3**	**0.009**	**0.098**
S-DW	43.6	0.041	0.103
S-Cos	41.4	0.042	0.105
S-ERG	41.6	0.042	0.107

(b) UNC28 (B=5000)

	Δ R$_{tot}$	$\Delta \mathcal{R}_{diam}$	Δ B$_R$
G (original)	22.4	0.006	0.009
Random	19.8	0.005	0.014
DW	22.2	0.006	0.004
Cos	19.1	0.005	0.102
ERG	**8.8**	**0.002**	**0.003**
S-DW	20.6	0.006	0.008
S-Cos	22.1	0.006	0.019
S-ERG	22.3	0.006	0.004

(c) Google+ (B=5000)

	Δ R$_{tot}$	$\Delta \mathcal{R}_{diam}$	Δ B$_R$
G (original)	276.4	0.078	0.51
Random	129.4	0.037	0.47
DW	274.1	0.078	0.51
Cos	86.8	0.025	0.47
ERG	**37.1**	**0.011**	**0.29**
S-DW	272.5	0.078	0.49
S-Cos	236.0	0.067	0.47
S-ERG	276.4	0.079	0.52

Gender is the protected attribute in all networks with two possible values SA ={male, female}. We select the largest connected component for all the datasets.

The original values of group social capital per gender are depicted in Table 1. As seen in the Table, our study unveils that the disadvantaged group according to the three defined measures (group isolation, group diameter and group control) corresponds to females in the three datasets. Hence, $S_d = S_{female}$.

In contrast, group extensions of traditional individual social capital metrics fail to model long-range interactions between nodes and hence are unable to uncover the extent of social capital disparities between groups. For example, in the Google+ dataset, females are 49% more isolated than men according to our proposed group isolation (R$_{tot}$) metric (564.1 vs 287.7 for females vs males). However, they are only 7% more isolated (8,073 vs 7,477 for females vs males) when using the geodesic distance and a traditional *farness centrality measure* [63]. Furthermore, according to the commonly used betweenness measure (number of shortest paths through a node) [34] on the UNC28 dataset, males —as opposed to females— are the disadvantaged group with 7% less control than females (3,767 vs 4,051).

4.2 Baselines

We compare edge augmentation by means of `ERG-Link` with three baselines:
(1) *Random*: an algorithm that adds edges at random to the graph.
(2) DW: an algorithm that adds edges using the dot product similarity of *DeepWalk* [56] embeddings as a distance between nodes. It is based on sampling random walks and training a BERT model to compute each node's embedding. Results correspond to the following hyper-parameters: 128 embedding dimension, 40 walk length, and 10 window size;
(3) *Cos*: a greedy algorithm that adds edges using the cosine similarity of the rows of the adjacency matrix [65] as a distance between nodes. This is an example of a classic method based on neighborhood similarity.

All the baselines correspond to an algorithm similar to Algorithm 1 (lines 2 and 6 remain the same) with one difference: instead of using the effective resistances to quantify the distances between nodes, the baselines consider cosine or DW distances. Note that we do not include any GNN-based method as a baseline because they require training a neural network to estimate the pairwise distances in the graph. This is computationally unfeasible in our task as it would entail retraining the neural network every time a new edge is added [82].

4.3 Experimental methodology

We set a budget B of a maximum of 5,000 new links to be added to the the UNC28 and Google+ datasets, which corresponds to approximately 0.05% of the number of all potential edges in the graph. We also run experiments with a maximum of 50 new edges for the Facebook dataset to showcase that even with an extremely low budget, `ERG-Link` is able to significantly improve the social capital of the disadvantaged group and reduce the structural unfairness in the graph. We compute the social capital for each group (male and female) and the structural group unfairness on both the original and the augmented graphs (after all edges have been added) based on the defined measures. We also compute them at each step of edge addition to shed light on the evolution of the structural group unfairness as new edges are added.

4.4 Structural group unfairness mitigation

Table 2 depicts the three structural group unfairness measures on the original graph G and after adding 50 edges to the Facebook dataset and 5,000 edges to the UNC28 and Google+ datasets. The groups are defined based on gender (male, female) and the disadvantaged group are females according to the three structural group unfairness measures, as depicted in Table 1 and top row of Table 2. The disparities in social capital between males and females are particularly large in the Google+ network with a Δ R$_{tot}$ of 276.4, meaning that females have significantly lower levels of information flow than males in this network. Δ B$_R$ also shows a difference of 0.51 on the control of the network, which is a large difference given that the values of B$_R$ are in the range $[0, 2 - 2/|\mathcal{V}|]$.

Regarding the results of the graph intervention algorithms, we observe how edge augmentation via `ERG-Link` outperforms all the baseline methods on the three datasets in terms of reducing structural group unfairness. Interestingly, the larger the unfairness in the original graph, the larger the improvement after the intervention with `ERG-Link`. For example, in the case of isolation disparity, Δ R$_{tot}$, the original values of 22.4, 41.62 and 276.4 improve by **60%**,

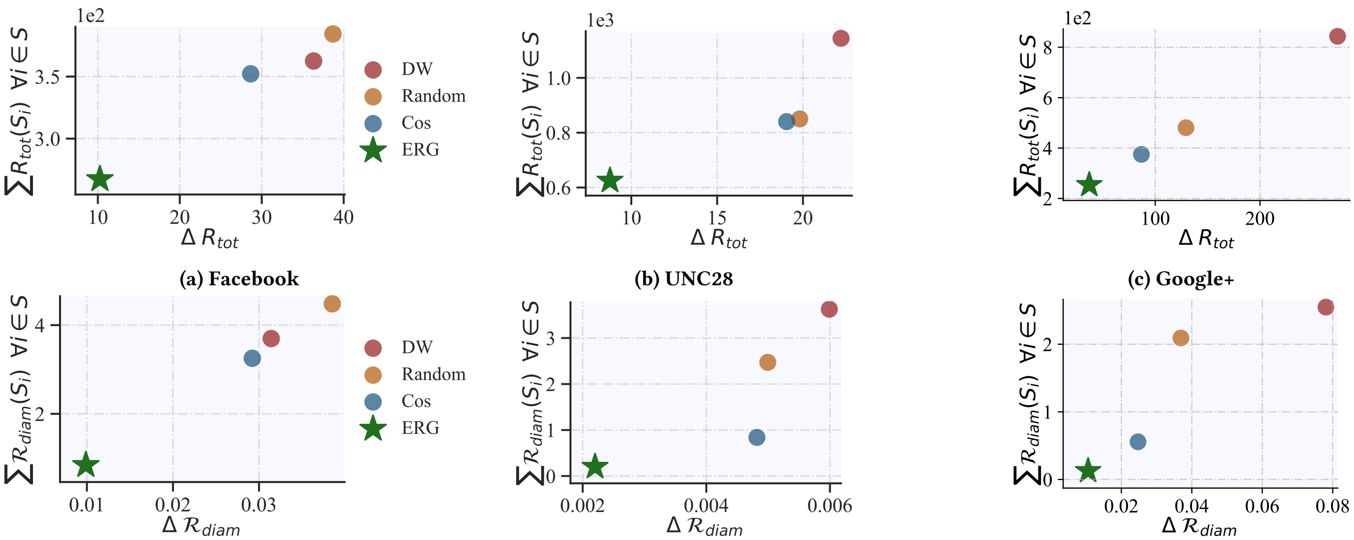

Figure 2: Pareto front of the structural group unfairness (X-axis) vs the sum of the group social capital of all the groups (Y-axis) using $\mathrm{R_{tot}}$ on the top row and $\mathcal{R}_{\mathbf{diam}}$ on the bottom row. From left to right: Facebook, UNC28 and Google+ datasets.

75% and **87%** after `ERG-Link`'s intervention in the UNC28, Facebook and Google+ networks, respectively. A similar behavior is observed for the other structural group unfairness measures.

For illustration purposes, we also report results of the three versions of Algorithm 1 using the three distances (DeepWalk, cosine and effective resistance) but where the added edges connect the nodes with the smallest —instead of the largest— pairwise distance, similar to how link recommendation algorithms work. We refer to these methods as the "Strong" version (for strong ties) and hence denote them with an "S-" before the abbreviation. The bottom part of Table 2 contains the results of edge augmentation with these variations. As expected, edge augmentation in this case does not significantly improve the structural group unfairness since such methods are not designed to improve the information flow in the most isolated nodes in the graph, but to connect nodes that are already structurally close to each other (foster strong ties).

Note how the edge augmentations when using DW and Cos distances still yield graphs with significant levels of structural group unfairness. Furthermore, in the case of using DW distances, the improvement in performance worsens as the graph gets larger (Google+). While using Cos distance for edge augmentation improves $\Delta\,\mathrm{R_{tot}}$ and $\Delta\,\mathcal{R}_{\mathrm{diam}}$, it is unable to always improve $\Delta\,\mathrm{B_R}$ due to the inherently more intricate nature of control disparity optimization. Unlike $\Delta\,\mathrm{R_{tot}}$ and $\Delta\,\mathcal{R}_{\mathrm{diam}}$, minimizing $\Delta\,\mathrm{B_R}$ requires the precise identification of network gaps which are difficult to detect using a cosine similarity distance. For completeness, the social capital metrics for each group (males and females) in the three data sets are shown in Table 4 in Appendix D.1. Note how females always have lower values of social capital ($\mathrm{R_{tot}}$, $\mathcal{R}_{\mathrm{diam}}$ and $\mathrm{B_R}$) than males, even after the graph interventions.

4.5 Overall social capital improvement

In this section, we illustrate how edge augmentation via `ERG-Link` is more effective than the baseline methods not only to reduce the structural group unfairness ($\Delta\,\mathrm{R_{tot}}$ and $\Delta\,\mathcal{R}_{\mathrm{diam}}$), but also to improve the overall social capital measures of group isolation $R_{tot}(S_i)$ and group diameter $R_{diam}(S_i)$ for all the groups in the graph. For each dataset, the structural group unfairness metric ($\Delta\,\mathrm{R_{tot}}$ or $\Delta\,\mathcal{R}_{\mathrm{diam}}$) is shown on the x-axis and the overall group isolation (sum of group isolation for all groups) or overall group diameter (sum of diameters for all groups) on the y-axis. For both axes, the lower the values, the better. Edge augmentation via `ERG-Link` clearly outperforms any other graph intervention strategy, providing evidence that it not only reduces inequalities in social capital between groups, but also improves the social capital of all groups by reducing their group isolation.

4.6 Evolution of structural group unfairness throughout the interventions

Fig. 3 illustrates the evolution of the structural group unfairness metrics for the Google+ dataset as new edges are added to the network with a total budget of 5,000 edges. As seen in the Figure, edge augmentation via `ERG-Link` quickly mitigates the group isolation and diameter disparities, even after the addition of a small number of edges. Furthermore, edge augmentation via `ERG-Link` exhibits a smoother and more consistent reduction in control disparity ($\Delta\,\mathrm{B_R}$), in contrast to the stair-step behavior observed when adding edges using the baseline methods.

Note that $\mathrm{B_R}$ is a finite resource to be allocated among the groups and cannot be globally maximized. Hence, the right-most graphs show how $\Delta\,\mathrm{B_R}$ is improved by decreasing $\mathrm{B_R}$ of the group with highest initial $\mathrm{B_R}$ and increasing it otherwise (top-right). The goal is to converge both controls to the optimal bound of $2 - 2/|\mathcal{V}|$ as indicated by the horizontal black line.

We also show in Appendix D.3 the evolution of the structural group unfairness and group social capital metrics when the budget B of edges to add is very small (50 edges) and large (5,000 edges) to study the efficacy of the interventions. Fig. 4 in Appendix D.3 depict

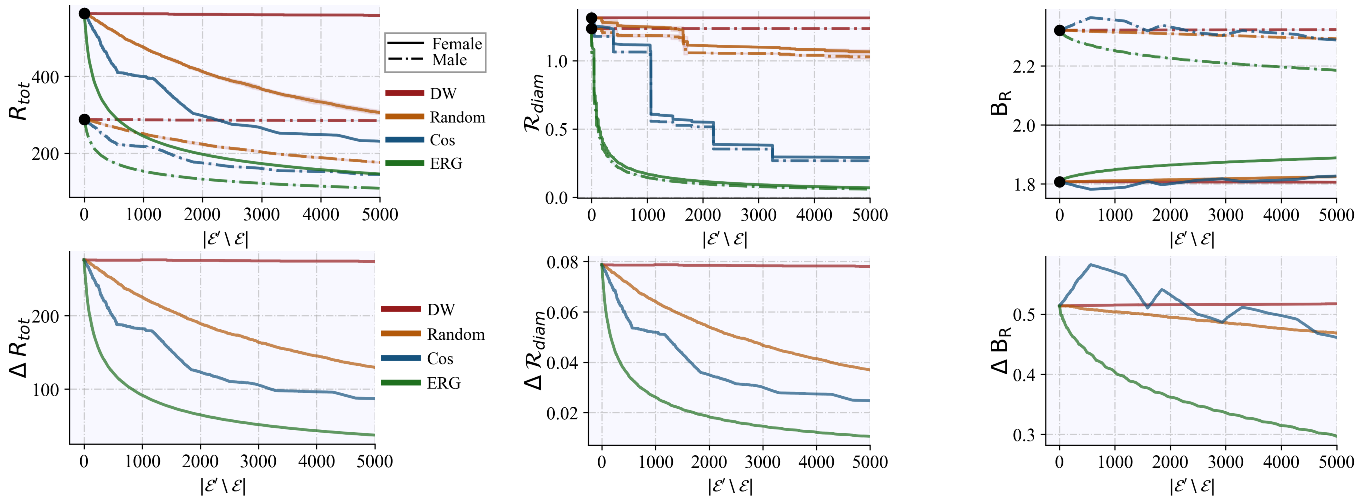

Figure 3: Evolution of the structural group unfairness metrics as the number of added edges increases on the Google+ dataset with a total budget of 5,000 new edges.

the distribution of the social capital metrics for the nodes in each of the groups in the original graph and the resulting graphs after edge augmentation. The results are consistent: for a fixed budget, edge augmentation via `ERG-Link` yields the best results both for the mitigation of structural group unfairness and the increase in the overall social capital of all the groups in the graph. Additionally, for a fixed structural group unfairness mitigation goal, edge augmentation via `ERG-Link` achieves it with a significantly smaller budget of added edges than any of the baseline methods.

5 DISCUSSION, CONCLUSION AND FUTURE WORK

In this paper, we have presented a novel method, based on the effective resistance, to measure and mitigate structural unfairness in groups within a social network, where the groups are defined according to the values of a protected attribute. We have proposed a graph intervention approach rooted in the literature of the philosophy of discrimination. We acknowledge that in the literature there is not a single universal and absolute interpretation of fairness [23]. At present, the distributive paradigm of fairness [61] dominates the political and philosophical discussions of social justice [83]. In social networks, this translates into the pursuit of equal social capital for all individuals in the network. Nevertheless, there are alternative interpretations of fairness [19] that remain unexplored in social networks. For example, a sufficitarian approach to fairness [53, 66] would entail guaranteeing a minimum threshold of social capital for all network members, while a prioritarian conception of fairness [19] would give priority to the individuals belonging to a vulnerable social group. Moreover, prominent voices in social philosophy have long defended that fairness can be understood not only as the redistribution of social capital among individuals, but also as the social recognition of identity groups [33]. In this paper, we focus on group fairness in graphs and combine both a prioritarian and a Rawlsian approach to fairness since our intervention aims to mitigate the social capital gap for the most disadvantaged group while increasing the social capital of all groups. Given the importance of social networks in the definition of the social fabric, the ultimate goal of our work is to spur a reflection on the potential use of social networks as reparative tools for social inequality [24].

While one could argue that defining groups in terms of protected attributes can be considered a form of discrimination, our approach does not aim to systematically increase the social capital of a pre-defined vulnerable social group, but to detect those that are structurally disadvantaged in the network (*i.e.*, those having the lowest levels of social capital) and implement a mechanism at scale that benefits an entire community. This type of mitigation of structural unfairness is particularly relevant since the potential disadvantages suffered by groups in social networks can add up to already existing social conditions that contribute to systemic injustice [84].

In terms of limitations, we have not validated the proposal with real users, who could be reluctant to accept graph interventions that connect them with nodes outside of their close-knit circle, despite the benefits that these *weak ties* bring as contributors to innovation [59] and social equity. Nonetheless, we envision our proposal as a complement to existing graph interventions in a hybrid setup that combines the addition of edges connecting nodes with both small (strong ties) and large (weak ties) effective resistances. We leave for future work such a validation. Also, in line with the body of work on intersectional bias mitigation which focuses on demographic dimensions [46], we have considered *gender* as a binary protected attribute in our experiments. We would have preferred to work with a non-binary approach to gender, yet available social network datasets only contain gender as a binary variable.

From a technical perspective, we plan to explore alternative edge augmentation algorithms to mitigate structural group unfairness using our effective resistance-based measures measures and we would like to incorporate additional node features corresponding to the characteristics of the individuals in the network. We also maintain ongoing discussions with several community organizations to define a project with participants that arrive in another country as refugees, for whom information access and connection with the local population of the hosting country are key.

ACKNOWLEDGMENTS

This work was supported by the European Commission under Horizon Europe Programme, grant number 101120237 - ELIAS, by Intel corporation, a nominal grant received at the ELLIS Unit Alicante Foundation from the Regional Government of Valencia in Spain (Convenio Singular signed with Generalitat Valenciana, Conselleria de Innovación, Industria, Comercio y Turismo, Dirección General de Innovación) and a grant by the Banc Sabadell Foundation.

REFERENCES

[1] Sameer Agarwal, Kristin Branson, and Serge Belongie. 2006. Higher Order Learning with Graphs. In *Proceedings of the 23rd International Conference on Machine Learning*. Association for Computing Machinery, New York, NY, USA, 17–24. https://doi.org/10.1145/1143844.1143847

[2] Adrián Arnaiz-Rodríguez, Ahmed Begga, Francisco Escolano, and Nuria Oliver. 2022. DiffWire: Inductive Graph Rewiring via the Lovász Bound. In *Proceedings of the First Learning on Graphs Conference (Proceedings of Machine Learning Research, Vol. 198)*. PMLR, 15:1–15:27. https://proceedings.mlr.press/v198/arnaiz-rodriguez22a.html

[3] Brian Ball and Mark EJ Newman. 2013. Friendship networks and social status. *Network Science* 1, 1 (2013), 16–30.

[4] Pradeep Kr Banerjee, Kedar Karhadkar, Yu Guang Wang, Uri Alon, and Guido Montúfar. 2022. Oversquashing in GNNs through the lens of information contraction and graph expansion. In *2022 58th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 1–8.

[5] Ashkan Bashardoust, Sorelle A Friedler, Carlos E Scheidegger, Blair D Sullivan, and Suresh Venkatasubramanian. 2023. Reducing Access Disparities in Networks using Edge Augmentation. In *Proceedings of the 2023 Conference on Fairness, Accountability, and Transparency*. ACM.

[6] Mitchell Black, Zhengchao Wan, Amir Nayyeri, and Yusu Wang. 2023. Understanding Oversquashing in GNNs through the Lens of Effective Resistance. In *Proceedings of the 40th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 202)*. PMLR, 2528–2547.

[7] Stephen P Borgatti. 2005. Centrality and network flow. *Social networks* 27, 1 (2005), 55–71.

[8] Stephen P Borgatti, Candace Jones, and Martin G Everett. 1998. Network measures of social capital. *Connections* 21, 2 (1998), 27–36.

[9] Stephen P Borgatti, Ajay Mehra, Daniel J Brass, and Giuseppe Labianca. 2009. Network analysis in the social sciences. *science* 323, 5916 (2009), 892–895.

[10] Enrico Bozzo and Massimo Franceschet. 2013. Resistance distance, closeness, and betweenness. *Social Networks* 35, 3 (2013), 460–469.

[11] Ulrik Brandes and Daniel Fleischer. 2005. Centrality measures based on current flow. In *Annual symposium on theoretical aspects of computer science*. Springer, 533–544.

[12] Ronald S Burt. 1987. Social contagion and innovation: Cohesion versus structural equivalence. *American journal of Sociology* 92, 6 (1987), 1287–1335.

[13] Ronald S Burt. 1999. The social capital of opinion leaders. *The Annals of the American Academy of Political and Social Science* 566, 1 (1999), 37–54.

[14] Ronald S Burt. 2000. The network structure of social capital. *Research in organizational behavior* 22 (2000), 345–423.

[15] Ronald S Burt. 2004. Structural holes and good ideas. *American journal of sociology* 110, 2 (2004), 349–399.

[16] Karen E Campbell, Peter V Marsden, and Jeanne S Hurlbert. 1986. Social resources and socioeconomic status. *Social networks* 8, 1 (1986), 97–117.

[17] Ashok K Chandra, Prabhakar Raghavan, Walter L Ruzzo, and Roman Smolensky. 1989. The electrical resistance of a graph captures its commute and cover times. In *Proceedings of the twenty-first annual ACM symposium on Theory of computing*. 574–586.

[18] Fan RK Chung. 1997. *Spectral graph theory*. Vol. 92. American Mathematical Soc.

[19] Mark Coeckelbergh. 2022. *The Political Philosophy of AI*. Polity.

[20] Ronald R Coifman, Stephane Lafon, Ann B Lee, Mauro Maggioni, Boaz Nadler, Frederick Warner, and Steven W Zucker. 2005. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proceedings of the National Academy of Sciences of the United States of America* 102, 21 (2005), 7426–7431.

[21] James S Coleman. 1988. Social capital in the creation of human capital. *American journal of sociology* 94 (1988), S95–S120.

[22] Council of Europe. 2010. European Convention on Human Rights. www.conventions.coe.int.

[23] Georgina Curto and Flavio Comim. 2023. SAF: Stakeholder's Agreement on Fairness in the Practice of Machine Learning Developent. *Science and Engineering Ethics* 29 (2023).

[24] J.L. Davis, A. Williams, and M.W. Yang. 2021. Algorithmic reparation. *Big Data and Society* 8, 2 (2021).

[25] Michela Del Vicario, Gianna Vivaldo, Alessandro Bessi, Fabiana Zollo, Antonio Scala, Guido Caldarelli, and Walter Quattrociocchi. 2016. Echo chambers: Emotional contagion and group polarization on facebook. *Scientific reports* 6, 1 (2016), 37825.

[26] Karel Devriendt and Renaud Lambiotte. 2022. Discrete curvature on graphs from the effective resistance. *Journal of Physics: Complexity* 3, 2 (2022), 025008.

[27] Francesco Di Giovanni, Lorenzo Giusti, Federico Barbero, Giulia Luise, Pietro Lio, and Michael Bronstein. 2023. On Over-Squashing in Message Passing Neural Networks: The Impact of Width, Depth, and Topology. In *International Conference on Machine Learning*.

[28] Peter G Doyle and J Laurie Snell. 1984. *Random walks and electric networks*. Vol. 22. American Mathematical Soc.

[29] Wendy Ellens, Floske M Spieksma, Piet Van Mieghem, Almerima Jamakovic, and Robert E Kooij. 2011. Effective graph resistance. *Linear algebra and its applications* 435, 10 (2011), 2491–2506.

[30] Benjamin Fish, Ashkan Bashardoust, Danah Boyd, Sorelle Friedler, Carlos Scheidegger, and Suresh Venkatasubramanian. 2019. Gaps in information access in social networks?. In *The World Wide Web Conference*. 480–490.

[31] Francois Fouss, Stephane Faulkner, Manuel Kolp, Alain Pirotte, Marco Saerens, et al. 2005. Web Recommendation System Based on a Markov-Chainmodel.. In *ICEIS (4)*. 56–63.

[32] Francois Fouss, Alain Pirotte, Jean-Michel Renders, and Marco Saerens. 2007. Random-walk computation of similarities between nodes of a graph with application to collaborative recommendation. *IEEE Transactions on knowledge and data engineering* 19, 3 (2007), 355–369.

[33] Nancy Fraser and Axel Honneth. 2003. *Redistribution or recognition? A political -philosophical exchange*. Verso Books.

[34] Linton C Freeman. 1977. A set of measures of centrality based on betweenness. *Sociometry* (1977), 35–41.

[35] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. 2018. Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. In *Proceedings of the 2018 world wide web conference*. 913–922.

[36] Arpita Ghosh, Stephen Boyd, and Amin Saberi. 2008. Minimizing effective resistance of a graph. *SIAM review* 50, 1 (2008), 37–66.

[37] Mark Granovetter. 1983. The strength of weak ties: A network theory revisited. *Sociological theory* (1983), 201–233.

[38] Didem Gündoğdu, Pietro Panzarasa, Nuria Oliver, and Bruno Lepri. 2019. The bridging and bonding structures of place-centric networks: Evidence from a developing country. *PloS one* 14, 9 (2019), e0221148.

[39] Marilena Hohmann, Karel Devriendt, and Michele Coscia. 2023. Quantifying ideological polarization on a network using generalized Euclidean distance. *Science Advances* 9, 9 (2023), eabq2044. https://doi.org/10.1126/sciadv.abq2044 arXiv:https://www.science.org/doi/pdf/10.1126/sciadv.abq2044

[40] Neema Iyer. 2021. *Alternate Realities, Alternate Internets: African Feminist Research for a Feminist Internet*. Springer International Publishing, Cham, 93–113. https://doi.org/10.1007/978-3-030-83734-1_6

[41] Matthew O Jackson. 2019. *The human network: How your social position determines your power, beliefs, and behaviors*. Vintage.

[42] Leo Katz. 1953. A new status index derived from sociometric analysis. *Psychometrika* 18, 1 (1953), 39–43.

[43] David Kempe, Jon Kleinberg, and Éva Tardos. 2003. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. 137–146.

[44] Thomas N Kipf and Max Welling. 2016. Variational Graph Auto-Encoders. In *NeurIPS Workshop on Bayesian Deep Learning*.

[45] Douglas J Klein and Milan Randić. 1993. Resistance distance. *Journal of mathematical chemistry* 12 (1993), 81–95.

[46] John P. Lalor, Yi Yang, Kendall Smith, Nicole Forsgren, and Ahmed Abbasi. 2022. Benchmarking Intersectional Biases in NLP. *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (2022), 3598–3609. https://doi.org/10.18653/V1/2022.NAACL-MAIN.263

[47] Jure Leskovec and Julian Mcauley. 2012. Learning to discover social circles in ego networks. *Advances in neural information processing systems* 25 (2012).

[48] Yaron Lipman, Raif M Rustamov, and Thomas A Funkhouser. 2010. Biharmonic distance. *ACM Transactions on Graphics (TOG)* 29, 3 (2010), 1–11.

[49] László Lovász. 1993. Random walks on graphs. *Combinatorics, Paul erdos is eighty* 2, 1-46 (1993), 4. https://web.cs.elte.hu/~lovasz/erdos.pdf

[50] Janine Nahapiet and Sumantra Ghoshal. 1998. Social capital, intellectual capital, and the organizational advantage. *Academy of management review* 23, 2 (1998), 242–266.

[51] Mark EJ Newman. 2003. Mixing patterns in networks. *Physical review E* 67, 2 (2003), 026126.

[52] Mark EJ Newman. 2005. A measure of betweenness centrality based on random walks. *Social networks* 27, 1 (2005), 39–54.

[53] Martha C Nussbaum. 2012. *Creating Capabilities*. Harvard University Press, Cambridge, Massachusetts and London, England.

[54] Aviv Ovadya and Luke Thorburn. 2023. Bridging Systems: Open Problems for Countering Destructive Divisiveness across Ranking, Recommenders, and Governance. *arXiv preprint arXiv:2301.09976* (2023).

[55] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1998. *The pagerank citation ranking: Bring order to the web*. Technical Report. technical report, Stanford University.

[56] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. 2014. Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. 701–710.

[57] Robert D Putnam. 2015. Bowling alone: America's declining social capital. In *The city reader*. Routledge, 188–196.

[58] Huaijun Qiu and Edwin R Hancock. 2007. Clustering and embedding using commute times. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 11 (2007), 1873–1890.

[59] Karthik Rajkumar, Guillaume Saint-Jacques, Iavor Bojinov, Erik Brynjolfsson, and Sinan Aral. 2022. A causal test of the strength of weak ties. *Science (American Association for the Advancement of Science)* 377, 6612 (2022), 1304–1310. https://doi.org/10.1126/science.abl4476

[60] Ladislav Rampášek, Michael Galkin, Vijay Prakash Dwivedi, Anh Tuan Luu, Guy Wolf, and Dominique Beaini. 2022. Recipe for a general, powerful, scalable graph transformer. *Advances in Neural Information Processing Systems* 35 (2022), 14501–14515.

[61] J Rawls. 1971. *A Theory of Justice*. Oxford University Press, Oxford.

[62] Veronica Red, Eric D Kelsic, Peter J Mucha, and Mason A Porter. 2011. Comparing community structure to characteristics in online collegiate social networks. *SIAM review* 53, 3 (2011), 526–543.

[63] Gert Sabidussi. 1966. The centrality index of a graph. *Psychometrika* 31, 4 (1966), 581–603.

[64] Fernando P Santos, Yphtach Lelkes, and Simon A Levin. 2021. Link recommendation algorithms and dynamics of polarization in online social networks. *Proceedings of the National Academy of Sciences* 118, 50 (2021), e2102141118.

[65] Akrati Saxena, George Fletcher, and Mykola Pechenizkiy. 2022. Fairsna: Algorithmic fairness in social network analysis. *arXiv preprint arXiv:2209.01678* (2022).

[66] Amartya Sen. 2001. *Development as freedom*. Oxford University Press.

[67] Daniel A Spielman and Nikhil Srivastava. 2008. Graph sparsification by effective resistances. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*. 563–568.

[68] Karen Stephenson and Marvin Zelen. 1989. Rethinking centrality: Methods and examples. *Social networks* 11, 1 (1989), 1–37.

[69] Jessica Su, Aneesh Sharma, and Sharad Goel. 2016. The effect of recommendations on network structure. In *Proceedings of the 25th international conference on World Wide Web*. 1157–1167.

[70] Simon Szreter. 2002. The state of social capital: Bringing back in power, politics, and history. *Theory and society* 31, 5 (2002), 573–621.

[71] Simon Szreter and Michael Woolcock. 2004. Health by association? Social capital, social theory, and the political economy of public health. *International journal of epidemiology* 33, 4 (2004), 650–667.

[72] Sotharith Tauch, William Liu, and Russel Pears. 2015. Measuring cascade effects in interdependent networks by using effective graph resistance. In *2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 683–688.

[73] Prasad Tetali. 1991. Random walks and the effective resistance of networks. *Journal of Theoretical Probability* 4 (1991), 101–109.

[74] Ali Tizghadam and Alberto Leon-Garcia. 2009. A graph theoretical approach to traffic engineering and network control problem. In *2009 21st International Teletraffic Congress*. IEEE, 1–8.

[75] Ali Tizghadam and Alberto Leon-Garcia. 2010. Betweenness centrality and resistance distance in communication networks. *IEEE Network* 24, 6 (2010), 10–16. https://doi.org/10.1109/MNET.2010.5634437

[76] Hanghang Tong, B Aditya Prakash, Tina Eliassi-Rad, Michalis Faloutsos, and Christos Faloutsos. 2012. Gelling, and melting, large graphs by edge manipulation. In *Proceedings of the 21st ACM international conference on Information and knowledge management*. 245–254.

[77] Jake Topping, Francesco Di Giovanni, Benjamin Paul Chamberlain, Xiaowen Dong, and Michael M. Bronstein. 2022. Understanding over-squashing and bottlenecks on graphs via curvature. In *International Conference on Learning Representations*. https://openreview.net/forum?id=7UmjRGzp-A

[78] United Nations. 1966. International Covenant on Civil and Political Rights.

[79] Xindi Wang, Onur Varol, and Tina Eliassi-Rad. 2022. Information access equality on generative models of complex networks. *Applied Network Science* 7, 1 (2022), 1–20.

[80] Stanley Wasserman and Katherine Faust. 1994. Social network analysis: Methods and applications. (1994).

[81] Michael Woolcock et al. 2001. The place of social capital in understanding social and economic outcomes. *Canadian journal of policy research* 2, 1 (2001), 11–17.

[82] Shiwen Wu, Fei Sun, Wentao Zhang, Xu Xie, and Bin Cui. 2022. Graph neural networks in recommender systems: a survey. *Comput. Surveys* 55, 5 (2022), 1–37.

[83] I.M. Young. 1990. *Justice and the politics of difference*. Princeton University Press, Princeton, N.J.

[84] Edward N. Zalta and Uri Nodelman (Eds.). 2011. *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/discrimination/

[85] Saijun Zhang, Steven G Anderson, and Min Zhan. 2011. The differentiated impact of bridging and bonding social capital on economic well-being: An individual level perspective. *J. Soc. & Soc. Welfare* 38 (2011), 119.

[86] Matteo Zignani, Sabrina Gaito, Gian Paolo Rossi, Xiaohan Zhao, Haitao Zheng, and Ben Zhao. 2014. Link and triadic closure delay: Temporal metrics for social network dynamics. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 8. 564–573.

A TABLE OF NOTATION

Table 3: Table of Notation

Symbol	Description	Definition
$G = (\mathcal{V}, \mathcal{E})$	Graph = (Nodes, Edges)	
$n = \|\mathcal{V}\|$	Number of nodes	
$\mathbf{A}$	Adjacency matrix: $\mathbf{A} \in \mathbb{R}^{n\times n}$	$A_{u,v} = 1$ if $(u, v) \in \mathcal{E}$ and 0 otherwise
v or u	Node $v \in \mathcal{V}$ or $u \in \mathcal{V}$	
e	Edge $e \in \mathcal{E}$	
d_v	Degree of node, i.e., number of neighbors v	$d_v = \sum_{u\in V} A_{v,u}$
$\mathbf{D}$	Degree diagonal matrix where d_v in D_{vv}	$\mathbf{D} = \text{diag}(d_0, \dots, d_{\mathcal{V}})$
$\text{vol}(G)$	Sum of the degrees of the graph	$\text{vol}(G) = \sum_{u\in V} d_u = 2\|\mathcal{E}\| = \text{Tr}[\mathbf{D}]$
$\mathcal{N}(u)$	Neighbors of u	$\mathcal{N}(u) = \{v : (u, v) \in \mathcal{E}\}$
S_i	Subset of nodes	$S \subseteq V$
SA	Set of sensitive attributes	$SA = \{sa_1, sa_2, \dots, sa_{\|SA\|}\}$
$SA(v)$	Value of the sensitive attribute of the node v	
sa_i	Specific value of a sensitive attribute, e.g., sa_i=*female*.	
S_i	Set of nodes defined by their sensitive attribute	$S_i = \{v \in V \| SA(v) = sa_i\}$
S_d	Set of nodes defined by their sensitive attribute with the highest level of isolation $\text{R}_{\text{tot}}(S_i)$	
$\mathbf{L}$	Graph Laplacian	$\mathbf{L} = \mathbf{D} - \mathbf{A} = \mathbf{\Phi}\mathbf{\Lambda}\mathbf{\Phi}^\top$
Λ	Eigenvalue matrix of $\mathbf{L}$	
Φ	Matrix of eigenvectors of $\mathbf{L}$	
λ_i	The i-th smallest eigenvalue of $\mathbf{L}$	
$\mathbf{f}_i$	Eigenvector associated with the i-th smallest eigenvalue of $\mathbf{L}$	
$\mathbf{L}^+$	The pseudo-inverse of $\mathbf{L}$	$\mathbf{L}^+ = \sum_{i>1} \lambda_i^{-1} \phi\phi^\top$
h_G	Cheeger constant	Eq. 9
$\mathbf{e}_u$	Unit vector with unit value at u and 0 elsewhere	
R_{uv}	Effective resistance between nodes u and v	$R_{uv} = (\mathbf{e}_u - \mathbf{e}_v)\mathbf{L}^+(\mathbf{e}_u - \mathbf{e}_v)$
$\mathbf{R}$	Effective resistance matrix where the i, j entry corresponds to R_{ij}	$\mathbf{R} = \mathbf{1}\text{diag}(\mathbf{L}^+)^\top + \text{diag}(\mathbf{L}^+)\mathbf{1}^\top - 2\mathbf{L}^+$
$\mathbf{Z}$	Commute Time Embedding matrix	$\mathbf{Z} = \sqrt{vol(G)}\mathbf{\Lambda}^{-1/2}\mathbf{\Phi}^\top$
$\mathbf{z}_u$	Commute times embedding of node $Z_{u,:}$	
$\text{CT}(u, v)$	Commute time	$\text{CT}(u, v) = \text{vol}(G)R_{u.v}$
R_{tot}	Total Effective Resistance of G	$\text{R}_{\text{tot}} = \frac{1}{2}\mathbf{1}^\top\mathbf{R}\mathbf{1}$
$\mathcal{R}_{\text{diam}}$	Resistance Diameter of G	$\mathcal{R}_{\text{diam}} = \max_{u,v\in\mathcal{V}} R_{u,v}$
$\text{R}_{\text{tot}}(u)$	Node Isolation or Total Effective Resistance	$\text{R}_{\text{tot}}(u) = \sum_{v\in V} R_{uv}$
$\mathcal{R}_{\text{diam}}(u)$	Node Resistance Diameter	$\mathcal{R}_{\text{diam}}(u) = \max_{v\in V} R_{uv}$
$\text{B}_\text{R}(u)$	Node Control or Resistance Betweenness	$\text{B}_\text{R}(u) = \sum_{v\in\mathcal{N}(u)} R_{uv}$
$\text{R}_{\text{tot}}(S_i)$	Group Isolation or Total Effective Resistance	$\text{R}_{\text{tot}}(S_i) = \|S\|^{-1} \sum_{u\in S} \text{R}_{\text{tot}}(u)$
$\mathcal{R}_{\text{diam}}(S_i)$	Group Resistance Diameter	$\mathcal{R}_{\text{diam}}(S) = \|S\|^{-1} \sum_{u\in S} \mathcal{R}_{\text{diam}}(u)$
$\text{B}_\text{R}(S_i)$	Group Control or average Betweenness	$\text{B}_\text{R}(S) = \|S\|^{-1} \sum_{u\in S} \text{B}_\text{R}(u)$

B EFFECTIVE RESISTANCE AND INFORMATION FLOW

In this section, we provide an overview of the graph diffusion concepts that constitute the theoretical foundation of the effective resistance as a measure of information flow in a graph.

B.1 Graph Diffusion Measures

Discrete information propagation. Information propagation in networks has been widely studied [12, 43], prominently by means of graph diffusion and Random Walks methods. A Random Walk (RW) on a graph is a Markov chain that starts at a given node i, and moves randomly to another node from its neighborhood with probability $1/D_{i,i}$. The RW transition probability matrix is given by $\mathbf{P} = \mathbf{D}^{-1}\mathbf{A}$ and defines the discrete probability of a random walker to move from node u to node v. $\mathbf{P}^k$ is the k-th power of the transition matrix P: the entry $(P^k)_{ij}$

denotes the probability of transitioning from node i to node j in exactly k steps. The graph's diffusion matrix is defined as $\mathbf{T} = \sum_{k=0}^{\infty} \theta_k \mathbf{P}^k$, and it represents the cumulative effect of multiple steps of a random walk on the graph. Each entry T_{ij} of $\mathbf{T}$ corresponds to the probability of transitioning from node i to node j over an infinite number of steps. θ_k is known as the teleport probability at step k in the random walk. It quantifies the likelihood that, at each step, the random walker will teleport to a random node instead of following an edge. Thus, the sequence $\{\theta_k\}$ is a series of teleport probabilities over the steps. The resulting T captures the cumulative probabilities of transitioning between nodes over an infinite number of steps in the random walk, such that the probability of co-occurrence of two nodes on a random walk corresponds to the probability of information flowing between these two nodes.

However, this approach to assess information flow between nodes in a graph has several limitations. First, it requires considering all the potential paths in a graph, which might not be computationally feasible for large graphs. To overcome this issue, a value of k is typically chosen, which limits the power of the method. Second, the teleport probabilities, θ_k, need to be defined for each k-hop. Several methods have studied how to approximate it, such as Independent Cascade [43], Katz [42], SIR or PageRank [55]. Independent Cascade or SIR methods [43] are based on infection models, where they sample guided random walks and, thus, usually rely on expensive Monte Carlo simulations leading to a sub-optimal probability of transition, unable to consider the topology of the entire graph.

Graph continuous diffusion metrics. Graph continuous diffusion metrics —such as the Heat kernel distance, [20], effective resistance (or *commute times* distance) [31, 32, 45, 58] or the bi-harmonic distance [48]— arise as a generalization of random walk metrics. Their mathematical foundations allow for a better characterization of the information flow and an intuitive interpretation of the diffusion processes in a network.

Diffusion metrics define distances based on fine-grained nuances of the topology of the graph that are not captured by simple geodesic distances. When two nodes can be reached by many paths, they should be *closer* than when they can be reached only by few paths of equal length. When two nodes can be reached by a set of edge-independent paths, they are *closer* than when they are reached by redundant paths. Similarly, when two nodes are separated by a shorter path, they are *closer* than when they are separated by a longer path [10].

In addition, these metrics provide a node embedding, *i.e.*, a numerical representation of each node in the graph that reflects its importance in the process of information diffusion. These embeddings capture the global structure of the network because they incorporate both the local and global geometry of the graph.

The continuous diffusion metrics can be computed using the pseudo-inverse (or Green's function) of the combinatorial graph Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{A}$, or the normalized Laplacian $\mathcal{L} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$ [32]. The pseudo-inverse, denoted as $\mathbf{L}^{+}$ is computed using the spectral decomposition: $\mathbf{L}^{+} = \sum_{i \geq 0} \lambda_i^{-1} \phi\phi^{\top}$ where $\mathbf{L} = \mathbf{\Phi}\mathbf{\Lambda}\mathbf{\Phi}^{\top}$, where λ_i is the i-smallest eigenvalue of the Laplacian corresponding to the ϕ_i eigenvector.

In this paper, we use the effective resistance, which is a continuous diffusion metric.

B.2 Effective Resistance and Commute Times

The Commute Time (CT) [58], $\mathsf{CT}(u, v)$, is the expected number of steps that a random walker needs to go from node u to v and come back to u. The Effective Resistance, R_{uv}, is the Commute Time divided by the volume of the graph [45]. In addition to providing a distance for all pairs of nodes —whether connected or not— R_{uv} may be viewed as an indicator of the criticality or importance of the edges in the flow of information throughout the network [67].

Intuitively, this distance captures how structurally similar and connected are two nodes in a graph. If two nodes are structurally similar to each other, then the effective resistance between them will be small. Conversely, if two nodes are weakly or not connected, then their effective resistance will be large. In addition, we can define a commute time embedding (CTE, $\mathbf{Z} = \sqrt{vol(G)}\mathbf{\Lambda}^{-1/2}\mathbf{\Phi}^{\top}$) of the nodes in the graph —similar to the idea of the *node's access signature* in Bashardoust et al. [5]—, where the Euclidean distance in such an embedding corresponds exactly to the commute times $\mathsf{CT}(u, v) = ||Z_{u,:} - Z_{v,:}||^2 = \mathbb{E}_u[v] + \mathbb{E}_v[u] = 2|\mathcal{E}|R_{uv}$. This distance is upper bounded by the geodesic distance, with equality in the case of the graph being a tree.

Note that the effective resistance does not rely on any parameter and it is an accurate metric to measure the graph's information flow [2, 17, 36], as explained next.

Information flow in a graph. The graph's information flow is given by the graph's conductance, which is measured leveraging the Cheeger Constant, h_G, of a graph [18]:

$$h_G = \min_{H \subseteq V} \frac{|\{e = (u, v) : u \in S, v \in \bar{H}\}|}{\min(vol(H), vol(\bar{H}))} \tag{9}$$

The larger h_G, the harder it is to disconnect the graph into separate communities. Therefore, to increase the information flow in the network, one could add edges to the original graph G, creating a new graph G', such that $h_{G'} > h_G$. In addition, by virtue of the Cheeger Inequality, h_G is bounded by the smallest non-zero eigenvalue of $\mathbf{L}$ defined as λ_2:

$$2h_G \leq \lambda_2 < \frac{h_G^2}{2} \tag{10}$$

Finally, the CT is bounded by λ_2 as per the Lovász Bound [49]

$$\left|\frac{\mathrm{CT}(u,v)}{\mathrm{vol}(G)} - \left(\frac{1}{d_u} + \frac{1}{d_v}\right)\right| \leq \frac{1}{\lambda_2}\frac{2}{d_{min}} \tag{11}$$

where vol(G) is the volume of G, *i.e.*, the sum of the degrees of the all nodes in the graph; d_u, d_v are the degrees of nodes u and v, respectively; and d_{min} is the minimum degree in the graph.

Therefore, a graph's information flow is bounded by λ_2 which is bounded by h_G. The intuition is that graphs with large $\lambda_2 \propto h_G$ have short CT distances and thus they have better information flow. Edge augmentation in a graph would lead to a new graph G' where $h_{G'} > h_G$, with smaller CT distances and therefore better information flow.

The effective resistance is also related to other ways of computing the information flow between two nodes in a graph, such as the Jacobian [6, 27, 77].

B.3 Theoretical Metrics Derived from Effective Resistance

In this section, we introduce several measures that are derived from the effective resistance. These measures have been proposed in previous work and constitute the grounds for the group social capital metrics that we propose in this paper (Sections 3.1.2 and 3.2).

Total effective resistance. $\mathrm{R_{tot}}$ [29] is the sum of all effective distances in the graph. A lower value of $\mathrm{R_{tot}}$ indicates ease of signal propagation across the entire network and hence larger information flow. $\mathrm{R_{tot}}$ is given by:

$$\mathrm{R_{tot}} = \mathrm{R_{tot}}(\mathcal{V}) = \frac{1}{2}\mathbf{1}^\top\mathbf{R}\mathbf{1} = \frac{1}{2}\sum_{(v,u)\in V} R_{uv} = n\sum_{2}^{n}\frac{1}{\lambda_n} = n\,\mathrm{Tr}(\mathbf{L}^\dagger) \tag{12}$$

The minimum $\mathrm{R_{tot}} = |V| - 1 = n - 1$ is achieved in a fully connected graph. Conversely, the maximum $\mathrm{R_{tot}}$ is achieved on a path graph (or linear graph) and $\mathrm{R_{tot}} = \sum_i^{n-1} i = \frac{1}{2}(n(n-1))$. Therefore, $\mathrm{R_{tot}}$ is —for connected graphs— in the range $[n-1, \frac{n(n-1)}{2}]$

Additionally, since the distance between u and v is the Euclidean distance in the embedding $\mathbf{Z}$, $\mathrm{R_{tot}}$ can be obtained as follows [36]:

$$\mathrm{R_{tot}} = \sum_{(v,u)\in V} ||Z_{u:} - Z_{v:}||^2 = n\sum_{u\in V}||Z_{u:}||^2 \tag{13}$$

Similarly to R_{uv}, $\mathrm{R_{tot}}$ is theoretically related to the connectivity of the graph defined by it smallest non-zero eigenvalue. Ellens et al. [29] demonstrated the relation between $\mathrm{R_{tot}}$ and λ_2:

$$\frac{n}{\lambda_2} < \mathrm{R_{tot}} \leq \frac{n(n-1)}{\lambda_2}. \tag{14}$$

Resistance diameter. The proposed group resistance diameter is based on the resistance diameter of a graph $\mathcal{R}_{\mathrm{diam}}$, which is the maximum effective resistance on the graph [17, 58]:

$$\mathcal{R}_{\mathrm{diam}} = \max_{u,v\in V} R_{uv} \tag{15}$$

$\mathcal{R}_{\mathrm{diam}} \propto \lambda_2$ [17, 18], since

$$\frac{1}{n\lambda_2} \leq \mathcal{R}_{\mathrm{diam}} \leq \frac{2}{\lambda_2} \tag{16}$$

and specifically [2, 58]:

$$h_G \leq \frac{\alpha^\epsilon}{\sqrt{\mathcal{R}_{\mathrm{diam}}\cdot\epsilon}} vol(S)^{\epsilon-1/2}, \tag{17}$$

By [58] we know that

$$\mathcal{R}_{\mathrm{diam}} \leq \frac{1}{\lambda_2} \text{ and } \mathcal{R}_{\mathrm{diam}} \leq \frac{1}{h_G^2} \tag{18}$$

In addition, $\mathcal{R}_{\mathrm{diam}}$ is related to the *cover time* of the graph, which is the expected time required for a random walk to visit every node at least once, *i.e.*, the expected time for a piece of information to reach the entire network. $\mathcal{R}_{\mathrm{diam}}$ can be used to estimate the cover time of the graph, as per [17]:

$$m\,\mathcal{R}_{\mathrm{diam}} \leq \text{cover time} \leq O(m\,\mathcal{R}_{\mathrm{diam}}\log n) \tag{19}$$

Resistance betweenness and curvature. As the effective resistance is an information distance, this metric can be used to propose alternative betweenness or criticality metrics to the shortest path betweenness [52]. In the literature, several effective resistance-based measures have been proposed to determine a node's criticality, such as: the current flow betweenness [10, 11, 52, 74, 75], the resistance curvature of a node [26, 77], and the information bottleneck property of a node [2, 4, 6]. Note that the last two definitions are mathematically equivalent [26].

In this work, we focus on the resistance curvature of a node and the information bottleneck, which have been shown to define how much information is squeezed into a node when the information flows in the graph [2, 4]. The resistance curvature of a node is expressed as [26]:

$$p_u = 1 - \frac{1}{2}\sum_{v\in\mathcal{N}(u)} R_{uv} \tag{20}$$

Therefore, it fulfills the following equality:

$$p_u = 1 - \frac{1}{2} \sum_{v \in \mathcal{N}(u)} R_{uv} = 1 - \frac{1}{2} \mathrm{B_R}(u) \rightarrow \mathrm{B_R}(u) = -2(p_u - 1). \tag{21}$$

Although the definition of a node's resistance curvature involves computing the sum of the effective resistances between the node and its neighboring nodes, the overall structure of the graph affects all R_{uv}'s and, consequently, the value of p_i and $\mathrm{B_R}(u)$. The curvature of a node is bounded by $1 - d_u/2 \leq p_u \leq 1/2$.

C GROUP SOCIAL CAPITAL METRICS AND EDGE AUGMENTATION ALGORITHM

In this section, we explore the properties and attributes of the proposed group social capital metrics and present an efficient version of the greedy edge augmentation algorithm described in the main paper.

C.1 Group Social Capital Metrics

C.1.1 Group Isolation and Isolation Disparity. Group Isolation is based on the previously explained notion of total effective resistance of a graph and its close connection with the current flow closeness centrality [52]. We propose to define the isolation of a node as its total effective resistance $\mathrm{R_{tot}}(u) = \sum_{v \in \mathcal{V}} R_{uv}$. The proposed group isolation is obtained as the expectation in $\mathrm{R_{tot}}(u)$ for all the nodes in the group:

$$\begin{aligned}
\mathrm{R_{tot}}(S_i) =& \mathbb{E}_{u \sim S_i}\left[\mathrm{R_{tot}}(u)\right] \\
=& |\mathcal{V}|\, \mathbb{E}_{u \sim S_i}\left[\frac{1}{|\mathcal{V}|} \sum_{v \in \mathcal{V}} R_{uv}\right] = |\mathcal{V}| \mathbb{E}_{u \sim S_i}\left[\mathbb{E}_{v \sim \mathcal{V}}\left[R_{uv}\right]\right] \\
=& |\mathcal{V}| \mathbb{E}_{u \sim S_i, v \sim \mathcal{V}}\left[R_{uv}\right],
\end{aligned} \tag{22}$$

and it is also proportional to the average R_{uv} of the nodes in S to all other nodes in the graph, multiplied by the size of that group.

$\mathrm{R_{tot}}(S_i)$ is computed as the expectation of $\mathrm{R_{tot}}(u)$ for all nodes in group S_i:

$$\begin{aligned}
\mathrm{R_{tot}}(S_i) =& |\mathcal{V}| \mathbb{E}_{u \sim S_i, v \sim \mathcal{V}}\left[R_{uv}\right] \\
=& |\mathcal{V}| \frac{1}{|S_i|} \sum_{u \in S_i} \frac{1}{|\mathcal{V}|} \sum_{v \in \mathcal{V}} R_{uv} = \frac{1}{|S_i|} \sum_{u \in S_i} \sum_{v \in \mathcal{V}} R_{uv} = \frac{1}{|S_i|} \sum_{u \in S_i} \mathrm{R_{tot}}(u) \\
=& \mathbb{E}_{u \sim S_i}\left[\mathrm{R_{tot}}(u)\right]
\end{aligned} \tag{23}$$

Finally, the cumulative group isolation across all groups in the graph fulfills the following equality with the total effective resistance of the graph.

$$\frac{1}{2} \sum_{i \in SA} |S_i|\, \mathrm{R_{tot}}(S_i) = \frac{1}{2} \sum_{i \in SA} |S_i| \frac{1}{|S_i|} \sum_{u \in S_i} \sum_{v \in \mathcal{V}} R_{uv} = \frac{1}{2} \sum_{v \in \mathcal{V}} \sum_{v \in \mathcal{V}} R_{uv} = \mathrm{R_{tot}} = n \operatorname{Tr}(\mathbf{L}\dagger)$$

Therefore, since $\mathrm{R_{tot}}$ is related to the information flow, $\mathrm{R_{tot}}(S_i)$ measures the ease of information flow through group S_i in the graph.

The optimal –yet extreme– scenario of maximum information flow in a graph is such where all nodes in the graph are connected. Hence, G will be a fully connected graph. In this scenario, the total effective resistance of the graph reaches its minimum $n-1$, where $n = |\mathcal{V}|$. The total effective resistance of all existing edges in the graph is $n-1$, and given that we have all connections, $\mathrm{R_{tot}}$ also sums up to $n-1$. Therefore, every pair of nodes will be separated by $R_{uv} = 2/n$. In this scenario, $\mathrm{R_{tot}}(u) = (2/n)(n-1) = 2 - 2/n$ for all nodes in the graph. It leads to an group isolation of $\mathrm{R_{tot}}(S_i) = 2 - 2/n$ for all the different groups in the graph and therefore a Isolation disparity $\Delta\, \mathrm{R_{tot}} = 0$.

Regarding group isolation disparity, Eq. (5) can be redefined using Eq. (23). It is equivalent to the equality for all groups of the mean R_{uv} between of the nodes in the group and all nodes in the graph:

$$\begin{aligned}
\mathrm{R_{tot}}(S_i) = \mathrm{R_{tot}}(S_j), \forall i, j \in SA \\
\mathbb{E}_{u \sim S_i, v \sim \mathcal{V}}\left[R_{uv}\right] = \mathbb{E}_{u \sim S_j, v \sim \mathcal{V}}\left[R_{uv}\right], \forall i, j \in SA \times SA
\end{aligned} \tag{24}$$

C.1.2 Group Control and Control Disparity. In this section, we provide the proof for the bounds associated with the proposed control metrics.

Bounds of $\mathrm{B_R}(u)$ *and* $\mathrm{B_R}(S_i)$. In this section, we show the proof of the bounds of the node and group control.

THEOREM C.1. *The control of a node is bounded by* $1 \leq \mathrm{B_R}(u) \leq d_u$, *being* d_u *the degree of node* u. *Equality on the upper bound holds when all the edges are cut edges,* i.e., *edges that if removed the graph would become disconnected.*

Proof. The resistance curvature of a node is known to be bounded by $1 - d_u/2 \leq p_u \leq 1/2$ [26], and the relation between the curvature and group control is given by the equality $p_u = 1 - \frac{1}{2}\,\mathrm{B_R}(u)$. Therefore, we obtain the bounds as

$$\begin{aligned} 1 - \frac{d_u}{2} \leq p_u \leq \frac{1}{2} \rightarrow \\ 1 - \frac{d_u}{2} \leq 1 - \frac{1}{2}\,\mathrm{B_R}(u) \leq -\frac{1}{2} \rightarrow \\ -d_u \leq -\,\mathrm{B_R}(u) \leq -1 \rightarrow \\ 1 \leq \mathrm{B_R}(u) \leq d_u \end{aligned}$$

□

Theorem C.2. *The control of a group is bounded by* $1 \leq \mathrm{B_R}(S_i) \leq \frac{vol(S_i)}{|S_i|}$*, being* $vol(S_i)$ *the sum of the degrees of node* u*. Thus,* $vol(S_i)/|S_i|$ *is the average degree of all the nodes in* S_i*. Equality on the upper bound holds when the subgraph with all nodes of* S_i *and their neighbors is a tree graph, i.e., a connected acyclic undirected graph.*

Proof. The control of a group is defined as $\mathrm{B_R}(S_i) = \mathbb{E}_{u \sim S_i}[\mathrm{B_R}(u)]$ and using Theorem C.1, we derive the bounds of $\mathrm{B_R}(S_i)$ as follows:

$$\begin{aligned} &1 \leq \mathrm{B_R}(u) \leq d_u \rightarrow \\ &1 \leq \mathbb{E}_{u \sim S_i}[\mathrm{B_R}(u)] \leq \mathbb{E}_{u \sim S_i}[d_u] \rightarrow \\ &1 \leq \mathrm{B_R}(S_i) \leq \frac{\mathrm{vol}(S_i)}{|S_i|} \end{aligned}$$

□

Control as an allocation problem. Here we delve into the properties and behavior of the group control ($\mathrm{B_R}(S_i)$) as a limited resource to be distributed in the network. The sum of R_{uv} for every edge always equals to $|\mathcal{V}| - 1$ [29]:

$$\sum_{(u,v) \in \mathcal{E}} R_{uv} = |V| - 1.$$

Therefore, the sum of the node controls in the graph is defined as

$$\sum_{u \in \mathcal{V}} B_R(u) = \sum_{u \in \mathcal{V}} \sum_{v \in \mathcal{N}(u)} R_{uv} = 2 \times \sum_{(u,v) \in \mathcal{E}} R_{uv} = 2|V| - 2,$$

and the expectation as:

$$\mathbb{E}_{u \sim V}[B_R(u)] = 2 - \frac{2}{|V|}$$

independently of the number of edges (density) of the graph.

As a consequence of the definition of group control (Eq. (4)), the weighted sum of group control for all groups in the graph and the weighted mean also remain constant at:

$$\sum_{S_i \in V} |S_i| \times B_R(S_i) = 2|V| - 2 \quad \text{and} \quad \frac{1}{|V|} \sum_{S_i \in V} |S_i| \times B_R(S_i) = 2 - \frac{2}{|V|}.$$

C.2 Efficient version of Algorithm

Algorithm 2 shows an efficient manner to update the pseudo-inverse of the Laplacian after adding one edge to the graph. Therefore, we avoid the computation of $\mathbf{L}^\dagger$ after every edge addition. $\mathbf{L}^\dagger$ is easily updated using the Woodbury's formula which is based on the values of $\mathbf{L}^\dagger$ (see Black et al. [6] for a proof).

D ADDITIONAL EXPERIMENTS

In this section, we complement the experimental results reported in Section 4, and we report results on additional experiments.

D.1 Group Social Capital Metrics

The experiments presented in the main paper illustrate how edge augmentation via `ERG-Link` is able to (1) significantly mitigate the structural group unfairness (Table 2); and (2) increase the social capital for all the groups in the graph (Fig. 2). Table 4 depicts the group social capital for each group after each intervention. Note that the optimal value of $\mathrm{B_R}$ is to get every group's control equal to $2 - 2/|\mathcal{V}|$.

Algorithm 2: ERP-Link

Data: Graph $G = (\mathcal{V}, \mathcal{E})$, a protected attribute SA, budget B of total number of edges to add
Result: Interviened Graph $G' = (\mathcal{V}', \mathcal{E}')$

$\mathbf{L} = \mathbf{D} - \mathbf{A}$;
$S_d = \text{argmax}_{S_i \forall i \in SA} \mathrm{R}_{\text{tot}}(S_i)$; `// Identify the most disadvantaged group`
$\mathbf{L}^\dagger = \sum_{i>0} \frac{1}{\lambda_i} \phi_i \phi_i^\top = \left(\mathbf{L} + \frac{\mathbf{1}\mathbf{1}^\top}{n}\right)^{-1} - \frac{\mathbf{1}\mathbf{1}^\top}{n}$; `// Pre-computation of` $\mathbf{L}^\dagger$
Repeat
  $\mathbf{R} = \mathbf{1}\,\text{diag}(\mathbf{L}^\dagger)^\top + \text{diag}(\mathbf{L}^\dagger)\mathbf{1}^\top - 2\mathbf{L}^\dagger$; `// Compute effective resistance`
  $C = \{(u, v) \mid u \in S_d \text{ or } v \in S_d, (u, v) \notin E\}$; `// Select edge candidates`
  $\mathcal{E}' = \mathcal{E}' \cup \arg\max_{(u,v) \in C} R_{uv}$; `// Add edge with maximum effective resistance from` C
  `// Fast update of` $\mathbf{L}$ `and` $\mathbf{L}^\dagger$
  $\mathbf{L} = \mathbf{L} + (\mathbf{e}_u - \mathbf{e}_v)(\mathbf{e}_u - \mathbf{e}_v)^\top$;
  $\mathbf{L}^\dagger = \mathbf{L}^\dagger - \frac{1}{1+R_{uv}} \times (\mathbf{L}^\dagger_{u,:} - \mathbf{L}^\dagger_{v,:}) \otimes (\mathbf{L}^\dagger_{u,:} - \mathbf{L}^\dagger_{v,:})$; `// updated by Woodbury`
Until $|\mathcal{E}' \setminus \mathcal{E}| = B$;
return G';

Table 4: Group Social Capital after Edge Augmentation. The best values are highlighted in bold. Note how edge augmentation via `ERG-Link` is able to not only increase the social capital of the disadvantaged group (females) but also of the rest of the groups (males).

(a) Facebook (50) Female

	$R_{tot} \downarrow$	$\mathcal{R}_{diam} \downarrow$	$\mathrm{B_R}$
G	221.4	2.29	1.927
Random	211.5	2.26	1.927
DW	199.4	1.87	1.929
Cos	190.4	1.64	1.918
ERG	**138.7**	**0.43**	1.933

(b) UNC28 (5,000) Female

	$R_{tot} \downarrow$	$\mathcal{R}_{diam} \downarrow$	$\mathrm{B_R}$
G	608.6	2.11	1.994
Random	435.0	1.23	1.992
DW	583.2	1.81	1.997
Cos	429.6	0.42	1.940
ERG	**316.8**	**0.10**	2.001

(c) Google+ (5,000) Female

	$R_{tot} \downarrow$	$\mathcal{R}_{diam} \downarrow$	$\mathrm{B_R}$
G	564.1	1.31	1.807
Random	305.1	1.07	1.824
DW	558.7	1.31	1.806
Cos	230.9	0.29	1.827
ERG	**145.5**	**0.07**	1.889

(d) Facebook (50) Male

	$R_{tot} \downarrow$	$\mathcal{R}_{diam} \downarrow$	$\mathrm{B_R}$
G	179.8	2.25	2.034
Random	172.8	2.22	2.035
DW	163.1	1.83	2.033
Cos	161.8	1.61	2.039
ERG	**128.5**	**0.42**	2.031

(e) UNC28 (5,000) Male

	$R_{tot} \downarrow$	$\mathcal{R}_{diam} \downarrow$	$\mathrm{B_R}$
G	586.3	2.11	2.004
Random	415.2	1.23	2.005
DW	561.0	1.81	2.001
Cos	410.6	0.41	2.043
ERG	**308.0**	**0.09**	1.998

(f) Google+ (5,000) Male

	$R_{tot} \downarrow$	$\mathcal{R}_{diam} \downarrow$	$\mathrm{B_R}$
G	287.7	1.24	2.321
Random	175.7	1.03	2.293
DW	284.6	1.24	2.323
Cos	144.1	0.27	2.288
ERG	**108.5**	**0.06**	2.185

D.2 Distribution of Social Capital by Group

For completeness, Fig. 4 illustrates the distributions of all effective resistances R_{uv} and the node's social capital metrics $-\mathrm{R}_{\text{tot}}(u)$, $\mathcal{R}_{\text{diam}}(u)$ and $\mathrm{B_R}(u)-$ in the graph before and after the edge augmentation interventions for each of the groups.

The plots correspond to two different edge augmentation experiments on the Facebook dataset, with budgets of B=50 and B=5,000 new edges. Edge augmentation via `ERG-Link` is able to drastically reduce all effective resistances of the graph, unlike the other methods and even for the small budget. Note how there is still a long tail in the distribution of effective resistances after edge augmentation with the baseline methods, which illustrates that there are still nodes that struggle to exchange information even after the intervention.

Regarding the node social capital metrics, edge augmentation via `ERG-Link` reduces all $\mathrm{R}_{\text{tot}}(u)$ and $\mathcal{R}_{\text{diam}}(u)$, which explains why $\mathrm{R}_{\text{tot}}(S_i)$ and $\mathrm{R}_{\text{tot}}(S_i)$ is improved for all groups in the graph, and particularly for the disadvantaged group. Thus, $\Delta\, \mathrm{R}_{\text{tot}}$ and $\Delta\, \mathcal{R}_{\text{diam}}$ are significantly reduced.

D.3 Evolution of Group Social Capital During the Interventions

In this section, we provide additional results regarding the evolution of group social capital metrics during the graph's intervention, as initially shown in Fig. 3. The goal is to analyze the effectiveness of the edge augmentation methods on both small ($B = 50$ edges) and large

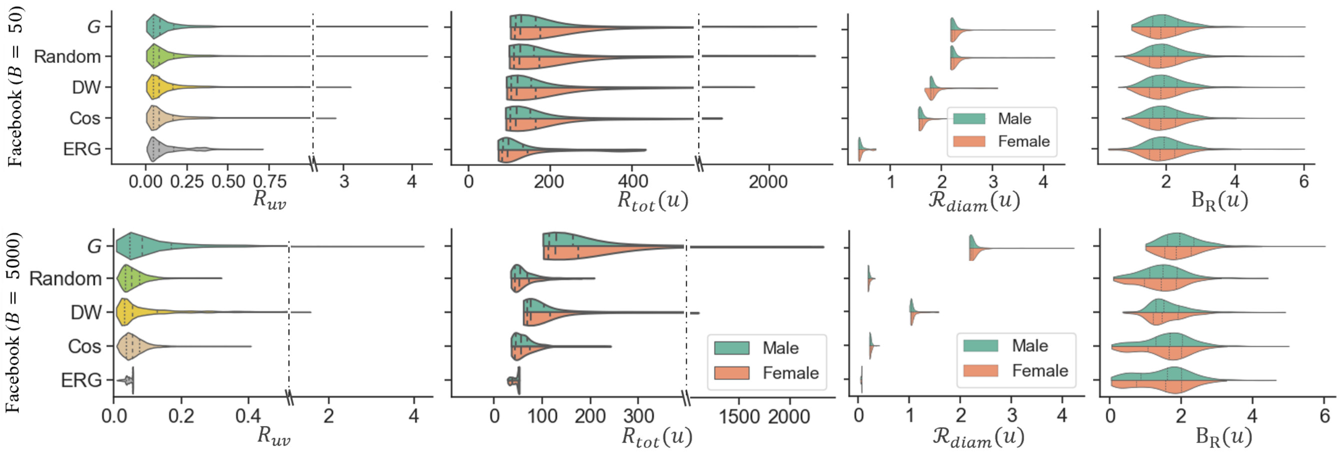

Figure 4: Distribution of R_{uv} and proposed social capital metrics on the Facebook dataset and after different graph interventions, 50 links in the top row and 5,000 in the bottom row. Columns show (left) distribution of all R_uv distances, (center-left) distribution of all $\mathrm{R_{tot}}(u)$, (center-right) distribution of all $\mathcal{R}_{\mathbf{diam}}(u)$, (right) distribution of all $\mathrm{B_R}(u)$. The distributions for the node metrics are shown for the two groups according to the protected attribute (gender).

($B = 5,000$ edges) budget scenarios. Figs. 5 and 6 show the evolution of both the group social capital metrics for each group and the structural group unfairness metrics on the Facebook dataset.

We observe in Fig. 5 how edge augmentation via ERG-Link is able to significantly improve the group social capital for all groups and reduce all structural unfairness metrics. In contrast, the baselines fail to do so.

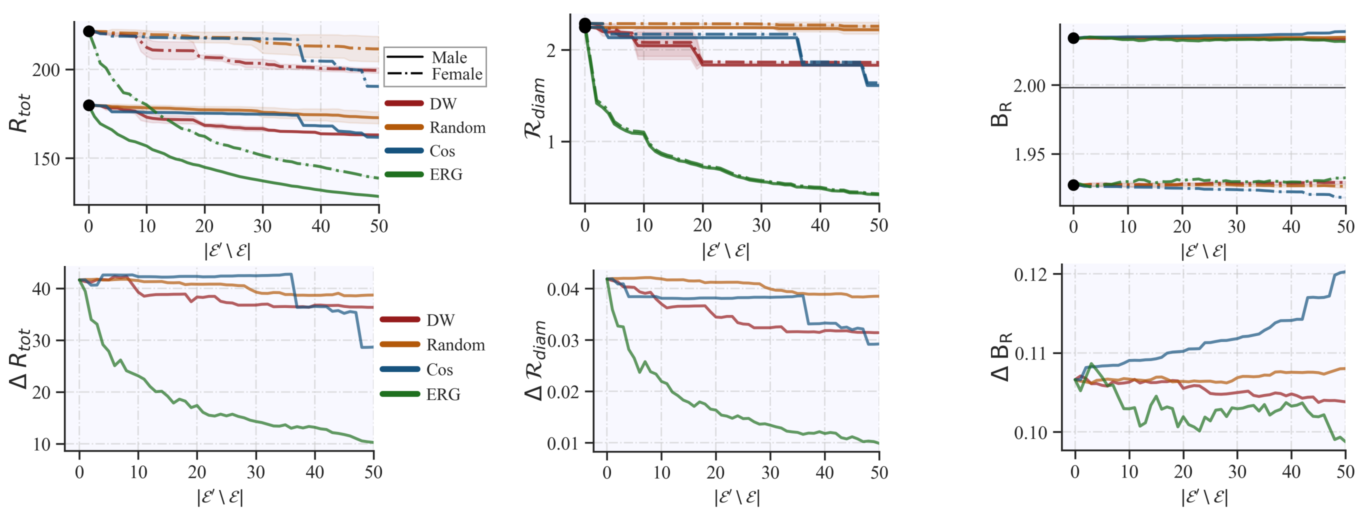

Figure 5: Evolution of group social capital and disparity metrics for both groups as the number of added edges increases on the Facebook dataset with a budget of 50 new edges.

Fig. 6 depicts the evolution of group social capital and disparity metrics when adding 5,000 edges to the Facebook dataset. The more edges we add to a graph, the denser the graph becomes and therefore the better the information flow. Thus, one can expect that after a large number of added edges, all methods behave similarly.

However, we observe some differences. First, the convergence to minimal isolation and diameter disparities is significantly faster when adding edges via ERG-Link than any other method. Second, the decrease in $\mathrm{R_{tot}}(S_i)$ is very significant for both groups (males and females) even after just adding a small number of edges by means of ERG-Link. Third, regarding group control and control disparity, edge augmentation via ERG-Link systematically reduces $\Delta\,\mathrm{B_R}$ while converging each group control to the optimal $\mathrm{B_R}(S_i) = 2 - 2/|\mathcal{V}|$.

We also show in Fig. 7 the evolution of $\mathrm{B_R}(S_i)$ and $\Delta\,\mathrm{B_R}$ on the UNC28 dataset with 1,000 edge additions along with the distribution of node's control $\mathrm{B_R}(u)$ after the intervention to showcase an scenario where ERG-Link is able to reach the optimal control disparity in the graph, allocating the same amount of control for each group, $\mathrm{B_R}(S_i) \approx 2 - 2/|\mathcal{V}|\ \forall\, i \in SA \rightarrow \Delta\,\mathrm{B_R} \approx 0$.

Last but not least, the different baselines do not reach a better behavior than the random method, neither on the group social capital metrics nor in terms of structural unfairness. After 5,000 edge additions, the random method significantly reduces the unfairness metrics $\Delta\,\mathrm{R_{tot}}$ and $\Delta\,\mathcal{R}_{\mathrm{diam}}$ while also achieving decent rated of $\mathrm{R_{tot}}(S_i)$ and $\mathcal{R}_{\mathrm{diam}}(S_i)$ since the budget is high enough to improve the information flow with no strategy. However, all baselines struggle to optimize $\Delta\,\mathrm{B_R}(S_i)$ and $\mathrm{B_R}(S_i)$. None of them is able to improve, and the cosine similarity approach even increased the Control Disparity.

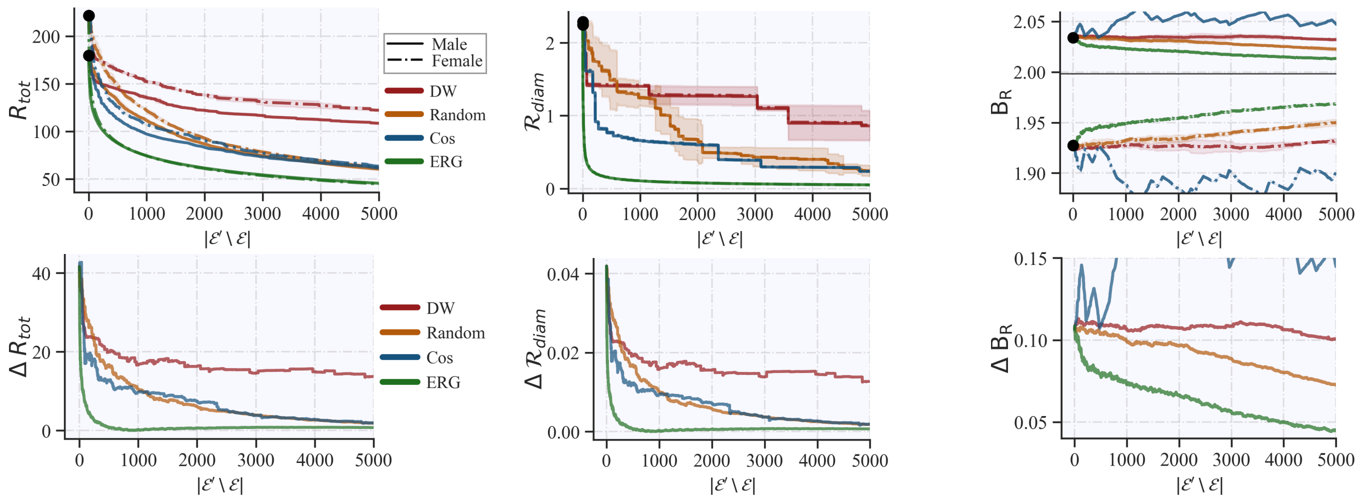

Figure 6: Evolution of group social capital metrics for both groups and fairness metrics as the number of added links increases, on Facebook dataset after adding 5,000 edges. Edge augmentation via ERG-Link exhibits a faster rate of convergence to the optimal scenario than the baselines.

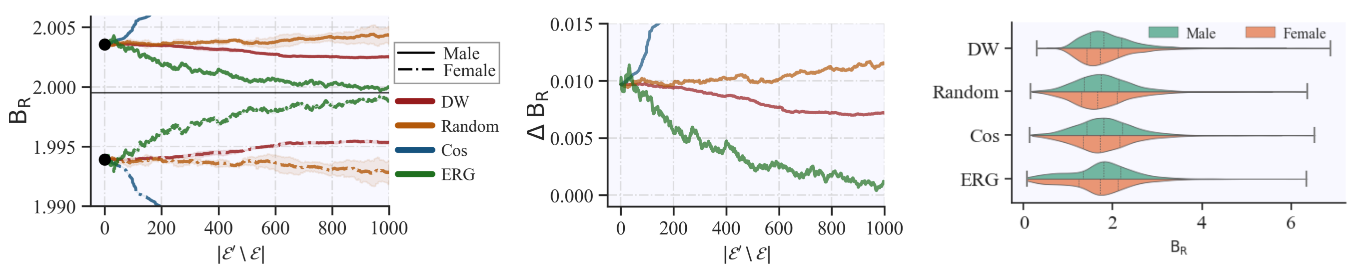

Figure 7: Behavior of node, group control and control disparity ($\mathrm{B_R}(u), \mathrm{B_R}(S_i), \Delta\,\mathrm{B_R}$). Illustration of the evolution of the group control (left, $\mathrm{B_R}(S_i)$) and control disparity (middle, $\Delta\,\mathrm{B_R}$) through an edge augmentation with $B = 1,000$ edges on the UNC dataset. Right-most figure: distribution of the nodes' control ($\mathrm{B_R}(u)$) for each group. Note how edge augmentation via ERG-Link yields a control for all groups approaching their optimal value of $2 - 2/|\mathcal{V}|$ by reducing the control of the privileged group (males) while increasing the control of the disadvantaged group (females).