Targeting Network Interventions with Social Norm

Dawen Meng *

June 2, 2023

Abstract

In this paper, I analyze a network game in the local-average setup where a player's payoff depends on his social distance to the social norm he confronts. It focuses on a social planner's optimal targeting interventions that change individual characteristics and/or network structure. I explore these intervention problems in various contexts depending on the objectives and constraints of the planner. First, I discuss the characteristic interventions with utilitarian and/or egalitarian planners. Then, I study the relative intervention problem in which a planner maximizes the ratio of the equilibrium social welfare over its first-best value subject to a quadratic budget constraint. In what follows, I turn to the structural intervention problem, where a social planner can design a network to maximize the social welfare subject to the unanimous approval constraints of players. I give a negative result that there exists no socially optimal and unanimously approved network if preferences are heterogeneous across players. Finally, I give a numerical algorithm to search for a socially optimal network among the entire network class in a extremely conformist society.

Keywords: Social norms, local-average model, characteristic intervention, structural intervention, conformism, individualism

1 Introduction

Network effects are ubiquitous in numerous facets of daily lives, ranging from the bandwagon psychology of consumers to the pandemic of an infectious disease, from the shaping of public opinions via word-of-mouth communication to financial contagions in a network of cross-collateralization. The past two decades have witnessed an explosive growth of research on network games, whereby

*E-mail: devinmeng@hotmail.com, School of Economics, Shanghai University of Finance and Economics.
Financial support from the National Science Foundation of China (NSFC-71301094) is gratefully acknowledged.

individuals are embedded in a social network and each one's payoff is affected by the behaviors of his peers. It focuses on how network structure influences individual behaviors, and provides predictions on network-specific equilibria. While traditional literature takes both the network structure and individual attributes as given, a recent branch of the network literature diverts its attention to a new problem of network intervention. It assumes that a social planner with certain objective can conduct ex-ante interventions by changing the individual characteristics from their status-quo levels (*characteristic intervention*), and/or restructure the present network (*structural intervention*), subject to certain constraints. The main goal of these problems is to understand how the planner can optimally intervene in view of primitives of the environment.

Examples of network interventions abound in both business practices and public policies. Sellers often foster the consumers' preferences by a "seeding" strategy. That is, they first sent advertisements and/or free samples to a few fashion leaders, then tipping a contagion of purchases as neighbors of these seeded individuals mimic them. In many instances, local governments adopt various network restructuring policies to increase the value of a community. For the purpose of improving economic development, attracting investment, and lowering crime rates, some policy makers advocate a gentrification process, which changes the composition of a neighborhood through the influx of more affluent and well-educated residents. The moving to opportunity (MTO) experiment sponsored by the U.S. government in 1990s relocated families from high- to low-poverty neighborhoods. It proved to be an effective tool in breaking delinquent networks and thus reducing local crime.

The existing literature on network interventions focuses mainly on the local-aggregate model, in which an individual's utility is affected by the sum of activities taken by his neighbors. Despite its clear intuitions in theoretical analysis, the local-aggregate model fails to be empirically applicable. The standard model adopted by empirical studies usually takes the linear-in-means form

$$x_{ig} = \beta z_{ig} + \gamma y_g + \theta \bar{x}_{ig} + \epsilon_{ig}, \tag{1}$$

where x_{ig} is the action taken by individual i belonging to group g, z_{ig} denotes the individual-specific characteristics, y_g is the common element within group g. N_g is the number of individuals within group g, $\bar{x}_{ig} \equiv \frac{1}{N_g-1}\sum_{j\neq i} x_{jg}$ is the social norm confronted by individual i (average activity of agents other than i), ϵ_{ig} denotes a noisy term, (β, γ, θ) are parameters to be estimated. Blume et al. (2015) and Boucher and Fortin (2016) interpret (1) as player i's best-response function in a complete information game. Its corresponding payoff function, however, is not of the local-aggregate form well-studied by theoretical literature, but rather of the local-average form. That is, individuals have a preference to conform to the average action of their neighbors in a social network. This paper aims to reconcile the theoretical research on network interventions and econometric studies on model identifications. It contributes to the existing literature in the following aspects.

First, I embed network interventions into a local-average model featuring individual taste for conformity. Conformism is a doctrine of mean advocating that the easiest and hence best life is attained by blending in with one's surroundings and refraining from any eccentric or out-of-the-ordinary behavior. It may be best expressed in the old saying, "*When in Rome, do as the Romans do.*" The network intervention problems with local-average model is in sharp contrast to those in the traditional local-aggregate scenario. The seminal paper by Galeotti et al. (2020) shows that as the budget increases, the optimal intervention is progressively similar to the first or the lowest principal component depending on whether the game features strategic substitutes or complements. These principal components reflect either global or local network structures. In the local-average model, however, I find that for very large budget, the optimal intervention is approximately equal across individuals regardless of the network structure. The intuition is simple. Since everyone dislikes her social distance away from her reference group, a utilitarian planner will allocate progressively equal portion of resources across individuals to eliminate their ex-ante disparities if budget permits.

Second, I study richer forms of network interventions. In previous researches, the planner are traditionally assumed to be a utilitarian welfare maximizer. In this paper, the planner possesses more diverse objectives. She aims at increasing the social welfare or decreasing the social variance depending on whether she is utilitarian or egalitarian. While previous researches pay their attention to the absolute value of social welfare, this paper also explores the relative intervention problem, whereby a planner aims to maximize the ratio of social welfare in equilibrium to its first-best counterpart. With local-aggregate payoffs, the first-best social welfare is usually not achievable in Nash equilibrium. So the welfare rate is strictly smaller than 100%, no matter how the planner may choose to intervene in it. Things are different in our local-average model. The planner can eliminate all individual heterogeneities by assigning all players the same preference. Then, players choose the same action in equilibrium, which leads to a zero social distance for everyone. This amount to a non-game theoretic problem free of any strategic interaction. The social welfare in equilibrium hence coincides with its first-best value.

Finally, this paper incorporates structural intervention into the traditional characteristic intervention framework, and considers the individuals' attitude towards a restructured network.

While some papers study the characteristic and structural interventions separately, I incorporate them in a unified framework. In this paper, the planner can adjust the individual preferences and design an optimal network simultaneously. An exception of previous studies is Kor and Zhou (2022), who also consider a joint intervention problem. These authors assume that the planner pays a quadratic cost in form of Frobenius norm when restructuring the original network. My paper, however, assumes that there is no physical cost of structural intervention. This assumption comes from the consideration that social networks, with the advent of the information era, are often made of logical rather than physical links. Advanced cyber technology facilitates network building, and thus allows the planner to save concrete infrastructure cost. But the individuals' attitudes towards a restructured network are nonnegligible. Anyone retains the right of voting for his/her most preferred network, at least of voting against a harmful network. The planner must abide by restrictions other than budget constraints when conducting structural intervention. This paper considers the optimal network design problem under these constraints. Once again, a significant difference in structural intervention separates our local-average model from the standard local-aggregate model. If there is no restructuring cost and the network game features strategic complements, a denser network is beneficial to everyone. So the planner need not consider the boycotting of voters when densifying a network. Obviously, the complete (densest) network is both socially optimal and unanimously approved. In our local-average scenario, however, any network structure alteration will inevitably affect the social distances faced by all players. Their attitudes towards a new network are therefore more diverse and complex, which deserves an in-depth discussion.

Related literature. This paper is related to three bodies of work. One deals with the network-based targeting problem, which investigates how the player provides personalized policies to players depending on their network positions. These personalized operations take different forms in specific contexts. Candogan et al. (2012), Bloch and Quérou (2013), Fainmesser and Galeotti (2016), Chen et al. (2018, 2022), and Meng et al. (2022), among many others, discuss the network-based price discrimination, in which players are charged discriminatory prices depending on their centrality in the network. In Galeotti and Goyal (2009), the targeting operation appears as advertising or free sampling strategy. Bimpikis et al. (2016) considers competitive targeting through advertising and the spreading of information, Liu (2019) studies targeting in terms of industrial policies in production networks, while King et al. (2019) make use of sectoral targeting to obtain efficient carbon tax reforms. More recently, several papers examine characteristic intervention as a new form of targeting operation, which are closely related to our paper. Demange (2017) considers a targeting characteristic intervention problem based directly on a given reaction function. He shows that the optimal intervention policies rely crucially on whether this reaction function is linear, concave or convex. Galeotti et al. (2020) studies a similar problem based on a local-aggregate payoff function. They decomposes the optimal intervention into orthogonal principal components of a network and shows that it is approximately simple for large budget. Meng et al.(2022) embeds the characteristic intervention problem into a mechanism design framework. They discuss how a profit-maximizing seller can intervene to change the distribution of individual types subject to a fixed budget constraint. Kor and Zhou (2023) extend the single-activity characteristic intervention model to the case of multiple activities.

My paper is also related to a body of literature on structural network intervention (the optimal network design), which refers to a planner's purposeful operations on a network to affect the behaviors of players and thus achieve certain desirable outcomes. Valente (2012) presents four types of structural operations: (*i*) individual intervention, which identifies key nodes of a network on the basis of certain centrality measures; (*ii*) segmentation intervention, which identifies key groups of players; (*iii*) induction interventions, in which an initial set of players are selected as "seeds", then cascades in information/behavioral diffusion are created; and (*iv*) alteration interventions that restructures a network by adding/deleting nodes and/or edges. Ballester et al. (2006), Ballester et al. (2010), Golub and Lever (2010), Zhou and Chen (2015), and Meng et al. (2022) focus on identifying the key players, key groups and/or key links. Hiller (2017), König et al. (2014) and Belhaj et al. (2016) address the efficient-network-searching problems of a planner in games featuring local complementarities and costly links. They show that for general cost functions, efficient networks belong to the class of nested split graphs. Li (2021) extends these results from simple to weighted and directed networks by showing that the optimal networks belong to the class of generalized nested split graphs. Sciabolazza et al. (2020) present a network structural intervention to foster scientific collaboration and productivity at a research university, and provide an experimental evaluation on the impact of it. Sun et al. (2021), Kor and Zhou (2022) provide a general local-aggregate framework to evaluate and compare the equilibrium effects of joint intervention that combines both structural and characteristic interventions.

The last body of literature related to the present paper is local-average (conformist) model in network games, which includes both theoretical and econometric studies. Patacchini and Zenou (2012) discuss how conformism and deterrence affect juvenile criminal activities, then test the model using a data set of adolescent friendship net. Liu et al. (2014) characterize Nash equilibrium of a model with peer effects, and give identification conditions for a general econometric network model that incorporates both local-aggregate and local-average endogenous peer effects. Blume et al. (2015) discuss the identification conditions in linear social interactions models with incomplete information. Boucher (2016) embed the local-average model into a network formation framework by exploring a model with both peer effects and self-selection. He also estimates the model using student-level data on participation in high school extracurricular activities. By a dynamic local-average model where every player wants to be assimilated to the majority norm, Olcina et al. (2017) show that the individual norms always converges to a steady state and the speed of convergence depends on both the network structure and the individual tastes for conformity. More recently, Golub and Morris (2020) characterize the convention equilibrium of a coordination game when players possess heterogeneous higher-order beliefs about the expectations of others. Ushchev and Zenou (2020) analyze the comparative statics properties of the local-average model as well as its welfare and policy implications.

The rest of the paper is organized as follows. Section 2 presents the baseline model. Section 3 explores the characteristic interventions. Section 4 discusses a unified framework incorporating both characteristic and structural interventions. Section 5 concludes.

2 The baseline model

Consider $n \geqslant 2$ agents who are embedded in a social network $\mathbf{g} \equiv \langle I, \mathcal{E} \rangle$, where $I \equiv \{1, \cdots, n\}$ denotes the set of vertexes, $\mathcal{E}$ is the set of edges. An irreducible adjacency matrix $\mathbf{G} = [g_{ij}]$ keeps track of the direct connections. By convention, we assume $g_{ij} = g_{ji} = 1$ if agents i and j are adjacent; and $g_{ij} = 0$ otherwise. Self-loops are not allowed, i.e., $g_{ii} = 0, \forall i \in I$. Agent i's utility function takes a standard linear-quadratic form:

$$u_i(\alpha_i, \theta, \mathbf{x}, \mathbf{g}) = \alpha_i x_i - \frac{1}{2} x_i^2 - \frac{\theta}{2} (x_i - \overline{x}_i)^2, \tag{2}$$

where $\alpha_i \in \mathbb{R}$ stands for agent i's preference, $x_i \in \mathbb{R}$ denotes the quality of goods he consumes, let $\boldsymbol{\alpha} \equiv (\alpha_i)_{i \in I}$ and $\mathbf{x} \equiv (x_i)_{i \in I}$.[1] $\overline{x}_i \equiv \sum_{j=1}^n \widehat{g}_{ij} x_j$ is the *social norm* confronted by agent i, namely, the average consumption within her neighborhood, where $\widehat{g}_{ij} \equiv g_{ij}/d_i$, $d_i \equiv \sum_{j=1}^n g_{ij}$

[1] Here, we extend the well-known model of Mussa and Rosen (1978) by allowing α_i and x_i to be either positive or negative. It means that certain quality is not desired by everybody. Quality sold at the same price are ranked from top to bottom by consumers with positive α_i and in the reverse order by those with negative α_i. The sign of x_i stands for the direction in which certain quality is intensified. Say, nutrition versus taste of food, stickiness versus informativeness of social networking service.

denotes the degree (number of neighbours) of a node i. $\theta > 0$ stands for an agent's taste for conformity. The first two terms $\alpha_i x_i - x_i^2/2$ represent the stand-alone utility obtained from one's own consumption when there is no interaction with others. The third term $-\theta(x_i - \overline{x}_i)^2/2$ captures the peer-group pressure faced by agent i, who suffers a loss from failing to conform to the social norm around him. We represent by

$$W(\theta, \boldsymbol{\alpha}, \mathbf{x}, \mathbf{g}) \equiv \sum_{i=1}^{n} u_i(\theta, \alpha_i, \mathbf{x}, \mathbf{g}) = \boldsymbol{\alpha}^\top \mathbf{x} - \frac{1}{2}\mathbf{x}^\top \mathbf{H}\boldsymbol{\alpha}$$

the aggregate social welfare, and by

$$V(\mathbf{x}, \mathbf{g}) \equiv \sum_{i=1}^{n} (x_i - \overline{x})^2 = \mathbf{x}^\top \mathbf{Q}\mathbf{x}$$

the social variance measuring the dispersion of individual consumptions, where $\mathbf{H} \equiv \mathbf{I} + \theta(\mathbf{I} - \widehat{\mathbf{G}})^\top(\mathbf{I} - \widehat{\mathbf{G}})$, $\widehat{\mathbf{G}} \equiv [\widehat{g}_{ij}]$ is the row-normalized adjacency matrix, $\overline{x} \equiv \frac{1}{n}\sum_{i=1}^{n} x_i$ denotes the global average of consumptions, $\mathbf{Q} \equiv \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^\top$ denotes the projection matrix onto the subspace orthogonal to the all-ones vector $\mathbf{1} \equiv (1, \cdots, 1)^\top$. Consumers in a society with larger $V(\mathbf{x}, \mathbf{g})$ are more diverse in their quality choices.

As a benchmark, we first consider the socially optimal outcomes. The social welfare-maximizing consumption is

$$\mathbf{x}^{FB}(\theta, \boldsymbol{\alpha}, \mathbf{g}) = \mathbf{H}^{-1}\boldsymbol{\alpha} = \arg\max_{\mathbf{x} \in \mathbb{R}^n} W(\theta, \boldsymbol{\alpha}, \mathbf{x}, \mathbf{g}).$$

Representing the first-best social welfare as functions of the network $\mathbf{g}$ and primitives $(\theta, \boldsymbol{\alpha})$, we get:

$$\mathbb{W}^{FB}(\theta, \boldsymbol{\alpha}, \mathbf{g}) \equiv W(\boldsymbol{\alpha}, \theta, \mathbf{x}^{FB}, \mathbf{g}) = \frac{1}{2}\boldsymbol{\alpha}^\top \mathbf{H}^{-1}\boldsymbol{\alpha}. \tag{3}$$

In a game-theoretic framework, each individual i chooses x_i to maximize his own utility taking the network structure $\mathbf{g}$ and the choices of other agents $\mathbf{x}_{-i} \equiv (x_j)_{j \neq i}$ as given. By computing the first-order condition with respect to x_i, we obtain the best-reply function for each i: $x_i = (1-\lambda)\alpha_i + \lambda\overline{x}_i$, where $\lambda \equiv \theta/(1+\theta) \in [0, 1)$. In matrix form, we get $\mathbf{x} = (1-\lambda)\boldsymbol{\alpha} + \lambda\widehat{\mathbf{G}}\mathbf{x}$, which guarantees a unique interior Nash equilibrium

$$\mathbf{x}^N(\lambda, \boldsymbol{\alpha}, \mathbf{g}) = \widehat{\mathbf{M}}\boldsymbol{\alpha}, \tag{4}$$

with $\widehat{\mathbf{M}} \equiv (1-\lambda)(\mathbf{I} - \lambda\widehat{\mathbf{G}})^{-1}$. The associated equilibrium payoff is

$$u_i^N(\lambda, \boldsymbol{\alpha}, \mathbf{G}) = \frac{1}{2}\left[\alpha_i^2 - \frac{1}{\lambda}\left(\alpha_i - \sum_{j=1}^{n} \widehat{m}_{ij}\alpha_j\right)^2\right], \tag{5}$$

where $\widehat{m}_{ij}$ is the $(i,j)-$entry of matrix $\widehat{\mathbf{M}}$. The social welfare and variance in Nash equilibrium are

$$\mathbb{W}^N(\lambda, \boldsymbol{\alpha}, \mathbf{g}) \equiv W(\boldsymbol{\alpha}, \theta, \mathbf{x}^N, \mathbf{g}) = \frac{1}{2}\boldsymbol{\alpha}^\top \mathbf{W} \boldsymbol{\alpha}$$

and

$$\mathbb{V}^N(\lambda, \boldsymbol{\alpha}, \mathbf{g}) \equiv V(\mathbf{x}^N, \mathbf{g}) = \boldsymbol{\alpha}^\top \mathbf{V} \boldsymbol{\alpha},$$

where $\mathbf{W} \equiv \mathbf{I} - \frac{1}{\lambda}(\mathbf{I} - \widehat{\mathbf{M}})^\top (\mathbf{I} - \widehat{\mathbf{M}}), \mathbf{V} \equiv \widehat{\mathbf{M}}^\top \mathbf{Q} \widehat{\mathbf{M}}$. We assume throughout this paper that matrix $\mathbf{W}$ is nonsingular.

3 Characteristic intervention

A planner aiming at optimizing her own objective at equilibrium can intervene to adjust the vector of preferences $\boldsymbol{\alpha}$, or to restructure the network $\mathbf{g}$ at an ex ante stage. Following the planner's choice of her interventions, all individuals choose their consumptions simultaneously, a Nash equilibrium is then fulfilled. In this section we discuss the optimal interventions on individual characters $\boldsymbol{\alpha}$ in various environments.

Problems of interest are to study the bounds of social welfare attained by the planner when she controls the social variance at a fixed level; or in duality, the possible bounds of social variance under a fixed level of social welfare. The specific forms of network intervention problem vary with the planner's underlying philosophy regarding the social welfare and social variance. Say, if she is both utilitarian and egalitarian, she avoids enlarging the social variance while maximizing the social welfare; she also tries to minimize the social variance while pegging the social welfare at a fixed level. The optimal intervention problems under unitary social variance and unitary social welfare are represented, respectively, as:

$$[\mathcal{P}^{w|v}_{max}/\mathcal{P}^{w|v}_{min}] : \max_{\boldsymbol{\alpha} \in \mathbb{R}^n} / \min \mathbb{W}^N(\lambda, \boldsymbol{\alpha}, \mathbf{g}), s.t. : \mathbb{V}^N(\lambda, \boldsymbol{\alpha}, \mathbf{g}) = 1, \tag{6}$$

and

$$[\mathcal{P}^{v|w}_{max}/\mathcal{P}^{v|w}_{min}] : \max_{\boldsymbol{\alpha} \in \mathbb{R}^n} / \min \mathbb{V}^N(\lambda, \boldsymbol{\alpha}, \mathbf{g}), s.t. : \mathbb{W}^N(\lambda, \boldsymbol{\alpha}, \mathbf{g}) = 1. \tag{7}$$

We denote by $(\boldsymbol{\alpha}^{j|k}_i, \Pi^{j|k}_i)$ the optimal solution and value pair to problem $[\mathcal{P}^{j|k}_i]$, for all $(i,j,k) \in \{min, max\} \times \{w, v\} \times \{w, v\}$.

Proposition 1 *Keeping the social variance at a unitary level, the planner can intervene to achieve:*

- *an infinite upper bound of social welfare* $\Pi_{max}^{w|v} = \infty$ *at* $\boldsymbol{\alpha}_{max}^{w|v}$ *satisfying*

$$\text{Proj}_{\boldsymbol{1}^{\perp}} \boldsymbol{\alpha}_{max}^{w|v} \in \mathscr{G}(\lambda_1(\boldsymbol{\Sigma}_{-n,-n}),\, \boldsymbol{W},\, \boldsymbol{V}),\, \left\| \text{Proj}_{\boldsymbol{1}} \boldsymbol{\alpha}_{max}^{w|v} \right\| = \infty,$$

where $\boldsymbol{1}^{\perp}$ *denotes the orthogonal complement space of* $\boldsymbol{1}$*,* $\mathscr{G}(\lambda, \boldsymbol{A}, \boldsymbol{B}) \equiv \{\boldsymbol{x} | \boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{B}\boldsymbol{x}\}$ *denotes the* $\lambda-$*generalized eigenspace of matrix* $\boldsymbol{A}$ *relative to* $\boldsymbol{B}$*,* $\lambda_i(\cdot)$ *denotes the ith largest eigenvalue of a matrix,* $\boldsymbol{\Sigma}_{-n,-n} \equiv (\widehat{\boldsymbol{V}}_{-n})^{-\frac{1}{2}} (\boldsymbol{U}_{-n})^{\top} \boldsymbol{W} \boldsymbol{U}_{-n} (\widehat{\boldsymbol{V}}_{-n})^{-\frac{1}{2}}$*,* $\nu_i \equiv \lambda_i(\boldsymbol{V})$*,* $\widehat{\boldsymbol{V}}_{-n} \equiv diag\{\nu_1, \cdots, \nu_{n-1}\}$*,* $\boldsymbol{U}_{-n} \equiv (\boldsymbol{u}_{\cdot i})_{i \neq n}$*,* $\boldsymbol{u}_{\cdot i}$ *denotes the orthonormal eigenvector of* $\boldsymbol{V}$ *associated with* ν_i*;*[2]

- *a lower bound* $\Pi_{min}^{w|v} = \lambda_{n-1}(\boldsymbol{\Sigma}_{-n,-n}) \in [\underline{\lambda}, \overline{\lambda}]$ *at* $\boldsymbol{\alpha}_{min}^{w|v} \in \mathscr{G}(\lambda_{n-1}(\boldsymbol{\Sigma}_{-n,-n}),\, \boldsymbol{W},\, \boldsymbol{V})$*,*[3] *where*

$$\underline{\lambda} = \begin{cases} \frac{\lambda_n(\boldsymbol{W})}{\lambda_1(\boldsymbol{V})} & \text{if } \lambda_n(\boldsymbol{W}) > 0 \\ \frac{\lambda_n(\boldsymbol{W})}{\lambda_{n-1}(\boldsymbol{V})} & \text{if } \lambda_n(\boldsymbol{W}) < 0 \end{cases}, \overline{\lambda} = \begin{cases} \frac{\lambda_2(\boldsymbol{W})}{\lambda_1(\boldsymbol{V})} & \text{if } \lambda_2(\boldsymbol{W}) > 0 \\ \frac{\lambda_2(\boldsymbol{W})}{\lambda_{n-1}(\boldsymbol{V})} & \text{if } \lambda_2(\boldsymbol{W}) < 0 \end{cases}.$$

Proof. See appendix A. ■

From any point $\boldsymbol{\alpha}$ satisfying $\boldsymbol{\alpha}^{\top}\mathbf{V}\boldsymbol{\alpha} = 1$, the objective function $\boldsymbol{\alpha}^{\top}\mathbf{W}\boldsymbol{\alpha}$ increases along the direction of $\mathbf{1}$, i.e., $(\boldsymbol{\alpha} + a\mathbf{1})^{\top}\mathbf{W}(\boldsymbol{\alpha} + a\mathbf{1}) = \boldsymbol{\alpha}^{\top}\mathbf{W}\boldsymbol{\alpha} + 2a\mathbf{1}^{\top}\boldsymbol{\alpha} + na^2 > \boldsymbol{\alpha}^{\top}\mathbf{W}\boldsymbol{\alpha}$ for an a with sufficiently large magnitude. While the constraint remains unchanged in the same direction since $(\boldsymbol{\alpha} + a\mathbf{1})^{\top}\mathbf{V}(\boldsymbol{\alpha} + a\mathbf{1}) = 1, \forall a \in \mathbb{R}$. It is clear that an unboundedly large value $\Pi_{max}^{w|v}$ is attained at $a = \infty$, while $\Pi_{min}^{w|v}$ is attained at $a = \bar{\alpha} \equiv -\frac{1}{n}\mathbf{1}^{\top}\boldsymbol{\alpha}$, and hence $\boldsymbol{\alpha} + a\mathbf{1} = \left(\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^{\top}\right)\boldsymbol{\alpha}$.

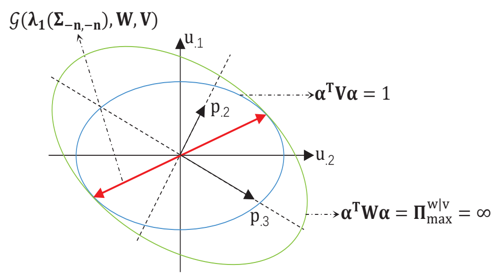
Figure 4 provides geometric illustrations of the above results for the case of $n = 3$. $\boldsymbol{\alpha}^{\top}\mathbf{V}\boldsymbol{\alpha} = 1$ represents an elliptic cylindrical shell whose axis lies in $\mathbf{1}$. $\boldsymbol{\alpha}^{\top}\mathbf{W}\boldsymbol{\alpha} = C$ represent either an ellipsoid or a hyperboloid, depending on the signs of constant C and eigenvalues $\lambda_i(\mathbf{W}), i = 2, 3$.[4] An infinitely large value $\Pi_{max}^{w|v} = \infty$ can be achieved at an extremely remote $\boldsymbol{\alpha}_{max}^{w|v}$ on the elliptic cylindrical shell. The projection of $\boldsymbol{\alpha}_{max}^{w|v}$ onto the orthogonal complement space of $\mathbf{1}$ lies in the steepest ascending direction, that is, the first generalized eigenspace of $\mathbf{W}$ to $\mathbf{V}$. Figures 1(a), 2(a) and 3(a) depict cross-sections of surfaces $\boldsymbol{\alpha}^{\top}\mathbf{V}\boldsymbol{\alpha} = 1$ and $\boldsymbol{\alpha}^{\top}\mathbf{W}\boldsymbol{\alpha} = \Pi_{max}^{w|v}$ for different eigenvalues $\lambda_i(\mathbf{W}), i = 2, 3$, when the cutting plane is perpendicular to $\mathbf{1}$, where $\mathbf{u}_{\cdot i}$ (resp. $\mathbf{p}_{\cdot i}$) denotes the ith principal components of matrix $\mathbf{V}$ (resp. $\mathbf{W}$).

[2] Notice that eigenvector $\mathbf{u}_{\cdot j}$ is not unique when the associated eigenvalue ν_i has a multiplicity exceeding one. Nevertheless, it follows from the uniqueness of the spectral decomposition in matrix theory that $\mathbf{U}_{-n}(\widehat{\mathbf{V}}_{-n})^{-1}(\mathbf{U}_{-n})^{\top}$ does not vary with the choice of $\mathbf{U}_{-n}$ and hence $\lambda_i(\mathbf{\Sigma}_{-n,-n}) = \lambda_i(\mathbf{U}_{-n}(\widehat{\mathbf{V}}_{-n})^{-1}(\mathbf{U}_{-n})^{\top}\mathbf{W}), \forall i \in \{1, \cdots, n-1\}$ is irrelevant to this choice.

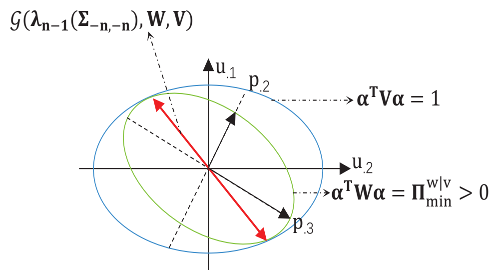
[3] A point deserving comment is that $\lambda_{n-1}(\mathbf{\Sigma}_{-n,-n})$ is not necessarily $\lambda_n(\mathbf{\Sigma})$, since the comparison between $\lambda_{n-1}(\mathbf{\Sigma}_{-n,-n})$ and 1 is ambiguous: if $\lambda_{n-1}(\mathbf{\Sigma}_{-n,-n}) < 1$, $\lambda_{n-1}(\mathbf{\Sigma}_{-n,-n}) = \lambda_n(\mathbf{\Sigma})$; if $\lambda_{n-1}(\mathbf{\Sigma}_{-n,-n}) > 1$, $\lambda_{n-1}(\mathbf{\Sigma}_{-n,-n}) > \lambda_n(\mathbf{\Sigma}) = 1$.

[4] Notice that $\lambda_1(\mathbf{W}) = 1$.

If $\lambda_2(\mathbf{W}) > \lambda_3(\mathbf{W}) > 0$, $\boldsymbol{\alpha}^\top \mathbf{W} \boldsymbol{\alpha} = C$ is a triaxial ellipsoid with its shortest axis lying in $\mathbf{1}$. The minimum value $\Pi_{min}^{w|v}$ is obtained when the ellipsoid is inscribed inside the elliptic cylinder. The minimizer $\boldsymbol{\alpha}_{min}^{w|v}$ lies within the orthogonal complement space of $\mathbf{1}$, and is in the direction of the slowest ascent, i.e., $\boldsymbol{\alpha}_{min}^{w|v} \in \mathscr{G}(\lambda_{n-1}(\boldsymbol{\Sigma}_{-n,-n}), \mathbf{W}, \mathbf{V})$ (Figure 4(a)). If $\lambda_2(\mathbf{W}) > 0 > \lambda_3(\mathbf{W})$, a minimum value $\Pi_{min}^{w|v} < 0$ is achieved when the two-sheeted hyperboloid $\boldsymbol{\alpha}^\top \mathbf{W} \boldsymbol{\alpha} = \Pi_{min}^{w|v}$ is tangent with the elliptic cylindrical shell $\boldsymbol{\alpha}^\top \mathbf{V} \boldsymbol{\alpha} = 1$ (Figure 4(b)). If $0 > \lambda_2(\mathbf{W}) > \lambda_3(\mathbf{W})$, a minimum value $\Pi_{min}^{w|v} < 0$ is achieved when the elliptic cylindrical shell $\boldsymbol{\alpha}^\top \mathbf{V} \boldsymbol{\alpha} = 1$ is circumscribed by a one-sheeted hyperboloid $\boldsymbol{\alpha}^\top \mathbf{W} \boldsymbol{\alpha} = \Pi_{min}^{w|v}$ (Figure 4(c)). In these cases, the minimizer $\boldsymbol{\alpha}_{min}^{w|v}$ lies within the orthogonal complement space of $\mathbf{1}$, and is in the direction of the steepest descent i.e., $\boldsymbol{\alpha}_{min}^{w|v} \in \mathscr{G}(\lambda_{n-1}(\boldsymbol{\Sigma}_{-n,-n}), \mathbf{W}, \mathbf{V})$. Figures 1(b), 2(b) and 3(b) depict the associated cross-sections cut by a plane passing through the origin and perpendicular to $\mathbf{1}$.

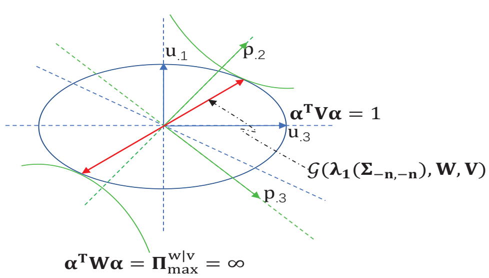


(a) $[\mathcal{P}_{max}^{w|v}]$

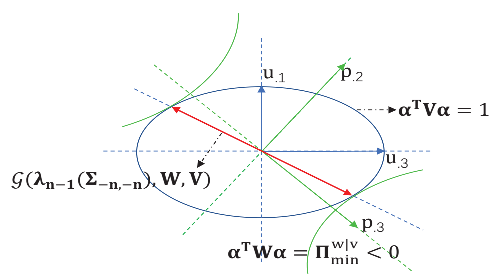


(b) $[\mathcal{P}_{min}^{w|v}]$

Figure 1. $\lambda_2(\mathbf{W}) > 0, \lambda_3(\mathbf{W}) > 0$



(a) $[\mathcal{P}_{max}^{w|v}]$

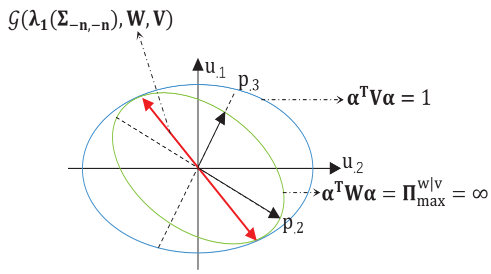


(b) $[\mathcal{P}_{min}^{w|v}]$

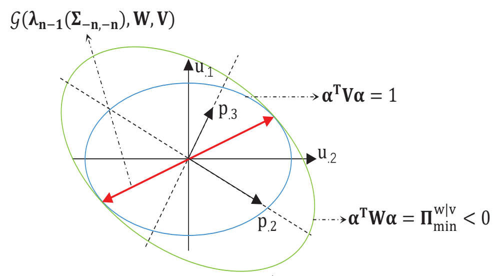
Figure 2. $\lambda_2(\mathbf{W}) > 0, \lambda_3(\mathbf{W}) < 0$

Proposition 2 *Keeping the social welfare at a unitary level, the planner can intervene to achieve the following bounds for social variance:*

- *if* $\lambda_n(\boldsymbol{W}) > 0$, *an optimal value* $\Pi_{max}^{v|w} = \lambda_1(\boldsymbol{W}^{-1}\boldsymbol{V})$ *is attained at a vector* $\boldsymbol{\alpha}_{max}^{v|w} \in \mathscr{G}(\lambda_1(\boldsymbol{W}^{-1}\boldsymbol{V}), \boldsymbol{V}, \boldsymbol{W})$; *if* $\lambda_n(\boldsymbol{W}) < 0$, $\Pi_{max}^{v|w} = \infty$;
- *the lower bound* $\Pi_{min}^{v|w} = 0$ *is attained at* $\boldsymbol{\alpha}_{min}^{v|w} \propto \boldsymbol{1}$.

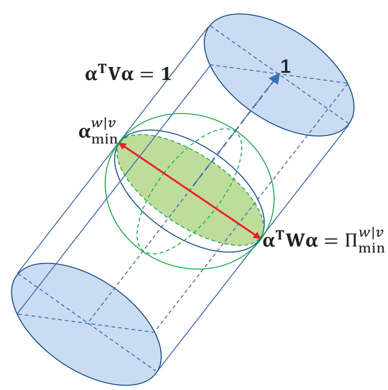


(a) $[\mathcal{P}_{max}^{w|v}]$

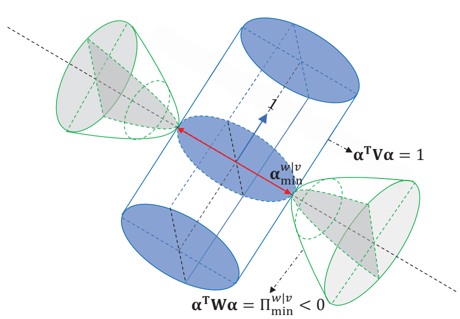


(b) $[\mathcal{P}_{min}^{w|v}]$

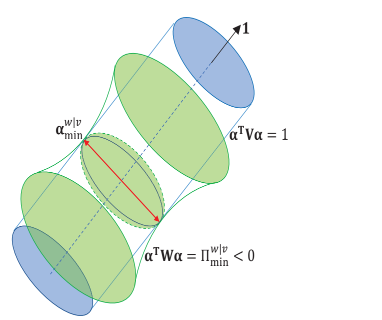
Figure 3. $\lambda_2(\mathbf{W}) < 0, \lambda_3(\mathbf{W}) < 0$



(a) $\lambda_2(\mathbf{W}) > 0, \lambda_3(\mathbf{W}) > 0, \Pi_{min}^{w|v} > 0$



(b) $\lambda_2(\mathbf{W}) > 0, \lambda_3(\mathbf{W}) < 0, \Pi_{min}^{w|v} < 0$

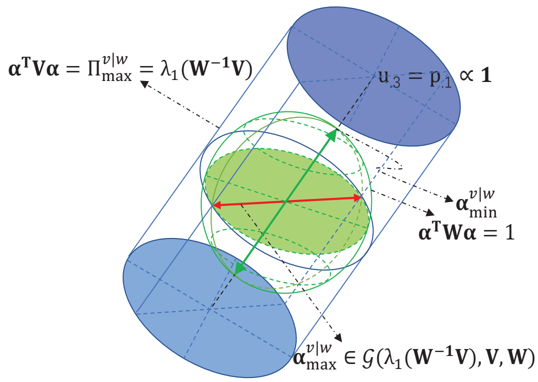


(c) $\lambda_2(\mathbf{W}) < 0, \lambda_3(\mathbf{W}) < 0, \Pi_{min}^{w|v} < 0$

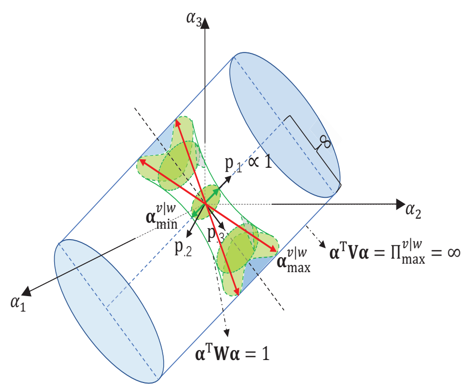
Figure 4. The optimal solutions to $[\mathcal{P}_{min}^{w|v}]$

Proof. See Appendix B. ■

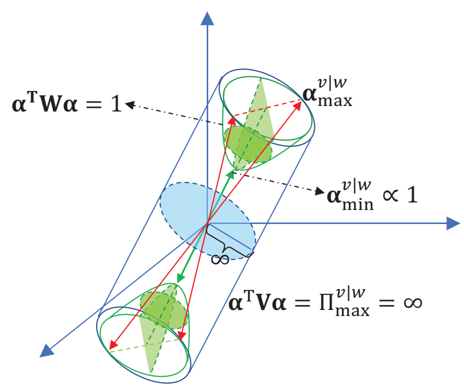
Figure 5 gives geometric illustrations for the $n = 3$ case. When $\lambda_3(\mathbf{W}) > 0$, the maximum value $\Pi_{max}^{v|w} = \lambda_1(\mathbf{W}^{-1}\mathbf{V})$ is achieved when the elliptic cylinder $\boldsymbol{\alpha}^\top \mathbf{V}\boldsymbol{\alpha} = \Pi_{max}^{v|w}$ is circumscribed around the elliptic shell $\boldsymbol{\alpha}^\top \mathbf{W}\boldsymbol{\alpha} = 1$ (Figure 5(a)). When $\lambda_3(\mathbf{W}) < 0$, the elliptic cylinder expands unboundedly on the hyperbolic shell $\boldsymbol{\alpha}^\top \mathbf{W}\boldsymbol{\alpha} = 1$, so $\Pi_{max}^{v|w} = \infty$. In both cases, the minimum value $\Pi_{min}^{v|w} = 0$ is achieved at $\boldsymbol{\alpha}_{min}^{v|w} \propto \mathbf{1}$, where the elliptic cylinder degenerates to its axis.



(a) $\lambda_2(\mathbf{W}) > 0, \lambda_3(\mathbf{W}) > 0, \Pi_{max}^{v|w} = \lambda_1(\mathbf{W}^{-1}\mathbf{V}), \Pi_{min}^{v|w} = 0$



(b) $\lambda_2(\mathbf{W}) > 0, \lambda_3(\mathbf{W}) < 0, \Pi_{max}^{v|w} = \infty, \Pi_{min}^{v|w} = 0$



(c) $\lambda_2(\mathbf{W}) < 0, \lambda_3(\mathbf{W}) < 0, \Pi_{max}^{v|w} = \infty, \Pi_{min}^{v|w} = 0$

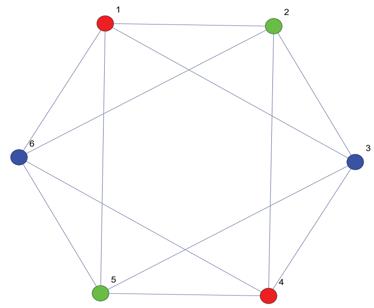
Figure 5. The optimal solution to $[\mathcal{P}_{min}^{v|w}]/[\mathcal{P}_{max}^{v|w}]$

Corollary 1 *For a $d-$regular network $\boldsymbol{G}$, we have: (i) if $\lambda_2(\boldsymbol{G}) + \lambda_n(\boldsymbol{G}) > 0$, both $\boldsymbol{\alpha}_{min}^{w|v}$ and $\boldsymbol{\alpha}_{max}^{v|w}$ are within space $\mathscr{S}(\lambda_2(\boldsymbol{G}), \boldsymbol{G})$, $\Pi_{min}^{w|v} = \psi(\lambda_2(\boldsymbol{G}))$, $\Pi_{max}^{v|w} = \frac{1}{\psi(\lambda_2(\boldsymbol{G}))}$; (ii) if $\lambda_2(\boldsymbol{G}) + \lambda_n(\boldsymbol{G}) < 0$, both $\boldsymbol{\alpha}_{min}^{w|v}$ and $\boldsymbol{\alpha}_{max}^{v|w}$ are within space $\mathscr{S}(\lambda_n(\boldsymbol{G}), \boldsymbol{G})$, $\Pi_{min}^{w|v} = \psi(\lambda_n(\boldsymbol{G}))$, $\Pi_{max}^{v|w} = \frac{1}{\psi(\lambda_n(\boldsymbol{G}))}$, where $\mathscr{S}(\lambda, \boldsymbol{A})$ denotes the $\lambda-$eigenspace of matrix $\boldsymbol{A}$, $\psi(x) \equiv \frac{d^2 - \lambda x^2}{d^2(1-\lambda)}$.*

Proof. See Appendix C. ■

For a regular network, $\boldsymbol{\alpha}_{min}^{w|v}$ and $\boldsymbol{\alpha}_{max}^{v|w}$ coincide with either the second or the least principal component of $\mathbf{G}$. The magnitudes comparison between the second and the smallest eigenvalues plays a crucial role in the determination of the optimal interventions. By the Perron-Frobenius

(a) $\lambda_1(\mathbf{G}) = 3, \lambda_2(\mathbf{G}) = 2.778, \lambda_n(\mathbf{G}) = -2.489$



(b) $\lambda_1(\mathbf{G}) = 4, \lambda_2(\mathbf{G}) = 0, \lambda_n(\mathbf{G}) = -2$

Figure 6. $\lambda_2(\mathbf{G}), \lambda_n(\mathbf{G})$ for different networks

Theorem, we have $|\lambda_i(\mathbf{G})| \leqslant \lambda_1(\mathbf{G}), \forall i \neq 1$ for a connected network. So the comparison between magnitudes of $\lambda_2(\mathbf{G})$ and $\lambda_n(\mathbf{G})$ hinges on the spectral gap $\lambda_1(\mathbf{G}) - \lambda_2(\mathbf{G})$, which reflects the degree of "cohesiveness" of a network. [5] If the *spectral gap* is sufficiently small, i.e., $\lambda_1(\mathbf{G}) - \lambda_2(\mathbf{G}) \approx 0$, and $\mathbf{G}$ is non-bipartite[6], then $\lambda_2(\mathbf{G}) + \lambda_n(\mathbf{G}) > 0$. If, on the contrary, the network has a large spectral gap, i.e., $\lambda_2(\mathbf{G}) \ll \lambda_1(\mathbf{G})$, then $\lambda_2(\mathbf{G}) + \lambda_n(\mathbf{G})$ is more inclined to be negative.

In Figure 6(a), there is a unique bridge between two isolated communities. $\lambda_2(\mathbf{G})$ turns out

[5]Consider a network consisting of two identical and isolated components, each has an adjacency matrix $\widetilde{\mathbf{G}}$. Then the overall network has an adjacency matrix

$$\mathbf{G} = \begin{bmatrix} \widetilde{\mathbf{G}} & \mathbf{0} \\ \mathbf{0} & \widetilde{\mathbf{G}} \end{bmatrix}.$$

The first eigenvalue and eigenvector are

$$\lambda_2(\mathbf{G}) = \max_{\|\mathbf{x}_1\|^2 + \|\mathbf{x}_2\|^2 = 1} \mathbf{x}_1^\top \widetilde{\mathbf{G}} \mathbf{x}_1 + \mathbf{x}_2^\top \widetilde{\mathbf{G}} \mathbf{x}_2 \text{ and } (\mathbf{v}^*, \mathbf{v}^*) = \arg \max_{\|\mathbf{x}_1\|^2 + \|\mathbf{x}_2\|^2 = 1} \mathbf{x}_1^\top \widetilde{\mathbf{G}} \mathbf{x}_1 + \mathbf{x}_2^\top \widetilde{\mathbf{G}} \mathbf{x}_2.$$

The second largest eigenvalue is

$$\lambda_2(\mathbf{G}) = \max_{\|\mathbf{y}_1\|^2 + \|\mathbf{y}_2\|^2 = 1, (\mathbf{y}_1, \mathbf{y}_2) \perp (\mathbf{v}^*, \mathbf{v}^*)} \mathbf{y}_1^\top \widetilde{\mathbf{G}} \mathbf{y}_1 + \mathbf{y}_2^\top \widetilde{\mathbf{G}} \mathbf{y}_2.$$

By choosing $(\mathbf{y}_1, \mathbf{y}_2) = (\mathbf{v}^*, -\mathbf{v}^*)$, we have $\lambda_2(\mathbf{G}) = \lambda_1(\mathbf{G}) = 2(\mathbf{v}^*)^\top \widetilde{\mathbf{G}} \mathbf{v}^*$. A network with two identical and weakly connected communities is akin to the 2-component network $\mathbf{G}$, it thus has a small spectral gap. If $\mathbf{G}$ is a regular graph, then its spectral gap is equal to the second smallest eigenvalue of its Laplacian matrix and known as *algebraic connectivity*. Abdi et al. (2021) shows that among all connected 3-regular graphs on $n \geqslant 10$ vertices, the path-form graph given in Figure 6(a) is the unique one with minimum algebraic connectivity.

[6]By graph theory, a connected graph $\mathbf{G}$ satisfies $|\lambda_n(\mathbf{G})| = \lambda_1(\mathbf{G})$ if and only if $\mathbf{G}$ is bipartite, $|\lambda_n(\mathbf{G})| < \lambda_1(\mathbf{G})$ for any non-bipartite graph.

to be almost as large as $\lambda_1(\mathbf{G})$. This yields a small spectral gap and thus $\lambda_2(\mathbf{G}) + \lambda_n(\mathbf{G}) > 0$. Figure 6(b) gives a complete tripartite network. This network is highly cohesive in the sense that there exists an edge between every pair of vertices from different independent groups labelled by different colors, but there is no inter-group edge. In this network, $\lambda_2(\mathbf{G})$ is considerably smaller than $\lambda_1(\mathbf{G})$, which leads to a large spectral gap and thus $\lambda_2(\mathbf{G}) + \lambda_n(\mathbf{G}) < 0$.

The above problems admit no physical cost of intervention. We now assume that the planner needs to pay a cost of $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^0\|$ when adjusting a vector of status quo preference $\boldsymbol{\alpha}^0$ to a vector $\boldsymbol{\alpha}$. This cost increases in the magnitude of the change and is separable across individuals. An utilitarian (resp. egalitarian) planner aims at maximizing the social welfare (resp. minimizing the social variance) subject to a budget constraint:

$$[\mathcal{P}^u] : \max_{\boldsymbol{\alpha}\in\mathbb{R}^n} \mathbb{W}^N(\lambda, \boldsymbol{\alpha}, \mathbf{g}), s.t. : \|\boldsymbol{\alpha} - \boldsymbol{\alpha}^0\| \leqslant C, \tag{8}$$

$$[\mathcal{P}^e] : \min_{\boldsymbol{\alpha}\in\mathbb{R}^n} \mathbb{V}^N(\lambda, \boldsymbol{\alpha}, \mathbf{g}), s.t. : \|\boldsymbol{\alpha} - \boldsymbol{\alpha}^0\| \leqslant C, \tag{9}$$

where C is a given budget.

Proposition 3 *From problem $[\mathcal{P}^u]$, we get the following results.*

(i) If $|\bar{\alpha}^0| \equiv \mathbf{1}^\top \boldsymbol{\alpha}^0/n \neq 0$, the maximum value attained is

$$W_{max} = \sum_{i=1}^{m} \tilde{w}_i \left(\frac{\mu^* \left\| \mathrm{Proj}_{\mathscr{S}(\tilde{w}_i, \boldsymbol{W})} \boldsymbol{\alpha}^0 \right\|}{\mu^* - \tilde{w}_i} \right)^2 ; \tag{10}$$

the optimal intervention policy $\boldsymbol{\alpha}^*$ satisfies

$$\mathbb{C}os \left\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, \boldsymbol{W}) \right\rangle = \frac{\|\boldsymbol{\alpha}^0\|}{C} \frac{|\tilde{w}_i|}{\mu^* - \tilde{w}_i} \mathbb{C}os \left\langle \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, \boldsymbol{W}) \right\rangle, \forall i \in I, \tag{11}$$

where $\mu^* > 1$, the shadow price of the budget constraint, is uniquely determined by

$$\sqrt{\sum_{i=1}^{m} \left(\frac{\tilde{w}_i \left\| \mathrm{Proj}_{\mathscr{S}(\tilde{w}_i, \boldsymbol{W})} \boldsymbol{\alpha}^0 \right\|}{\mu^* - \tilde{w}_i} \right)^2} = C, \tag{12}$$

$1 = \tilde{w}_1 > \tilde{w}_2 > \cdots > \tilde{w}_m$ are distinct eigenvalues of matrix $\boldsymbol{W}$, $\mathscr{S}(\tilde{w}_i, \boldsymbol{W})$ is the $\tilde{w}_i-$eigenspace of $\boldsymbol{W}$, $\mathrm{Proj}_{\boldsymbol{X}}\boldsymbol{x}$ denotes the projection of a vector $\boldsymbol{x}$ onto space $\boldsymbol{X}$, $\mathbb{C}os\langle\cdot,\cdot\rangle$ denotes the cosine of angle between a vector and a linear space.

(ii) If $|\bar{\alpha}^0| = 0$, and

$$C \leqslant C^* \equiv \sqrt{\sum_{i=2}^{m} \left(\frac{\tilde{w}_i \left\| \mathrm{Proj}_{\mathscr{S}(\tilde{w}_i, \boldsymbol{W})} \boldsymbol{\alpha}^0 \right\|}{1 - \tilde{w}_i} \right)^2},$$

the optimal social welfare attained is

$$W_{max} = \sum_{i=2}^{m} \tilde{w}_i \left(\frac{\mu^* \left\| \text{Proj}_{\mathscr{S}(\tilde{w}_i, \boldsymbol{W})} \boldsymbol{\alpha}^0 \right\|}{\mu^* - \tilde{w}_i} \right)^2; \tag{13}$$

the optimal intervention policy $\boldsymbol{\alpha}^*$ *satisfies*

$$\mathbb{C}os \left\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, \boldsymbol{W}) \right\rangle = \frac{\|\boldsymbol{\alpha}^0\|}{C} \frac{|\tilde{w}_i|}{\mu^* - \tilde{w}_i} \mathbb{C}os \left\langle \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, \boldsymbol{W}) \right\rangle, \forall i = 1, \cdots, m; \tag{14}$$

the optimal shadow price $\mu^* > 1$ *is determined by*

$$\sqrt{\sum_{i=2}^{m} \left(\frac{\tilde{w}_i \left\| \text{Proj}_{\mathscr{S}(\tilde{w}_i, \boldsymbol{W})} \boldsymbol{\alpha}^0 \right\|}{\mu^* - \tilde{w}_i} \right)^2} = C.$$

(iii) If $|\bar{\alpha}^0| = 0$ *and* $C > C^*$*, the optimal value attained is*

$$W_{max} = C^2 - (C^*)^2 + \sum_{i=2}^{m} \tilde{w}_i \left(\frac{\left\| \text{Proj}_{\mathscr{S}(\tilde{w}_i, \boldsymbol{W})} \boldsymbol{\alpha}^0 \right\|}{1 - \tilde{w}_i} \right)^2; \tag{15}$$

the optimal intervention policy $\boldsymbol{\alpha}^*$ *satisfies*

$$\mathbb{C}os \left\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_1, \boldsymbol{W}) \right\rangle = \sqrt{1 - \left(\frac{C^*}{C} \right)^2} \tag{16}$$

and

$$\mathbb{C}os \left\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, \boldsymbol{W}) \right\rangle = \frac{\|\boldsymbol{\alpha}^0\|}{C} \frac{|\tilde{w}_i|}{\mu^* - \tilde{w}_i} \mathbb{C}os \left\langle \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, \boldsymbol{W}) \right\rangle, \forall i \geqslant 2; \tag{17}$$

the optimal shadow price is $\mu^* = 1$.

Proof. See Appendix D. ■

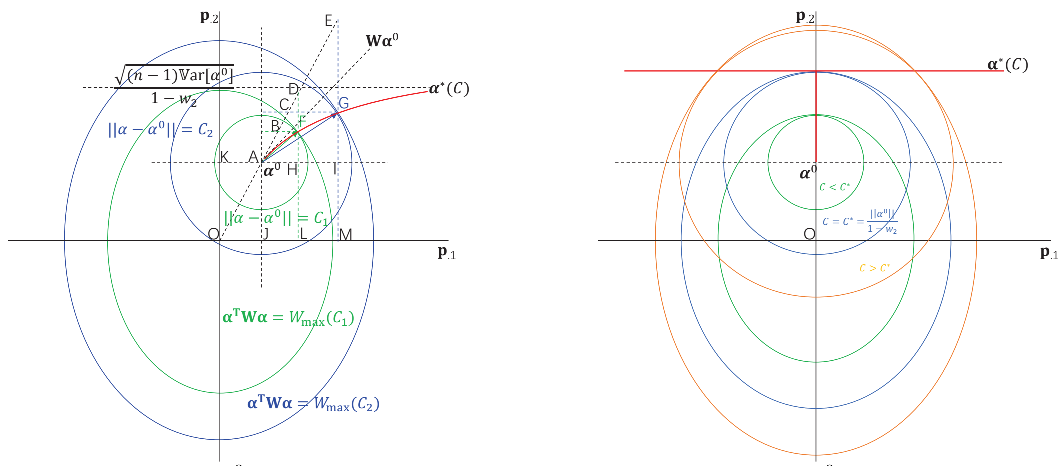
The cosine similarity between the optimal intervention and the ith eigenspace $\mathbb{C}os\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, \mathbf{W})\rangle$ measures the extent to which the principal components associated with $\tilde{w}_i$ are represented in the optimal intervention. Expression (11) shows that this hinges on two factors. The first factor, $\mathbb{C}os \left\langle \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, \mathbf{W}) \right\rangle$ reflects a status-quo effect: how much the initial condition influences the optimal intervention for a given budget. The smaller is the angle between $\boldsymbol{\alpha}^0$ and $\mathscr{S}(\tilde{w}_i, \mathbf{W})$, the greater portion of resources are allocated towards the direction of $\mathscr{S}(\tilde{w}_i, \mathbf{W})$. The second factor $\gamma_i(C) \equiv |\tilde{w}_i|/(\mu^* - \tilde{w}_i)$ is determined jointly by the network topology (via its ith eigenvalue $\tilde{w}_i$), and the total budget C (via the shadow price μ^*). It increases in C, and the increasing rate is the greatest for the first principal component. Figure 7(a) provides a graphical illustration for the case of $n = 2$ and $|\bar{\alpha}^0| \neq 0$.[7] Notice that

$$\gamma_i(C) = \frac{C \cdot \mathbb{C}os \left\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, \mathbf{W}) \right\rangle}{\|\boldsymbol{\alpha}^0\| \, \mathbb{C}os \left\langle \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, \mathbf{W}) \right\rangle}$$

[7] A two-person connected network $\mathbf{G}$ must be 1−regular, with $\lambda_1(\mathbf{G}) = 1, \lambda_2(\mathbf{G}) = -1$ and $\widehat{\mathbf{M}} = (1 - \lambda)(\mathbf{I} - \lambda\mathbf{G})^{-1}$. The least eigenvalue is positive: $\lambda_2(\mathbf{W}) = ((1 - \lambda)/(1 + \lambda))^2 > 0$.

at the optimum. So we find: $\gamma_1(C_1) = |AH|/|OJ| = |AD|/|OA| < \gamma_1(C_2) = |AI|/|OJ| = |AE|/|OA|$, $\gamma_2(C_1) = |FH|/|AJ| = |AB|/|OA| < \gamma_2(C_2) = |GI|/|AJ| = |AC|/|OA|$, and $\Delta\gamma_1 \equiv \gamma_1(C_2) - \gamma_1(C_1) = |DE|/|OA| > |BC|/|OA| = \Delta\gamma_2 \equiv \gamma_2(C_2) - \gamma_2(C_1)$ for $C_2 > C_1$.

This proposition also shows that when the status-quo vector $\boldsymbol{\alpha}^0$ is orthogonal to the first principal component $\mathbf{p}_{\cdot 1} \propto \mathbf{1}$, there exists a critical level of budget C^*. When the budget is below C^*, the optimal intervention remains perpendicular to $\mathbf{p}_{\cdot 1}$, while it is parallel to $\mathbf{p}_{\cdot 1}$ when the budget exceeds C^*. It means that the planner will leave the direction $\mathbf{p}_{\cdot 1}$ unintervened when her budget is insufficient. When budget permits, however, she will allocate all extra resources in the direction $\mathbf{p}_{\cdot 1}$. This can be illustrated by Figure 7(b). The trajectory of $\boldsymbol{\alpha}^*$(C) is of $T-$shape. In what follows, we give limit forms of the optimal interventions in the cases of extremely small and large budgets.



(a) $\bar{\alpha}^0 \neq 0$ (b) $\bar{\alpha}^0 = 0$

Figure 7. Trajectory of the optimal solution $\boldsymbol{\alpha}^*$(C)

Corollary 2 *(i) The optimal intervention* $\boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0$ *tends to the direction of* $\boldsymbol{W\alpha}^0$ *as the budget becomes small, i.e.,* $\lim_{C\to 0}\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \boldsymbol{W\alpha}^0\rangle = 0$. *(ii) For large budget, it converges to the first eigenspace, i.e.,* $|\lim_{C\to\infty} \cos\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \boldsymbol{1}\rangle| = 1$.

Proof. See Appendix E. ■

This corollary shows that, with a large budget, the welfare-maximizing intervention $\boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0$ is approximately simple in the sense that the budget is distributed equally across all individuals. If the total budget is small, however, the allocation of budget is proportional to the tangent direction $\mathbf{W}\boldsymbol{\alpha}^0$, which depends on both the network structure $\mathbf{G}$ and the status quo vector $\boldsymbol{\alpha}^0$. To facilitate further analysis, we denote by $\boldsymbol{\alpha}^s \equiv \boldsymbol{\alpha}^0 + sign(\bar{\alpha}^0)C\mathbf{1}/\sqrt{n}$ the simple approximation of the optimal intervention for large budget, and by $\boldsymbol{\alpha}_w^t \equiv \boldsymbol{\alpha}^0 + C\mathbf{W}\boldsymbol{\alpha}^0/\|\mathbf{W}\boldsymbol{\alpha}^0\|$ the tangent approximation for small budget.

The following proposition characterizes solution to the variance-minimization problem $[\mathcal{P}^e]$.

Proposition 4 *(i) if $C < \sqrt{(n-1)\mathbb{V}ar(\boldsymbol{\alpha}^0)}$, then the minimum value attained is*

$$V_{min} = \sum_{i=1}^{r-1} \tilde{\nu}_i \left(\frac{\mu^\dagger \left\| \mathrm{Proj}_{\mathscr{S}(\tilde{\nu}_i, \boldsymbol{V})} \boldsymbol{\alpha}^0 \right\|}{\mu^\dagger - \tilde{\nu}_i} \right)^2 ; \tag{18}$$

the optimal intervention policy $\boldsymbol{\alpha}^\dagger - \boldsymbol{\alpha}^0$ satisfies

$$\mathbb{C}os \left\langle \boldsymbol{\alpha}^\dagger - \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{\nu}_i, \boldsymbol{V}) \right\rangle = \frac{\boldsymbol{\alpha}^0}{C} \frac{\tilde{\nu}_i}{\mu^\dagger - \tilde{\nu}_i} \mathbb{C}os \left\langle \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{\nu}_i, \boldsymbol{V}) \right\rangle, \forall i = 1, \cdots, r, \tag{19}$$

where $\mathbb{V}ar(\boldsymbol{\alpha}^0) = \frac{1}{n-1}(\boldsymbol{\alpha}^0)^\top \boldsymbol{Q}\boldsymbol{\alpha}^0$ denotes the sample variance of vector $\boldsymbol{\alpha}^0$, $\tilde{\nu}_1 > \cdots > \tilde{\nu}_{r-1} > \tilde{\nu}_r = 0$ are distinct eigenvalues of $\boldsymbol{V}$, $\mu^\dagger < 0$ is the shadow price of the planner's budget determined by

$$\sqrt{\sum_{i=1}^{r} \left[\frac{\tilde{\nu}_i \left\| \mathrm{Proj}_{\mathscr{S}(\tilde{\nu}_i, \boldsymbol{V})} \boldsymbol{\alpha}^0 \right\|}{\mu^\dagger - \tilde{\nu}_i} \right]^2} = C; \tag{20}$$

(ii) if $C \geqslant \sqrt{(n-1)\mathbb{V}ar(\boldsymbol{\alpha}^0)}$, then we can obtain $V_{min} = 0$ by choosing an arbitrary $\boldsymbol{\alpha}^\dagger \in \{\boldsymbol{z} \in \mathbb{R}^n : \boldsymbol{z} \propto \boldsymbol{1}\} \cap \mathcal{O}(\boldsymbol{\alpha}^0, C)$, where $\mathcal{O}(\boldsymbol{\alpha}^0, C) \equiv \{\boldsymbol{z} \in \mathbb{R}^n : \|\boldsymbol{z} - \boldsymbol{\alpha}^0\| \leqslant C\}$ denotes the neighborhood of $\boldsymbol{\alpha}^0$ with a radius C;

(iii) $\boldsymbol{\alpha}^\dagger - \boldsymbol{\alpha}^0$ points to the direction of $-\boldsymbol{V}\boldsymbol{\alpha}^0$ as the total budget tends to zero, i.e., $\lim_{C\to 0}\langle \boldsymbol{\alpha}^\dagger - \boldsymbol{\alpha}^0, -\boldsymbol{V}\boldsymbol{\alpha}^0\rangle = 0$.

Proof. See Appendix F. ■

Figure 8 depicts the cross-section, within the orthogonal complements of $\mathbf{1}$, of a cylindrical surface $\boldsymbol{\alpha}^\top \mathbf{V} \boldsymbol{\alpha} = V_{min}$ and a sphere $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^0\| \leqslant C$. Elliptical cylinder $\boldsymbol{\alpha}^\top \mathbf{V} \boldsymbol{\alpha} = V_{min}$ shrinks as the budget C increases. When C exceeds $\|\mathbf{Q}\boldsymbol{\alpha}^0\| = \sqrt{(n-1)\mathbb{V}ar(\boldsymbol{\alpha}^0)}$, the cylinder is degenerate to its axis $\mathbf{1}$, we thus have $V_{min} = 0$. In the opposite extreme case where C approaches zero, $\boldsymbol{\alpha}^\dagger - \boldsymbol{\alpha}^0$ tends to the opposite tangent direction $-\mathbf{V}\boldsymbol{\alpha}^0$. We denote by $\boldsymbol{\alpha}_v^t \equiv \boldsymbol{\alpha}^0 - C\mathbf{V}\boldsymbol{\alpha}^0 / \|\mathbf{V}\boldsymbol{\alpha}^0\|$ the tangent approximation of the optimal variance-minimizing intervention under small budget.

Approximating the optimal intervention by the simple and tangent interventions allows us to avoid computational complexity. The simple intervention $\boldsymbol{\alpha}^s$, in particular, is very easy to implement since it depends on neither the network structure nor the status-quo preferences. Meanwhile, some discrepancy between the optimal value and its approximations may inevitably exist. We denote by $W^*(C) \equiv [\boldsymbol{\alpha}^*(C)]^\top \mathbf{W} [\boldsymbol{\alpha}^*(C)]$ and $V^*(C) \equiv [\boldsymbol{\alpha}^\dagger(C)]^\top \mathbf{V} [\boldsymbol{\alpha}^\dagger(C)]$ the optimal social

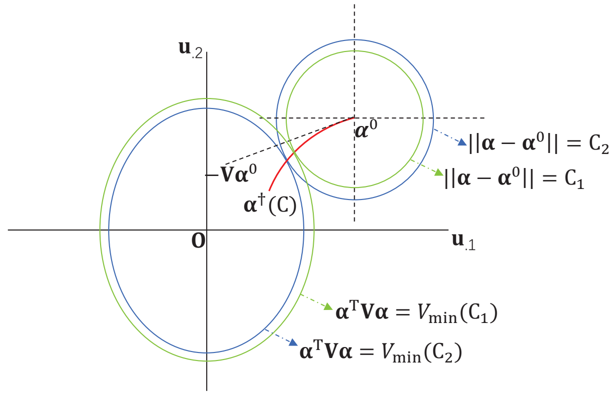


Figure 8. Trajectory of the optimal solution $\boldsymbol{\alpha}^{\dagger}(C)$

welfare and social variance. $W^s(C) \equiv [\boldsymbol{\alpha}^s(C)]^\top \mathbf{W}[\boldsymbol{\alpha}^s(C)]$ and $W^t(C) \equiv [\boldsymbol{\alpha}_w^t(C)]^\top \mathbf{W}[\boldsymbol{\alpha}_w^t(C)]$ represent the levels of social welfare under simple and tangent approximations, respectively. We use $\Delta_w^s(C,\boldsymbol{\alpha}^0) \equiv W^*(C) - W^s(C)$ and $\Delta_w^t(C,\boldsymbol{\alpha}^0) \equiv W^*(C) - W^t(C)$ to represent the errors of simple and tangent approximations. Similarly, $V^t(C) \equiv [\boldsymbol{\alpha}_v^t(C)]^\top \mathbf{V}[\boldsymbol{\alpha}_v^t(C)]$ represents the social variance under tangent intervention, $\Delta_v^t(C,\boldsymbol{\alpha}^0) \equiv V^t(C) - V^*(C)$ represents the error on variance under simple approximation. The following proposition characterizes the limit behaviors of these errors under extremely large and small budgets, and thus measures the accuracies of these approximations.

Proposition 5 • *The error of social welfare under simple approximation is*

$$\Delta_w^s(\infty,\boldsymbol{\alpha}^0) = \lim_{C\to\infty} \Delta_w^s(C,\boldsymbol{\alpha}^0) = \sum_{i=2}^{m} \frac{\tilde{w}_i^2 \left\| \mathrm{Proj}_{\mathscr{S}(\tilde{w}_i,\boldsymbol{W})}\boldsymbol{\alpha}^0 \right\|^2}{1-\tilde{w}_i}; \tag{21}$$

if the sample variance $\mathbb{V}ar(\boldsymbol{\alpha}^0)$ *is fixed at* ξ*, then* $\Delta_w^s(\infty,\boldsymbol{\alpha}^0)$ *attains its maximum value* $\frac{(n-1)\xi\tilde{w}_{i^*}^2}{1-\tilde{w}_{i^*}}$ *at* $\boldsymbol{\alpha}^0 \in \mathscr{S}(\tilde{w}_{i^*}, \boldsymbol{W})$*, and its minimum value* $\frac{(n-1)\xi\tilde{w}_{i_*}^2}{1-\tilde{w}_{i_*}}$ *at* $\boldsymbol{\alpha}^0 \in \mathscr{S}(\tilde{w}_{i_*}, \boldsymbol{W})$*, where* $i^* = \arg\max_{j\neq 1} \frac{\tilde{w}_j^2}{1-\tilde{w}_j}$, $i_* = \arg\min_{j\neq 1} \frac{\tilde{w}_j^2}{1-\tilde{w}_j}$;

- *With extremely small budget, the tangent approximation errors of social welfare and social variance are*

$$\Delta_w^t(C,\boldsymbol{\alpha}^0) = \begin{cases} 0 & \text{if } \boldsymbol{\alpha}^0 \in \bigcup_{i=1}^m \mathscr{S}(\tilde{w}_i, \boldsymbol{W}) \\ \frac{12\left[\|\boldsymbol{W}\boldsymbol{\alpha}^0\|^2 \|\boldsymbol{W}^2\boldsymbol{\alpha}^0\|^2 - \left\|\boldsymbol{W}^{\frac{3}{2}}\boldsymbol{\alpha}^0\right\|^4\right]}{\|\boldsymbol{W}\boldsymbol{\alpha}^0\|^5} C^3 + o(C^3) & \text{if } \boldsymbol{\alpha}^0 \notin \bigcup_{i=1}^m \mathscr{S}(\tilde{w}_i, \boldsymbol{W}) \end{cases},$$

and

$$\Delta_v^t(C,\boldsymbol{\alpha}^0) = \begin{cases} 0 & \text{if } \boldsymbol{\alpha}^0 \in \bigcup_{i=1}^r \mathscr{S}(\tilde{\nu}_i, \boldsymbol{V}) \\ \frac{12\left[\|\boldsymbol{V}\boldsymbol{\alpha}^0\|^2 \|\boldsymbol{V}^2\boldsymbol{\alpha}^0\|^2 - \left\|\boldsymbol{V}^{\frac{3}{2}}\boldsymbol{\alpha}^0\right\|^4\right]}{\|\boldsymbol{V}\boldsymbol{\alpha}^0\|^5} C^3 + o(C^3) & \text{if } \boldsymbol{\alpha}^0 \notin \bigcup_{i=1}^r \mathscr{S}(\tilde{\nu}_i, \boldsymbol{V}) \end{cases}.$$

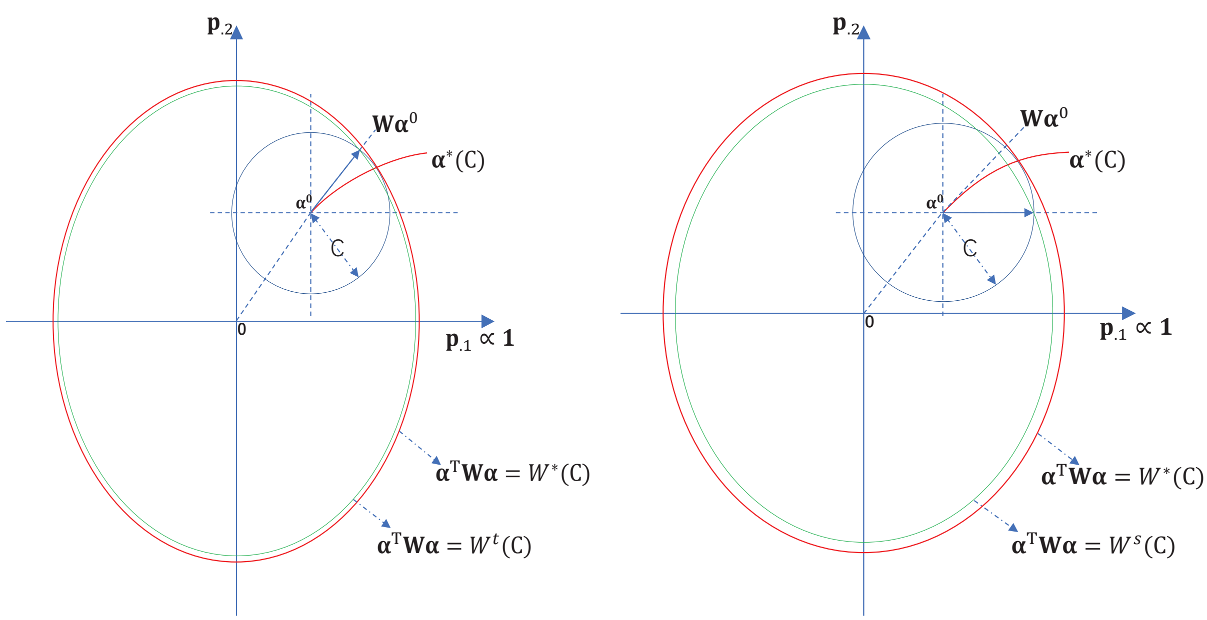
Proof. See Appendix G. ■

We illustrate the above results by Figures 9 and 10. The red and green curves represent, respectively, the optimal and the approximated values. The gaps between them measure the errors of approximations. Figure 9(a) (resp. 10) shows that, when the status quo vector $\boldsymbol{\alpha}^0$ lies within none of the eigenspaces of $\mathbf{W}$ (resp. $\mathbf{V}$), the tangent approximation error for both social welfare (resp. social variance) vanishes in the order of $O(C^3)$ as C tends to zero. Figure (9(b)) shows that the error of simple approximation persists as $C \to \infty$, i.e., $\Delta_w^s(\infty, \boldsymbol{\alpha}^0) > 0$. In the case where $\mathbf{W}$ is positive definite (i.e., $\tilde{w}_m > 0$), we have $\frac{\tilde{w}_2^2}{1-\tilde{w}_2} > \cdots > \frac{\tilde{w}_m^2}{1-\tilde{w}_m}$. So $\Delta_w^s(\infty, \boldsymbol{\alpha}^0)$ achieves its maximum (resp. minimum) at $\boldsymbol{\alpha}^0 \in \mathscr{S}(\tilde{w}_2, \mathbf{W})$ (resp. $\boldsymbol{\alpha}^0 \in \mathscr{S}(\tilde{w}_m, \mathbf{W})$).

Since

$$\lim_{C\to\infty} \frac{C^2}{W^*(C)} = \lim_{\mu\to\tilde{w}_1=1} \frac{\sum_{j=1}^{m} \left(\frac{\tilde{w}_j \left\| \mathrm{Proj}_{\mathscr{S}(\tilde{w}_i,\mathbf{W})} \boldsymbol{\alpha}^0 \right\|}{\mu - \tilde{w}_j} \right)^2}{\sum_{j=1}^{m} \tilde{w}_j \left(\frac{\left\| \mathrm{Proj}_{\mathscr{S}(\tilde{w}_i,\mathbf{W})} \boldsymbol{\alpha}^0 \right\|}{\mu - \tilde{w}_j} \right)^2} = 1$$

The relative approximation error $\frac{\Delta_w^s(C,\boldsymbol{\alpha}^0)}{W^*(C)}$ is an infinitesimal of order $O(1/C^2)$. With controllably small error, the simple interventions achieves most of the optimal social welfare for sufficiently large budget.



(a) tangent approximation for small budget

(b) simple approximation for large budget

Figure 9. Tangent and simple approximations in $[\mathcal{P}^u]$

In the game-theoretic framework, the first-best social welfare $\mathbb{W}^{FB}(\theta, \boldsymbol{\alpha}, \mathbf{g}) = \frac{1}{2}\boldsymbol{\alpha}^\top \mathbf{H}^{-1} \boldsymbol{\alpha}$ is usually unachievable. A problem of interest is how to achieve the largest possible proportion

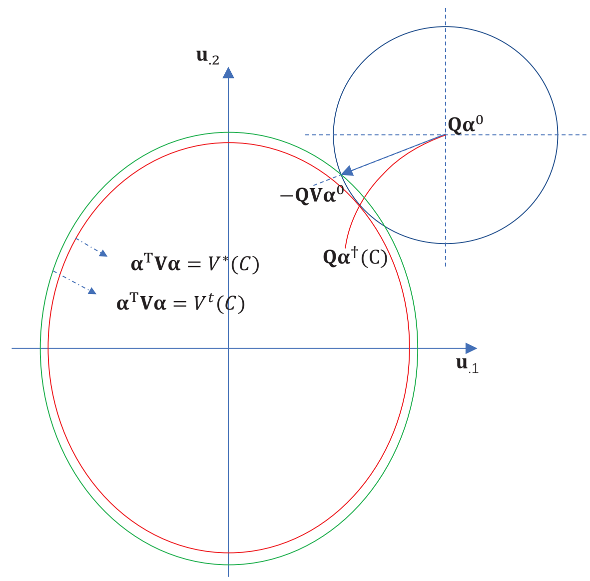


Figure 10. Tangent approximation for small budget in $[\mathcal{P}^e]$

of $\mathbb{W}^{FB}(\theta, \boldsymbol{\alpha}, \mathbf{g})$ through ex-ante intervention. We proceed to discuss the *relative intervention problem*

$$[\mathcal{P}^r] : \max_{\boldsymbol{\alpha}\in\mathbb{R}^n} \frac{\boldsymbol{\alpha}^\top \mathbf{W}\boldsymbol{\alpha}}{\boldsymbol{\alpha}^\top \mathbf{H}^{-1}\boldsymbol{\alpha}}, \text{s.t.} : \|\boldsymbol{\alpha} - \boldsymbol{\alpha}^0\| \leqslant C. \tag{22}$$

To facilitate discussion, we prepare some notations. Let $R^*(C)$ and $\boldsymbol{\alpha}^r(C)$ be the optimal value and solution to problem $[\mathcal{P}^r]$, respectively. Let $1 = \gamma_1 \geqslant \gamma_2 \geqslant \cdots \geqslant \gamma_n$ be the generalized eigenvalues of matrix $\mathbf{W}$ relative to $\mathbf{H}^{-1}$, among which $1 = \tilde{\gamma}_1 > \cdots > \tilde{\gamma}_m$ are distinct values. $N_i \equiv \{j \in I | \gamma_j = \tilde{\gamma}_i\}$, $n_i \equiv \#N_i$ denotes the multiplicity of γ_i. $\mathscr{G}(\tilde{\gamma}_i, \mathbf{W}, \mathbf{H}^{-1}) \equiv \{\mathbf{x} : \mathbf{W}\mathbf{x} = \tilde{\gamma}_i \mathbf{H}^{-1}\mathbf{x}\}$ denotes the generalized eigenspace, of $\mathbf{W}$ to $\mathbf{H}^{-1}$, associated with $\tilde{\gamma}_i$. Let $\overline{C}_0 \equiv \infty$ and $\overline{C}_t \equiv \mathscr{D}\left(\boldsymbol{\alpha}^0, \oplus_{i=1}^t \mathscr{G}(\tilde{\gamma}_i, \mathbf{W}, \mathbf{H}^{-1})\right), \forall t \in \{1, \cdots, m\}$, where $\mathscr{D}(\mathbf{y}, \mathbf{Z}) \equiv \inf\{\|\mathbf{y} - \mathbf{z}\| : \mathbf{z} \in \mathbf{Z}\}$ is the Euclidean distance from a vector $\mathbf{y}$ to a linear space $\mathbf{Z}$, symbol $\oplus$ denotes the direct sum of linear spaces. It is obvious that $\overline{C}_t \geqslant \overline{C}_{t+1}, \forall t \in \{0, \cdots, m-1\}$. Since $\mathbf{W}\mathbf{1} = \mathbf{H}^{-1}\mathbf{1}$, we have $\{\mathbf{y}|\mathbf{y} \propto \mathbf{1}\} \subseteq \mathscr{G}(\tilde{\gamma}_1, \mathbf{W}, \mathbf{H}^{-1})$. It follows that $\overline{C}_1 \leqslant \sqrt{(\boldsymbol{\alpha}^0)^\top \mathbf{Q}\boldsymbol{\alpha}^0} = \sqrt{(n-1)\mathbb{V}ar(\boldsymbol{\alpha}^0)}$, equality holds only when $n_1 = 1$.[8] Since all the m eigenspaces span the entire Euclidean space, i.e., $\oplus_{i=1}^t \mathscr{G}(\tilde{\gamma}_i, \mathbf{W}, \mathbf{H}^{-1}) = \mathbb{R}^n$, we have $\overline{C}_m = 0$.

Proposition 6 *We have the following results regarding problem $[\mathcal{P}^r]$:*

- *$R^*(C) \in [\tilde{\gamma}_t, 1]$ is attained at $\boldsymbol{\alpha}^r \in \oplus_{i=1}^t \mathscr{G}(\tilde{\gamma}_i, \boldsymbol{W}, \boldsymbol{H}^{-1})$, whenever $C \geqslant \overline{C}_t$, $\forall t \in \{1, \cdots, m\}$. In particular, $R^*(C) = \tilde{\gamma}_1 = 1$ whenever $C \geqslant \overline{C}_1$.*

[8] By Perron-Frobenius Theorem, 1 is a simple eigenvalue for both matrices $\mathbf{W}$ and $\mathbf{H}^{-1}$. However, it is not necessarily true that $\tilde{\gamma}_1 = 1$ is a simple generalized eigenvalue of $\mathbf{W}$ to $\mathbf{H}^{-1}$.

- *If* $C < \overline{C}_1$*, then* $\boldsymbol{\alpha}^r(C) = \Delta(C)\boldsymbol{\alpha}^0$*, and*

$$R^*(C) = \frac{(\boldsymbol{\alpha}^0)^\top \Delta(C)\, \boldsymbol{W}\Delta(C)\boldsymbol{\alpha}^0}{(\boldsymbol{\alpha}^0)^\top \Delta(C)\boldsymbol{H}^{-1}\Delta(C)\boldsymbol{\alpha}^0},$$

where $\Delta(C) \equiv [\boldsymbol{I} + \xi(C)\boldsymbol{H}^{-1} - \eta(C)\,\boldsymbol{W}]^{-1}$*,* $\xi(C)$ *and* $\eta(C)$ *are determined jointly by*[9]

$$\left[\boldsymbol{\alpha}^0\right]^\top \Delta(C)\left[\boldsymbol{\alpha}^0\right] = \left\|\boldsymbol{\alpha}^0\right\|^2 - C^2 \quad (23)$$

$$\left[\boldsymbol{\alpha}^0\right]^\top [\Delta(C)]^2\left[\boldsymbol{\alpha}^0\right] = \left\|\boldsymbol{\alpha}^0\right\|^2 - C^2. \quad (24)$$

Proof. See Appendix H. ∎

For large budget $C \geqslant \overline{C}_1$, the planner can obtain the first-best social welfare $W^{FB}(\theta, \boldsymbol{\alpha}, \mathbf{g})$ by choosing an $\boldsymbol{\alpha}^r$ within the first generalized eigenspace $\mathscr{G}(1, \mathbf{W}, \mathbf{H}^{-1})$. For small budget, the first-best efficiency is unachievable. To avoid the computational complexity, we approximate the optimal intervention by allocating the total budget in the tangent direction of the optimal trajectory $\boldsymbol{\alpha}^r(C)$ at $C = 0$. That is,

$$\boldsymbol{\alpha}^t(C) = \boldsymbol{\alpha}^0 + \frac{d\boldsymbol{\alpha}^r(0)/dC}{\|d\boldsymbol{\alpha}^r(0)/dC\|}C = \boldsymbol{\alpha}^0 + C\frac{\Delta'(0)\boldsymbol{\alpha}}{\|\Delta'(0)\boldsymbol{\alpha}\|}.$$

We proceed to discuss the accuracy of the tangent approximation.

Proposition 7 *For small budget, the error of tangent approximation is an infinitesimal with higher order than* C^2*, i.e.,* $\Delta R^t(C) \equiv R^*(C) - R^t(C) = o(C^2)$.

Proof. See Appendix I. ∎

Figure 11 depicts the optimal relative intervention $\boldsymbol{\alpha}^r(C)$ and its tangent approximation $\boldsymbol{\alpha}^t(C)$ for $n = 2$, where $\mathbf{p}_{\cdot i}$ (resp. $\mathbf{q}_{\cdot i}$) denotes the ith principal component of matrix $\mathbf{W}$ (resp. $\mathbf{H}^{-1}$).[10] Figure 12 depicts $R^*(C)$, $R^t(C)$ and a stepwise function $m(C) = \tilde{\gamma}_i, \forall C \in [\overline{C}_i, \overline{C}_{i-1}), \forall i \in \{1, \cdots, m\}$, for a six-node graph and $\lambda = 0.6$. In what follows, we proceed to discuss the relative intervention problem for two special network classes: the regular network $\mathscr{R}$, and complete multipartite network, $\mathscr{K}_{p_1,\cdots,p_s}$, where $\{p_i\}_{i=1}^s$ are partition numbers.

Proposition 8 *If* $\boldsymbol{G}$ *is a regular or complete bipartite network, i.e.,* $\boldsymbol{G} \in \mathscr{R} \cup \mathscr{K}_{p,q}$*, and* $\boldsymbol{\alpha}^0 \in \oplus_{i=1}^t \mathscr{G}(\tilde{\gamma}_i, \boldsymbol{W}, \boldsymbol{H}^{-1})$*, then* $\boldsymbol{\alpha}^r(C) \in \oplus_{i=1}^t \mathscr{G}(\tilde{\gamma}_i, \boldsymbol{W}, \boldsymbol{H}^{-1}), \forall t \in \{1, \cdots, m\}$.

[9] This proposition relies on an implicit hypothesis that the status quo vector $\boldsymbol{\alpha}^0$ guarantees the solvability of equation system (23) and (24). We give a counterexample. Let $\boldsymbol{\alpha}^0$ be a common eigenvector of matrices $\mathbf{W}$ and $\mathbf{H}$, corresponding to eigenvalues w and h, respectively. Then, (23) and (24) take the form $\frac{h}{h+\xi-\eta h w}\|\boldsymbol{\alpha}^0\|^2 = \|\boldsymbol{\alpha}^0\|^2 - C^2$ and $\left(\frac{h}{h+\xi-\eta h w}\right)^2 \|\boldsymbol{\alpha}^0\|^2 = \|\boldsymbol{\alpha}^0\|^2 - C^2$. This system is either unidentified ($C = 0$), or inconsistent ($C > 0$).

[10] A two-node connected graph must be complete, so $\mathbf{W}$ and $\mathbf{H}^{-1}$ share the same set of principal components. For the $n \geqslant 3$ case, it is not necessarily true that $\mathbf{p}_{\cdot j} \propto \mathbf{q}_{\cdot j}, \forall j \geqslant 2$.

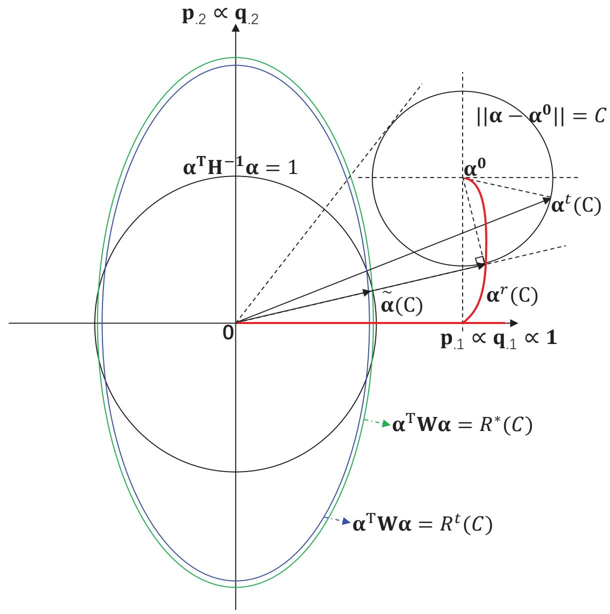
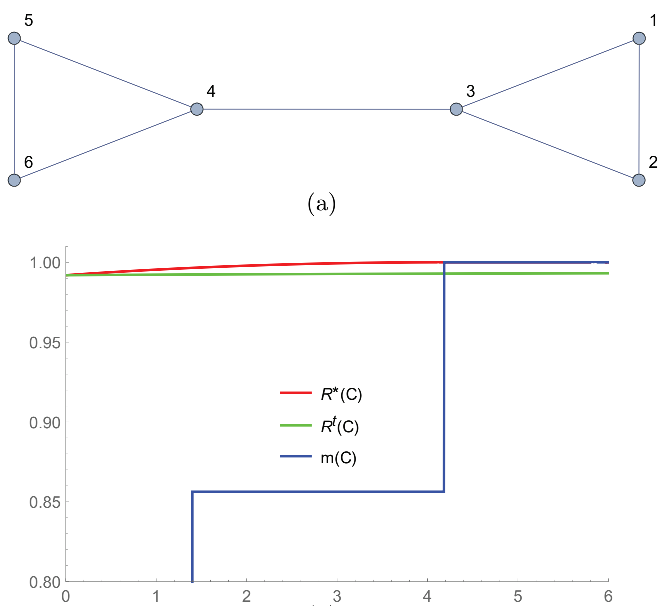


Figure 11. tangent approximation of relative intervention.



(a)

(b)

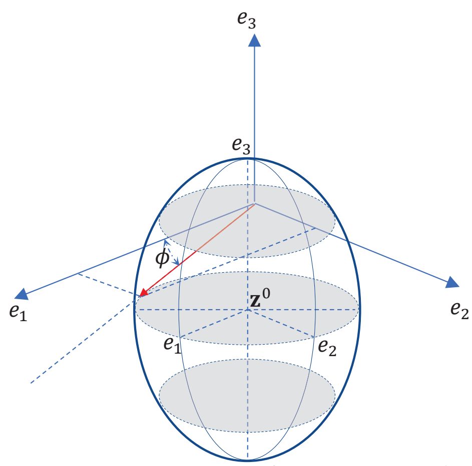
Figure 12. Tangent approximations in $[\mathcal{P}^r]$

Proof. See Appendix J. ■

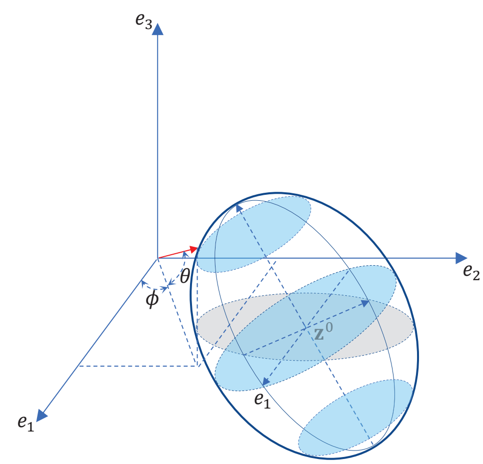
Notice that $\oplus_{i=1}^{m}\mathscr{G}(\tilde{\gamma}_i, \mathbf{W}, \mathbf{H}^{-1}) = \mathbb{R}^n$, so the above result holds obviously for $t = m$. However, if $\mathbf{G}$ is neither regular nor complete bipartite, matrices $\mathbf{W}$ and $\mathbf{H}$ don't necessarily commute. So $\mathbf{\Gamma} \equiv \mathbf{H}^{1/2}\mathbf{W}\mathbf{H}^{1/2}$ and $\mathbf{H}$ fail to share the same set of principal components (or eigenspaces for the case of multiple eigenvalues).[11] Figure 13 depicts the case with three agents and distinct eigenvalues $1 = \gamma_1 > \gamma_2 > \gamma_3$. Deforming the spherical region $\mathcal{O}(\boldsymbol{\alpha}^0, C) \equiv \{\boldsymbol{\alpha} : \|\boldsymbol{\alpha} - \boldsymbol{\alpha}^0\| \leqslant C\}$ by means of affine transformations $\mathbf{z} \equiv \mathbb{S}^\top \mathbf{H}^{-1/2}\boldsymbol{\alpha}$ and $\mathbf{z}^0 \equiv \mathbb{S}^\top \mathbf{H}^{-1/2}\boldsymbol{\alpha}^0$, we obtain the elliptical region

$$\mathcal{O}_{\widehat{\mathbf{H}}}(\mathbf{z}^0, C) \equiv \left\{\mathbf{z} : \sqrt{(\mathbf{z} - \mathbf{z}^0)^\top \widehat{\mathbf{H}}(\mathbf{z} - \mathbf{z}^0)} \leqslant C\right\},$$

where $\equiv \mathbb{S}^\top \mathbf{H}\mathbb{S}$. If $\mathbf{W}$ and $\mathbf{H}$ commute, $\widehat{\mathbf{H}}$ is a diagonal matrix. Then $\mathcal{O}_{\widehat{\mathbf{H}}}(\mathbf{z}^0, C)$ is a right ellipsoid, whose axes are parallel to the canonical bases $\{\mathscr{S}(\gamma_i, \widehat{\mathbf{\Gamma}})\}_{i=1}^3$. If $\mathbf{z}^0 \in \oplus_{i=1}^2 \mathscr{S}(\gamma_i, \widehat{\mathbf{\Gamma}})$, (equivalently, $\boldsymbol{\alpha}^0 \in \oplus_{i=1}^2 \mathscr{G}(\gamma_i, \mathbf{W}, \mathbf{H}^{-1})$), $\mathcal{O}_{\widehat{\mathbf{H}}}(\mathbf{z}^0, C)$ intersects the coordinate plane e_1Oe_2 at its its equator. The optimal intervention policy $\mathbf{z}^r$ must also located within e_1Oe_2 (equivalently, $\boldsymbol{\alpha}^r \in \oplus_{i=1}^2 \mathscr{G}(\gamma_i, \mathbf{W}, \mathbf{H}^{-1})$), with an polar angle ϕ from the first principal component e_1. The optimal value attained is $R^*(C) = \cos^2(\phi)\gamma_1 + \sin^2(\phi)\gamma_2$. If $\mathbf{W}$ and $\mathbf{H}$ don't commute, $\widehat{\mathbf{H}}$ is non-diagonal. Then $\mathcal{O}_{\widehat{\mathbf{H}}}(\mathbf{z}^0, C)$ is obliquely positioned in the canonical coordinate system $\{e_i\}_{i=1}^3$, with only its minor axis in alignment with e_1. In this case, $\mathbf{z}^r \notin \oplus_{i=1}^2 \mathscr{S}(\gamma_i, \widehat{\mathbf{\Gamma}})$ even if $\mathbf{z}^0 \in \oplus_{i=1}^2 \mathscr{S}(\gamma_i, \widehat{\mathbf{\Gamma}})$. The direction of $\mathbf{z}^r$ is determined by its zenithal and azimuthal angles (θ, ϕ). The optimal value attained is $R^*(C) = \cos^2(\theta)\cos^2(\phi)\gamma_1 + \cos^2(\theta)\sin^2(\phi)\gamma_2 + \sin^2(\theta)\gamma_3$.



(a) $\mathbf{WH} = \mathbf{HW}, \boldsymbol{\alpha}^r \in \oplus_{i=1}^2 \mathscr{G}(\tilde{\gamma}_i, \mathbf{W}, \mathbf{H}^{-1})$



(b) $\mathbf{WH} \neq \mathbf{HW}, \boldsymbol{\alpha}^r \notin \oplus_{i=1}^2 \mathscr{G}(\tilde{\gamma}_i, \mathbf{W}, \mathbf{H}^{-1})$

Figure 13. $\boldsymbol{\alpha}^0 \in \oplus_{i=1}^2 \mathscr{G}(\tilde{\gamma}_i, \mathbf{W}, \mathbf{H}^{-1})$

From the above analysis, we see that the first generalized eigenspace $\mathscr{G}(1, \mathbf{W}(\lambda), \mathbf{H}^{-1}(\lambda))$,

[11] For any network $\mathbf{G}$, the first principal component of $\mathbf{\Gamma}$ and $\mathbf{H}$ are both $\frac{1}{\sqrt{n}}\mathbf{1}$, but their remaining principal components don't coincide with each other.

and its Euclidean distance to $\boldsymbol{\alpha}^0$, play important roles in the relative intervention problem for a given λ. Next, I establish bounds on the $\lambda-$intersection of these eigenspaces.

Lemma 1 *$\lambda-$intersection of the first generalized eigenspaces $\bigcap_{\lambda\in(0,1)} \mathscr{G}\left(1, \boldsymbol{W}(\lambda), \boldsymbol{H}^{-1}(\lambda)\right)$ satisfies:*

$$\left[\mathbb{N}_\ell(\widehat{\boldsymbol{G}}) \bigcap \mathbb{N}_r(\widehat{\boldsymbol{G}})\right] \oplus \left\{\boldsymbol{x}|\boldsymbol{x} \propto \boldsymbol{1}\right\} \subseteq \bigcap_{\lambda\in(0,1)} \mathscr{G}\left(1, \boldsymbol{W}(\lambda), \boldsymbol{H}^{-1}(\lambda)\right) \subseteq \mathbb{N}_r(\widehat{\boldsymbol{G}}) \oplus \left\{\boldsymbol{x}|\boldsymbol{x} \propto \boldsymbol{1}\right\},$$

where $\mathbb{N}_\ell(\widehat{\boldsymbol{G}}) \equiv \left\{\boldsymbol{x}\middle|\widehat{\boldsymbol{G}}^\top \boldsymbol{x} = \boldsymbol{0}\right\}$ and $\mathbb{N}_r(\widehat{\boldsymbol{G}}) \equiv \left\{\boldsymbol{x}\middle|\widehat{\boldsymbol{G}}\boldsymbol{x} = \boldsymbol{0}\right\}$ denotes the left and right null spaces of matrix $\widehat{\boldsymbol{G}}$.

Proof. See Appendix K. ■

For graphs of general form, the left and right null spaces don't coincide with each other necessarily. See the tree graph given in Figure 14. $dim\mathbb{N}_r(\widehat{\mathbf{G}}) = n - rank(\widehat{\mathbf{G}}) = 7$, $dim[\mathbb{N}_r(\widehat{\mathbf{G}}) \cap \mathbb{N}_\ell(\widehat{\mathbf{G}})] = n - rank(\widehat{\mathbf{G}}^\top, \widehat{\mathbf{G}}) = 6$. So $\mathbb{N}_r(\widehat{\mathbf{G}}) \cap \mathbb{N}_\ell(\widehat{\mathbf{G}}) \subset \mathbb{N}_r(\widehat{\mathbf{G}})$.[12] The following lemma shows that $\mathbb{N}_\ell(\widehat{\mathbf{G}}) = \mathbb{N}_r(\widehat{\mathbf{G}}) = \mathbb{N}(\widehat{\mathbf{G}})$ for regular or complete multipartite graph, so $\bigcap_{\lambda\in(0,1)} \mathscr{G}\left(1, \mathbf{W}(\lambda), \mathbf{H}^{-1}(\lambda)\right)$ sandwiched between $[\mathbb{N}_\ell(\widehat{\mathbf{G}}) \bigcap \mathbb{N}_r(\widehat{\mathbf{G}})] \oplus \{\mathbf{x}|\mathbf{x} \propto \mathbf{1}\}$ and $\mathbb{N}_r(\widehat{\mathbf{G}}) \oplus \{\mathbf{x}|\mathbf{x} \propto \mathbf{1}\}$ equals $\mathbb{N}(\mathbf{G}) \oplus \{\mathbf{x}\,|\mathbf{x} \propto \mathbf{1}\}$.

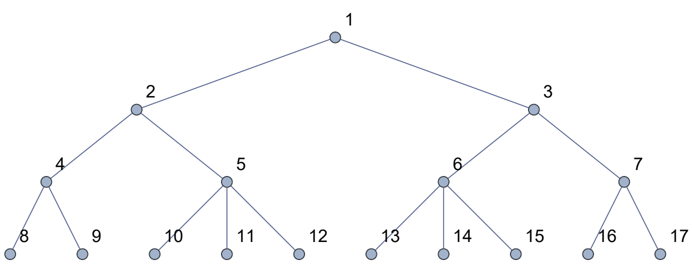


Figure 14. $\mathbb{N}_\ell(\widehat{\mathbf{G}}) \cap \mathbb{N}_r(\widehat{\mathbf{G}}) \subset \mathbb{N}_r(\widehat{\mathbf{G}})$

Lemma 2 *If $\boldsymbol{G} \in \mathscr{R} \cup \mathscr{K}_{p_1,\cdots,p_s}$, then*

$$\bigcap_{\lambda\in(0,1)} \mathscr{G}\left(1, \boldsymbol{W}(\lambda), \boldsymbol{H}^{-1}(\lambda)\right) = \mathbb{N}(\boldsymbol{G}) \oplus \{\boldsymbol{x}\,|\boldsymbol{x} \propto \boldsymbol{1}\}.$$

Proof. See Appendix L. ■

With Lemmas 1 and 2 in hand, we have the following result guaranteeing the first-best relative intervention.

[12]Seven linear dependent vectors $\{e_8-e_9, e_{10}-e_{11}, e_{10}-e_{12}, e_{13}-e_{14}, e_{13}-e_{15}, e_{16}-e_{17}, e_2+e_{13}+e_{17}-e_3-e_8-e_{10}\}$ constitutes a basis of $\mathbb{N}_r(\widehat{\mathbf{G}})$, among which only the six vectors associated with the pendant (terminal) vertexes, $\{e_8-e_9, e_{10}-e_{11}, e_{10}-e_{12}, e_{13}-e_{14}, e_{13}-e_{15}, e_{16}-e_{17}\}$, constitutes a basis of $\mathbb{N}_r(\widehat{\mathbf{G}}) \cap \mathbb{N}_\ell(\widehat{\mathbf{G}})$.

Proposition 9 *If $\boldsymbol{G} \in \mathscr{R} \cup \mathscr{K}_{p_1,\cdots,p_s}$, and*

$$C \geqslant \sqrt{(n-1)\mathbb{V}ar(\boldsymbol{\alpha}^0) - \left\|\mathrm{Proj}_{\mathscr{S}(0,\boldsymbol{G})}\boldsymbol{\alpha}^0\right\|^2},$$

then $R^*(C,\lambda) = 1$ is attained at $\boldsymbol{\alpha}^r(C,\lambda) \in \mathbb{N}(\boldsymbol{G}) \oplus \{\boldsymbol{x}|\boldsymbol{x} \propto \boldsymbol{1}\}$, $\forall\lambda \in (0,1)$.

Proof. See Appendix M. ■

4 Structural intervention

In this section, we turn to another form of network intervention: the structural intervention. In the characteristic intervention problems, the planner can only adjust the individual preferences taking the network structure as given. Now, we assume that the planner can also restructure endogenously the network. The planner does not need to pay any physical cost when building a network. But he has to consider the individuals' attitudes towards it. That is, whether a network provided by the planner can be agreed upon unanimously by all players. A problem of interest is the compatibility between social optimality and individual choices. We first give a negative result in an extremely conformist society. It shows that except in the trivial case of homogenous preferences, the socially optimal network provided by the planner will inevitably be boycotted by some individuals. Let $\mathscr{G}(I)$ denote the class of networks on players set I, $\mathscr{R}(I)$ is the subclass of regular networks on I, $\mathscr{O}(\lambda,\boldsymbol{\alpha},I) \equiv \arg\max_{\mathbf{g}\in\mathscr{G}(I)} \mathbb{W}^N(\lambda,\boldsymbol{\alpha},\mathbf{g})$ denotes the set of socially optimal networks on I given $\boldsymbol{\alpha}$ and λ, $\mathscr{O}_i(\lambda,\boldsymbol{\alpha},I) \equiv \arg\max_{\mathbf{g}\in\mathscr{G}(I)} u_i^N(\lambda,\boldsymbol{\alpha},\mathbf{g})$ is player i's most preferred network class.

Proposition 10 *(i) If $\boldsymbol{\alpha} \propto \boldsymbol{1}$, $\mathscr{O}_i(1,\boldsymbol{\alpha},I) = \mathscr{O}(1,\boldsymbol{\alpha},I) = \mathscr{G}(I), \forall i \in I$; (ii) if $\boldsymbol{\alpha} \not\propto \boldsymbol{1}$, $\mathscr{R}(I) \subseteq \mathscr{O}(1,\boldsymbol{\alpha},I)$, and $\mathscr{O}(1,\boldsymbol{\alpha},I) \cap \mathscr{O}_i(1,\boldsymbol{\alpha},I) = \emptyset, \forall i \notin \bar{I}(\boldsymbol{\alpha}) \equiv \{i \in I : \alpha_i = \bar{\alpha}\}$; $\mathscr{O}(1,\boldsymbol{\alpha},I) = \mathscr{O}_i(1,\boldsymbol{\alpha},I), \forall i \in \bar{I}(\boldsymbol{\alpha})$.*

Proof. See Appendix N. ■

This proposition shows that if all players have homogenous preferences, both the planner and individuals feel indifferent among all possible networks. In the case of heterogenous preferences, however, an utilitarian planner will choose a network, such that the degree-weighted average of individual characteristics equals their simple average value, as the socially optimal one. The median players ($i \in \bar{I}(\boldsymbol{\alpha})$), if any, will agree upon this choice. However, the non-median players will unanimously boycott it, because anyone of them prefers to residing at the center of a star network.

Given the above negative result, we next examine the impacts of marginal alterations. Let $\mathbf{G}^{[+ij]} \equiv \mathbf{G} + e_i e_j^\top + e_j e_i^\top$ represent the extended network when a new edge (i, j) is added to $\mathbf{G}$, $\Delta u_i^{[+i'j']}(\lambda, \boldsymbol{\alpha}, \mathbf{G}) \equiv u_i^N(\lambda, \boldsymbol{\alpha}, \mathbf{G}^{[+i'j']}) - u_i^N(\lambda, \boldsymbol{\alpha}, \mathbf{G})$ is the incremental utility of agent i upon new edge (i', j'). Let

$$\mathcal{P}(\lambda, \boldsymbol{\alpha}, \mathbf{G}) \equiv \left\{(i', j') \in \mathcal{E}^c : \Delta u_i^{[+i'j']}(\lambda, \boldsymbol{\alpha}, \mathbf{G}) \geqslant 0, \forall i, \ \textit{with at least one strict inequality}\right\}$$

denote the set of new edges upon which the original network is Pareto improved, where $\mathcal{E}^c$ denotes the the set of the null edges (unlinked pairs). The following proposition states that in a totally conformist society the marginal extension of $\mathbf{G}$ will inevitably hurt some players, and thus cannot receive unanimous approval.

Proposition 11 *$\mathcal{P}(1, \boldsymbol{\alpha}, \boldsymbol{G}) = \emptyset, \forall \boldsymbol{G} \in \mathscr{G}(I)$.*

Proof. See Appendix O. ■

We next turn to the opposite extreme case with almost individualism agents ($\lambda \approx 0$). We begin with a lemma.

Lemma 3 *If $\boldsymbol{\alpha} \not\propto \mathbf{1}$, λ is close to zero, and $dis\left[(i, j), k, \boldsymbol{G}\right] = s$, then*

$$\Delta u_k^{[+ij]}(\lambda, \boldsymbol{\alpha}, \boldsymbol{G}) = \begin{cases} (\alpha_k - \bar{\alpha}_k)\left(\frac{\hat{g}_{ki}^{[s]}(\alpha_j - \bar{\alpha}_i)}{1+d_i} + \frac{\hat{g}_{kj}^{[s]}(\alpha_i - \bar{\alpha}_j)}{1+d_j}\right)\lambda^{s+1} + o(\lambda^{s+1}) & if \quad s \geqslant 1 \\ \frac{\lambda(\alpha_j - \bar{\alpha}_i)}{2(d_i+1)}\left[2\alpha_i - \left(1 + \frac{d_i}{d_i+1}\right)\bar{\alpha}_i - \frac{\alpha_j}{d_i+1}\right] + o(\lambda) & if \quad k = i \\ \frac{\lambda(\alpha_i - \bar{\alpha}_j)}{2(d_j+1)}\left[2\alpha_j - \left(1 + \frac{d_j}{d_j+1}\right)\bar{\alpha}_j - \frac{\alpha_i}{d_j+1}\right] + o(\lambda) & if \quad k = j \end{cases},$$

where $dis\left[(i, j), k, \boldsymbol{G}\right] \equiv \min\{\rho_{ik}(\boldsymbol{G}), \rho_{jk}(\boldsymbol{G})\}$ is the geodesic distance from node k to edge (i, j), $\rho_{ij}(\boldsymbol{G}) \equiv \min\{s \in I : g_{ij}^{[s]} > 0\}$ is the distance between nodes i and j, $g_{ij}^{[s]} \equiv e_i^\top \boldsymbol{G}^s e_j$ denotes the (i, j) element of matrix $\boldsymbol{G}^s$, similarly, $\hat{g}_{ij}^{[s]} \equiv e_i^\top \widehat{\boldsymbol{G}}^s e_j$ is the (i, j) element of matrix $\widehat{\boldsymbol{G}}^s$.

Proof. See Appendix P. ■

This lemma shows that a newly added edge produces a ripple effect. It has the most prominent impact on the pair of nodes directly involved. The impact propagates outward with its magnitude decaying geometrically with the geodesic distance. The impact on an $s-$step node is of order λ^{s+1}. We now weaken the Pareto-improvement set $\mathcal{P}(\lambda, \boldsymbol{\alpha}, \mathbf{G})$ by allowing a loss tolerance of order one when players vote for a new link:

$$\mathcal{P}^\epsilon(\lambda, \boldsymbol{\alpha}, \mathbf{G}) \equiv \left\{(i', j') \in \mathcal{E}^c : \Delta u_i^{[+i'j']}(\lambda, \boldsymbol{\alpha}, \mathbf{G}) \geqslant -\epsilon\lambda, \forall i \in I\right\}.$$

We assume that a player will give his consent to the new link if his loss is within the range of tolerance. Then a positive result on unanimous approval is obtained.

Proposition 12 *For any pair $(i,j) \in \mathcal{E}^c$ and $\forall \epsilon > 0$, if conditions*

$$(\alpha_j - \bar{\alpha}_i)\left[2\alpha_i - \left(1 + \frac{d_i}{d_i+1}\right)\bar{\alpha}_i - \frac{\alpha_j}{d_i+1}\right] \geqslant 0 \tag{25}$$

and

$$(\alpha_i - \bar{\alpha}_j)\left[2\alpha_j - \left(1 + \frac{d_j}{d_j+1}\right)\bar{\alpha}_j - \frac{\alpha_i}{d_j+1}\right] \geqslant 0 \tag{26}$$

hold, then there exists a small λ such that $(i,j) \in \mathcal{P}^{\epsilon}(\lambda, \boldsymbol{\alpha}, \boldsymbol{G})$.

Proof. See Appendix Q. ■

The case of extreme individualism society is in sharp contrast to the totally conformist case discussed in Proposition 11. When $\lambda \approx 1$, a new link has substantial impacts on all individuals across the network, regardless of their locations. So, a new link needs to receive unanimous approval from all players. Things are simpler for the $\lambda \approx 0$ case. The distant nodes receive much less impact than the close ones. When a small loss tolerance is allowed, a new link will receive consent from all remote players, one only needs to consider the attitudes of the pair of directly involved nodes. The unanimous approval conditions are therefore much easier to satisfy than in the totally conformist case. Obviously, conditions (25) and (26) admits a large degree of freedom since the variables involved outnumbered largely the inequalities.[13] So for any $(i,j) \in \mathcal{E}^c$, it is easy to find a vector $\boldsymbol{\alpha}$ guaranteeing unanimous approval.

As shown in Proposition 11, the planner's goal of welfare optimality is generally incompatible with the agents' individual rationalities in an extremely conformist society. We now allow the planner to neglect the attitudes of the agents. That is, the planner are given a mandate to build a network, and no individual has the power to veto his decision. From Proposition 10, we see that in a totally conformist society, a utilitarian planner can restrict his attention to the class of regular networks to maximize the social welfare. Nevertheless, the regular class $\mathscr{R}(I)$ itself may still be very large for a large n. We proceed to give an algorithm to search the optimal network within the regular class. Expanding $\widehat{\mathbf{M}}(\lambda, \mathbf{G})$ around $\lambda = 1$ yields

$$\widehat{\mathbf{M}}(\lambda, \mathbf{G}) \quad = \quad \mathbf{1}\mathbf{1}^{\top}\widehat{\mathbb{D}} + \sum_{k=0}^{\infty} \frac{(-1)^k}{k!}\mathbb{U}_{-1}\widehat{\boldsymbol{\Lambda}}_{-1}^{(k)}(1)(\mathbb{U}_{-1})^{\top}\widehat{\mathbb{D}}(1-\lambda)^{k+1},$$

where $\widehat{\mathbb{D}}$, $\widehat{\boldsymbol{\Lambda}}_{-1}(\lambda)$ and $\mathbb{U}_{-1}$ are given in Appendix N,

$$\widehat{\boldsymbol{\Lambda}}_{-1}^{(k)}(1) \equiv diag\left\{\left[\left.\frac{1}{1-\lambda_i\lambda}\right|_{\lambda=1}^{(k)}\right]_{i=2}^{n}\right\} = diag\left\{\left[\frac{k!\lambda_i^k}{(1-\lambda_i)^{k+1}}\right]_{i=2}^{n}\right\}$$

[13] The variables contained in (25) and (26) are α_i, α_j, $(\alpha_k)_{k\in\mathcal{N}_i\cup\mathcal{N}_j}$, $\mathcal{N}_s \equiv \{k : g_{ks} = 1\}$ denotes the neighbourhood of vertex s. So the degree of freedom of (25) and (26) is $2 + \#(\mathcal{N}_i \cup \mathcal{N}_i) - 2 = d_i + d_j - \#(\mathcal{N}_i \cap \mathcal{N}_j)$.

denotes the $k-$th order derivative of $\widehat{\mathbf{\Lambda}}_{-1}(\lambda)$ with respect to λ, evaluated at 1. Then,

$$\widehat{\mathbf{M}}(\lambda,\mathbf{G}) = \mathbf{I} - \mathbf{B}_0 + \sum_{k=1}^{\infty} \mathbf{B}_k (1-\lambda)^k,$$

where

$$\mathbf{B}_k \equiv \begin{cases} \mathbf{I} - \mathbf{1}\mathbf{1}^\top \widehat{\mathbb{D}} & k=0 \\ \mathbb{U}_{-1} diag\left\{ \left(\frac{(-\lambda_i)^{k-1}}{(1-\lambda_i)^k}\right)\Big|_{i=2}^{n} \right\} (\mathbb{U}_{-1})^\top \widehat{\mathbb{D}} & k \geqslant 1 \end{cases}.$$

Expanding $\mathbf{W}^N(\lambda,\mathbf{G})$ with respect to λ around 1 yields:

$$\mathbf{W}^N(\lambda,\mathbf{G}) = \mathbf{I} - \frac{1}{\lambda}(\mathbf{I}-\widehat{\mathbf{M}})^\top(\mathbf{I}-\widehat{\mathbf{M}}) = \sum_{k=0}^{\infty} \mathbf{\Omega}_k (1-\lambda)^k,$$

where

$$\mathbf{\Omega}_k = \begin{cases} \mathbf{I} - \mathbf{B}_0^\top \mathbf{B}_0 & k=0 \\ -\mathbf{B}_0^\top \mathbf{B}_0 + \mathbf{B}_0^\top \mathbf{B}_1 + \mathbf{B}_1^\top \mathbf{B}_0 & k=1 \\ -\mathbf{B}_0^\top \mathbf{B}_0 + \mathbf{B}_0^\top \left(\sum_{s=1}^{k} \mathbf{B}_s\right) + \left(\sum_{s=1}^{k} \mathbf{B}_s\right)^\top \mathbf{B}_0 - \sum_{s=1}^{k-1}\sum_{t=1}^{k-s} \mathbf{B}_s^\top \mathbf{B}_t & k \geqslant 2 \end{cases}.$$

The associated social welfare is

$$\mathbb{W}^N(\lambda,\boldsymbol{\alpha},\mathbf{G}) = \frac{1}{2}\sum_{k=0}^{\infty} \left(\boldsymbol{\alpha}^\top \mathbf{\Omega}_k \boldsymbol{\alpha}\right)(1-\lambda)^k. \tag{27}$$

When λ is sufficiently close to one, an utilitarian planner has lexicographic preferences over terms in (27). That is, she first orders the constant term $\boldsymbol{\alpha}^\top \mathbf{\Omega}_0 \boldsymbol{\alpha} = \boldsymbol{\alpha}^\top \left(\mathbf{I} - \mathbf{B}_0^\top \mathbf{B}_0\right)\boldsymbol{\alpha}$, then $\boldsymbol{\alpha}^\top \mathbf{\Omega}_1 \boldsymbol{\alpha}$, then $\boldsymbol{\alpha}^\top \mathbf{\Omega}_2 \boldsymbol{\alpha}$, and so on.

The problem

$$[\mathscr{P}^I]: \max_{\boldsymbol{\alpha}\in\mathcal{O}(\boldsymbol{\alpha}^0,C),\mathbf{G}\in\mathscr{G}(I)} \mathbb{W}^N(\lambda,\boldsymbol{\alpha},\mathbf{G})$$

encompasses both characteristic and structural interventions. It can be solved approximately by an inductive procedure.

- *Round* 0. Solve problem

  $$[\mathscr{P}_1^I]: \max_{\boldsymbol{\alpha}\in\mathcal{O}(\boldsymbol{\alpha}^0,C),\mathbf{G}\in\mathscr{G}(I)} \boldsymbol{\alpha}^\top \left(\mathbf{I} - \mathbf{B}_0^\top \mathbf{B}_0\right)\boldsymbol{\alpha} = \max_{\boldsymbol{\alpha}\in\mathcal{O}(\boldsymbol{\alpha}^0,C)} \max_{\mathbf{G}\in\mathscr{G}(I)} n\left[\bar{\alpha}^2 - (\bar{\alpha}-\bar{\alpha}_d)^2\right],$$

  where $\bar{\alpha}, \bar{\alpha}_d$ are simple average and degree-weighted average of $\boldsymbol{\alpha}$. To guarantee $\bar{\alpha} = \bar{\alpha}_d$, the $0-$round optimal network is chosen among the regular class, $\mathcal{G}_0(I) = \mathscr{R}(I)$, and the optimal characteristic vector is the solution to $\max_{\boldsymbol{\alpha}\in\mathcal{O}(\boldsymbol{\alpha}^0,C)} n\bar{\alpha}^2 = \max_{\boldsymbol{\alpha}\in\mathcal{O}(\boldsymbol{\alpha}^0,C)} \frac{1}{n}\boldsymbol{\alpha}^\top \mathbf{1}\mathbf{1}^\top \boldsymbol{\alpha}$. A Lagrangian function with multiplier ω can be written as

  $$\mathscr{L}(\omega,\boldsymbol{\alpha}) = \frac{1}{n}\boldsymbol{\alpha}^\top \mathbf{1}\mathbf{1}^\top \boldsymbol{\alpha} + \omega\left[C^2 - (\boldsymbol{\alpha}-\boldsymbol{\alpha}^0)^\top(\boldsymbol{\alpha}-\boldsymbol{\alpha}^0)\right].$$

The first-order condition is

$$\frac{1}{n}\mathbf{1}\mathbf{1}^\top \boldsymbol{\alpha} = \omega(\boldsymbol{\alpha} - \boldsymbol{\alpha}^0). \tag{28}$$

Inserting (28) into $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^0\| = C$ yields $\omega = \frac{\sqrt{n}|\bar{\alpha}|}{C}$. Therefore $\boldsymbol{\alpha}^* = \boldsymbol{\alpha}^0 + \frac{C\bar{\alpha}}{|\bar{\alpha}|\sqrt{n}}\mathbf{1}$.

- *Round* $k, k \geqslant 1$. For $\mathbf{G} \in \mathscr{R}(I)$, $\widehat{\mathbb{D}} = \frac{1}{n}\mathbf{I}$, then we have

$$\mathbf{B}_k \equiv \begin{cases} \mathbf{Q} & k = 0 \\ \frac{1}{n}\mathbb{U}_{-1} diag\left\{ \left(\frac{(-\lambda_i)^{k-1}}{(1-\lambda_i)^k} \right)\Big|_{i=2}^{n} \right\} (\mathbb{U}_{-1})^\top & k \geqslant 1 \end{cases}.$$

It follows that $\mathbf{B}_0^\top \mathbf{B}_t = \mathbf{B}_t, \forall t \geqslant 0$, $\boldsymbol{\Omega}_0 = -\mathbf{Q} = -\frac{1}{n}\mathbb{U}_{-1}(\mathbb{U}_{-1})^\top$,

$$\boldsymbol{\Omega}_1 = 2\mathbf{B}_0^\top \mathbf{B}_1 - \mathbf{B}_0^\top \mathbf{B}_0 = 2\mathbf{B}_1 - \mathbf{Q} = \frac{1}{n}\mathbb{U}_{-1} diag\left\{\frac{1+\lambda_i}{1-\lambda_i}\right\}_{n=2}^{n} (\mathbb{U}_{-1})^\top,$$

and

$$\begin{aligned} \boldsymbol{\Omega}_k &= -\mathbf{B}_0^\top \mathbf{B}_0 + 2\mathbf{B}_0^\top \sum_{t=1}^{k} \mathbf{B}_t - \sum_{s=1}^{k-1}\sum_{t=1}^{k-s} \mathbf{B}_s^\top \mathbf{B}_t \\ &= \frac{1}{n}\mathbb{U}_{-1}\left\{ diag\left[-1 + 2\sum_{t=1}^{k} \frac{(-\lambda_i)^{t-1}}{(1-\lambda_i)^t} - \sum_{s=1}^{k-1}\sum_{t=1}^{k-s} \frac{(-\lambda_i)^{s+t-2}}{(1-\lambda_i)^{s+t}} \right]_{i=2}^{n} \right\} (\mathbb{U}_{-1})^\top \\ &= \frac{1}{n}\mathbb{U}_{-1} diag\left\{ \frac{(-\lambda_i)^{k-1}(k+\lambda_i)}{(1-\lambda_i)^k} \right\}_{i=2}^{n} (\mathbb{U}_{-1})^\top \forall k \geqslant 2. \end{aligned}$$

In summary, $\boldsymbol{\Omega}_k = \frac{1}{n}\mathbb{U}_{-1} diag\left\{\eta_k(\lambda_i)\right\}_{i=2}^{n} (\mathbb{U}_{-1})^\top$, where

$$\eta_k(x) \equiv \begin{cases} -1 & k = 0 \\ \frac{1+x}{1-x} & k = 1 \\ \frac{(-x)^{k-1}(k+x)}{(1-x)^k} & k \geqslant 2 \end{cases}.$$

Since $\mathbb{U}_{-1}\mathbf{1} = \mathbf{0}$ for $\mathbf{G} \in \mathscr{R}(I)$, $(\boldsymbol{\alpha}^*)^\top \boldsymbol{\Omega}_k(\boldsymbol{\alpha}^*) = \left(\boldsymbol{\alpha}^0 + \frac{C\bar{\alpha}}{|\bar{\alpha}|\sqrt{n}}\mathbf{1}\right)^\top \boldsymbol{\Omega}_k \left(\boldsymbol{\alpha}^0 + \frac{C\bar{\alpha}}{|\bar{\alpha}|\sqrt{n}}\mathbf{1}\right) = (\boldsymbol{\alpha}^0)^\top \boldsymbol{\Omega}_k(\boldsymbol{\alpha}^0)$. Let $\tilde{\lambda}_1 > \cdots > \tilde{\lambda}_p$ be distinct eigenvalues of $\widehat{\mathbf{G}}$, $I_i \equiv \{j : \lambda_j = \tilde{\lambda}_i\}$. Since $\{\mathbb{u}_{\cdot j}/\sqrt{n}\}_{j \in I_i}$ constitute a set of orthonormal base of eigenspace $\mathscr{S}(\tilde{\lambda}_i, \widehat{\mathbf{G}})$, we have

$$(\boldsymbol{\alpha}^*)^\top \boldsymbol{\Omega}_k(\boldsymbol{\alpha}^*) = \sum_{i=2}^{p} \eta_k(\tilde{\lambda}_i) \frac{1}{n} \sum_{j \in I_i} \left(\mathbb{u}_{\cdot j}^\top \boldsymbol{\alpha}^0\right)^2 = \sum_{i=1}^{p} \eta_k(\tilde{\lambda}_i) \left\| \mathrm{Proj}_{\mathscr{S}(\tilde{\lambda}_i, \widehat{\mathbf{G}})} \boldsymbol{\alpha}^0 \right\|^2, \forall k \geqslant 0.$$

We can thus write the round$-k$ intervention problem as

$$[\mathscr{P}_k^I] : \max_{\mathbf{G} \in \mathcal{G}_{k-1}(I)} (\boldsymbol{\alpha}^0)^\top \boldsymbol{\Omega}_{k-1}(\boldsymbol{\alpha}^0).$$

The optimal network class is then $\mathcal{G}_k(I) = \arg\max_{\mathbf{G} \in \mathcal{G}_{k-1}(I)} (\boldsymbol{\alpha}^0)^\top \boldsymbol{\Omega}_{k-1}(\boldsymbol{\alpha}^0)$.

The following proposition summarizes the above analysis.

$\mathbf{G}$	$\mathbf{G}_1$	$\mathbf{G}_2$	$\mathbf{G}_3$	$\mathbf{G}_4$	$\mathbf{G}_5$
$\delta_1(\mathbf{G})$	257.778	90.667	30.733	91.277	63.222

Table 1. $\delta_1(\mathbf{G})$ within class $\mathscr{R}(I)$

Proposition 13 *In an extremely conformist society i.e., $\lambda \approx 1$, the solution $(\boldsymbol{\alpha}^*, \boldsymbol{G}^*)$ to problem $[\mathscr{P}^I]$ entails:*

- *a characteristic intervention satisfying $\boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0 \propto \boldsymbol{1}$;*
- *the optimal network $\boldsymbol{G}^*$ found inductively within network class*

$$\mathcal{G}_k(I) = \arg \max_{\boldsymbol{G} \in \mathcal{G}_{k-1}(I)} \delta_k(\boldsymbol{G}),$$

with the initial class $\mathcal{G}_0(I) = \mathscr{R}(I)$, and

$$\delta_k(\boldsymbol{G}) \equiv \sum_{i=2}^{p} \eta_k(\tilde{\lambda}_i) \left\| \mathrm{Proj}_{\mathscr{S}(\tilde{\lambda}_i, \widehat{\boldsymbol{G}})} \boldsymbol{\alpha}^0 \right\|^2,$$

$1 = \tilde{\lambda}_1 > \cdots > \tilde{\lambda}_p$ *are distinct eigenvalues of $\widehat{\boldsymbol{G}}$,*

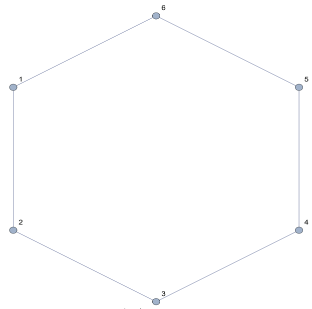
$$\eta_k(x) \equiv \begin{cases} \frac{1+x}{1-x} & k = 1 \\ \frac{(-x)^{k-1}(k+x)}{(1-x)^k} & k \geqslant 2 \end{cases}.$$

An example. To illustrate this algorithm, let us consider a numerical example where $n = 6$ and $\boldsymbol{\alpha}^0 = (2, 3, 4, -5, -6, -3)^\top$. The 0−round optimization gives $\boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0 \propto \mathbf{1}$, and $\mathcal{G}_0(I) = \mathscr{R}(I)$, as shown in Figure 15. In round−1, we proceed to calculate index $\delta_1(\mathbf{G})$ within class $\mathscr{R}(I)$ (see Table-1). It is obvious that the socially optimal network is $\mathbf{G}^* = \mathbf{G}_1$.

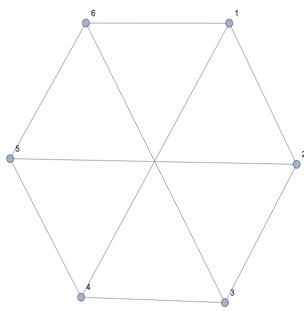
5 Conclusion

This paper contributes to the growing literature on network interventions, in which a planner can adjust purposely the individuals' characteristics and/or the structure of a network. My analysis is based on a local-average model, in which each agent seeks to minimize the social distance between his own action and the average action of his peers (his social norm).

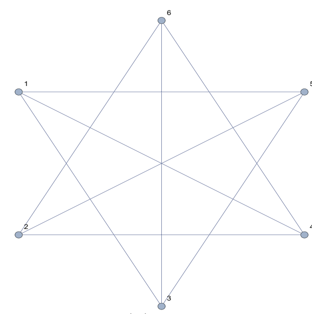
I first explore the characteristic intervention problems of a planner with various objectives and constraints. She optimizes the social welfare subject to a fixed level of social variance, or in duality, optimizes the social variance subject to a fixed level of social welfare. This paper gives the optimal interventions and the optimal values attained in these setups. I also consider the



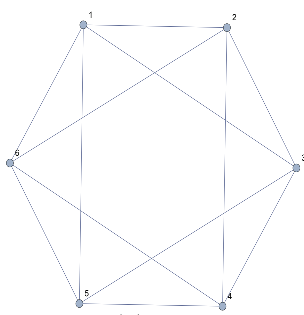
(a) $\mathbf{G}_1$



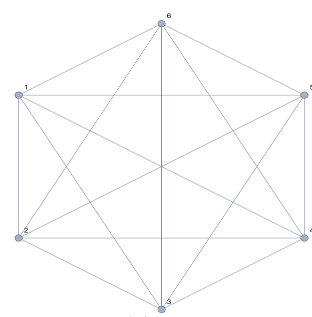
(b) $\mathbf{G}_2$



(c) $\mathbf{G}_3$



(d) $\mathbf{G}_4$



(e) $\mathbf{G}_5$

Figure 15. Regular class $\mathscr{R}(I)$

welfare-maximizing and variance-minimizing problems under quadratic budget constraints. It is shown that the total budget and the standalone preference vector have joint impacts on the optimal interventions. Moreover, for extremely large (resp. small) budgets, the approximated simple (resp. tangent) interventions achieve most of the optimal efficiencies. I further identify the order of errors, and thus the accuracies, of these approximations. I proceed to discuss the relative characteristic intervention problem, in which a planner seeks to maximize the ratio of equilibrium social welfare to its first-best counterpart subject to a quadratic budget constraint. I give conditions under which a 100% ratio is achieved.

I then discuss the joint intervention problem incorporating both characteristic and structural interventions. The alteration of network incurs no physical cost, but the attitudes of individuals matter for the planner's network design problem. I begin with a negative result stating that there exists no socially optimal and unanimously approved network if players have heterogeneous preferences. In what follows, I discuss the impact of adding a new link to an existing network. It is shown that in a totally conformist society, any new link will inevitably hurt someone, and thus fails to meet the unanimous approval condition. In the opposite case with extreme individualism, however, a more optimistic result is obtained when a weaker approval condition is adopted. I show that the impact on individual payoff decays geometrically as the geodesic distance from the new edge to the node affected expands. Therefore, if a one-order loss tolerance is allowed by every player, and a new link is not detrimental to the pair of players directly involved, then this link receives unanimous approval. Lastly, a searching algorithm that leads to the socially optimal network is provided.

Appendix A. Proof of Proposition 1

Let $\nu_i \equiv \lambda_i(\mathbf{V})$, $\mathbf{U}$ is an orthogonal matrix whose ith column $\mathbf{u}_{\cdot i}$ is the orthonormal eigenvector of $\mathbf{V}$ associated with ν_i. It follows immediately from $\widehat{\mathbf{M}}\mathbf{1} = \mathbf{1}$ and $\mathbf{Q}\mathbf{1} = \mathbf{0}$ that the least eigenvalue is $\nu_n = 0$, and $\mathbf{1}$ is its associated eigenvector. Since $\mathbf{G}$ is irreducible, by the Perron-Frobenius theorem, $\mathbf{1}$ is the only eigenvector (up to a scalar multiple) corresponding to $\nu_n = 0$, i.e., $\mathbf{u}_{\cdot n} = \pm\frac{1}{\sqrt{n}}\mathbf{1}$. Let $\widehat{\mathbf{V}} \equiv diag\{\nu_1, \cdots, \nu_{n-1}, 1\}$, $\mathbf{y} \equiv \widehat{\mathbf{V}}^{1/2}\mathbf{U}^\top\boldsymbol{\alpha}$. Problems $[\mathcal{P}^{w|v}_{max}/\mathcal{P}^{w|v}_{min}]$ are then transformed into

$$\min_{\mathbf{y}\in\mathbb{R}^n} / \max_{\mathbf{y}\in\mathbb{R}^n} \mathbf{y}^\top\boldsymbol{\Sigma}\mathbf{y}, s.t. : \mathbf{y}_{-n}^\top\mathbf{y}_{-n} = 1, \tag{29}$$

where $\boldsymbol{\Sigma} \equiv [\sigma_{ij}]_{n\times n} \equiv \widehat{\mathbf{V}}^{-\frac{1}{2}}\mathbf{U}^\top\mathbf{W}\mathbf{U}\widehat{\mathbf{V}}^{-\frac{1}{2}}$, $\mathbf{y}_{-n} \equiv (y_i)_{i\neq n}$ is a subvector obtained by deleting the last entry y_n from $\mathbf{y}$. Since $\mathbf{W}\mathbf{1} = \mathbf{1}$, we have $\sigma_{j,n} \equiv e_j^\top\boldsymbol{\Sigma}e_n = \pm\frac{1}{\sqrt{n\nu_j}}\mathbf{u}_{\cdot j}^\top\mathbf{1} = 0, \forall j \neq n$, and $\sigma_{n,n} \equiv e_n^\top\boldsymbol{\Sigma}e_n = e_n^\top\mathbf{U}^\top\mathbf{W}\mathbf{U}e_n = 1$, where e_i is the ith standard base vector with only its ith entry equal to 1 and the rest 0s. Problems $[\mathcal{P}^{w|v}_{max}/\mathcal{P}^{w|v}_{min}]$ are then rewritten as

$$\max_{\mathbf{y}\in\mathbb{R}^n} / \min \mathbf{y}_{-n}^\top\boldsymbol{\Sigma}_{-n,-n}\mathbf{y}_{-n} + y_n^2, s.t. : \mathbf{y}_{-n}^\top\mathbf{y}_{-n} = 1, \tag{30}$$

where

$$\boldsymbol{\Sigma}_{-n,-n} \equiv (\sigma_{ij})_{i,j\neq n} = \left(\frac{\mathbf{u}_{\cdot i}^\top\mathbf{W}\mathbf{u}_{\cdot j}}{\sqrt{\nu_i\nu_j}}\right)_{i,j\neq n} = (\widehat{\mathbf{V}}_{-n})^{-\frac{1}{2}}(\mathbf{U}_{-n})^\top\mathbf{W}\mathbf{U}_{-n}(\widehat{\mathbf{V}}_{-n})^{-\frac{1}{2}}$$

denotes the principal submatrix obtained by deleting the last row and column from $\boldsymbol{\Sigma}$, $\widehat{\mathbf{V}}_{-n} \equiv diag\{\nu_1, \cdots, \nu_{n-1}\}$, $\mathbf{U}_{-n} \equiv (\mathbf{u}_{\cdot j})_{j\neq n}$.

The max-problem $[\mathcal{P}^{w|v}_{max}]$. Since y_n is not included in constraint $\mathbf{y}_{-n}^\top\mathbf{y}_{-n} = 1$, we can thus obtain the maximum value $\Pi^{w|v}_{max} = \infty$ by choosing $y_n^* = \pm\infty$ and $\mathbf{y}^*_{-n} \in \mathscr{S}(\lambda_1(\boldsymbol{\Sigma}_{-n,-n}), \boldsymbol{\Sigma}_{-n,-n})$, where $\mathscr{S}(\lambda, \mathbf{A}) \equiv \{\mathbf{x}|\mathbf{A}\mathbf{x} = \lambda\mathbf{x}\}$ represents the $\lambda-$eigenspace of matrix $\mathbf{A}$. Substituting $\mathbf{y}^*_{-n} = (\widehat{\mathbf{V}}_{-n})^{\frac{1}{2}}(\mathbf{U}_{-n})^\top\boldsymbol{\alpha}^*$, $y_n^* = \pm\frac{1}{\sqrt{n}}\mathbf{1}^\top\boldsymbol{\alpha}^*$, and $\boldsymbol{\Sigma}_{-n,-n} \equiv (\widehat{\mathbf{V}}_{-n})^{-\frac{1}{2}}(\mathbf{U}_{-n})^\top\mathbf{W}\mathbf{U}_{-n}(\widehat{\mathbf{V}}_{-n})^{-\frac{1}{2}}$ into $\boldsymbol{\Sigma}_{-n,-n}\mathbf{y}^*_{-n} = \lambda_1(\boldsymbol{\Sigma}_{-n,-n})\mathbf{y}^*_{-n}$, we get

$$(\widehat{\mathbf{V}}_{-n})^{-\frac{1}{2}}(\mathbf{U}_{-n})^\top\mathbf{W}\mathbf{U}_{-n}(\widehat{\mathbf{V}}_{-n})^{-\frac{1}{2}}(\widehat{\mathbf{V}}_{-n})^{\frac{1}{2}}(\mathbf{U}_{-n})^\top\boldsymbol{\alpha}^* = \lambda_1(\boldsymbol{\Sigma}_{-n,-n})(\widehat{\mathbf{V}}_{-n})^{\frac{1}{2}}(\mathbf{U}_{-n})^\top\boldsymbol{\alpha}^*. \tag{31}$$

Since $\mathbf{U}_{-n}(\mathbf{U}_{-n})^\top = \mathbf{U}\mathbf{U}^\top - \mathbf{u}_{\cdot n}\mathbf{u}_{\cdot n}^\top = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^\top \equiv \mathbf{Q}$, $\mathbf{1}^\top\mathbf{W} = \mathbf{1}^\top$, and $\mathbf{V}\mathbf{Q} = \mathbf{V}$, (31) implies

$$\begin{bmatrix} (\mathbf{U}_{-n})^\top \\ \frac{1}{\sqrt{n}}\mathbf{1}^\top \end{bmatrix} \mathbf{W}\mathbf{Q}\boldsymbol{\alpha}^* = \lambda_1(\boldsymbol{\Sigma}_{-n,-n}) \begin{bmatrix} \widehat{\mathbf{V}}_{-n} & \\ & 0 \end{bmatrix} \begin{bmatrix} (\mathbf{U}_{-n})^\top \\ \frac{1}{\sqrt{n}}\mathbf{1}^\top \end{bmatrix} \boldsymbol{\alpha}^*.$$

It follows immediately that

$$\mathbf{W}\mathbf{Q}\boldsymbol{\alpha}^* = \lambda_1(\boldsymbol{\Sigma}_{-n,-n})\mathbf{U} \begin{bmatrix} \widehat{\mathbf{V}}_{-n} & \\ & 0 \end{bmatrix} \mathbf{U}^\top\boldsymbol{\alpha}^* = \lambda_1(\boldsymbol{\Sigma}_{-n,-n})\mathbf{V}\boldsymbol{\alpha}^* = \lambda_1(\boldsymbol{\Sigma}_{-n,-n})\mathbf{V}\mathbf{Q}\boldsymbol{\alpha}^*. \tag{32}$$

Therefore, the optimal intervention satisfies:

$$\mathrm{Proj}_{\mathbf{1}^\perp}\boldsymbol{\alpha}^{w|v}_{max} \in \mathscr{G}(\lambda_1(\mathbf{\Sigma}_{-n,-n}),\mathbf{W},\mathbf{V}), \left\|\mathrm{Proj}_{\mathbf{1}}\boldsymbol{\alpha}^{w|v}_{max}\right\| = \infty,$$

where $\mathbf{1}^\perp$ denotes the orthogonal complement space of $\mathbf{1}$.

The min-problem $[\mathcal{P}^{w|v}_{min}]$. The minimum value $\Pi^{w|v}_{min} = \lambda_{n-1}(\mathbf{\Sigma}_{-n,-n})$ is achieved by choosing $y^*_n = 0$ and $\mathbf{y}^*_{-n} \in \mathscr{S}(\lambda_{n-1}(\mathbf{\Sigma}_{-n,-n}),\mathbf{\Sigma}_{-n,-n})$. Substituting $\mathbf{y}^*_{-n} = (\widehat{\mathbf{V}}_{-n})^{\frac{1}{2}}(\mathbf{U}_{-n})^\top\boldsymbol{\alpha}^*$, $y^*_n = \pm\frac{1}{\sqrt{n}}\mathbf{1}^\top\boldsymbol{\alpha}^* = 0$, and $\mathbf{\Sigma}_{-n,-n} \equiv (\widehat{\mathbf{V}}_{-n})^{-\frac{1}{2}}(\mathbf{U}_{-n})^\top\mathbf{W}\mathbf{U}_{-n}(\widehat{\mathbf{V}}_{-n})^{-\frac{1}{2}}$ into $\mathbf{\Sigma}_{-n,-n}\mathbf{y}^*_{-n} = \lambda_{n-1}(\mathbf{\Sigma}_{-n,-n})\mathbf{y}^*_{-n}$, we get

$$(\widehat{\mathbf{V}}_{-n})^{-\frac{1}{2}}(\mathbf{U}_{-n})^\top\mathbf{W}\mathbf{U}_{-n}(\widehat{\mathbf{V}}_{-n})^{-\frac{1}{2}}(\widehat{\mathbf{V}}_{-n})^{\frac{1}{2}}(\mathbf{U}_{-n})^\top\boldsymbol{\alpha}^* = \lambda_{n-1}(\mathbf{\Sigma}_{-n,-n})(\widehat{\mathbf{V}}_{-n})^{\frac{1}{2}}(\mathbf{U}_{-n})^\top\boldsymbol{\alpha}^*. \quad (33)$$

Since $\mathbf{U}_{-n}(\mathbf{U}_{-n})^\top = \mathbf{Q}$, $\mathbf{1}^\top\boldsymbol{\alpha}^* = 0$ and $\mathbf{1}^\top\mathbf{W} = \mathbf{1}^\top$, (33) implies

$$\begin{bmatrix} (\mathbf{U}_{-n})^\top \\ \frac{1}{\sqrt{n}}\mathbf{1}^\top \end{bmatrix}\mathbf{W}\boldsymbol{\alpha}^* = \lambda_{n-1}(\mathbf{\Sigma}_{-n,-n})\begin{bmatrix} \widehat{\mathbf{V}}_{-n} & \\ & 0 \end{bmatrix}\begin{bmatrix} (\mathbf{U}_{-n})^\top \\ \frac{1}{\sqrt{n}}\mathbf{1}^\top \end{bmatrix}\boldsymbol{\alpha}^*.$$

It follows immediately that

$$\mathbf{W}\boldsymbol{\alpha}^* = \lambda_{n-1}(\mathbf{\Sigma}_{-n,-n})\mathbf{U}\begin{bmatrix} \widehat{\mathbf{V}}_{-n} & \\ & 0 \end{bmatrix}\mathbf{U}^\top\boldsymbol{\alpha}^* = \lambda_{n-1}(\mathbf{\Sigma}_{-n,-n})\mathbf{V}\boldsymbol{\alpha}^*.$$

Therefore, $\boldsymbol{\alpha}^* \in \mathscr{G}(\lambda_{n-1}(\mathbf{\Sigma}_{-n,-n}),\mathbf{W},\mathbf{V})$.

We proceed to give bounds for $\lambda_{n-1}(\mathbf{\Sigma}_{-n,-n})$. From

$$\lambda_{n-1}((\mathbf{U}_{-n})^\top\mathbf{W}\mathbf{U}_{-n})\mathbf{I}_{n-1} \le (\mathbf{U}_{-n})^\top\mathbf{W}\mathbf{U}_{-n} \le \lambda_1((\mathbf{U}_{-n})^\top\mathbf{W}\mathbf{U}_{-n})\mathbf{I}_{n-1},$$

[14] we have

$$\begin{aligned}\lambda_{n-1}\left((\mathbf{U}_{-n})^\top\mathbf{W}\mathbf{U}_{-n}\right)(\widehat{\mathbf{V}}_{-n})^{-1} &\leqslant \mathbf{\Sigma}_{-n,-n} \equiv (\widehat{\mathbf{V}}_{-n})^{-1/2}(\mathbf{U}_{-n})^\top\mathbf{W}\mathbf{U}_{-n}(\widehat{\mathbf{V}}_{-n})^{-1/2} \\ &\leqslant \lambda_1\left((\mathbf{U}_{-n})^\top\mathbf{W}\mathbf{U}_{-n}\right)(\widehat{\mathbf{V}}_{-n})^{-1}.\end{aligned}$$

It follows that[15]

$$\begin{aligned}\underline{\lambda} &\equiv \lambda_{n-1}\left(\lambda_{n-1}((\mathbf{U}_{-n})^\top\mathbf{W}\mathbf{U}_{-n})(\widehat{\mathbf{V}}_{-n})^{-1}\right) \leqslant \lambda_{n-1}\left(\mathbf{\Sigma}_{-n,-n}\right) \\ &\leqslant \lambda_{n-1}\left(\lambda_1((\mathbf{U}_{-n})^\top\mathbf{W}\mathbf{U}_{-n})(\widehat{\mathbf{V}}_{-n})^{-1}\right) \equiv \overline{\lambda}.\end{aligned}$$

Since $(\mathbf{U}_{-n})\mathbf{U}^\top_{-n} = \mathbf{Q}$, and $\mathbf{Q}$ is idempotent, i.e., $\mathbf{Q}^2 = \mathbf{Q}$, we have

$$\mathbb{\lambda}\left((\mathbf{U}_{-n})^\top\mathbf{W}\mathbf{U}_{-n}\right) = \overline{\mathbb{\lambda}}\left(\mathbf{W}\mathbf{U}_{-n}(\mathbf{U}_{-n})^\top\right) = \overline{\mathbb{\lambda}}(\mathbf{W}\mathbf{Q}) = \overline{\mathbb{\lambda}}(\mathbf{Q}\mathbf{W}\mathbf{Q}),$$

[14] Here $\mathbf{A} \ge \mathbf{B}$ iff $\mathbf{A} - \mathbf{B}$ is nonnegative definite.

[15] If $\mathbf{A} \ge \mathbf{B}$, then the Weyl inequality implies $\lambda_i(\mathbf{A}) = \lambda_i(\mathbf{A} - \mathbf{B} + \mathbf{B}) \geqslant \lambda_i(\mathbf{B}) + \lambda_n(\mathbf{A} - \mathbf{B}) \geqslant \lambda_i(\mathbf{B}), \forall i$.

where $\mathbb{\lambda}(\cdot)$ denotes the spectrum (the set of eigenvalues) of a matrix, $\bar{\mathbb{\lambda}}(\cdot)$ denotes the set of non-zero eigenvalues. Let $w_i \equiv \lambda_i(\mathbf{W})$ be the ith largest eigenvalue of matrix $\mathbf{W}$, and $\mathbf{p}_{\cdot i}$ be the orthonormal eigenvector associated with w_i, let $\mathbf{P} \equiv (\mathbf{p}_{\cdot i})_{i\in I}$. Since $\mathbf{G}$ is irreducible, it is easy to find that $w_1 = 1$ has multiplicity one, and $\mathbf{p}_{\cdot 1} = \frac{1}{\sqrt{n}}\mathbf{1}$ is unique (only up to a sign). We have

$$\mathbf{p}_{\cdot i}^{\top}\mathbf{QWQp}_{\cdot j} = \begin{cases} 0 & \text{if} \quad i \neq j \\ 0 & \text{if} \quad i = j = 1 \\ w_i & \text{if} \quad i = j \neq 1 \end{cases},$$

then $\mathbf{P}^{\top}\mathbf{QWQP} = diag\{0, w_2, \cdots, w_n\}$, so $\mathbb{\lambda}((\mathbf{U}_{-n})^{\top}\mathbf{WU}_{-n}) = \bar{\mathbb{\lambda}}(\mathbf{P}^{\top}\mathbf{QWQP}) = \{w_2, \cdots, w_n\}$. Therefore, $\lambda_1((\mathbf{U}_{-n})^{\top}\mathbf{WU}_{-n}) = \lambda_2(\mathbf{W}), \lambda_{n-1}((\mathbf{U}_{-n})^{\top}\mathbf{WU}_{-n}) = \lambda_n(\mathbf{W})$.

If $\lambda_1((\mathbf{U}_{-n})^{\top}\mathbf{WU}_{-n}) = \lambda_2(\mathbf{W}) > 0$,

$$\overline{\lambda} \equiv \lambda_{n-1}\left[\lambda_1((\mathbf{U}_{-n})^{\top}\mathbf{WU}_{-n})(\widehat{\mathbf{V}}_{-n})^{-1}\right] = \lambda_2(\mathbf{W})\lambda_{n-1}\left((\widehat{\mathbf{V}}_{-n})^{-1}\right) = \frac{\lambda_2(\mathbf{W})}{\lambda_1(\mathbf{V})};$$

if $\lambda_2(\mathbf{W}) < 0$, $\overline{\lambda} = \lambda_2(\mathbf{W})\lambda_1((\widehat{\mathbf{V}}_{-n})^{-1}) = \frac{\lambda_2(\mathbf{W})}{\lambda_{n-1}(\mathbf{V})}$. [16] Similarly, we have

$$\underline{\lambda} \equiv \lambda_{n-1}\left(\lambda_n(\mathbf{W})(\widehat{\mathbf{V}}_{-n})^{-1}\right) = \begin{cases} \frac{\lambda_n(\mathbf{W})}{\lambda_1(\mathbf{V})} & \text{if } \lambda_n(\mathbf{W}) > 0 \\ \frac{\lambda_n(\mathbf{W})}{\lambda_{n-1}(\mathbf{V})} & \text{if } \lambda_n(\mathbf{W}) < 0 \end{cases}.$$

Appendix B. Proof of Proposition 2

Let $|\widehat{\mathbf{W}}| \equiv diag\{1, |w_2|, \cdots, |w_n|\}$, $\mathbf{y} \equiv |\widehat{\mathbf{W}}|^{\frac{1}{2}}\mathbf{P}^{\top}\boldsymbol{\alpha}$, where $w_i = \lambda_i(\mathbf{W})$ is the ith largest eigenvalue of $\mathbf{W}$, $\mathbf{p}_{\cdot i}$ is its associated orthonormal eigenvector, $\mathbf{P} \equiv (\mathbf{p}_{\cdot i})_{i\in I}$.

- If $w_n > 0$, then programs $[\mathcal{P}^{v|w}_{min}]/[\mathcal{P}^{v|w}_{max}]$ are transformed into

  $$[\mathscr{P}^{v|w}_{min}]/[\mathscr{P}^{v|w}_{max}] : \max_{\mathbf{y}}/\min \mathbf{y}^{\top}|\widehat{\mathbf{W}}|^{-\frac{1}{2}}\mathbf{P}^{\top}\mathbf{VP}|\widehat{\mathbf{W}}|^{-\frac{1}{2}}\mathbf{y}, s.t. : \mathbf{y}^{\top}\mathbf{y} = 1.$$

  Since $\mathbf{p}_{\cdot 1} = \pm\frac{1}{\sqrt{n}}\mathbf{1}$, $\mathbf{V1} = \mathbf{0}$, we can write $|\widehat{\mathbf{W}}|^{-\frac{1}{2}}\mathbf{P}^{\top}\mathbf{VP}|\widehat{\mathbf{W}}|^{-\frac{1}{2}}$ as the block-diagonal form

  $$|\widehat{\mathbf{W}}|^{-\frac{1}{2}}\mathbf{P}^{\top}\mathbf{VP}|\widehat{\mathbf{W}}|^{-\frac{1}{2}} = \begin{bmatrix} 0 & \mathbf{0}^{\top} \\ \mathbf{0} & (|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}}(\mathbf{P}_{-1})^{\top}\mathbf{VP}_{-1}(|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}} \end{bmatrix},$$

  where $|\widehat{\mathbf{W}}|_{-1} \equiv diag\{|w_2|, \cdots, |w_n|\}$, $\mathbf{P}_{-1} \equiv (\mathbf{p}_{\cdot i})_{i\neq 1}$. Then $[\mathscr{P}^{v|w}_{min}]$ and $[\mathscr{P}^{v|w}_{max}]$ are rewritten as

  $$[\mathscr{P}^{v|w}_{min}]/[\mathscr{P}^{v|w}_{max}] : \max_{\mathbf{y}}/\min(\mathbf{y}_{-1})^{\top}(|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}}(\mathbf{P}_{-1})^{\top}\mathbf{VP}_{-1}(|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}}\mathbf{y}_{-1},$$
  $$s.t. : y_1^2 + (\mathbf{y}_{-1})^{\top}\mathbf{y}_{-1} = 1,$$

  where $\mathbf{y}_{-1} \equiv (y_i)_{i\neq 1}$.

[16] For a $m \times m$ matrix $\mathbf{A}$, and a scalar a, $\lambda_i(a\mathbf{A}) = a\lambda_i(\mathbf{A})$ if $a > 0$; and $\lambda_i(a\mathbf{A}) = a\lambda_{m-i+1}(\mathbf{A})$ if $a < 0$.

(i) The maximum value is

$$\begin{aligned}\Pi^{v|w}_{max} &= \lambda_1\left((|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}}(\mathbf{P}_{-1})^\top\mathbf{V}\mathbf{P}_{-1}(|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}}\right)\\ &= \lambda_1\left(\mathbf{P}_{-1}(|\widehat{\mathbf{W}}|_{-1})^{-1}(\mathbf{P}_{-1})^\top\mathbf{V}\right)\\ &= \lambda_1\left[\left(\mathbf{W}^{-1}-\frac{1}{n}\mathbf{1}\mathbf{1}^\top\right)\mathbf{V}\right] = \lambda_1(\mathbf{W}^{-1}\mathbf{V}).\end{aligned}$$

The corresponding optimal solution $\mathbf{y}^*$ is such that $y_1^* = 0$, $\|\mathbf{y}^*_{-1}\| = 1$ and

$$\mathbf{y}^*_{-1} \in \mathscr{S}\left(\lambda_1(\mathbf{W}^{-1}\mathbf{V}), (|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}}(\mathbf{P}_{-1})^\top\mathbf{V}\mathbf{P}_{-1}(|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}}\right).$$

Inserting $(|\widehat{\mathbf{W}}|_{-1})^{\frac{1}{2}}(\mathbf{P}_{-1})^\top\boldsymbol{\alpha} = \mathbf{y}^*_{-1}$ into

$$(|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}}(\mathbf{P}_{-1})^\top\mathbf{V}\mathbf{P}_{-1}(|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}}\mathbf{y}^*_{-1} = \lambda_1(\mathbf{W}^{-1}\mathbf{V})\mathbf{y}^*_{-1}$$

yields

$$(\mathbf{P}_{-1})^\top\mathbf{V}\boldsymbol{\alpha} = \lambda_1(\mathbf{W}^{-1}\mathbf{V})|\widehat{\mathbf{W}}|_{-1}(\mathbf{P}_{-1})^\top\boldsymbol{\alpha}.$$

Since $y_1^* = \pm\frac{1}{\sqrt{n}}\mathbf{1}^\top\boldsymbol{\alpha} = 0$, we have

$$\begin{bmatrix}\frac{1}{\sqrt{n}}\mathbf{1}^\top\\ (\mathbf{P}_{-1})^\top\end{bmatrix}\mathbf{V}\boldsymbol{\alpha} = \lambda_1(\mathbf{W}^{-1}\mathbf{V})\begin{bmatrix}1 & \\ & |\widehat{\mathbf{W}}|_{-1}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{n}}\mathbf{1}^\top\\ (\mathbf{P}_{-1})^\top\end{bmatrix}\boldsymbol{\alpha}.$$

Hence,

$$\mathbf{V}\boldsymbol{\alpha} = \lambda_1(\mathbf{W}^{-1}\mathbf{V})\mathbf{P}|\widehat{\mathbf{W}}|\mathbf{P}^\top\boldsymbol{\alpha} = \lambda_1(\mathbf{W}^{-1}\mathbf{V})\mathbf{W}\boldsymbol{\alpha}.$$

Therefore, the optimal intervention policy is $\boldsymbol{\alpha}^{v|w}_{max} \in \mathscr{G}(\lambda_1(\mathbf{W}^{-1}\mathbf{V}), \mathbf{V}, \mathbf{W})$.

(ii) The minimum value is $\Pi^{v|w}_{min} = 0$, the optimal solution is such that $y_1^* = \pm 1$, $\mathbf{y}^*_{-1} = \mathbf{0}$, accordingly, $\boldsymbol{\alpha}^{v|w}_{min} = \pm\frac{1}{\sqrt{n}}\mathbf{1}$.

- If $w_n < 0$ then programs $[\mathcal{P}^{v|w}_{min}]/[\mathcal{P}^{v|w}_{max}]$ are transformed into

$$\begin{aligned}[\mathscr{P}^{v|w}_{min}]/[\mathscr{P}^{v|w}_{max}] : \max_{\mathbf{y}}/\min (\mathbf{y}_{-1})^\top(|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}}(\mathbf{P}_{-1})^\top\mathbf{V}\mathbf{P}_{-1}(|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}}\mathbf{y}_{-1},\\ s.t.: (\mathbf{y}_K)^\top\mathbf{y}_K - (\mathbf{y}_{-K})^\top\mathbf{y}_{-K} = 1,\end{aligned}$$

where $K \equiv \{i \in I : w_i > 0\}, -K \equiv N\backslash K = \{i \in I : w_i < 0\}$,[17] $\mathbf{y}_K \equiv (y_j)_{j\in K}, \mathbf{y}_{-K} \equiv (y_j)_{j\notin K}$.

[17]Remember that $w_i \neq 0, \forall i \in I$, since $\mathbf{W}$ is nonsingular.

(*i*) Choosing a vector $\mathbf{y}$ satisfying: $(y_1)^2 = 1 + \|\mathbf{y}_{-K}\|^2$, $\|\mathbf{y}_{-K}\| = \infty$, and $y_j = 0, \forall j \in K \backslash \{1\}$, we have $(\mathbf{y}_K)^\top \mathbf{y}_K - (\mathbf{y}_{-K})^\top \mathbf{y}_{-K} = y_1^2 - \|\mathbf{y}_{-K}\|^2 = 1$ and

$$\begin{aligned}
\Pi_{max}^{v|w} &\geqslant (\mathbf{y}_{-1})^\top (|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}} (\mathbf{P}_{-1})^\top \mathbf{V}\mathbf{P}_{-1} (|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}} \mathbf{y}_{-1} &(34)\\
&= \left\{ \begin{array}{c} \frac{(\mathbf{y}_{-1})^\top (|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}} (\mathbf{P}_{-1})^\top \mathbf{V}\mathbf{P}_{-1} (|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}} \mathbf{y}_{-1}}{(\mathbf{y}_{-1})^\top (|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}} (\mathbf{P}_{-1})^\top \mathbf{P}_{-1} (|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}} \mathbf{y}_{-1}} \\ \times \frac{(\mathbf{y}_{-1})^\top (|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}} (\mathbf{P}_{-1})^\top \mathbf{P}_{-1} (|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}} \mathbf{y}_{-1}}{(\mathbf{y}_{-1})^\top \mathbf{y}_{-1}} \|\mathbf{y}_{-1}\|^2 \end{array} \right\} &(35)\\
&\geqslant \min_{\mathbf{1}^\top \mathbf{z}=0} \frac{\mathbf{z}^\top \mathbf{V}\mathbf{z}}{\mathbf{z}^\top \mathbf{z}} \min_{\mathbf{y}_{-1}} \frac{(\mathbf{y}_{-1})^\top (|\widehat{\mathbf{W}}|_{-1})^{-1} \mathbf{y}_{-1}}{(\mathbf{y}_{-1})^\top \mathbf{y}_{-1}} \|\mathbf{y}_{-1}\|^2 &(36)\\
&= \lambda_{n-1}(\mathbf{V}) \frac{1}{|w_2|} \|\mathbf{y}_{-1}\|^2 = \infty. &(37)
\end{aligned}$$

(35) $\Rightarrow$ (36) follows from $\mathbf{z} \equiv \mathbf{P}_{-1}(|\widehat{\mathbf{W}}|_{-1})^{-\frac{1}{2}} \mathbf{y}_{-1} \bot \mathbf{1}$ and $(\mathbf{P}_{-1})^\top \mathbf{P}_{-1} = \mathbf{I}_{n-1}$, (37) follows from $\lambda_{n-1}(\mathbf{V}) > 0$ and $\|\mathbf{y}_{-1}\| = \|\mathbf{y}_{-K}\| = \infty$.

(*ii*) $\Pi_{min}^{v|w} = 0$ is attained by choosing $\mathbf{y}^* = e_1 \equiv (1, 0 \cdots, 0)^\top$, or equivalently, $\boldsymbol{\alpha}_{min}^{v|w} = \pm \frac{1}{\sqrt{n}} \mathbf{1}$.

Appendix C. Proof of Corollary 1

Since $\mathbf{G}$ is a $d-$regular network, it is easy to find that matrices $\mathbf{V}$, $\mathbf{W}$ and $\mathbf{G}$ commute in pairs. Hence, they are simultaneously diagonalized by a common orthogonal matrix $\mathbf{S} \equiv (\mathbf{s}_{\cdot i})_{i \in I}$, where $\mathbf{s}_i$ denotes the normalized eigenvector of $\mathbf{G}$ associated with $\lambda_i(\mathbf{G})$, $\mathbf{s}_{\cdot 1} = \pm \mathbf{1}/\sqrt{n}, \lambda_1(\mathbf{G}) = d$. Let $\boldsymbol{\Lambda} \equiv diag\{[\lambda_i(\mathbf{G})]_{i \in I}\}$, $\widetilde{\mathbf{V}} \equiv \mathbf{S}^\top \mathbf{V}\mathbf{S}$, then we have

$$\begin{aligned}
\widetilde{\mathbf{V}} &= (1-\lambda)^2 \mathbf{S}^\top \left(\mathbf{I} - \frac{\lambda}{d}\mathbf{G}\right)^{-1} \mathbf{Q} \left(\mathbf{I} - \frac{\lambda}{d}\mathbf{G}\right)^{-1} \mathbf{S} &(38)\\
&= (1-\lambda)^2 \mathbf{S}^\top \mathbf{Q} \left(\mathbf{I} - \frac{\lambda}{d}\mathbf{G}\right)^{-1} \left(\mathbf{I} - \frac{\lambda}{d}\mathbf{G}\right)^{-1} \mathbf{Q}\mathbf{S} &(39)\\
&= (1-\lambda)^2 \left(\mathbf{S} - \frac{1}{\sqrt{n}} \mathbf{1} e_1^\top\right)^\top \left(\mathbf{I} - \frac{\lambda}{d}\mathbf{G}\right)^{-2} \left(\mathbf{S} - \frac{1}{\sqrt{n}} \mathbf{1} e_1^\top\right) &(40)\\
&= (1-\lambda)^2 \mathbf{S}^\top \left(\mathbf{I} - \frac{\lambda}{d}\mathbf{G}\right)^{-2} \mathbf{S} - e_1 e_1^\top &(41)\\
&= (1-\lambda)^2 \left(\mathbf{I} - \frac{\lambda}{d}\boldsymbol{\Lambda}\right)^{-2} - e_1 e_1^\top, &(42)
\end{aligned}$$

where $\boldsymbol{\Lambda} \equiv diag\{\lambda_i(\mathbf{G})_{i \in I}\}$. (38) $\Rightarrow$ (39) follows from the facts that $\mathbf{Q} \equiv \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^\top$ is idempotent, and $\mathbf{Q}$ and $\mathbf{G}$ commute, i.e., $\mathbf{Q} = \mathbf{Q}^2$ and $\mathbf{Q}\mathbf{G} = \mathbf{G}\mathbf{Q} = \mathbf{G} - \frac{d}{n}\mathbf{1}\mathbf{1}^\top$; (39) $\Rightarrow$ (40) follows from $\mathbf{Q}\mathbf{S} = \mathbf{S} - \frac{1}{\sqrt{n}}\mathbf{1}e_1^\top$; (40) $\Rightarrow$ (41) follows from

$$\frac{(1-\lambda)^2}{\sqrt{n}} e_1 \mathbf{1}^\top \left(\mathbf{I} - \tfrac{\lambda}{d}\mathbf{G}\right)^{-2} \mathbf{S} = \frac{(1-\lambda)^2}{\sqrt{n}} \mathbf{S}^\top \left(\mathbf{I} - \tfrac{\lambda}{d}\mathbf{G}\right)^{-2} \mathbf{1} e_1^\top = e_1 e_1^\top,$$
$$(1-\lambda)^2 \tfrac{1}{n} e_1 \mathbf{1}^\top \left(\mathbf{I} - \tfrac{\lambda}{d}\mathbf{G}\right)^{-2} \mathbf{1} e_1^\top = e_1 e_1^\top.$$

Diagonal matrix $\widehat{\mathbf{V}}$ is obtained by replacing the zero element in $\widetilde{\mathbf{V}}$ with one, i.e., $\widehat{\mathbf{V}} = (1-\lambda)^2\left(\mathbf{I}-\frac{\lambda}{d}\mathbf{\Lambda}\right)^{-2}$.[18] We thus have

$$\begin{aligned}
\mathbf{W} &= \mathbf{S}\left\{\mathbf{I}-\frac{1}{\lambda}\left[\mathbf{I}-(1-\lambda)\left(\mathbf{I}-\frac{\lambda}{d}\mathbf{\Lambda}\right)^{-1}\right]^2\right\}\mathbf{S}^\top \\
\mathbf{\Sigma} &= \widehat{\mathbf{V}}^{-\frac{1}{2}}\mathbf{S}^\top\mathbf{W}\mathbf{S}\widehat{\mathbf{V}}^{-\frac{1}{2}} = \frac{1}{(1-\lambda)^2}\left(\mathbf{I}-\frac{\lambda}{d}\mathbf{\Lambda}\right)^2\left\{\mathbf{I}-\frac{1}{\lambda}\left[\mathbf{I}-(1-\lambda)\left(\mathbf{I}-\frac{\lambda}{d}\mathbf{\Lambda}\right)^{-1}\right]^2\right\}, \\
\mathbf{W}^{-1}\mathbf{V} &= \mathbf{S}\left\{\mathbf{I}-\frac{1}{\lambda}\left[\mathbf{I}-(1-\lambda)\left(\mathbf{I}-\frac{\lambda}{d}\mathbf{\Lambda}\right)^{-1}\right]^2\right\}^{-1}\left[(1-\lambda)^2\left(\mathbf{I}-\frac{\lambda}{d}\mathbf{\Lambda}\right)^{-2}-e_1e_1^\top\right]\mathbf{S}^\top.
\end{aligned}$$

Then,

$$\lambda_{n-1}(\mathbf{\Sigma}_{-n,-n}) = \min_{i\neq 1}\psi(\lambda_i(\mathbf{G})), \lambda_1(\mathbf{W}^{-1}\mathbf{V}) = \max_{i\neq 1}\frac{1}{\psi(\lambda_i(\mathbf{G}))}$$

where

$$\psi(x) \equiv \frac{1}{(1-\lambda)^2}\left(1-\frac{\lambda x}{d}\right)^2\left\{1-\frac{1}{\lambda}\left[1-(1-\lambda)\left(1-\frac{\lambda x}{d}\right)^{-1}\right]^2\right\} = \frac{d^2-\lambda x^2}{d^2(1-\lambda)}.$$

Since $|\lambda_i(\mathbf{G})| \leqslant d, \forall i \in I$, the least eigenvalue of $\mathbf{W}$ is positive, i.e., $\lambda_n(\mathbf{W}) = \min_{i\in I} w(\lambda_i(\mathbf{G})) > 0$, where

$$w(x) \equiv 1-\frac{1}{\lambda}\left[1-(1-\lambda)\left(1-\frac{\lambda}{d}x\right)^{-1}\right]^2 = \frac{(1-\lambda)(d^2-\lambda x^2)}{(d-\lambda x)^2}.$$

From Propositions 1 and 2, we have: $\Pi^{w|v}_{min} = \lambda_{n-1}(\mathbf{\Sigma}_{-n,-n})$, $\Pi^{v|w}_{max} = \lambda_1(\mathbf{W}^{-1}\mathbf{V})$.

- if $\lambda_2(\mathbf{G})+\lambda_n(\mathbf{G}) > 0$,

$$\begin{aligned}
\Pi^{w|v}_{min} &= \lambda_{n-1}(\mathbf{\Sigma}_{-n,-n}) = \min_{i\neq 1}\psi(\lambda_i(\mathbf{G})) = \psi(\lambda_2(\mathbf{G})), \\
\Pi^{v|w}_{max} &= \lambda_1(\mathbf{W}^{-1}\mathbf{V}) = \max_{i\neq 1}\frac{1}{\psi(\lambda_i(\mathbf{G}))} = \frac{1}{\psi(\lambda_2(\mathbf{G}))}.
\end{aligned}$$

  The optimal interventions $\boldsymbol{\alpha}^{w|v}_{min}$ and $\boldsymbol{\alpha}^{v|w}_{max}$ are both within the second eigenspace:

$$\boldsymbol{\alpha}^{w|v}_{min} \in \mathscr{S}(\lambda_2(\mathbf{G}),\mathbf{G}), \boldsymbol{\alpha}^{v|w}_{max} \in \mathscr{S}(\lambda_2(\mathbf{G}),\mathbf{G}).$$

- if $\lambda_2(\mathbf{G})+\lambda_n(\mathbf{G}) < 0$,

$$\Pi^{w|v}_{min} = \psi(\lambda_n(\mathbf{G})), \Pi^{v|w}_{max} = \frac{1}{\psi(\lambda_n(\mathbf{G}))}.$$

  The optimal interventions $\boldsymbol{\alpha}^{w|v}_{min}$ and $\boldsymbol{\alpha}^{v|w}_{max}$ are both within the least eigenspace $\mathscr{S}(\lambda_n(\mathbf{G}),\mathbf{G})$.

[18] Here, the diagonal elements of $\widehat{\mathbf{V}}$ is not necessarily in descending order as defined in Appendix **A**.

Appendix D. Proof of Proposition 3

Let $w_i = \lambda_i(\mathbf{W})$ be the ith eigenvalue of matrix $\mathbf{W}$, among which $1 = \tilde{w}_1 > \cdots > \tilde{w}_m$ are m distinct values, $S_i \equiv \{j \in I : w_j = \tilde{w}_i\}$, $n_i \equiv \#S_i$, $\widehat{\mathbf{W}} \equiv diag\{1, w_1, \cdots, w_n\}$, $\mathbf{P} \equiv (\mathbf{p}_{\cdot i})_{i\in I}$, $\mathbf{P}_i \equiv (\mathbf{p}_{\cdot j})_{j\in S_i}$, $\mathbf{p}_{\cdot i}$ is the orthonormal eigenvector associated with w_i. Since $\mathbf{G}$ is connected, by the Perron-Frobenius theorem, we have $S_1 = \{1\}$, $n_1 = 1$ and $\mathbf{P}_1 = \mathbf{p}_{\cdot 1} = \pm \mathbf{1}/\sqrt{n}$.

Using orthogonal transformations $\widehat{\boldsymbol{\alpha}} \equiv \mathbf{P}^\top \boldsymbol{\alpha}$ and $\widehat{\boldsymbol{\alpha}}^0 \equiv \mathbf{P}^\top \boldsymbol{\alpha}^0$, we rewrite the original problem $[\mathcal{P}^u]$ as:

$$\max_{\widehat{\boldsymbol{\alpha}} \in \mathbb{R}^n} \widehat{\boldsymbol{\alpha}}^\top \widehat{\mathbf{W}} \widehat{\boldsymbol{\alpha}}, s.t. : \|\widehat{\boldsymbol{\alpha}} - \widehat{\boldsymbol{\alpha}}^0\| \leqslant C. \tag{43}$$

The corresponding Lagrangian with multiplier μ is

$$\mathscr{L}(\widehat{\boldsymbol{\alpha}}, \mu) = \widehat{\boldsymbol{\alpha}}^\top \widehat{\mathbf{W}} \widehat{\boldsymbol{\alpha}} + \mu \left[C^2 - (\widehat{\boldsymbol{\alpha}} - \widehat{\boldsymbol{\alpha}}^0)^\top (\widehat{\boldsymbol{\alpha}} - \widehat{\boldsymbol{\alpha}}^0) \right].$$

Taking our observation that the constraint is binding at the optimum together with the standard results on the Karush-Kuhn-Tucker conditions, we write the first and second order conditions holding with a positive μ as:

$$\frac{\partial \mathscr{L}}{\partial \hat{\alpha}_i} = 2\left[w_i \hat{\alpha}_i - \mu \left(\hat{\alpha}_i - \hat{\alpha}_i^0 \right) \right] = 0, \forall i \in I, \tag{44}$$

$$\frac{\partial^2 \mathscr{L}}{\partial \hat{\alpha}_i^2} = 2(w_i - \mu) \leqslant 0, \forall i \in I. \tag{45}$$

There are now three cases to consider.

(i) If $|\hat{\alpha}_1^0| = \sqrt{n}|\bar{\alpha}^0| \neq 0$, (44) implies $\mu^* \neq w_1$. Then

$$\hat{\alpha}_i^* = \frac{\mu^* \hat{\alpha}_i^0}{\mu^* - w_i}, \forall i \in I \tag{46}$$

holds with a nonzero denominator. Since

$$\left\| \mathrm{Proj}_{\mathscr{S}(\tilde{w}_i, \mathbf{W})} \boldsymbol{\alpha}^0 \right\|^2 = (\boldsymbol{\alpha}^0)^\top \mathbf{P}_i \mathbf{P}_i^\top \boldsymbol{\alpha}^0 = \sum_{j \in S_i} (\hat{\alpha}_j^0)^2,$$

the Lagrangian multiplier μ^* is pinned down by

$$\|\widehat{\boldsymbol{\alpha}}^* - \widehat{\boldsymbol{\alpha}}^0\| = \sqrt{\sum_{i=1}^n \left(\frac{w_i \hat{\alpha}_i^0}{\mu^* - w_i} \right)^2} = \sqrt{\sum_{i=1}^m \left(\frac{\tilde{w}_i \left\| \mathrm{Proj}_{\mathscr{S}(\tilde{w}_i, \mathbf{W})} \boldsymbol{\alpha}^0 \right\|}{\mu^* - \tilde{w}_i} \right)^2} = C. \tag{47}$$

From (46), we get

$$\mathbf{p}_{\cdot i}^\top (\boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0) = \hat{\alpha}_i^* - \hat{\alpha}_i^0 = \frac{w_i}{\mu^* - w_i} \hat{\alpha}_i^0 = \frac{w_i}{\mu^* - w_i} \mathbf{p}_{\cdot i}^\top \boldsymbol{\alpha}^0,$$

then

$$\cos\langle \mathbf{p}_{\cdot i}, \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0 \rangle = \frac{\|\boldsymbol{\alpha}^0\|}{C} \frac{w_i}{\mu^* - w_i} \cos\langle \mathbf{p}_{\cdot i}, \boldsymbol{\alpha}^0 \rangle, \forall i \in I.$$

It follows that

$$\begin{aligned}\mathbb{C}os\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, \mathbf{W})\rangle &= \cos\left\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathrm{Proj}_{\mathscr{S}(\tilde{w}_i,\mathbf{W})}(\boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0)\right\rangle \\ &= \cos\left\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathbf{P}_i\mathbf{P}_i^\top\left(\boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0\right)\right\rangle \\ &= \sqrt{\sum_{j\in S_i}\left(\frac{\mathbf{p}_{\cdot j}^\top(\boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0)}{\|\boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0\|}\right)^2} \\ &= \sqrt{\sum_{j\in S_i}\cos^2\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathbf{p}_{\cdot j}\rangle} \\ &= \frac{\|\boldsymbol{\alpha}^0\|}{C}\frac{|\tilde{w}_i|}{\mu^* - \tilde{w}_i}\mathbb{C}os\langle \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, \mathbf{W})\rangle, \forall i = 1, \cdots m. \end{aligned} \tag{48}$$

The optimal value attained is

$$W_{max} = \sum_{i=1}^n w_i(\hat{\alpha}_i)^2 = \sum_{i=1}^n w_i\left(\frac{\mu^*\hat{\alpha}_i^0}{\mu^* - w_i}\right)^2 = \sum_{i=1}^m \tilde{w}_i\left(\frac{\mu^*\left\|\mathrm{Proj}_{\mathscr{S}(\tilde{w}_i,\mathbf{W})}\boldsymbol{\alpha}^0\right\|}{\mu^* - \tilde{w}_i}\right)^2.$$

(*ii*) If $|\hat{\alpha}_1^0| = \sqrt{n}|\bar{\alpha}^0| = 0$, and

$$C \leqslant C^* \equiv \sqrt{\sum_{i\neq 1}\left(\frac{\tilde{w}_i}{1 - \tilde{w}_i}\right)^2\left\|\mathrm{Proj}_{\mathscr{S}(\tilde{w}_i,\mathbf{W})}\boldsymbol{\alpha}^0\right\|^2},$$

we have

$$\hat{\alpha}_1^* = 0, \hat{\alpha}_i^* = \frac{\mu^*\hat{\alpha}_i^0}{\mu^* - w_i}, \forall i \geqslant 2, \tag{49}$$

where $\mu^* > 1$ is determined by

$$C = \sqrt{\sum_{i\neq 1}\left(\frac{\tilde{w}_i}{\mu - \tilde{w}_i}\right)^2\left\|\mathrm{Proj}_{\mathscr{S}(\tilde{w}_i,\mathbf{W})}\boldsymbol{\alpha}^0\right\|^2}. \tag{50}$$

From (49), we get $(\boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0)\perp\mathbf{1}$ and

$$\mathbb{C}os\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, \mathbf{W})\rangle = \frac{\tilde{w}_i}{\mu^* - \tilde{w}_i}\frac{\|\boldsymbol{\alpha}^0\|}{C}\mathbb{C}os\langle \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, \mathbf{W})\rangle, \forall i \neq 1. \tag{51}$$

The optimal value attained is

$$W_{max} = \sum_{i=1}^n w_i(\hat{\alpha}_i)^2 = \sum_{i=2}^m \tilde{w}_i\left(\frac{\mu^*\left\|\mathrm{Proj}_{\mathscr{S}(\tilde{w}_i,\mathbf{W})}\boldsymbol{\alpha}^0\right\|}{\mu^* - \tilde{w}_i}\right)^2.$$

(*iii*) If $|\hat{\alpha}_1^0| = \sqrt{n}|\bar{\alpha}^0| = 0$, and $C > C^*$, we have $\mu^* = w_1 = 1$, $|\hat{\alpha}_1^*| = \sqrt{C^2 - (C^*)^2}$, and

$$\hat{\alpha}_i^* = \frac{\hat{\alpha}_i^0}{1 - w_i}, \forall i \neq 1. \tag{52}$$

In analogy to the derivation of (48), (52) implies

$$\mathbb{C}os\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, \mathbf{W})\rangle = \frac{\tilde{w}_i}{1-\tilde{w}_i}\frac{\|\boldsymbol{\alpha}^0\|}{C}\mathbb{C}os\langle \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, \mathbf{W})\rangle, \forall i \neq 1. \tag{53}$$

From $|\hat{\alpha}_1^*| = \sqrt{C^2 - (C^*)^2}$, we get

$$\mathbb{C}os\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_1, \mathbf{W})\rangle = \frac{|\hat{\alpha}_1^*|}{C} = \sqrt{\frac{C^2 - (C^*)^2}{C^2}}. \tag{54}$$

The maximum valued attained is

$$W_{max} = \sum_{i=1}^{n} w_i(\hat{\alpha}_i)^2 = C^2 - (C^*)^2 + \sum_{i=2}^{m} \tilde{w}_i \left(\frac{\left\|\text{Proj}_{\mathscr{S}(\tilde{w}_i,\mathbf{W})}\boldsymbol{\alpha}^0\right\|}{1-\tilde{w}_i}\right)^2.$$

Appendix E: Proof of Corollary 2

- (47) (resp. (50)) implies $\mu^* \to \infty$ as $C \to 0$ when $\bar{\alpha}^0 \neq 0$ (resp.$\bar{\alpha}^0 = 0$), then

$$\begin{aligned}\lim_{C\to 0} \cos\left\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathbf{p}_{\cdot j}\right\rangle &= \lim_{C\to 0} \cos\left\langle \widehat{\boldsymbol{\alpha}}^* - \widehat{\boldsymbol{\alpha}}^0, e_j\right\rangle \\ &= \lim_{\mu\to\infty} \frac{\frac{w_j\hat{\alpha}_j^0}{\mu-w_j}}{\sqrt{\sum_{i=1}^n\left(\frac{w_i\hat{\alpha}_i^0}{\mu-w_i}\right)^2}} = \frac{w_j\hat{\alpha}_j^0}{\sqrt{\sum_{i=1}^n\left(w_i\hat{\alpha}_i^0\right)^2}}, \forall j. \end{aligned} \tag{55}$$

  It follows that: as $C \to 0$, $\mathbf{P}^\top(\boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0) \propto \widehat{\mathbf{W}}\widehat{\boldsymbol{\alpha}}^0 = \mathbf{P}^\top\mathbf{W}\mathbf{P}\mathbf{P}^\top\boldsymbol{\alpha}^0 = \mathbf{P}^\top\mathbf{W}\boldsymbol{\alpha}^0$, therefore $\lim_{C\to 0}\left\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathbf{W}\boldsymbol{\alpha}^0\right\rangle = 0$.

- If $|\hat{\alpha}_1^0| \neq 0$, (47) implies $\mu^* \to \tilde{w}_1 = 1$ as $C \to \infty$, then we get $\lim_{C\to\infty}\mathbb{C}os\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_1, W)\rangle = 1$, and $\lim_{C\to\infty}\mathbb{C}os\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, W)\rangle = 0, \forall i \neq 1$ from (48). When $|\hat{\alpha}_1^0| = 0$, we also have $\lim_{C\to\infty}\mathbb{C}os\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_1, W)\rangle = 1$, and $\lim_{C\to\infty}\mathbb{C}os\langle \boldsymbol{\alpha}^* - \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{w}_i, W)\rangle = 0, \forall i \neq 1$ from (53) and (54).

Appendix F. Proof of Proposition 4

- First, we discuss the case with $C < \mathscr{D}(\boldsymbol{\alpha}^0, \mathscr{S}(\tilde{\nu}_r, \mathbf{V})) = \sqrt{(\boldsymbol{\alpha}^0)^\top\mathbf{Q}\boldsymbol{\alpha}^0} = \sqrt{(n-1)\mathbb{V}ar(\boldsymbol{\alpha}^0)}$, where $\mathscr{S}(\tilde{\nu}_r, \mathbf{V}) \equiv \{\mathbf{x}|\mathbf{x} \propto 1\}$ denotes the eigenspace associated with $\tilde{\nu}_r = 0$, $\mathscr{D}(\cdot,\cdot)$ denotes the point-to-space distance. Under transformations $\tilde{\boldsymbol{\alpha}} \equiv \mathbf{U}^\top\boldsymbol{\alpha}$ and $\tilde{\boldsymbol{\alpha}}^0 \equiv \mathbf{U}^\top\boldsymbol{\alpha}^0$, the original problem $[\mathcal{P}^e]$ is rewritten as

$$\min_{\tilde{\boldsymbol{\alpha}}\in\mathbb{R}^n} \tilde{\boldsymbol{\alpha}}^\top\widetilde{\mathbf{V}}\tilde{\boldsymbol{\alpha}}, \text{s.t.} : \|\tilde{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}}^0\| \leqslant C, \tag{56}$$

  where $\mathbf{U} \equiv (\mathbf{u}_{\cdot i})$, $\widetilde{\mathbf{V}} \equiv diag\{\nu_1, \cdots, \nu_n\}$, ν_i and $\mathbf{u}_{\cdot i}$ are the ith eigenvalue and its corresponding eigenvector of matrix $\mathbf{V}$. Let $\tilde{\nu}_1 > \cdots > \tilde{\nu}_r = 0$ denotes r distinct eigenvalues,

$T_i \equiv \{j \in I : \nu_j = \tilde{\nu}_i\}$ is the set of indexes associated with $\tilde{\nu}_i$. The corresponding Lagrangian with multiplier μ is

$$\mathscr{L}(\tilde{\boldsymbol{\alpha}}, \mu) = \tilde{\boldsymbol{\alpha}}^\top \widetilde{\mathbf{V}} \tilde{\boldsymbol{\alpha}} + \mu \left[C^2 - (\tilde{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}}^0)^\top (\tilde{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}}^0)\right].$$

Taking our observation that the constraint is binding at the optimum, we have the first and second order conditions holding with a negative μ:

$$\begin{aligned} \frac{\partial \mathscr{L}}{\partial \tilde{\alpha}_i} &= 2\left[\nu_i \tilde{\alpha}_i - \mu\left(\tilde{\alpha}_i - \tilde{\alpha}_i^0\right)\right] = 0, \forall i \in I, \\ \frac{\partial^2 \mathscr{L}}{\partial \tilde{\alpha}_i^2} &= 2(\nu_i - \mu) > 0, \forall i \in I. \end{aligned}$$

Hence, the optimal solution $\tilde{\boldsymbol{\alpha}}^\dagger$ and the Lagrange multiplier $\mu\dagger$ satisfies

$$\tilde{\alpha}_i^\dagger = \tfrac{\mu \tilde{\alpha}_i^0}{\mu - \nu_i}, \forall i, \nu_1 \geqslant \cdots \geqslant \nu_{n-1} > \nu_n = 0 > \mu, \tag{57}$$

$$\|\tilde{\boldsymbol{\alpha}}^\dagger - \tilde{\boldsymbol{\alpha}}^0\| = \sqrt{\textstyle\sum_{i=1}^n \left(\frac{\nu_i \tilde{\alpha}_i^0}{\mu^\dagger - \nu_i}\right)^2} = C. \tag{58}$$

Substituting $\nu_j = \tilde{\nu}_i, \forall j \in T_i$ and $\sum_{j \in T_i} (\tilde{\alpha}_j^0)^2 = \|\mathrm{Proj}_{\mathscr{S}(\tilde{\nu}_i, \mathbf{V})} \boldsymbol{\alpha}^0\|^2$ into (58), we obtain

$$\sqrt{\sum_{i=1}^r \left(\frac{\tilde{\nu}_i \left\|\mathrm{Proj}_{\mathscr{S}(\tilde{\nu}_i, \mathbf{V})} \boldsymbol{\alpha}^0\right\|}{\mu^\dagger - \tilde{\nu}_i}\right)^2} = C. \tag{59}$$

(57) implies $\tilde{\alpha}_i^\dagger - \tilde{\alpha}_i^0 = \frac{\nu_i \tilde{\alpha}_i^0}{\mu - \nu_i}$. It follows that

$$\cos\left\langle \boldsymbol{\alpha}^\dagger - \boldsymbol{\alpha}^0, \mathbf{u}_{\cdot i}\right\rangle = \frac{\nu_i}{\mu^\dagger - \nu_i} \frac{\|\boldsymbol{\alpha}^0\|}{C} \cos\left\langle \boldsymbol{\alpha}^0, \mathbf{u}_{\cdot i}\right\rangle, \forall i = 1, \cdots, n. \tag{60}$$

Since $\sum_{j \in T_i} \cos^2 \langle \mathbf{x}, \mathbf{u}_{\cdot i}\rangle = \mathbb{C}os^2 \langle \mathbf{x}, \mathscr{S}(\tilde{\nu}_i, \mathbf{V})\rangle, \forall \mathbf{x}$ and $\nu_j = \tilde{\nu}_i, \forall j \in T_i$, (60) implies

$$\mathbb{C}os\left\langle \boldsymbol{\alpha}^\dagger - \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{\nu}_i, \mathbf{V})\right\rangle = \frac{\tilde{\nu}_i}{\mu^\dagger - \tilde{\nu}_i} \frac{\|\boldsymbol{\alpha}^0\|}{C} \mathbb{C}os\left\langle \boldsymbol{\alpha}^0, \mathscr{S}(\tilde{\nu}_i, \mathbf{V})\right\rangle, \forall i = 1, \cdots, r. \tag{61}$$

The minimum value attained is

$$V_{min} = \sum_{i=1}^n \nu_i \tilde{\alpha}_i^2 = \sum_{i=1}^n \nu_i \left[\frac{\mu^\dagger (\mathbf{u}_{\cdot i}^\top \boldsymbol{\alpha}^0)}{\mu^\dagger - \nu_i}\right]^2 = \sum_{i=1}^r \tilde{\nu}_i \left[\frac{\mu^\dagger \|\mathrm{Proj}_{\mathscr{S}(\tilde{\nu}_i, \mathbf{V})} \boldsymbol{\alpha}^0\|}{\mu^\dagger - \tilde{\nu}_i}\right]^2.$$

- If $C \geqslant \mathscr{D}(\boldsymbol{\alpha}^0, \mathscr{S}(\tilde{\nu}_r, \mathbf{V})) = \sqrt{(n-1)\mathbb{V}ar(\boldsymbol{\alpha}^0)}$, then $\mathcal{O}(\boldsymbol{\alpha}^0, C) \cap \mathscr{S}(\tilde{\nu}_r, \mathbf{V}) \neq \emptyset$. We can thus obtain a minimum $V_{min} = 0$ by choosing an arbitrary $\boldsymbol{\alpha}^\dagger \in \mathcal{O}(\boldsymbol{\alpha}^0, C) \cap \mathscr{S}(\tilde{\lambda}_r, \mathbf{V})$.

- (58) suggests $\mu \to -\infty$ as $C \to 0$. Therefore, we have

$$\begin{aligned} \lim_{C \to 0} \cos\left\langle \boldsymbol{\alpha}^\dagger - \boldsymbol{\alpha}^0, \mathbf{u}_{\cdot j}\right\rangle &= \lim_{C \to 0} \cos\left\langle \tilde{\boldsymbol{\alpha}}^\dagger - \tilde{\boldsymbol{\alpha}}^0, e_j\right\rangle \\ &= \lim_{\mu \to -\infty} \frac{\frac{\nu_j \tilde{\alpha}_j^0}{\mu - \nu_j}}{\sqrt{\sum_{i=1}^n \left(\frac{\nu_i \tilde{\alpha}_i^0}{\mu - \nu_i}\right)^2}} = -\frac{\nu_j \tilde{\alpha}_j^0}{\sqrt{\sum_{i=1}^n \left(\nu_i \tilde{\alpha}_i^0\right)^2}}, \forall j, \end{aligned} \tag{62}$$

where e_j is the jth standard base vector of space $\mathbb{R}^n$. (62) implies $\lim_{C\to 0}\mathbf{U}^\top(\boldsymbol{\alpha}^\dagger - \boldsymbol{\alpha}^0) \propto -\widetilde{\mathbf{V}}\mathbf{U}^\top\boldsymbol{\alpha}^0$, then we have $\lim_{C\to 0}(\boldsymbol{\alpha}^\dagger - \boldsymbol{\alpha}^0) \propto -\mathbf{U}\widetilde{\mathbf{V}}\mathbf{U}^\top\boldsymbol{\alpha}^0 = -\mathbf{V}\boldsymbol{\alpha}^0$. That is, $\lim_{C\to 0}\left\langle \boldsymbol{\alpha}^\dagger - \boldsymbol{\alpha}^0, -\mathbf{V}\boldsymbol{\alpha}^0\right\rangle = 0$.

Appendix G. Proof of Proposition 5

- $W^*(C)$, $W^s(C)$ and $W^t(C)$ are represented as

$$\begin{aligned}
W^*(C) &= \sum_{i=1}^n w_i(\hat\alpha_i^*)^2 = \sum_{i=1}^n w_i\left(\frac{\mu\hat\alpha_i^0}{\mu - w_i}\right)^2 = \sum_{i=1}^m \tilde w_i\left(\frac{\mu\left\|\mathrm{Proj}_{\mathscr{S}(\tilde w_i,\mathbf{W})}\boldsymbol{\alpha}^0\right\|}{\mu - \tilde w_i}\right)^2,\\
W^t(C) &= \left(\boldsymbol{\alpha}^0 + C\frac{\mathbf{W}\boldsymbol{\alpha}^0}{\|\mathbf{W}\boldsymbol{\alpha}^0\|}\right)^\top \mathbf{W}\left(\boldsymbol{\alpha}^0 + C\frac{\mathbf{W}\boldsymbol{\alpha}^0}{\|\mathbf{W}\boldsymbol{\alpha}^0\|}\right),\\
W^s(C) &= w_1\left(|\hat\alpha_1^0| + C(\mu)\right)^2 + \sum_{\ell\neq 1} w_\ell(\hat\alpha_\ell^0)^2\\
&= \left(\left\|\mathrm{Proj}_{\mathscr{S}(\tilde w_1,\mathbf{W})}\boldsymbol{\alpha}^0\right\| + C(\mu)\right)^2 + \sum_{j=2}^m \tilde w_j\left\|\mathrm{Proj}_{\mathscr{S}(\tilde w_j,\mathbf{W})}\boldsymbol{\alpha}^0\right\|^2
\end{aligned}$$

where

$$C(\mu) \equiv \sqrt{\sum_{i=1}^n\left(\frac{w_i\hat\alpha_i^0}{\mu - w_i}\right)^2} = \sqrt{\sum_{j=1}^m\left(\frac{\tilde w_j\left\|\mathrm{Proj}_{\mathscr{S}(\tilde w_j,\mathbf{W})}\boldsymbol{\alpha}^0\right\|}{\mu - \tilde w_j}\right)^2}.$$

The Taylor expansion of $\Delta_w^s(C(\mu),\boldsymbol{\alpha}^0) \equiv W^*(C(\mu)) - W^s(C(\mu))$ around $\mu = w_1 = \tilde w_1 = 1$ is

$$\Delta_w^s(C(\mu),\boldsymbol{\alpha}^0) = \left\{\begin{array}{c} \displaystyle\sum_{j=2}^m \frac{(\tilde w_j\|\mathrm{Proj}_{\mathscr{S}(\tilde w_j,\mathbf{W})}\boldsymbol{\alpha}^0\|)^2}{1-\tilde w_j} - \sum_{j=2}^m \frac{(\tilde w_j\|\mathrm{Proj}_{\mathscr{S}(\tilde w_j,\mathbf{W})}\boldsymbol{\alpha}^0\|)^2}{(1-\tilde w_j)^2}(\mu-1)\\ +o(\mu-1)\end{array}\right\}.$$

Therefore,

$$\Delta_w^s(\infty,\boldsymbol{\alpha}^0) = \lim_{\mu\to 1}\Delta_w^s(C(\mu),\boldsymbol{\alpha}^0) = \sum_{j=2}^m\frac{\tilde w_j^2\left\|\mathrm{Proj}_{\mathscr{S}(\tilde w_i,\mathbf{W})}\boldsymbol{\alpha}^0\right\|^2}{1-\tilde w_j},$$

Since $\sum_{i=2}^m \left\|\mathrm{Proj}_{\mathscr{S}(\tilde w_i,\mathbf{W})}\boldsymbol{\alpha}^0\right\|^2 = (\boldsymbol{\alpha}^0)^\top\mathbf{Q}(\boldsymbol{\alpha}^0) = (n-1)\mathbb{V}ar(\boldsymbol{\alpha}^0) = (n-1)\xi$, we have

$$\frac{(n-1)\xi\tilde w_{i_*}^2}{1-\tilde w_{i_*}} \leqslant \Delta_w^s(\infty,\boldsymbol{\alpha}^0) \leqslant \frac{(n-1)\xi\tilde w_{i^*}^2}{1-\tilde w_{i^*}},$$

where $i^* = \arg\max_{j\neq 1}\frac{\tilde w_j^2}{1-\tilde w_j}$, $i_* = \arg\min_{j\neq 1}\frac{\tilde w_j^2}{1-\tilde w_j}$. $\Delta_w^s(\infty,\boldsymbol{\alpha}^0)$ achieves its maximum $\frac{(n-1)\xi\tilde w_{i^*}^2}{1-\tilde w_{i^*}}$ at $\boldsymbol{\alpha}^0 \in \mathscr{S}(\tilde w_{i^*},\mathbf{W})$, and its minimum $\frac{(n-1)\xi\tilde w_{i_*}^2}{1-\tilde w_{i_*}}$ at $\boldsymbol{\alpha}^0 \in \mathscr{S}(\tilde w_{i_*},\mathbf{W})$.

- If $\boldsymbol{\alpha}^0 \in \mathscr{S}(\tilde{w}_i, \mathbf{W})$ for some i, we have $\|\mathrm{Proj}_{\mathscr{S}(\tilde{w}_i,\mathbf{W})}\boldsymbol{\alpha}^0\| = \|\boldsymbol{\alpha}^0\|$, $\|\mathrm{Proj}_{\mathscr{S}(\tilde{w}_j,\mathbf{W})}\boldsymbol{\alpha}^0\| = 0, \forall j \neq i$, and $C = \frac{|\tilde{w}_i|\|\boldsymbol{\alpha}^0\|}{\mu^* - \tilde{w}_i}$. Hence,

$$
\begin{aligned}
W^*(C) &= \tilde{w}_i \left(\frac{\mu^* \|\boldsymbol{\alpha}^0\|}{\mu^* - \tilde{w}_i} \right)^2, \\
W^t(C) &= \left(\boldsymbol{\alpha}^0 + \frac{C\tilde{w}_i \boldsymbol{\alpha}^0}{|\tilde{w}_i|\|\boldsymbol{\alpha}^0\|} \right)^\top \mathbf{W} \left(\boldsymbol{\alpha}^0 + \frac{C\tilde{w}_i \boldsymbol{\alpha}^0}{|\tilde{w}_i|\|\boldsymbol{\alpha}^0\|} \right) = \tilde{w}_i \left(\frac{\mu^* \|\boldsymbol{\alpha}^0\|}{\mu^* - \tilde{w}_i} \right)^2.
\end{aligned}
$$

Therefore,

$$
\Delta_w^t(C, \boldsymbol{\alpha}^0) \equiv W^*(C) - W^t(C) = 0, \forall \boldsymbol{\alpha}^0 \in \bigcup_{i=1}^{m} \mathscr{S}(\tilde{w}_i, \mathbf{W}). \tag{63}
$$

Next, we discuss the case with $\boldsymbol{\alpha}^0 \notin \mathscr{S}(\tilde{w}_i, \mathbf{W}), \forall i$. Differentiating $W^*(C)$ and $W^t(C)$ with respect to C yields:

$$
\begin{aligned}
\frac{dW^*(C)}{dC} &= \frac{dW^*(C)/d\mu}{dC/d\mu} = 2\mu \sqrt{\sum_{i=1}^{n} \left(\frac{w_i \hat{\alpha}_i^0}{\mu - w_i} \right)^2} \\
\frac{dW^t(C)}{dC} &= 2\frac{(\boldsymbol{\alpha}^0)^\top \mathbf{W}^2}{\|\mathbf{W}\boldsymbol{\alpha}^0\|} \left(\boldsymbol{\alpha}^0 + C\frac{\mathbf{W}\boldsymbol{\alpha}^0}{\|\mathbf{W}\boldsymbol{\alpha}^0\|} \right)
\end{aligned}
$$

$$
\begin{aligned}
\frac{d^2W^*(C)}{dC^2} &= \frac{\frac{d}{d\mu}\frac{dW^*(C)}{dC}}{dC/d\mu} = 2\left[\mu - \frac{\sum_{i=1}^n \left(\frac{w_i \hat{\alpha}_i^0}{\mu - w_i} \right)^2}{\sum_{i=1}^n \frac{(w_i \hat{\alpha}_i^0)^2}{(\mu - w_i)^3}} \right] = 2\frac{\sum_{i=1}^n \frac{w_i^3 (\hat{\alpha}_i^0)^2}{(\mu - w_i)^3}}{\sum_{i=1}^n \frac{(w_i \hat{\alpha}_i^0)^2}{(\mu - w_i)^3}} \\
\frac{d^2W^t(C)}{dC^2} &= 2\frac{(\boldsymbol{\alpha}^0)^\top \mathbf{W}^3 \boldsymbol{\alpha}^0}{(\boldsymbol{\alpha}^0)^\top \mathbf{W}^2 \boldsymbol{\alpha}^0} \\
\frac{d^3W^*(C)}{dC^3} &= \frac{\frac{d}{d\mu}\frac{d^2W^*(C)}{dC^2}}{dC/d\mu} = \frac{6\sqrt{\sum_{i=1}^n \left(\frac{w_i \hat{\alpha}_i^0}{\mu - w_i} \right)^2} \left\{ 1 - \frac{\sum_{i=1}^n \frac{(w_i \hat{\alpha}_i^0)^2}{(\mu - w_i)^4} \sum_{i=1}^n \frac{(w_i \hat{\alpha}_i^0)^2}{(\mu - w_i)^2}}{\left[\sum_{i=1}^n \frac{(w_i \hat{\alpha}_i^0)^2}{(\mu - w_i)^3} \right]^2} \right\}}{\sum_{i=1}^n \frac{(w_i \hat{\alpha}_i^0)^2}{(\mu - w_i)^3}} \\
\frac{d^kW^t(C)}{dC^k} &= 0, \forall k \geqslant 3.
\end{aligned}
$$

$$\begin{aligned}
\Delta_w^t(0,\boldsymbol{\alpha}^0) &= \lim_{\mu\to\infty}\sum_{i=1}^n\left(\frac{\mu\hat{\alpha}_i^0}{\mu-w_i}\right)^2 w_i-(\boldsymbol{\alpha}^0)^\top\mathbf{W}\boldsymbol{\alpha}^0=(\widehat{\boldsymbol{\alpha}}^0)^\top\widehat{\mathbf{W}}\widehat{\boldsymbol{\alpha}}^0-(\boldsymbol{\alpha}^0)^\top\mathbf{W}\boldsymbol{\alpha}^0=0,\\
\frac{\partial\Delta_w^t(0,\boldsymbol{\alpha}^0)}{\partial C} &= \lim_{\mu\to\infty}2\mu\sqrt{\sum_{i=1}^n\left(\frac{w_i\hat{\alpha}_i^0}{\mu-w_i}\right)^2}-2\sqrt{(\boldsymbol{\alpha}^0)^\top\mathbf{W}^2\boldsymbol{\alpha}^0}=0,\\
\frac{\partial^2\Delta_w^t(0,\boldsymbol{\alpha})}{\partial C^2} &= \lim_{\mu\to\infty}2\frac{\sum_{i=1}^n\frac{w_i^3(\hat{\alpha}_i^0)^2}{(\mu-w_i)^3}}{\sum_{i=1}^n\frac{(w_i\hat{\alpha}_i^0)^2}{(\mu-w_i)^3}}-\frac{2(\boldsymbol{\alpha}^0)^\top\mathbf{W}^3\boldsymbol{\alpha}^0}{(\boldsymbol{\alpha}^0)^\top\mathbf{W}^2\boldsymbol{\alpha}^0}=0,\\
\frac{\partial^3\Delta_w^t(0,\boldsymbol{\alpha}^0)}{\partial C^3} &= \lim_{\mu\to\infty}\frac{6\sqrt{\sum_{i=1}^n\left(\frac{w_i\hat{\alpha}_i^0}{\mu-w_i}\right)^2}\left[1-\frac{\sum_{i=1}^n\frac{(w_i\hat{\alpha}_i^0)^2}{(\mu-w_i)^4}\sum_{i=1}^n\frac{(w_i\hat{\alpha}_i^0)^2}{(\mu-w_i)^2}}{\left[\sum_{i=1}^n\frac{(w_i\hat{\alpha}_i^0)^2}{(\mu-w_i)^3}\right]^2}\right]}{-\sum_{i=1}^n\frac{(w_i\hat{\alpha}_i^0)^2}{(\mu-w_i)^3}},\\
&= \frac{6\sum_{i=1}^n\sum_{j=1}^n(\hat{\alpha}_i^0\hat{\alpha}_j^0w_iw_j)^2(w_i-w_j)^2}{\left[\sum_{i=1}^n(w_i\hat{\alpha}_i^0)^2\right]^{5/2}} \qquad (64)\\
&= \frac{12\left[\left(\sum_{i=1}^n w_i^4(\hat{\alpha}_i^0)^2\right)\left(\sum_{i=1}^n(w_i\hat{\alpha}_i^0)^2\right)-\left(\sum_{i=1}^n w_i^3(\hat{\alpha}_i^0)^2\right)^2\right]}{\left[\sum_{i=1}^n(w_i\hat{\alpha}_i^0)^2\right]^{\frac{5}{2}}} \qquad (65)\\
&= \frac{12\left[\|\mathbf{W}\boldsymbol{\alpha}^0\|^2\|\mathbf{W}^2\boldsymbol{\alpha}^0\|^2-\|\mathbf{W}^{3/2}\boldsymbol{\alpha}^0\|^4\right]}{\|\mathbf{W}\boldsymbol{\alpha}^0\|^5}. \qquad (66)
\end{aligned}$$

(64) ⇒ (65) follows from

$$\begin{aligned}
&\sum_{i=1}^n\sum_{j=1}^n(\hat{\alpha}_i^0\hat{\alpha}_j^0w_iw_j)^2(w_i-w_j)^2\\
=\;&\sum_{i=1}^n\sum_{j=1}^n(\hat{\alpha}_i^0\hat{\alpha}_j^0)^2w_i^4w_j^2+\sum_{i=1}^n\sum_{j=1}^n(\hat{\alpha}_i^0\hat{\alpha}_j^0)^2w_i^2w_j^4-2\sum_{i=1}^n\sum_{j=1}^n(\hat{\alpha}_i^0\hat{\alpha}_j^0)^2w_i^3w_j^3\\
=\;&2\sum_{i=1}^n\left(\hat{\alpha}_i^0w_i^2\right)^2\sum_{j=1}^n\left(\hat{\alpha}_j^0\right)^2w_j^2-\left(\sum_{i=1}^n(\hat{\alpha}_i^0)^2w_i^3\right)^2.
\end{aligned}$$

(65) ⇒ (66) follows from

$$\begin{aligned}
\sum_{i=1}^n w_i^4(\hat{\alpha}_i^0)^2 &= \sum_{i=1}^n(\boldsymbol{\alpha}^0)^\top\mathbf{p}_{\cdot i}w_i^4\mathbf{p}_{\cdot i}^\top\boldsymbol{\alpha}^0=(\boldsymbol{\alpha}^0)^\top\mathbf{W}^4\boldsymbol{\alpha}^0=\left\|\mathbf{W}^2\boldsymbol{\alpha}^0\right\|^2,\\
\sum_{i=1}^n w_i^2(\hat{\alpha}_i^0)^2 &= \sum_{i=1}^n(\boldsymbol{\alpha}^0)^\top\mathbf{p}_{\cdot i}w_i^2\mathbf{p}_{\cdot i}^\top\boldsymbol{\alpha}^0=(\boldsymbol{\alpha}^0)^\top\mathbf{W}^2\boldsymbol{\alpha}^0=\left\|\mathbf{W}\boldsymbol{\alpha}^0\right\|^2,\\
\sum_{i=1}^n w_i^3(\hat{\alpha}_i^0)^2 &= \sum_{i=1}^n(\boldsymbol{\alpha}^0)^\top\mathbf{p}_{\cdot i}w_i^3\mathbf{p}_{\cdot i}^\top\boldsymbol{\alpha}^0=(\boldsymbol{\alpha}^0)^\top\mathbf{W}^3\boldsymbol{\alpha}^0=\left\|\mathbf{W}^{\frac{3}{2}}\boldsymbol{\alpha}^0\right\|^2.
\end{aligned}$$

Taylor expansion of $\Delta_w^t(C,\boldsymbol{\alpha}^0)$ around zero is

$$\Delta_w^t(C,\boldsymbol{\alpha}^0)=\frac{12\left[\|\mathbf{W}\boldsymbol{\alpha}^0\|^2\|\mathbf{W}^2\boldsymbol{\alpha}^0\|^2-\|\mathbf{W}^{3/2}\boldsymbol{\alpha}^0\|^4\right]}{\|\mathbf{W}\boldsymbol{\alpha}^0\|^5}C^3+o(C^3),\forall\boldsymbol{\alpha}^0\notin\bigcup_{i=1}^m\mathscr{S}(\tilde{w}_i,\mathbf{W}). \qquad (67)$$

- The optimal value and its tangent approximation attained in problem $[\mathcal{P}^e]$ are

$$V^*(C) = \sum_{i=1}^{n} \nu_i \left(\frac{\mu^\dagger \tilde{\alpha}_i^0}{\mu^\dagger - \nu_i}\right)^2 = \sum_{i=1}^{r} \tilde{\nu}_i \left(\frac{\mu^\dagger \|\mathrm{Proj}_{\mathscr{S}(\tilde{\nu}_i,\mathbf{V})}\boldsymbol{\alpha}^0\|}{\mu^\dagger - \tilde{\nu}_i}\right)^2,$$

$$V^t(C) = \left(\boldsymbol{\alpha}^0 - \frac{C}{\|\mathbf{V}\boldsymbol{\alpha}^0\|}\mathbf{V}\boldsymbol{\alpha}^0\right)^\top \mathbf{V}\left(\boldsymbol{\alpha}^0 - \frac{C}{\|\mathbf{V}\boldsymbol{\alpha}^0\|}\mathbf{V}\boldsymbol{\alpha}^0\right),$$

where $\sqrt{\sum_{i=1}^r \left(\frac{\tilde{\nu}_i \|\mathrm{Proj}_{\mathscr{S}(\tilde{\nu}_i,\mathbf{V})}\boldsymbol{\alpha}^0\|}{\mu^\dagger - \tilde{\nu}_i}\right)^2} = C$. In analogy to (63) and (67), we have

$$\begin{aligned}\Delta_v^t(C,\boldsymbol{\alpha}^0) &\equiv V^t(C) - V^*(C)\\ &= \begin{cases} 0 & \text{if } \boldsymbol{\alpha}^0 \in \bigcup_{i=1}^r \mathscr{S}(\tilde{\nu}_i,\mathbf{V}) \\ \frac{12\left[\|\mathbf{V}\boldsymbol{\alpha}^0\|^2\|\mathbf{V}^2\boldsymbol{\alpha}^0\|^2 - \|\mathbf{V}^{3/2}\boldsymbol{\alpha}^0\|^4\right]}{\|\mathbf{V}\boldsymbol{\alpha}^0\|^5} C^3 + o(C^3), & \text{if } \boldsymbol{\alpha}^0 \notin \bigcup_{i=1}^r \mathscr{S}(\tilde{\nu}_i,\mathbf{V}) \end{cases}.\end{aligned}$$

Appendix H. Proof of Proposition 6

- If $C \geqslant \overline{C}_t \equiv \mathscr{D}(\boldsymbol{\alpha}^0, \oplus_{i=1}^t \mathscr{G}(\tilde{\gamma}_i, \mathbf{W}, \mathbf{H}^{-1}))$, then $\oplus_{i=1}^t \mathscr{G}(\tilde{\gamma}_i, \mathbf{W}, \mathbf{H}^{-1}) \cap \mathcal{O}(\boldsymbol{\alpha}^0, C) \neq \emptyset$, where $\mathcal{O}(\boldsymbol{\alpha}^0, C) \equiv \{\boldsymbol{\alpha} : \|\boldsymbol{\alpha} - \boldsymbol{\alpha}^0\| \leqslant C\}$ denotes a sphere of radius C, centered at $\boldsymbol{\alpha}^0$. An arbitrary $\boldsymbol{\alpha} \in \oplus_{i=1}^t \mathscr{G}(\tilde{\gamma}_i, \mathbf{W}, \mathbf{H}^{-1}) \cap \mathcal{O}(\boldsymbol{\alpha}^0, C)$ can be decomposed uniquely as $\boldsymbol{\alpha} = \sum_{i=1}^t \boldsymbol{\alpha}^i$, where $\boldsymbol{\alpha}^i \in \mathscr{G}(\tilde{\gamma}_i, \mathbf{W}, \mathbf{H}^{-1}), \forall i$. It follows that

$$[\boldsymbol{\alpha}^i]^\top \mathbf{W}[\boldsymbol{\alpha}^j] = \begin{cases} \tilde{\gamma}_i [\boldsymbol{\alpha}^i]^\top \mathbf{H}^{-1}[\boldsymbol{\alpha}^i] & \text{if } \quad i = j \\ 0 & \text{if } \quad i \neq j \end{cases}.$$

Hence,

$$R^*(C) \geqslant \frac{\boldsymbol{\alpha}^\top \mathbf{W}\boldsymbol{\alpha}}{\boldsymbol{\alpha}^\top \mathbf{H}^{-1}\boldsymbol{\alpha}} = \frac{\sum_{i=1}^t \tilde{\gamma}_i [\boldsymbol{\alpha}^i]^\top \mathbf{H}^{-1}[\boldsymbol{\alpha}^i]}{\sum_{i=1}^t [\boldsymbol{\alpha}^i]^\top \mathbf{H}^{-1}[\boldsymbol{\alpha}^i]} \in [\tilde{\gamma}_t, \tilde{\gamma}_1].$$

In particular, if $C \geqslant \overline{C}_1 \equiv \mathscr{D}(\boldsymbol{\alpha}^0, \mathscr{G}(\tilde{\gamma}_1, \mathbf{W}, \mathbf{H}^{-1}))$, $R^*(C) = \tilde{\gamma}_1 = 1$.

- We next discuss the case with $C < \overline{C}_1$. Notice that in $[\mathcal{P}^r]$, only the direction of $\boldsymbol{\alpha}$ matters. Without loss of generality, we can scale up $\boldsymbol{\alpha}$ on a elliptical surface and optimize within a region $\{\tilde{\boldsymbol{\alpha}} : \tilde{\boldsymbol{\alpha}}^\top \mathbf{H}^{-1}\tilde{\boldsymbol{\alpha}} = 1, \exists \boldsymbol{\alpha} \in \mathcal{O}(\boldsymbol{\alpha}^0, C), \ni \boldsymbol{\alpha} \propto \tilde{\boldsymbol{\alpha}}\}$. This region can be regarded as the shadow region on elliptical surface $\tilde{\boldsymbol{\alpha}}^\top \mathbf{H}^{-1}\tilde{\boldsymbol{\alpha}} = 1$, in which all light rays emitting towards the origin are blocked by sphere $\mathcal{O}(\boldsymbol{\alpha}^0, C)$. Therefore, we can confine our searching within the intersection of an elliptical surface $\boldsymbol{\alpha}^\top \mathbf{H}^{-1}\boldsymbol{\alpha} = 1$ and a conical surface $\cos\langle\tilde{\boldsymbol{\alpha}}, \boldsymbol{\alpha}^0\rangle = \frac{\sqrt{\|\boldsymbol{\alpha}^0\|^2 - C^2}}{\|\boldsymbol{\alpha}^0\|}$. The original problem $[\mathcal{P}^r]$ is therefore transformed into

$$[\mathscr{P}^r] : \max_{\tilde{\boldsymbol{\alpha}}} \tilde{\boldsymbol{\alpha}}^\top \mathbf{W}\tilde{\boldsymbol{\alpha}}, s.t. : \tilde{\boldsymbol{\alpha}}^\top \mathbf{H}^{-1}\tilde{\boldsymbol{\alpha}} \leqslant 1, \tilde{\boldsymbol{\alpha}}^\top \boldsymbol{\alpha}^0 \geqslant \sqrt{\|\boldsymbol{\alpha}^0\|^2 - C^2}\|\tilde{\boldsymbol{\alpha}}\|. \tag{68}$$

With multipliers λ and μ, we can write the Lagrangian as:

$$\mathscr{L}(\tilde{\boldsymbol{\alpha}}, \lambda, \mu) = \tilde{\boldsymbol{\alpha}}^\top \mathbf{W}\tilde{\boldsymbol{\alpha}} + \lambda(1 - \tilde{\boldsymbol{\alpha}}^\top \mathbf{H}^{-1}\tilde{\boldsymbol{\alpha}}) + \mu\left(\tilde{\boldsymbol{\alpha}}^\top \boldsymbol{\alpha}^0 - \sqrt{\|\boldsymbol{\alpha}^0\|^2 - C^2}\sqrt{\tilde{\boldsymbol{\alpha}}^\top \tilde{\boldsymbol{\alpha}}}\right)$$

Following the first order condition

$$\mathscr{L}_{\tilde{\boldsymbol{\alpha}}} = 2\mathbf{W}\tilde{\boldsymbol{\alpha}} - 2\lambda\mathbf{H}^{-1}\tilde{\boldsymbol{\alpha}} + \mu\tilde{\boldsymbol{\alpha}} - \mu\sqrt{\frac{\|\boldsymbol{\alpha}^0\|^2 - C^2}{\|\tilde{\boldsymbol{\alpha}}\|^2}}\tilde{\boldsymbol{\alpha}} = 0$$

and using simplified notations

$$\xi \equiv \frac{2\lambda}{\mu\sqrt{\phi}}, \eta \equiv \frac{2}{\mu\sqrt{\phi}}, \phi \equiv \frac{\|\boldsymbol{\alpha}^0\|^2 - C^2}{\|\tilde{\boldsymbol{\alpha}}\|^2}, \Delta(C) \equiv \left[\mathbf{I} + \xi(C)\mathbf{H}^{-1} - \eta(C)\mathbf{W}\right]^{-1},$$

we obtain

$$\begin{aligned} \tilde{\boldsymbol{\alpha}}(C) &= \frac{1}{\sqrt{\phi(C)}}\Delta(C)\boldsymbol{\alpha}^0, && (69) \\ R^*(C) &= \frac{1}{\phi(C)}(\boldsymbol{\alpha}^0)^\top\Delta(C)\mathbf{W}\Delta(C)(\boldsymbol{\alpha}^0). && (70) \end{aligned}$$

Functions $(\xi(C), \eta(C), \phi(C)) \in \mathbb{R}_+^3$ are determined by equations

$$\begin{aligned} \phi &= \left[\boldsymbol{\alpha}^0\right]^\top \Delta(C)\mathbf{H}^{-1}\Delta(C)\left[\boldsymbol{\alpha}^0\right], && (71) \\ \left[\boldsymbol{\alpha}^0\right]^\top \Delta(C)\left[\boldsymbol{\alpha}^0\right] &= \|\boldsymbol{\alpha}^0\|^2 - C^2, && (72) \\ \left[\boldsymbol{\alpha}^0\right]^\top [\Delta(C)]^2\left[\boldsymbol{\alpha}^0\right] &= \|\boldsymbol{\alpha}^0\|^2 - C^2, && (73) \end{aligned}$$

which are implied by

$$[\tilde{\boldsymbol{\alpha}}]^\top\mathbf{H}^{-1}\tilde{\boldsymbol{\alpha}} = 1, (\boldsymbol{\alpha}^0)^\top\tilde{\boldsymbol{\alpha}} = \sqrt{\|\boldsymbol{\alpha}^0\|^2 - C^2}\|\tilde{\boldsymbol{\alpha}}\|, \text{ and } \phi\|\tilde{\boldsymbol{\alpha}}\|^2 = \|\boldsymbol{\alpha}^0\|^2 - C^2.$$

Accordingly, the optimal solution to the original problem $[\mathcal{P}^r]$ is obtained by regressing $\boldsymbol{\alpha}^0$ on $\tilde{\boldsymbol{\alpha}}$:

$$\boldsymbol{\alpha}^r(C) = \tilde{\boldsymbol{\alpha}}(r)\left[\tilde{\boldsymbol{\alpha}}^\top(r)\tilde{\boldsymbol{\alpha}}(r)\right]^{-1}\tilde{\boldsymbol{\alpha}}^\top(r)\boldsymbol{\alpha}^0 = \Delta(C)\boldsymbol{\alpha}^0. \quad (74)$$

Appendix I. Proof of Proposition 7

Evaluating (72) and (73) at $C = 0$, we have $\xi(0) = \eta(0) = 0$, and thus $\Delta(0) = \mathbf{I}$. Differentiating (72) and (73) with respect to C in consecutive orders yields:

$$\left[\boldsymbol{\alpha}^0\right]^\top \Delta^{(k)}(C)\left[\boldsymbol{\alpha}^0\right] = \begin{cases} -2C & k = 1 \\ -2 & k = 2 \\ 0 & k \geqslant 3 \end{cases}, \quad (75)$$

$$\left[\boldsymbol{\alpha}^0\right]^\top \left[\Delta^2(C)\right]^{(k)}\left[\boldsymbol{\alpha}^0\right] = \left[\boldsymbol{\alpha}^0\right]^\top \sum_{s=0}^{k} C_k^s \Delta^{(s)}(C)\Delta^{(k-s)}(C)\left[\boldsymbol{\alpha}^0\right] = \begin{cases} -2C & k = 1 \\ -2 & k = 2 \\ 0 & k \geqslant 3 \end{cases}. \quad (76)$$

Evaluating (75) and (76) at $C = 0$ and $k = 2, 3$, we have

$$\left[\boldsymbol{\alpha}^0\right]^\top \Delta''(0)(\boldsymbol{\alpha}^0) = -2,$$
$$\left[\boldsymbol{\alpha}^0\right]^\top \Delta^{(3)}(0)(\boldsymbol{\alpha}^0) = 0,$$
$$2\left[\boldsymbol{\alpha}^0\right]^\top \Delta(0)\Delta''(0)(\boldsymbol{\alpha}^0) + 2\left[\boldsymbol{\alpha}^0\right]^\top [\Delta'(0)]^2(\boldsymbol{\alpha}^0) = -2,$$
$$6\left[\boldsymbol{\alpha}^0\right]^\top \Delta'(0)\Delta''(0)(\boldsymbol{\alpha}^0) + 2\left[\boldsymbol{\alpha}^0\right]^\top \Delta(0)[\Delta^{(3)}(0)](\boldsymbol{\alpha}^0) = 0.$$

It follows that

$$\|\Delta'(0)\boldsymbol{\alpha}^0\|^2 = (\boldsymbol{\alpha}^0)^\top[\Delta'(0)]^2(\boldsymbol{\alpha}^0) = 1, \tag{77}$$

and

$$(\boldsymbol{\alpha}^0)^\top\Delta'(0)\Delta''(0)(\boldsymbol{\alpha}^0) = 0. \tag{78}$$

So, the tangent intervention is

$$\boldsymbol{\alpha}^t(C) = \boldsymbol{\alpha}^0 + \frac{d\boldsymbol{\alpha}^r(0)/dC}{\|d\boldsymbol{\alpha}^r(0)/dC\|}C = \left[\mathbf{I} + \Delta'(0)\right]\boldsymbol{\alpha}^0. \tag{79}$$

$(\boldsymbol{\alpha}^0)^\top\Delta'(0)(\boldsymbol{\alpha}^0) = 0$ implies the proportionality condition

$$\frac{\eta'(0)}{(\boldsymbol{\alpha}^0)^\top\mathbf{H}^{-1}\boldsymbol{\alpha}^0} = \frac{\xi'(0)}{(\boldsymbol{\alpha}^0)^\top\mathbf{W}\boldsymbol{\alpha}^0} \tag{80}$$

and thus $\Delta'(0)\boldsymbol{\alpha}^0$ is computed as:

$$\Delta'(0)\boldsymbol{\alpha}^0 = \left[\eta'(0)\mathbf{W} - \xi'(0)\mathbf{H}^{-1}\right]\boldsymbol{\alpha}^0 = \frac{\left((\boldsymbol{\alpha}^0)^\top\mathbf{H}^{-1}(\boldsymbol{\alpha}^0)\mathbf{W} - (\boldsymbol{\alpha}^0)^\top\mathbf{W}(\boldsymbol{\alpha}^0)\mathbf{H}^{-1}\right)\boldsymbol{\alpha}^0}{\|\left((\boldsymbol{\alpha}^0)^\top\mathbf{H}^{-1}(\boldsymbol{\alpha}^0)\mathbf{W} - (\boldsymbol{\alpha}^0)^\top\mathbf{W}(\boldsymbol{\alpha}^0)\mathbf{H}^{-1}\right)\boldsymbol{\alpha}^0\|}. \tag{81}$$

Expanding $\boldsymbol{\alpha}^r(C)$, to the cubic term, around zero yields

$$\boldsymbol{\alpha}^r(C) = \left[\mathbf{I} + \Delta'(0)C + \frac{1}{2}\Delta''(0)C^2 + \frac{1}{6}\Delta^{(3)}(0)C^3 + o(C^3)\right]\boldsymbol{\alpha}^0. \tag{82}$$

Using (79) and (82), we have

$$\begin{aligned}
[\boldsymbol{\alpha}^r(C)]^\top\mathbf{W}[\boldsymbol{\alpha}^r(C)] &= \left[\begin{array}{c}(\boldsymbol{\alpha}^0)^\top\mathbf{W}(\boldsymbol{\alpha}^0) + 2(\boldsymbol{\alpha}^0)^\top\Delta'(0)\mathbf{W}(\boldsymbol{\alpha}^0)C \\ +(\boldsymbol{\alpha}^0)^\top\left[\Delta'(0)\mathbf{W}\Delta'(0) + \Delta''(0)\mathbf{W}\right](\boldsymbol{\alpha}^0)C^2\end{array}\right] + o(C^2), \\
[\boldsymbol{\alpha}^r(C)]^\top\mathbf{H}^{-1}[\boldsymbol{\alpha}^r(C)] &= \left[\begin{array}{c}(\boldsymbol{\alpha}^0)^\top\mathbf{H}^{-1}(\boldsymbol{\alpha}^0) + 2(\boldsymbol{\alpha}^0)^\top\Delta'(0)\mathbf{H}^{-1}(\boldsymbol{\alpha}^0)C \\ +(\boldsymbol{\alpha}^0)^\top\left[\Delta'(0)\mathbf{H}^{-1}\Delta'(0) + \Delta''(0)\mathbf{H}^{-1}\right](\boldsymbol{\alpha}^0)C^2\end{array}\right] + o(C^2), \\
\left[\boldsymbol{\alpha}^t(C)\right]^\top\mathbf{W}\left[\boldsymbol{\alpha}^t(C)\right] &= \left[\begin{array}{c}(\boldsymbol{\alpha}^0)^\top\mathbf{W}(\boldsymbol{\alpha}^0) + 2(\boldsymbol{\alpha}^0)^\top\Delta'(0)\mathbf{W}(\boldsymbol{\alpha}^0)C \\ +(\boldsymbol{\alpha}^0)^\top\Delta'(0)\mathbf{W}\Delta'(0)(\boldsymbol{\alpha}^0)C^2\end{array}\right], \\
\left[\boldsymbol{\alpha}^t(C)\right]^\top\mathbf{H}^{-1}\left[\boldsymbol{\alpha}^t(C)\right] &= \left[\begin{array}{c}(\boldsymbol{\alpha}^0)^\top\mathbf{H}^{-1}(\boldsymbol{\alpha}^0) + 2(\boldsymbol{\alpha}^0)^\top\Delta'(0)\mathbf{H}^{-1}(\boldsymbol{\alpha}^0)C \\ +(\boldsymbol{\alpha}^0)^\top\Delta'(0)\mathbf{H}^{-1}\Delta'(0)(\boldsymbol{\alpha}^0)C^2\end{array}\right].
\end{aligned}$$

Hence, the error of tangent approximation is

$$\begin{aligned}\Delta R^t(C) &\equiv R^*(C) - R^t(C) \\ &= \frac{[\boldsymbol{\alpha}^r(C)]^\top \mathbf{W}[\boldsymbol{\alpha}^r(C)]}{[\boldsymbol{\alpha}^r(C)]^\top \mathbf{H}^{-1}[\boldsymbol{\alpha}^r(C)]} - \frac{[\boldsymbol{\alpha}^t(C)]^\top \mathbf{W}[\boldsymbol{\alpha}^t(C)]}{[\boldsymbol{\alpha}^t(C)]^\top \mathbf{H}^{-1}[\boldsymbol{\alpha}^t(C)]} \\ &= \frac{\rho(C)}{[\boldsymbol{\alpha}^r(C)]^\top \mathbf{H}^{-1}[\boldsymbol{\alpha}^r(C)] \times [\boldsymbol{\alpha}^t(C)]^\top \mathbf{H}^{-1}[\boldsymbol{\alpha}^t(C)]},\end{aligned}$$

where

$$\begin{aligned}\rho(C) &\equiv \Big[[\boldsymbol{\alpha}^r(C)]^\top \mathbf{W}[\boldsymbol{\alpha}^r(C)] \cdot [\boldsymbol{\alpha}^t(C)]^\top \mathbf{H}^{-1}[\boldsymbol{\alpha}^t(C)] - [\boldsymbol{\alpha}^r(C)]^\top \mathbf{H}^{-1}[\boldsymbol{\alpha}^r(C)] \cdot [\boldsymbol{\alpha}^t(C)]^\top \mathbf{W}[\boldsymbol{\alpha}^t(C)]\Big] \\ &= \Big[(\boldsymbol{\alpha}^0)^\top \Delta''(0)\mathbf{W}(\boldsymbol{\alpha}^0) \cdot (\boldsymbol{\alpha}^0)^\top \mathbf{H}^{-1}(\boldsymbol{\alpha}^0) - (\boldsymbol{\alpha}^0)^\top \Delta''(0)\mathbf{H}^{-1}(\boldsymbol{\alpha}^0) \cdot (\boldsymbol{\alpha}^0)^\top \mathbf{W}(\boldsymbol{\alpha}^0)\Big] C^2 + o(C^2).\end{aligned}$$

The coefficient of C^2 in $\rho(C)$ is

$$\begin{aligned}&(\boldsymbol{\alpha}^0)^\top \Delta''(0)\mathbf{W}(\boldsymbol{\alpha}^0) \times (\boldsymbol{\alpha}^0)^\top \mathbf{H}^{-1}(\boldsymbol{\alpha}^0) - (\boldsymbol{\alpha}^0)^\top \Delta''(0)\mathbf{H}^{-1}(\boldsymbol{\alpha}^0) \times (\boldsymbol{\alpha}^0)^\top \mathbf{W}(\boldsymbol{\alpha}^0) \\ = \; &\frac{1}{t}(\boldsymbol{\alpha}^0)^\top \Delta''(0)\left(\eta'(0)\mathbf{W} - \xi'(0)\mathbf{H}^{-1}\right)\boldsymbol{\alpha}^0 = \frac{1}{t}(\boldsymbol{\alpha}^0)^\top \Delta''(0)\Delta'(0)\boldsymbol{\alpha}^0 = 0,\end{aligned}$$

where $t \equiv \eta'(0)/(\boldsymbol{\alpha}^0)^\top \mathbf{H}^{-1}\boldsymbol{\alpha}^0 = \xi'(0)/(\boldsymbol{\alpha}^0)^\top \mathbf{W}\boldsymbol{\alpha}^0$, the last equality follows from (78). As a result, $\Delta R^t(C) = o(C^2)$.

Appendix J. Proof of Proposition 8

- First, we show that matrices $\mathbf{W}$ and $\mathbf{H}$ commute, i.e., $\mathbf{WH} = \mathbf{HW}$. It is obvious for a regular network $\mathbf{G} \in \mathscr{R}$. Next, we show this result holds for a complete bipartite network. If $\mathbf{G} \in \mathscr{K}_{p,q}$, we have

$$\mathbf{G} = \begin{bmatrix} \mathbf{0} & \mathbf{1}_p\mathbf{1}_q^\top \\ \mathbf{1}_q\mathbf{1}_p^\top & \mathbf{0} \end{bmatrix} \text{ and } \widehat{\mathbf{G}} = \begin{bmatrix} \mathbf{0} & \frac{1}{q}\mathbf{1}_p\mathbf{1}_q^\top \\ \frac{1}{p}\mathbf{1}_q\mathbf{1}_p^\top & \mathbf{0} \end{bmatrix},$$

where $\mathbf{1}_s$ denotes an $s-$dimensional all-ones vector.

$$\begin{aligned}\widehat{\mathbf{M}} &= (1-\lambda)\begin{bmatrix} \mathbf{I}_p & -\frac{\lambda}{q}\mathbf{1}_p\mathbf{1}_q^\top \\ -\frac{\lambda}{p}\mathbf{1}_q\mathbf{1}_p^\top & \mathbf{I}_q \end{bmatrix}^{-1} &(83)\\ &= (1-\lambda)\begin{bmatrix} \frac{\lambda^2}{pq}\mathbf{1}_p\mathbf{1}_q^\top\left(\mathbf{I}_q - \frac{\lambda^2}{q}\mathbf{1}_q\mathbf{1}_q^\top\right)^{-1}\mathbf{1}_q\mathbf{1}_p^\top & -\frac{\lambda}{q}\mathbf{1}_p\mathbf{1}_q^\top\left(\mathbf{I}_q - \frac{\lambda^2}{q}\mathbf{1}_q\mathbf{1}_q^\top\right)^{-1} \\ -\frac{\lambda}{p}\left(\mathbf{I}_q - \frac{\lambda^2}{q}\mathbf{1}_q\mathbf{1}_q^\top\right)^{-1}\mathbf{1}_q\mathbf{1}_p^\top & \left(\mathbf{I}_q - \frac{\lambda^2}{q}\mathbf{1}_q\mathbf{1}_q^\top\right)^{-1} \end{bmatrix} &(84)\\ &= (1-\lambda)\begin{bmatrix} \mathbf{I}_p + \frac{\lambda^2}{p(1-\lambda^2)}\mathbf{1}_p\mathbf{1}_p^\top & \frac{\lambda}{q(1-\lambda^2)}\mathbf{1}_p\mathbf{1}_q^\top \\ \frac{\lambda}{p(1-\lambda^2)}\mathbf{1}_q\mathbf{1}_p^\top & \mathbf{I}_q + \frac{\lambda^2}{q(1-\lambda^2)}\mathbf{1}_q\mathbf{1}_q^\top \end{bmatrix}, &(85)\end{aligned}$$

where $\mathbf{I}_s$ denotes the $s \times s$ identity matrix, (83) $\Rightarrow$ (84) follows from the partitioned inverse formula of square matrix,[19] (84) $\Rightarrow$ (85) follows from

$$\left(\mathbf{I}_q - \frac{\lambda^2}{q}\mathbf{1}_q\mathbf{1}_q^\top\right)^{-1} = \mathbf{I}_q + \frac{\lambda^2}{q(1-\lambda^2)}\mathbf{1}_q\mathbf{1}_q^\top,$$

which is implied by the Sherman-Morrison formula.[20] Then we have $(\mathbf{I}-\widehat{\mathbf{M}})^\top(\mathbf{I}-\widehat{\mathbf{M}}) = \lambda^2\mathbf{I}+\mathbf{A}$ and $(\mathbf{I}-\widehat{\mathbf{G}})^\top(\mathbf{I}-\widehat{\mathbf{G}}) = \mathbf{I}+\mathbf{B}$, where

$$\mathbf{A} \equiv \begin{bmatrix} \alpha\mathbf{1}_p\mathbf{1}_p^\top & \beta\mathbf{1}_p\mathbf{1}_q^\top \\ \beta\mathbf{1}_q\mathbf{1}_p^\top & \gamma\mathbf{1}_q\mathbf{1}_q^\top \end{bmatrix}, \mathbf{B} \equiv \begin{bmatrix} a\mathbf{1}_p\mathbf{1}_p^\top & b\mathbf{1}_p\mathbf{1}_q^\top \\ b\mathbf{1}_q\mathbf{1}_p^\top & c\mathbf{1}_q\mathbf{1}_q^\top \end{bmatrix},$$

$\alpha \equiv \frac{\lambda^2[q-p\lambda(2+\lambda)]}{p^2(1+\lambda)^2}; \beta \equiv -\frac{(p+q)\lambda^2}{pq(1+\lambda)^2}; \gamma \equiv \frac{\lambda^2[p-q\lambda(2+\lambda)]}{q^2(1+\lambda)^2}; a \equiv \frac{q}{p^2}, b \equiv -\frac{p+q}{pq}, c \equiv \frac{p}{q^2}$. Since

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} (\alpha ap+\beta bq)\,\mathbf{1}_p\mathbf{1}_p^\top & (\alpha bp+\beta cq)\,\mathbf{1}_p\mathbf{1}_q^\top \\ (a\beta p+b\gamma q)\,\mathbf{1}_q\mathbf{1}_p^\top & (\beta bp+\gamma cq)\,\mathbf{1}_q\mathbf{1}_q^\top \end{bmatrix}, \\ \mathbf{BA} &= \begin{bmatrix} (\alpha ap+\beta bq)\,\mathbf{1}_p\mathbf{1}_p^\top & (a\beta p+b\gamma q)\,\mathbf{1}_p\mathbf{1}_q^\top \\ (\alpha bp+\beta cq)\,\mathbf{1}_q\mathbf{1}_p^\top & (\beta bp+\gamma cq)\,\mathbf{1}_q\mathbf{1}_q^\top \end{bmatrix}, \end{aligned}$$

and $\alpha bp + \beta cq = a\beta p + b\gamma q = \frac{-(p+q)\lambda^2\left[p^2+q^2-\lambda pq(2+\lambda)\right]}{p^2q^2(1+\lambda)^2}$, we have $\mathbf{AB} = \mathbf{BA}$. So $(\mathbf{I}-\widehat{\mathbf{M}})^\top(\mathbf{I}-\widehat{\mathbf{M}})$ and $(\mathbf{I}-\widehat{\mathbf{G}})^\top(\mathbf{I}-\widehat{\mathbf{G}})$, and thus $\mathbf{W}$ and $\mathbf{H}$ commute.

- Since symmetric matrices $\mathbf{W}$ and $\mathbf{H}$ commute, they can be simultaneously diagonalized by a common orthogonal matrix $\mathbb{S} \equiv (\mathbb{s}_{\cdot j})_{j\in I}$, the jth column $\mathbb{s}_{\cdot j}$ is an orthonormal eigenvector, corresponding to γ_j, the jth eigenvalue of $\mathbf{\Gamma} \equiv \mathbf{H}^{1/2}\mathbf{W}\mathbf{H}^{1/2}$.[21] We denote by $\widehat{\mathbf{\Gamma}} \equiv \mathbb{S}^\top\mathbf{\Gamma}\mathbb{S} = diag\{\gamma_1, \cdots, \gamma_n\}$, $\widehat{\mathbf{H}} \equiv \mathbb{S}^\top\mathbf{H}\mathbb{S}$. With transformations $\mathbf{z} \equiv \mathbb{S}^\top\mathbf{H}^{-1/2}\boldsymbol{\alpha}$ and

[19]If an $n \times n$ matrix $\mathbf{\Sigma}$ is partitioned as

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}$$

and $\mathbf{\Sigma}_{11}$ is nonsingular, where $\mathbf{\Sigma}_{ij} \in \mathbb{R}^{n_i \times n_j}, n_1 + n_2 = n$, then we have

$$\mathbf{\Sigma}^{-1} = \begin{bmatrix} \mathbf{\Sigma}_{11}^{-1} + \mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22.1}^{-1}\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1} & -\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22.1}^{-1} \\ -\mathbf{\Sigma}_{22.1}^{-1}\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1} & \mathbf{\Sigma}_{22.1}^{-1} \end{bmatrix}.$$

[20]Suppose $\mathbf{A} \in \mathbb{R}^{n\times n}$ is a nonsingular matrix, $\mathbf{u}$ and $\mathbf{v}$ are $n-$dimensional column vectors, $1 + \mathbf{v}^\top\mathbf{A}^{-1}\mathbf{u} \neq 0$, then we have

$$(\mathbf{A}+\mathbf{u}\mathbf{v}^\top)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^\top\mathbf{A}^{-1}}{1+\mathbf{v}^\top\mathbf{A}^{-1}\mathbf{u}}.$$

[21]If $\mathbf{G} \in \mathscr{R}$, matrices $\mathbf{G}, \mathbf{W}$ and $\mathbf{H}$ are simultaneously diagonalizable. Matrices $\mathbb{S}$ and $\mathbf{S}$ (see Appendix C) consist of the same set of column vectors, but their orderings are not necessarily the same.

$\mathbf{z}^0 \equiv \mathbb{S}^\top \mathbf{H}^{-1/2}\boldsymbol{\alpha}^0$, the original problem $[\mathcal{P}^r]$ is rewritten as

$$[\mathscr{P}^r] : \max_{\mathbf{z}} \frac{\mathbf{z}^\top \widehat{\boldsymbol{\Gamma}} \mathbf{z}}{\mathbf{z}^\top \mathbf{z}}, s.t. : (\mathbf{z} - \mathbf{z}^0)^\top \widehat{\mathbf{H}} (\mathbf{z} - \mathbf{z}^0) \leqslant C^2. \tag{86}$$

Since $\boldsymbol{\alpha}^0 \in \oplus_{i=1}^t \mathscr{G}(\tilde{\gamma}_i, \mathbf{W}, \mathbf{H}^{-1})$, we have $\mathbf{z}^0 \equiv \mathbb{S}^\top \mathbf{H}^{-1/2}\boldsymbol{\alpha}^0 = (z_j^0)_{j\in I} \in \oplus_{i=1}^t \mathscr{S}(\tilde{\gamma}_i, \widehat{\boldsymbol{\Gamma}})$ by Lemma 4. That is, $z_j^0 = 0, \forall j \notin K \equiv \cup_{i=1}^t N_i$. Problem $[\mathscr{P}^r]$ is then transformed into:

$$\max_{\mathbf{z}} \varphi(\mathbf{z}), s.t. : \left(\mathbf{z}_K - \mathbf{z}_K^0\right)^\top \widehat{\mathbf{H}}_{K,K} \left(\mathbf{z}_K - \mathbf{z}_K^0\right) \leqslant C^2 - \mathbf{z}_{-K}^\top \widehat{\mathbf{H}}_{-K,-K} \mathbf{z}_{-K}, \tag{87}$$

where $\varphi(\mathbf{z}) \equiv \mathbf{z}^\top \widehat{\boldsymbol{\Gamma}} \mathbf{z} / \mathbf{z}^\top \mathbf{z}$, notations $\mathbf{x}_I$ and $\mathbf{A}_{I,J}$ represent $I-$subvector of vector $\mathbf{x}$ and $(I,J)-$block of matrix $\mathbf{A}$.

Suppose that the optimal solution $\mathbf{z}^*$ to (87) satisfies $z_\ell^* \neq 0$ for some $\ell \in I \backslash K$. Construct a vector $\tilde{\mathbf{z}} \equiv (\tilde{z}_j)_{j \in I}$ as:

$$\tilde{z}_j = \begin{cases} z_j^* & \forall j \in K \\ 0 & \forall j \notin K \end{cases}.$$

Then $\tilde{\mathbf{z}}$ satisfies constraint in (87):

$$\begin{aligned} \left(\tilde{\mathbf{z}}_K - \mathbf{z}_K^0\right)^\top \widehat{\mathbf{H}}_{K,K} \left(\tilde{\mathbf{z}}_K - \mathbf{z}_K^0\right) &= \left(\mathbf{z}_K^* - \mathbf{z}_K^0\right)^\top \widehat{\mathbf{H}}_{K,K} \left(\mathbf{z}_K^* - \mathbf{z}_K^0\right) \\ &\leqslant C^2 - (\mathbf{z}_{-K}^*)^\top \widehat{\mathbf{H}}_{-K,-K} \mathbf{z}_{-K}^* \\ &< C^2 = C^2 - \tilde{\mathbf{z}}_{-K}^\top \widehat{\mathbf{H}}_{-K,-K} \tilde{\mathbf{z}}_{-K}. \end{aligned}$$

Since $\|\mathbf{z}_{-K}^*\| \geqslant |z_\ell^*| > 0$, we have

$$\begin{aligned} \varphi(\tilde{\mathbf{z}}) - \varphi(\mathbf{z}^*) &= \frac{(\mathbf{z}_K^*)^\top \widehat{\boldsymbol{\Gamma}}_{K,K} \mathbf{z}_K^*}{\|\mathbf{z}_K^*\|^2} - \frac{(\mathbf{z}_K^*)^\top \widehat{\boldsymbol{\Gamma}}_{K,K} \mathbf{z}_K^* + (\mathbf{z}_{-K}^*)^\top \widehat{\boldsymbol{\Gamma}}_{-K,-K} \mathbf{z}_{-K}^*}{\|\mathbf{z}_K^*\|^2 + \|\mathbf{z}_{-K}^*\|^2} \\ &= \frac{\|\mathbf{z}_{-K}^*\|^2}{\|\mathbf{z}_K^*\|^2 + \|\mathbf{z}_{-K}^*\|^2} \left[\underbrace{\frac{(\mathbf{z}_K^*)^\top \widehat{\boldsymbol{\Gamma}}_{K,K} \mathbf{z}_K^*}{(\mathbf{z}_K^*)^\top \mathbf{z}_K^*}}_{=ⓐ} - \underbrace{\frac{(\mathbf{z}_{-K}^*)^\top \widehat{\boldsymbol{\Gamma}}_{-K,-K} \mathbf{z}_{-K}^*}{(\mathbf{z}_{-K}^*)^\top \mathbf{z}_{-K}^*}}_{=ⓑ} \right] > 0. \end{aligned}$$

The last inequality follows from ⓐ $\geqslant \min_{1 \leqslant i \leqslant t} \tilde{\gamma}_i >$ ⓑ. This contradicts the fact that $\mathbf{z}^*$ is the optimal solution. Therefore, we have $z_j^* = 0, \forall j \in I \backslash K$. That is, $\mathbf{z}^* \in \oplus_{i=1}^t \mathscr{S}(\tilde{\gamma}_i, \widehat{\boldsymbol{\Gamma}})$. By Lemma 4, we have $\boldsymbol{\alpha}^r(C) = \mathbf{H}^{1/2} \mathbb{S} \mathbf{z}^* \in \oplus_{i=1}^t \mathscr{G}(\tilde{\gamma}_i, \mathbf{W}, \mathbf{H}^{-1})$.

Lemma 4 *$\boldsymbol{\alpha} \in \oplus_{i=1}^t \mathscr{G}(\tilde{\gamma}_i, \boldsymbol{W}, \boldsymbol{H}^{-1})$ if and only if $\boldsymbol{z} \equiv \mathbb{S}^\top \boldsymbol{H}^{-1/2} \boldsymbol{\alpha} \in \oplus_{i=1}^t \mathscr{S}(\tilde{\gamma}_i, \boldsymbol{\Gamma})$.*

Proof. (i) $\boldsymbol{\alpha} \in \oplus_{i=1}^t \mathscr{G}(\tilde{\gamma}_i, \mathbf{W}, \mathbf{H}^{-1})$ implies $\boldsymbol{\alpha} = \sum_{i=1}^t \boldsymbol{\alpha}^i, \boldsymbol{\alpha}^i \in \mathscr{G}(\tilde{\gamma}_i, \mathbf{W}, \mathbf{H}^{-1})$, i.e., $\mathbf{W}\boldsymbol{\alpha}^i = \tilde{\gamma}_i \mathbf{H}^{-1} \boldsymbol{\alpha}^i$. It follows that $\widehat{\boldsymbol{\Gamma}} \mathbb{S}^\top \mathbf{H}^{-1/2} \boldsymbol{\alpha}^i = \tilde{\gamma}_i \mathbb{S}^\top \mathbf{H}^{-1/2} \boldsymbol{\alpha}^i$, then $\mathbb{S}^\top \mathbf{H}^{-1/2} \boldsymbol{\alpha}^i \in \mathscr{S}(\tilde{\gamma}_i, \widehat{\boldsymbol{\Gamma}})$. As a result, $\mathbf{z} \equiv \mathbb{S}^\top \mathbf{H}^{-1/2} \boldsymbol{\alpha} = \sum_{i=1}^t \mathbb{S}^\top \mathbf{H}^{-1/2} \boldsymbol{\alpha}^i \in \oplus_{i=1}^t \mathscr{S}(\tilde{\gamma}_i, \widehat{\boldsymbol{\Gamma}})$. (ii) Conversely, if $\mathbf{z} \equiv \mathbb{S}^\top \mathbf{H}^{-1/2} \boldsymbol{\alpha} \in$

$\oplus_{i=1}^{t}\mathscr{S}(\tilde{\gamma}_i,\widehat{\mathbf{\Gamma}})$. It can be decomposed uniquely as $\mathbf{z}=\sum_{i=1}^{t}\mathbb{S}^\top\mathbf{H}^{-1/2}\boldsymbol{\alpha}^i$, $\mathbb{S}^\top\mathbf{H}^{-1/2}\boldsymbol{\alpha}^i\in\mathscr{S}(\tilde{\gamma}_i,\widehat{\mathbf{\Gamma}})$, i.e., $\widehat{\mathbf{\Gamma}}\mathbb{S}^\top\mathbf{H}^{-1/2}\boldsymbol{\alpha}^i=\tilde{\gamma}_i\mathbb{S}^\top\mathbf{H}^{-1/2}\boldsymbol{\alpha}^i$. It follows that $\mathbf{W}\boldsymbol{\alpha}^i=\tilde{\gamma}_i\mathbf{H}^{-1}\boldsymbol{\alpha}^i$, then $\boldsymbol{\alpha}^i\in\mathscr{G}(\tilde{\gamma}_i,\mathbf{W},\mathbf{H}^{-1})$. As a result, $\boldsymbol{\alpha}\equiv\sum_{i=1}^{t}\boldsymbol{\alpha}^i\in\oplus_{i=1}^{t}\mathscr{G}(\tilde{\gamma}_i,\mathbf{W},\mathbf{H}^{-1})$.

■

Appendix K. Proof of Lemma 1

Any $\mathbf{x}\in[\mathbb{N}_\ell(\mathbf{G})\cap\mathbb{N}_r(\mathbf{G})]\oplus\{\mathbf{x}|\mathbf{x}\propto\mathbf{1}\}$ can be decomposed as $\mathbf{x}=\mathbf{x}^1+\mathbf{x}^2$, with $\widehat{\mathbf{G}}\mathbf{x}^1=\widehat{\mathbf{G}}^\top\mathbf{x}^1=0$ and $\mathbf{x}^2\propto\mathbf{1}$. Hence, $\widehat{\mathbf{M}}\mathbf{x}^1=\widehat{\mathbf{M}}^\top\mathbf{x}^1=(1-\lambda)\mathbf{x}^1$. It follows that $\mathbf{W}\mathbf{x}^1=\mathbf{H}^{-1}\mathbf{x}^1=(1-\lambda)\mathbf{x}^1$, and $\mathbf{W}\mathbf{x}^2=\mathbf{H}^{-1}\mathbf{x}^2=\mathbf{x}^2$. So we have $\mathbf{W}\mathbf{x}=\mathbf{H}^{-1}\mathbf{x},\forall\lambda\in(0,1)$. That is,

$$[\mathbb{N}_\ell(\mathbf{G})\cap\mathbb{N}_r(\mathbf{G})]\oplus\{\mathbf{x}|\mathbf{x}\propto\mathbf{1}\}\subseteq\bigcap_{\lambda\in(0,1)}\mathscr{G}(1,\mathbf{W}(\lambda),\mathbf{H}^{-1}(\lambda)). \tag{88}$$

Expanding matrix functions $\mathbf{W}(\lambda)$ and $\mathbf{H}^{-1}(\lambda)$ around zero yields:

$$\mathbf{W}(\lambda)=\mathbf{I}-\left(\mathbf{I}-\widehat{\mathbf{G}}\right)^\top\sum_{s=1}^{\infty}\sum_{p+q=s-1}\left(\widehat{\mathbf{G}}^\top\right)^p\left(\widehat{\mathbf{G}}\right)^q\left(\mathbf{I}-\widehat{\mathbf{G}}\right)\lambda^s$$
$$\mathbf{H}^{-1}(\lambda)=\mathbf{I}-\left(\mathbf{I}-\widehat{\mathbf{G}}\right)^\top\sum_{s=1}^{\infty}\left(\widehat{\mathbf{G}}^\top+\widehat{\mathbf{G}}-\widehat{\mathbf{G}}^\top\widehat{\mathbf{G}}\right)^{s-1}\left(\mathbf{I}-\widehat{\mathbf{G}}\right)\lambda^s$$

Hence,

$$\mathbf{x}^\top\left[\mathbf{H}^{-1}(\lambda)-\mathbf{W}(\lambda)\right]\mathbf{x}=\left\|(\mathbf{I}-\widehat{\mathbf{G}})\widehat{\mathbf{G}}\mathbf{x}\right\|^2\lambda^2+o(\lambda^2).$$

If $\mathbf{x}\notin\mathbb{N}_r(\mathbf{G})\oplus\{\mathbf{x}|\mathbf{x}\propto\mathbf{1}\}$, we have $\|(\mathbf{I}-\widehat{\mathbf{G}})\widehat{\mathbf{G}}\mathbf{x}\|>0$, then $\mathbf{x}^\top\left[\mathbf{H}^{-1}(\lambda)-\mathbf{W}(\lambda)\right]\mathbf{x}>0$, and thus $\mathbf{x}\notin\mathscr{G}(1,\mathbf{W}(\lambda),\mathbf{H}^{-1}(\lambda))$ whenever λ is close to zero. It follow that

$$\bigcap_{\lambda\in(0,1)}\mathscr{G}\left(1,\mathbf{W}(\lambda),\mathbf{H}^{-1}(\lambda)\right)\subseteq\mathbb{N}_r(\mathbf{G})\oplus\{\mathbf{x}|\mathbf{x}\propto\mathbf{1}\}. \tag{89}$$

The proof is completed.

Appendix L. Proof of Lemma 2

First, we show that

$$\mathbb{N}_\ell(\widehat{\mathbf{G}})=\mathbb{N}_r(\widehat{\mathbf{G}})=\mathbb{N}(\mathbf{G})\equiv\{\mathbf{x}|\mathbf{G}\mathbf{x}=0\} \tag{90}$$

for $\mathbf{G}\in\mathscr{R}\cup\mathscr{K}_{p_1,\cdots,p_s}$. It is obvious for $\mathbf{G}\in\mathscr{R}$. If $\mathbf{G}\in\mathscr{K}_{p_1,\cdots,p_s}$, the adjacency matrices $\mathbf{G}$ and $\widehat{\mathbf{G}}$ are

$$\mathbf{G}=\left[\begin{array}{cccc} O_{p_1} & \mathbf{1}_{p_1}\mathbf{1}_{p_2}^\top & \cdots & \mathbf{1}_{p_1}\mathbf{1}_{p_s}^\top \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{1}_{p_s}\mathbf{1}_{p_1}^\top & \mathbf{1}_{p_s}\mathbf{1}_{p_2}^\top & \cdots & O_{p_s} \end{array}\right]$$

and

$$\widehat{\mathbf{G}} = \begin{bmatrix} O_{p_1} & \frac{1}{n-p_1}\mathbf{1}_{p_1}\mathbf{1}_{p_2}^\top & \cdots & \frac{1}{n-p_1}\mathbf{1}_{p_1}\mathbf{1}_{p_s}^\top \\ \cdots & \cdots & \cdots & \cdots \\ \frac{1}{n-p_s}\mathbf{1}_{p_s}\mathbf{1}_{p_1}^\top & \frac{1}{n-p_s}\mathbf{1}_{p_s}\mathbf{1}_{p_2}^\top & \cdots & O_{p_s} \end{bmatrix},$$

where O_p denotes a $p \times p$ all-zeros matrix, $\mathbf{1}_q$ is a $q-$dimensional all-ones vector. Suppose that

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_s \end{bmatrix} \in \mathbb{N}_r(\mathbf{G}),$$

$\{\mathbf{x}_j\}_{j=1}^s$ denote s consecutive p_j-subvectors of $\mathbf{x}$. $\widehat{\mathbf{G}}\mathbf{x} = 0$ implies

$$\frac{1}{n-p_i}\sum_{j\neq i}\delta_j = 0, \forall i = 1, \cdots, s, \tag{91}$$

where $\delta_j \equiv \mathbf{1}_{p_j}^\top \mathbf{x}_j$. It follows that $\delta_i = 0, \forall i$, then

$$\mathbf{G}\mathbf{x} = \left(\sum_{j\neq i}\delta_j \mathbf{1}_{p_i}\right)_{i=1}^s = \mathbf{0}, \widehat{\mathbf{G}}^\top \mathbf{x} = \left(\sum_{j\neq i}\frac{\delta_j}{n-p_j}\mathbf{1}_{p_i}\right)_{i=1}^s = \mathbf{0}.$$

Conversely, if $\widehat{\mathbf{G}}^\top \mathbf{x} = 0$, then $\sum_{j\neq i}\frac{\delta_j}{n-p_j} = 0, \forall i = 1, \cdots, s$. This implies $\delta_j = 0, \forall j$. So $\widehat{\mathbf{G}}\mathbf{x} = \mathbf{G}\mathbf{x} = \mathbf{0}$. As a result, $\mathbb{N}_\ell(\widehat{\mathbf{G}}) = \mathbb{N}_r(\widehat{\mathbf{G}}) = \mathbb{N}(\mathbf{G})$. It follows directly from Lemma 1 that

$$\bigcap_{\lambda\in(0,1)} \mathscr{G}(1, \mathbf{W}(\lambda), \mathbf{H}^{-1}(\lambda)) = \mathbb{N}(G) \oplus \{\mathbf{x}\,|\mathbf{x} \propto 1\}, \forall \mathbf{G} \in \mathscr{R} \cup \mathscr{K}_{p_1,\cdots,p_s}.$$

Appendix M. Proof of Proposition 8

Following Lemmas 1 and 2, we have:

$$\begin{aligned} & \mathscr{D}\left(\boldsymbol{\alpha}^0, \bigcap_{\lambda\in(0,1)} \mathscr{G}(1, \mathbf{W}(\lambda), \mathbf{H}^{-1}(\lambda))\right) \\ = & \left\|\boldsymbol{\alpha}^0 - \mathrm{Proj}_{\mathscr{S}(0,\mathbf{G})\oplus\{\mathbf{x}|\mathbf{x}\propto 1\}}\boldsymbol{\alpha}^0\right\| \\ = & \sqrt{(\boldsymbol{\alpha}^0)^\top\left(\mathbf{I} - \mathbf{s}_{\cdot 1}\mathbf{s}_{\cdot 1}^\top - \sum_{j\in N_0}\mathbf{s}_{\cdot j}\mathbf{s}_{\cdot j}^\top\right)(\boldsymbol{\alpha}^0)} \\ = & \sqrt{(n-1)\mathbb{V}ar(\boldsymbol{\alpha}^0) - \left\|\mathrm{Proj}_{\mathscr{S}(0,\mathbf{G})}\boldsymbol{\alpha}^0\right\|^2} \\ \geqslant & \mathscr{D}\left(\boldsymbol{\alpha}^0, \mathscr{G}(1, \mathbf{W}(\lambda), \mathbf{H}^{-1}(\lambda))\right), \forall \lambda \in (0,1), \forall \mathbf{G} \in \mathscr{R} \cup \mathscr{K}_{p_1,\cdots,p_s}, \end{aligned}$$

where $\mathbf{s}_{\cdot j}$ is the orthonormal eigenvector associated with the jth eigenvalue of $\mathbf{G}$, $\mathbf{s}_{\cdot 1} = \pm\frac{1}{\sqrt{n}}\mathbf{1}$, $N_0 \equiv \{j|\lambda_j(\mathbf{G}) = 0\}$. As a result, if $C \geqslant \sqrt{(n-1)\mathbb{V}ar(\boldsymbol{\alpha}^0) - \|\mathrm{Proj}_{\mathscr{S}(0,\mathbf{G})}\boldsymbol{\alpha}^0\|^2}$, we have

$\mathscr{O}(\boldsymbol{\alpha}^0, C) \cap \mathscr{G}\left(1, \mathbf{W}(\lambda), \mathbf{H}^{-1}(\lambda)\right) \neq \emptyset, \forall \lambda \in (0,1)$, then the first-best value in the relative intervention problem $R^*(C, \lambda) = 1$ is guaranteed.

Appendix N. Proof of Proposition 10

Let $\widehat{\mathbb{D}} \equiv diag\left\{\bar{d}_1, \cdots, \bar{d}_n\right\}, \bar{d}_i \equiv d_i / \sum_{i=1}^n d_i$. For vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we define $\langle \mathbf{x}, \mathbf{y} \rangle_{\widehat{\mathbb{D}}} \equiv \mathbf{x}^\top \widehat{\mathbb{D}} \mathbf{y}$, it is the Euclidean inner product weighted by the vertex degrees of network $\mathbf{G}$. It is immediate that

$$\langle \widehat{\mathbf{G}}\mathbf{x}, \mathbf{y} \rangle_{\widehat{\mathbb{D}}} = \langle \mathbf{x}, \widehat{\mathbf{G}}\mathbf{y} \rangle_{\widehat{\mathbb{D}}} = \frac{1}{\sum_{i=1}^n d_i} \mathbf{x}^\top \mathbf{G} \mathbf{y}, \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

So $\widehat{\mathbf{G}}$ is a self-adjoint operator relative to inner product $\langle \cdot, \cdot \rangle_{\widehat{\mathbb{D}}}$. Therefore $\widehat{\mathbf{G}}$ is diagonalizable, and there is a basis of eigenvectors orthogonal to each other under $\langle \cdot, \cdot \rangle_{\widehat{\mathbb{D}}}$. That is, there exists a matrix $\mathbb{U} \equiv (\mathbb{u}_{\cdot i})_{i \in I}$ such that $\mathbb{U}^{-1} \widehat{\mathbf{G}} \mathbb{U} = diag\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$, and $\mathbb{U}^\top \widehat{\mathbb{D}} \mathbb{U} = \mathbf{I}$. Since $\mathbf{G}$ is a connected graph, by Perron-Frobenius Theorem, we have $1 = \lambda_1 > \lambda_2 \geqslant \cdots \geqslant \lambda_n$ and $\mathbb{u}_{\cdot 1} = \pm \mathbf{1}$. Then we have

$$\widehat{\mathbf{M}} = \mathbb{U} \widehat{\boldsymbol{\Lambda}}(\lambda) \mathbb{U}^\top \widehat{\mathbb{D}} = \mathbf{1}\mathbf{1}^\top \widehat{\mathbb{D}} + (1-\lambda) \mathbb{U}_{-1} \widehat{\boldsymbol{\Lambda}}_{-1}(\lambda) (\mathbb{U}_{-1})^\top \widehat{\mathbb{D}}, \tag{92}$$

where $\mathbb{U}_{-1} \equiv (\mathbb{u}_{\cdot j})_{j \neq 1}$,

$$\widehat{\boldsymbol{\Lambda}}(\lambda) \equiv \begin{bmatrix} 1 & & & \\ & \frac{1}{1-\lambda\lambda_2} & & \\ & & \ddots & \\ & & & \frac{1}{1-\lambda\lambda_n} \end{bmatrix}, \widehat{\boldsymbol{\Lambda}}_{-1}(\lambda) \equiv \begin{bmatrix} \frac{1}{1-\lambda\lambda_2} & & \\ & \ddots & \\ & & \frac{1}{1-\lambda\lambda_n} \end{bmatrix}.$$

As $\lambda \to 1$, $\widehat{\mathbf{M}} \to \mathbf{1}\mathbf{1}^\top \widehat{\mathbb{D}} = \mathbf{1}\bar{\mathbf{d}}^\top$, where $\bar{\mathbf{d}} \equiv (\bar{d}_1, \cdots, \bar{d}_n)^\top$. The limit of aggregate social welfare is

$$\mathbb{W}^N(1, \boldsymbol{\alpha}, \mathbf{G}) = \lim_{\lambda \to 1} \frac{1}{2} \boldsymbol{\alpha}^\top \left[\mathbf{I} - \frac{1}{\lambda} (\mathbf{I} - \widehat{\mathbf{M}})^\top (\mathbf{I} - \widehat{\mathbf{M}}) \right] \boldsymbol{\alpha} = \frac{n}{2} \left[\bar{\alpha}^2 - (\bar{\alpha} - \bar{\alpha}_d)^2 \right], \tag{93}$$

where $\bar{\alpha}_d \equiv \boldsymbol{\alpha}^\top \bar{\mathbf{d}}$ is the degree-weighted average of α_is. Therefore, the class of socially optimal networks is $\mathscr{O}(1, \boldsymbol{\alpha}, I) = \arg\max_{\mathbf{G} \in \mathscr{G}(I)} \mathbb{W}^N(1, \boldsymbol{\alpha}, \mathbf{G}) = \arg\min_{\mathbf{G} \in \mathscr{G}(I)} (\bar{\alpha} - \bar{\alpha}_d)^2$. Since $\bar{\alpha} = \bar{\alpha}_d, \forall \mathbf{G} \in \mathscr{R}(I)$, we find that a regular network is socially optimal, i.e., $\mathscr{R}(I) \subseteq \mathscr{O}(1, \boldsymbol{\alpha}, I), \forall \boldsymbol{\alpha} \in \mathbb{R}^n$.

The limit of individual equilibrium payoff is

$$u_i^N(1, \boldsymbol{\alpha}, \mathbf{G}) = \lim_{\lambda \to 1} \frac{1}{2} \boldsymbol{\alpha}^\top \left[e_i e_i^\top - \frac{1}{\lambda} (\mathbf{I} - \widehat{\mathbf{M}})^\top e_i e_i^\top (\mathbf{I} - \widehat{\mathbf{M}}) \right] \boldsymbol{\alpha} = \frac{1}{2} \left[\alpha_i^2 - (\alpha_i - \bar{\alpha}_d)^2 \right]. \tag{94}$$

An agent i's most preferred network class is $\mathscr{O}_i(1, \boldsymbol{\alpha}, I) = \arg\min_{\mathbf{G} \in \mathscr{G}(I)} (\alpha_i - \bar{\alpha}_d)^2$. Let $\mathcal{S}(i, I)$ denotes the class of star networks centered at node i, $\bar{I}(\boldsymbol{\alpha}) \equiv \{i \in I : \alpha_i = \bar{\alpha}\}$ be the set of median players. We have the following results:

- If $\boldsymbol{\alpha} \propto \mathbf{1}$, then $\alpha_i = \bar{\alpha} = \bar{\alpha}_d, \forall i \in I, \forall \mathbf{G} \in \mathscr{G}(I)$, both the planner and agents feel indifferent among all possible networks. Therefore, $\mathscr{O}_i(1, \boldsymbol{\alpha}, I) = \mathscr{O}(1, \boldsymbol{\alpha}, I) = \mathscr{G}(I), \forall i \in I$.

- If $\boldsymbol{\alpha} \not\propto \mathbf{1}$, then $I \backslash \bar{I}(\boldsymbol{\alpha}) \neq \emptyset$. For $\forall i \in I \backslash \bar{I}(\boldsymbol{\alpha})$, we have

$$|\alpha_i - \bar{\alpha}_d| \;=\; \left| \alpha_i - \left(\frac{1}{2}\alpha_i + \frac{1}{2(n-1)} \sum_{j \neq i} \alpha_j \right) \right| = \frac{n|\alpha_i - \bar{\alpha}|}{2(n-1)} < |\alpha_i - \bar{\alpha}|, \forall \mathbf{G} \in \mathcal{S}(i, I).$$

So, any non-median voter i strictly prefers $\mathcal{S}(i, I)$ to networks within $\mathscr{O}(1, \boldsymbol{\alpha}, I)$. The socially optimal network class $\mathscr{O}(1, \boldsymbol{\alpha}, I)$ is therefore not top-ranked among i's preference profile, i.e., $\mathscr{O}(1, \boldsymbol{\alpha}, I) \cap \mathscr{O}_i(1, \boldsymbol{\alpha}, I) = \emptyset, \forall i \in I \backslash \bar{I}(\boldsymbol{\alpha})$. For $i \in \bar{I}(\boldsymbol{\alpha})$ (if $\bar{I}(\boldsymbol{\alpha}) \neq \emptyset$), $\mathscr{O}(1, \boldsymbol{\alpha}, I) = \mathscr{O}_i(1, \boldsymbol{\alpha}, I)$ trivially holds.

Appendix O. Proof of Proposition 11

Suppose there exists a new link $(i', j') \in \mathcal{P}(1, \boldsymbol{\alpha}, \mathbf{G})$, then

$$\begin{aligned}
\Delta u_i^{[+i'j']}(1, \boldsymbol{\alpha}, \mathbf{G}) &\equiv u_i^N(1, \boldsymbol{\alpha}, \mathbf{G}^{[+i'j']}) - u_i^N(1, \boldsymbol{\alpha}, \mathbf{G}) \\
&= \frac{1}{2}(\alpha_i - \bar{\alpha}_d)^2 - \frac{1}{2}\left(\alpha_i - \frac{\sum_{i=1}^n d_i \bar{\alpha}_d + \alpha_{i'} + \alpha_{j'}}{\sum_{i=1}^n d_i + 2} \right)^2 \\
&= \frac{1}{2}\left(\frac{\alpha_{i'} + \alpha_{j'} - 2\bar{\alpha}_d}{\sum_{i=1}^n d_i + 2} \right) \left[2(\alpha_i - \bar{\alpha}_d) + \frac{2\bar{\alpha}_d - \alpha_{i'} - \alpha_{j'}}{\sum_{i=1}^n d_i + 2} \right] \geqslant 0, \forall i \in I,
\end{aligned}$$

strict inequality holds for at least one $j \in I$. Therefore, the $\bar{d}_i-$weighted average of $\Delta u_i^{[+i'j']}(1, \boldsymbol{\alpha}, \mathbf{G})$ is

$$\frac{\sum_{i=1}^n d_i \Delta u_i^{[+i'j']}(1, \boldsymbol{\alpha}, \mathbf{G})}{\sum_{i=1}^n d_i} = -\frac{1}{2}\left(\frac{\alpha_{i'} + \alpha_{j'} - 2\bar{\alpha}_d}{\sum_{i=1}^n d_i + 2} \right)^2 > 0,$$

a contradiction. Therefore, we have $\mathcal{P}(1, \boldsymbol{\alpha}, \mathbf{G}) = \emptyset, \forall \mathbf{G} \in \mathscr{G}(I)$.

Appendix P. Proof of Lemma 3

Let $\mathbf{E}_{ij} = (e_i, e_j)$ and $\mathbf{E}_{ji} = (e_j, e_i)$, then $\mathbf{G}^{+ij} = \mathbf{G} + \mathbf{E}_{ij}\mathbf{E}_{ji}^\top$,

$$\begin{aligned}
\widehat{\mathbf{G}}^{[+ij]} &= diag\left\{ \frac{1}{d_1}, \cdots, \frac{1}{d_i + 1}, \cdots, \frac{1}{d_j + 1}, \cdots, \frac{1}{d_n} \right\} \mathbf{G}^{+ij} \\
&= \left[\mathbf{D}^{-1} + \mathbf{E}_{ij} \begin{pmatrix} \frac{1}{d_i} - \frac{1}{d_i + 1} & 0 \\ 0 & \frac{1}{d_j} - \frac{1}{d_j + 1} \end{pmatrix} \mathbf{E}_{ij}^\top \right] \left(\mathbf{G} + \mathbf{E}_{ij}\mathbf{E}_{ji}^\top \right) \\
&= \widehat{\mathbf{G}} + \mathbf{E}_{ij}\boldsymbol{\Delta}_{ij},
\end{aligned}$$

where

$$\mathbf{\Delta}_{ij} \equiv \begin{bmatrix} \frac{1}{d_i+1} & 0 \\ 0 & \frac{1}{d_j+1} \end{bmatrix} \mathbf{E}_{ji}^\top - \begin{bmatrix} \frac{1}{d_i} - \frac{1}{d_i+1} & 0 \\ 0 & \frac{1}{d_j} - \frac{1}{d_j+1} \end{bmatrix} \mathbf{E}_{ij}^\top \mathbf{G}.$$

Applying the Binomial formula for matrix, we get

$$\left(\widehat{\mathbf{G}}^{[+ij]}\right)^s = \left(\widehat{\mathbf{G}} + \mathbf{E}_{ij}\mathbf{\Delta}_{ij}\right)^s = \sum_{(\ell_1,\cdots,\ell_s)\in\{0,1\}^s} \prod_{t=1}^{s} \widehat{\mathbf{G}}^{\ell_t} (\mathbf{E}_{ij}\mathbf{\Delta}_{ij})^{1-\ell_t}. \tag{95}$$

- *Case* I. $dis((i,j),k,\mathbf{G}) = s \geqslant 1$. Condition $dis((i,j),k,\mathbf{G}) = s$ implies $g_{ki}^{[\tau]} = g_{kj}^{[\tau]} = 0, \forall \tau \leqslant s-1$. So $e_k^\top \widehat{\mathbf{G}}^\tau \mathbf{E}_{ij} = (\hat{g}_{ki}^{[\tau]}, \hat{g}_{kj}^{[\tau]}) = (0,0), \forall \tau \leqslant s-1$.[22] (95) implies

$$\begin{aligned} A_{k,t}^{[+ij]} &\equiv e_k^\top (\widehat{\mathbf{G}}^{[+ij]})^t \boldsymbol{\alpha} = e_k^\top \widehat{\mathbf{G}}^t \boldsymbol{\alpha} \equiv A_{k,t}, \forall t \leqslant s, & (96)\\ A_{k,s+1}^{[+ij]} &\equiv e_k^\top (\widehat{\mathbf{G}}^{[+ij]})^{s+1} \boldsymbol{\alpha} \\ &= e_k^\top \widehat{\mathbf{G}}^s (\widehat{\mathbf{G}} + \mathbf{E}_{ij}\mathbf{\Delta}_{ij}) \boldsymbol{\alpha} \\ &= A_{k,s+1} + e_k^\top \widehat{\mathbf{G}}^s \mathbf{E}_{ij}\mathbf{\Delta}_{ij} \\ &= A_{k,s+1} + \hat{g}_{ki}^{[s]} \frac{\alpha_j - \bar{\alpha}_i}{d_i+1} + \hat{g}_{kj}^{[s]} \frac{\alpha_i - \bar{\alpha}_j}{d_j+1}. & (97) \end{aligned}$$

Substituting the power-series $\widehat{\mathbf{M}} = (1-\lambda)\sum_{k=0}^{\infty} \widehat{\mathbf{G}}^k$ into

$$u_k^N(\lambda, \boldsymbol{\alpha}, \mathbf{G}) = \frac{1}{2}\left[\alpha_k^2 - \frac{1}{\lambda}\left(\alpha_k - \sum_{j=1}^{n} \widehat{m}_{kj}\alpha_j\right)^2\right]$$

then expanding $u_k^N(\lambda, \boldsymbol{\alpha}, \mathbf{G})$ around zero yields:

$$\begin{aligned} u_k^N(\lambda, \boldsymbol{\alpha}, \mathbf{G}) &= \frac{1}{2}\left[\alpha_k^2 - \frac{1}{\lambda}\left(\alpha_k - (1-\lambda) e_k^\top \sum_{q=0}^{\infty} \lambda^q \widehat{\mathbf{G}}^q \boldsymbol{\alpha}\right)^2\right] \\ &= \frac{A_{k,0}^2}{2} - \frac{1}{2}\sum_{q=0}^{\infty} \sum_{\substack{s+t=q \\ 0\leqslant s,t\leqslant q}} (A_{k,s} - A_{k,s+1})(A_{k,t} - A_{k,t+1})\lambda^{q+1}. \end{aligned} \tag{98}$$

In analogy, we have

$$u_k^N(\lambda, \boldsymbol{\alpha}, \mathbf{G}^{[+ij]}) = \frac{\left(A_{k,0}^{[+ij]}\right)^2}{2} - \frac{1}{2}\sum_{q=0}^{\infty} \sum_{\substack{s+t=q \\ 0\leqslant s,t\leqslant q}} \left(A_{k,s}^{[+ij]} - A_{k,s+1}^{[+ij]}\right)\left(A_{k,t}^{[+ij]} - A_{k,t+1}^{[+ij]}\right)\lambda^{q+1}. \tag{99}$$

[22] Notice that for any pair p,q, $\hat{g}_{p,q}^{[\tau]} = 0$ if and only if $g_{p,q}^{[\tau]} = 0$. If $\hat{g}_{p,q}^{[\tau]} = \sum_{i_1,\cdots,i_{\tau-1}} \prod_{k=0}^{\tau-1} \hat{g}_{i_k,i_{k+1}} = \sum_{i_1,\cdots,i_{\tau-1}} \prod_{k=0}^{\tau-1} \frac{1}{d_{i_k}} g_{i_k,i_{k+1}} = 0$ where $i_0 = p, i_\tau = q$, then $\prod_{k=0}^{\tau-1} g_{i_k,i_{k+1}} = 0$, for any sequence $\{i_k\}_{k=1}^{\tau-1}$. It follows that $g_{p,q}^{[\tau]} = \sum_{i_1,\cdots,i_{\tau-1}} \prod_{k=0}^{\tau-1} g_{i_k,i_{k+1}} = 0$. Conversely, if $g_{p,q}^{[\tau]} = 0$, then $\hat{g}_{p,q}^{[\tau]} = 0$.

From (96), we find that the constant terms, and the coefficients of λ^ℓ, in (98) and (99) are identical for $\ell \leqslant s$. We have

$$\begin{aligned}\Delta u_k^{[+ij]}(\lambda,\boldsymbol{\alpha},\mathbf{G}) &= \left[\begin{array}{c}(A_{k,0}-A_{k,1})\,(A_{k,s}-A_{k,s+1})\\ -\left(A_{k,0}^{[+ij]}-A_{k,1}^{[+ij]}\right)\left(A_{k,s}^{[+ij]}-A_{k,s+1}^{[+ij]}\right)\end{array}\right]\lambda^{s+1}+o(\lambda^{s+1})\\ &= (A_{k,0}-A_{k,1})\left(A_{k,s+1}^{[+ij]}-A_{k,s+1}\right)\lambda^{s+1}+o(\lambda^{s+1}) \qquad (100)\\ &= (\alpha_k-\bar{\alpha}_k)\left(\hat{g}_{ki}^{[s]}\frac{\alpha_j-\bar{\alpha}_i}{1+d_i}+\hat{g}_{kj}^{[s]}\frac{\alpha_i-\bar{\alpha}_j}{1+d_j}\right)\lambda^{s+1}+o(\lambda^{s+1}), \qquad (101)\end{aligned}$$

(100) $\Rightarrow$ (101) follows from $A_{k,0}=\alpha_k, A_{k,1}=e_k^\top\widehat{\mathbf{G}}\boldsymbol{\alpha}=\bar{\alpha}_k$, and (97).

- *Case* II. $dis((i,j),k,\mathbf{G})=0$. In this case, $k\in\{i,j\}$, then we have

$$\Delta u_k^{[+ij]}(\lambda,\boldsymbol{\alpha},\mathbf{G})=\frac{1}{2}\left[(A_{k,0}-A_{k,1})^2-\left(A_{k,0}^{[+ij]}-A_{k,1}^{[+ij]}\right)^2\right]\lambda+o(\lambda). \qquad (102)$$

Inserting $A_{k,0}=A_{k,0}^{[+ij]}=\alpha_k$, $A_{k,1}=\bar{\alpha}_k$, $A_{k,1}^{[+ij]}=e_k^\top\widehat{\mathbf{G}}^{+ij}\boldsymbol{\alpha}=\frac{d_i\bar{\alpha}_i}{d_i+1}+\frac{\alpha_j}{d_i+1}$ into (102), we get:

$$\begin{aligned}\Delta u_i^{[+ij]}(\lambda,\boldsymbol{\alpha},\mathbf{G}) &= \frac{\lambda(\alpha_j-\bar{\alpha}_i)}{2(d_i+1)}\left[2\alpha_i-\left(1+\frac{d_i}{d_i+1}\right)\bar{\alpha}_i-\frac{\alpha_j}{d_i+1}\right]+o(\lambda), \qquad (103)\\ \Delta u_j^{[+ij]}(\lambda,\boldsymbol{\alpha},\mathbf{G}) &= \frac{\lambda(\alpha_i-\bar{\alpha}_j)}{2(d_j+1)}\left[2\alpha_j-\left(1+\frac{d_j}{d_j+1}\right)\bar{\alpha}_j-\frac{\alpha_i}{d_j+1}\right]+o(\lambda). \qquad (104)\end{aligned}$$

Appendix Q: Proof of Proposition 12

For an arbitrary $(i,j)\in\mathcal{E}^c$ and $k\notin\{i,j\}$, (101) implies that $\Delta u_k^{[+ij]}(\lambda,\boldsymbol{\alpha},\mathbf{G})=o(\lambda)$. So $\Delta u_k^{[+ij]}(\lambda,\boldsymbol{\alpha},\mathbf{G})\geqslant-\epsilon\lambda$ holds for an arbitrary $\epsilon>0$, whenever λ is sufficiently close to zero. (103) and (104) shows that $\Delta u_i^{[+ij]}(\lambda,\boldsymbol{\alpha},\mathbf{G})\geqslant-\epsilon\lambda$ and $\Delta u_j^{[+ij]}(\lambda,\boldsymbol{\alpha},\mathbf{G})\geqslant-\epsilon\lambda$ hold for an arbitrary $\epsilon>0$ and sufficiently small λ, provided that the following conditions are met:

$$(\alpha_j-\bar{\alpha}_i)\left[2\alpha_i-\left(1+\tfrac{d_i}{d_i+1}\right)\bar{\alpha}_i-\tfrac{\alpha_j}{d_i+1}\right]\geqslant 0, \qquad (105)$$

$$(\alpha_i-\bar{\alpha}_j)\left[2\alpha_j-\left(1+\tfrac{d_j}{d_j+1}\right)\bar{\alpha}_j-\tfrac{\alpha_i}{d_j+1}\right]\geqslant 0. \qquad (106)$$

References

[1] Abdi, M., Ghorbani, E., and Imrich,W., (2021), "Regular graphs with minimum spectral gap," *European Journal of Combinatorics*, 95, 103328.

[2] Ballester, C., Calvó-Armengol, A., and Zenou, Y., (2006), "Who's who in networks, wanted: the key player," *Econometrica*, 74, 1403–1417.

[3] Ballester, C., Calvó-Armengol, A., and Zenou, Y., (2010), "Delinquent networks," *Journal of the European Economic Association*, 8, 34–61

[4] Belhaj, M., Bervoets, S., and Deroïan, F., (2016), "Efficient networks in games with local complementarities," *Theoretical Economics*, 11, 357–380.

[5] Bimpikis, K., A. Ozdaglar, and E. Yildiz (2016), "Competitive targeted advertising over networks," *Operations Research*, 64, 705–720.

[6] Bloch, F., Jackson, M., and Tebaldi, P., (2021), "Centrality measures in networks," Available at SSRN: https://ssrn.com/abstract=2749124.

[7] Bloch, F. and N. Quérou (2013), "Pricing in social networks," *Games and Economic Behaviour*, 80, 243–261.

[8] Blume, L., Brock, W., Durlauf, S., and Jayaraman, R., (2015), "Linear social interactions models," *Journal of Political Economy*, 123, 444–496.

[9] Boucher, V. (2016), "Conformism and self-selection in social networks," *Journal of Public Economics*, 136, 30–44.

[10] Candogan, O., Bimpikis, K., Ozdaglar, A., (2012), "Optimal pricing in networks with externalities," *Operations Research*, 60, 883–905

[11] Cvetković, D., Doob, M., and Sachs, H., (1980), "Spectra of graphs: theory and application," Academic Press, Inc., New York–London.

[12] Demange, G., (2017), "Optimal targeting strategies in a network under complementarities," *Games and Economic Behaviour*, 105, 84–103.

[13] Chen, Y., and Zhou, J., 2015, "Key leaders in social networks," *Journal of Economic Theory*, 157, 212–235.

[14] Chen, Y., Zenou, Y., Zhou, J., (2018), "Competitive pricing strategies in social networks," *Rand Journal of Economics*, 49, 672–705.

[15] Fainmesser, I., Galeotti, A., (2016), "Pricing Network Effects," *The Review of Economic Studies*, 83, 165–198

[16] Galeotti, A., Golub, B., Goyal, S., (2020) "Targeting Interventions in Networks," *Econometrica*, 88, 2445–2471

[17] Galeotti, A., Goyal, S., (2009), "Influencing the Influencers: A Theory of Strategic Diffusion," *The RAND Journal of Economics* , 40, 509–532

[18] Golub, B., Jackson, M., (2012), "How homophily affects the speed of learning and best-response dynamics," *Quarterly Journal of Economics*, 127, 1287–338.

[19] Golub, B. and Lever, C., (2010), "The leverage of weak ties: How linking groups affect inequality," available at: http://bengolub.net/wp-content/uploads/2020/05/intergroup.pdf.

[20] Golub, B., and Morris, S., (2017), "Expectations, networks, and conventions," working paper, available at: https://arxiv.org/abs/2009.13802.

[21] Hiller, T., (2017), "Peer Effects in Endogenous Networks," *Games and Economic Behavior*, 105, 349–367.

[22] Jackson, M., (2008), Social and Economic Networks, Princeton: *Princeton University Press.*

[23] King, M., B. Tarbush, and A. Teytelboym (2019), "Targeted carbon tax reforms," *European Economic Review*, 119, 526–547.

[24] König, M., Tessone, C., and Zenou, Y., (2014), "Nestedness in networks: A theoretical model and some applications," *Theoretical Economics*, 9, 695–752.

[25] Kor R., and Zhou, J. (2022), "Welfare and Distributional Effects of Joint Intervention in Networks," working paper, available at https://arxiv.org/abs/2206.03863.

[26] Lee, L., Liu, X., Patacchini, E., and Zenou, Y., (2021), "Who is the key player? a network analysis of juvenile delinquency," *Journal of Business and Economic Statistics*, 849–857.

[27] Li, X., (2021), "Designing weighted and directed networks under complementarities," *working paper*, Available at SSRN: https://ssrn.com/abstract=3299331

[28] Liu, E., (2019), "Industrial policies in production networks," *The Quarterly Journal of Economics*, 134, 1883–1948.

[29] Meng, D., Sun, L., and Tian, G., (2022), "Dynamic mechanism design on social networks," *Game and Economic Behavior*, 131, 84–120.

[30] Mussa, M., and Rosen, S., (1978), "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301-317.

[31] Olcina, G., Panebianco, F., and Zenou, Y., (2017), "Conformism, social norms and the dynamics of assimilation," IZA Discussion Paper No. 11436.

[32] Patacchini, E., Zenou, Y., (2012), "Juvenile Delinquency and Conformism," *Journal of Law, Economics, & Organization*, 28, 1–31

[33] Sciabolazza, V., Vacca, R., and McCarty, C., (2020), "Connecting the dots: implementing and evaluating a network intervention to foster scientific collaboration and productivity," Social Networks, 61, 181–195.

[34] Sun, Y., Zhao, W., and Zhou, J., (2021), "Structural interventions in networks," working paper, Available at SSRN: https://ssrn.com/abstract=3777029.

[35] Valente, T., (2012),"Network Interventions," *Science*, 337, 49–53.

[36] Ushchev, P., Zenou, Y., (2018), "Price competition in product variety networks," *Games Economic Behaviors*, 110, 226–247.

[37] Ushchev, P., and Zenou, Y., (2020), "Social norms in networks," *Journal of Economic Theory*, 185, 104969.