Supplementary Material

A Proofs Omitted from Section 3

A.1 Proof of Proposition 1: Convexity Analysis

We decompose Proposition 1 in a series of small lemmas containing specifically created counterexamples to show the non-convexity of the objectives.

We recall that a matrix-valued function f is said to be matrix-convex if and only if it satisfies the following inequality for all $\lambda \in [0,1]$ and matrices $\mathbf{A}_1$ and $\mathbf{A}_2$:

$$f(\lambda \mathbf{A}_1 + (1-\lambda)\mathbf{A}_2) \leq \lambda f(\mathbf{A}_1) + (1-\lambda) f(\mathbf{A}_2) \tag{9}$$

Lemma 1 (P-directed)**.** *The objective function of Eq.* (3) *is not a matrix-convex function.*

Proof. Consider a vector of opinions $\mathbf{s}^\top = (0\ 1\ 1)$. Let $\mathbf{A}_1$ and $\mathbf{A}_2$ be two adjacency matrices of two connected graphs:

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1/100 & 0 & 99/100 \\ 1 & 0 & 0 \end{bmatrix}, \text{ and } \mathbf{A}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1/3 & 0 & 2/3 \\ 0 & 1 & 0 \end{bmatrix}.$$

Setting $\lambda = 0.5$, we can compute the two terms in the inequality to obtain $f((\mathbf{A}_1 + \mathbf{A}_2)/2) = 1.66$ and $(f(\mathbf{A}_1) + f(\mathbf{A}_2))/2 = 1.65$. Thus, the inequality in Eq. (9) is violated, and the objective function in Eq. (3) is not matrix-convex. $\square$

Lemma 2 (D-directed)**.** *The objective function of Eq.* (4) *is not a matrix-convex function.*

Proof. Consider a vector of opinions $\mathbf{s}^\top = (1\,0\,-1)$. Let $\mathbf{A}_1$ and $\mathbf{A}_2$ be two adjacency matrices of two connected graphs:

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 1 & 0 \\ 2/3 & 0 & 1/3 \\ 1 & 0 & 0 \end{bmatrix}, \text{ and } \mathbf{A}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1/3 & 0 & 2/3 \\ 0 & 1 & 0 \end{bmatrix}.$$

Setting $\lambda = 0.5$, we can compute the two terms in the inequality to obtain $f((\mathbf{A}_1 + \mathbf{A}_2)/2) = 0.45$ and $(f(\mathbf{A}_1) + f(\mathbf{A}_2))/2 = 0.43$. Thus, the inequality in Eq. (9) is violated, and the objective function in Eq. (4) is not matrix-convex. $\square$

Lemma 3 (PD-directed)**.** *The objective function of Eq.* (5) *is not a matrix-convex function.*

Proof. Consider a vector of opinions $\mathbf{s}^\top = (1\,0\,-1)$. Let $\mathbf{A}_1$ and $\mathbf{A}_2$ be two adjacency matrices of two connected graphs:

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1/100 & 0 & 99/100 \\ 1 & 0 & 0 \end{bmatrix}, \text{ and } \mathbf{A}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1/3 & 0 & 2/3 \\ 0 & 1 & 0 \end{bmatrix}.$$

Setting $\lambda = 0.5$, we can compute the two terms in the inequality to obtain $f((\mathbf{A}_1 + \mathbf{A}_2)/2) = 0.82$ and $(f(\mathbf{A}_1) + f(\mathbf{A}_2))/2 = 0.80$. Thus, the inequality in Eq. (9) is violated, and the objective function in Eq. (5) is not matrix-convex. $\square$

Lemma 4 (P-undirected)**.** *The objective function of Eq.* (6) *is not a matrix-convex function.*

Proof. Consider a vector of opinions $\mathbf{s}^\top = (0\ 1\ 1)$. Let $\mathbf{A}_1$ and $\mathbf{A}_2$ be two adjacency matrices of two connected graphs:

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 1/1000 & 0 \\ 1/1000 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \text{ and } \mathbf{A}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Setting $\lambda = 0.5$, we can compute the two terms in the inequality to obtain $f((\mathbf{A}_1 + \mathbf{A}_2)/2) = 1.25$ and $(f(\mathbf{A}_1) + f(\mathbf{A}_2))/2 = 1.11$. Thus, the inequality in Eq. (9) is violated, and the objective function in Eq. (6) is not matrix-convex. $\square$

Lemma 5 (D-undirected)**.** *The objective function of Eq.* (7) *is not a matrix-convex function.*

Proof. Consider a vector of opinions $\mathbf{s}^\top = (1\,0\,-1)$. Let $\mathbf{A}_1$ and $\mathbf{A}_2$ be two adjacency matrices of two connected graphs:

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \text{ and } \mathbf{A}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Setting $\lambda = 0.5$, we can compute the two terms in the inequality to obtain $f((\mathbf{A}_1 + \mathbf{A}_2)/2) = 0.5$ and $(f(\mathbf{A}_1) + f(\mathbf{A}_2))/2 = 0.49$. Thus, the inequality in Eq. (9) is violated, and the objective function in Eq. (7) is not matrix-convex. $\square$

A.2 Proof of Proposition 2: Convexity Analysis

See the proof of Theorem 1 in Musco *et al.* [25].

A.3 Proof of Theorem 1

Proof. As $\hat{\mathbf{L}}$ is a minimizer of $f(\hat{\mathbf{s}}, \mathbf{L})$, we have $f(\hat{\mathbf{s}}, \hat{\mathbf{L}}) \leq f(\hat{\mathbf{s}}, \mathbf{L}^*)$. If $f(\mathbf{s}, \hat{\mathbf{L}}) \leq f(\hat{\mathbf{s}}, \hat{\mathbf{L}})$ holds, we have $f(\mathbf{s}, \hat{\mathbf{L}}) \leq f(\hat{\mathbf{s}}, \mathbf{L}^*)$, leading to $f(\mathbf{s}, \hat{\mathbf{L}}) - f(\mathbf{s}, \mathbf{L}^*) \leq f(\hat{\mathbf{s}}, \mathbf{L}^*) - f(\mathbf{s}, \mathbf{L}^*) \leq K\|\mathbf{s} - \hat{\mathbf{s}}\|$. Otherwise, since f is K-Lipschitz continuous, we have

$$\begin{aligned} f(\mathbf{s}, \hat{\mathbf{L}}) - f(\mathbf{s}, \mathbf{L}^*) &\leq |f(\mathbf{s}, \hat{\mathbf{L}}) - f(\hat{\mathbf{s}}, \hat{\mathbf{L}})| + |f(\hat{\mathbf{s}}, \mathbf{L}^*) - f(\mathbf{s}, \mathbf{L}^*)| \\ &\leq 2K\|\mathbf{s} - \hat{\mathbf{s}}\|. \end{aligned}$$

$\square$

A.4 Proof of Proposition 3: Lipschitz Constants

We decompose Proposition 3 in the main text in a series of lemmas: we present their statement and proofs to bound the Lipschitz constants of each objective function studied. We begin by deriving the gradient of the objective function with respect to the opinion vector to obtain a representation of the Lipschitz constant in its infinitesimal form.

We recall that the Lipschitz constant is associated with the spectral norm of the matrix in the gradient, which corresponds to its maximum eigenvalue. More formally, a function f is Lipschitz continuous if there exists a constant K such that for all $\mathbf{s}, \hat{\mathbf{s}}$ in its domain, $|f(\mathbf{s}) - f(\hat{\mathbf{s}})| \leq K\|\mathbf{s} - \hat{\mathbf{s}}\|$. Equivalently, in its infinitesimal form, the gradient of f is bounded. Since we are dealing with quadratic forms, we recall its gradient from Eq. 81 in [32]:

$$\nabla_{\mathbf{s}} \mathbf{s}^\top \mathbf{M} \mathbf{s} = (\mathbf{M} + \mathbf{M}^\top)\mathbf{s}. \tag{10}$$

Lemma 6 (Polarization for directed graphs (P-Dir))**.** *Given a continuous differentiable function* $f(\mathbf{s}) = f(\mathbf{s}, \mathbf{L} = \mathbf{I} - \mathbf{A}) := \mathbf{s}^\top (2\mathbf{I} - \mathbf{A})^{-\top} (2\mathbf{I} - \mathbf{A})^{-1} \mathbf{s}$, *where* $\mathbf{A}$ *is a fixed row-stochastic matrix, and* $\mathbf{s} \in \mathbb{R}^n$, $f(\mathbf{s})$ *is Lipschitz continuous with* $K = 2$.

Proof. Using the gradient of a quadratic in Eq. (10), the gradient of f with respect to $\mathbf{s}$ is

$$\nabla_{\mathbf{s}} f(\mathbf{s}) = 2(2\mathbf{I} - \mathbf{A})^{-\top}(2\mathbf{I} - \mathbf{A})^{-1}\mathbf{s}.$$

The product $(2\mathbf{I} - \mathbf{A})^{-\top}(2\mathbf{I} - \mathbf{A})^{-1}$ has a spectral norm that is bounded by the product of the spectral norms of the two terms by the submultiplicative property of matrix norms. Moreover, eigenvalues of $(2\mathbf{I} - \mathbf{A})^{-1}$ are within $[\frac{1}{3}, 1]$ because the eigenvalues of $2\mathbf{I} - \mathbf{A}$ are within $[1, 3]$, using the row-stochasticity of $\mathbf{A}$. Hence the maximum is 1. Thus, the Lipschitz constant K for f is bounded by 2. $\square$

Lemma 7 (Disagreement for directed graphs (D-Dir))**.** *Given a continuous differentiable function* $f(\mathbf{s}) = f(\mathbf{s}, \mathbf{L} = \mathbf{I} - \mathbf{A}) := \frac{1}{2}\mathbf{s}^\top(2\mathbf{I} - \mathbf{A})^{-\top}(\mathbf{I} + \mathbf{D}^{\text{in}} - 2\mathbf{A})(2\mathbf{I} - \mathbf{A})^{-1}\mathbf{s}$*, where* $\mathbf{A}$ *is a fixed row-stochastic matrix,* $\mathbf{D}^{\text{in}}$ *is a fixed diagonal matrix with non-negative entries corresponding to the in-degree, and* $\mathbf{s} \in \mathbb{R}^n$*,* $f(\mathbf{s})$ *is Lipschitz continuous with* $K = 1 + \Delta(G)$*.*

Proof. Using Eq. (10), and the fact that $(\mathbf{ABC})^\top = \mathbf{C}^\top\mathbf{B}^\top\mathbf{A}^\top$, the gradient of f with respect to $\mathbf{s}$ is

$$\nabla_{\mathbf{s}} f(\mathbf{s}) = (2\mathbf{I} - \mathbf{A})^{-\top}(\mathbf{I} + \mathbf{D}^{\text{in}} - 2\mathbf{A})(2\mathbf{I} - \mathbf{A})^{-1}\mathbf{s}.$$

The matrices $(2\mathbf{I} - \mathbf{A})^{-\top}$, $(2\mathbf{I} - \mathbf{A})^{-1}$ are diagonally-dominant and positive definite, and their spectral norm is bounded by the product of the spectral norms of the two terms by the submultiplicative property of matrix norms. Moreover, eigenvalues of $(2\mathbf{I} - \mathbf{A})^{-1}$ are within $[\frac{1}{3}, 1]$ because the eigenvalues of $2\mathbf{I} - \mathbf{A}$ are within $[1, 3]$, using the row-stochasticity of $\mathbf{A}$. Hence the maximum is 1. In the other term ($\mathbf{D}^{\text{in}}$), the degree matrix is diagonal and hence contains the positive eigenvalues, and we denote its maximum value as $\Delta(G)$. The adjacency matrix $\mathbf{A}$ is row-stochastic, hence its minimum eigenvalue is 1, so the negative sign in front of it makes all eigenvalues negative therefore this term does not contribute to the maximum eigenvalue. Thus, the Lipschitz constant K for f is bounded by $1 + \Delta(G)$. $\square$

Lemma 8 (Polarization plus Disagreement for Directed graphs (PD-Dir))**.** *Given a continuously differentiable function*

$$f(\mathbf{s}) = \frac{1}{2}\mathbf{s}^\top(2\mathbf{I} - \mathbf{A})^{-\top}(2\mathbf{I} + \mathbf{D}^{\text{in}} - 2\mathbf{A})(2\mathbf{I} - \mathbf{A})^{-1}\mathbf{s},$$

where $\mathbf{A}$ *is a fixed row-stochastic matrix,* $\mathbf{D}^{\text{in}}$ *is a fixed diagonal matrix with non-negative entries corresponding to the in-degree, and* $\mathbf{s} \in \mathbb{R}^n$*, the function* $f(\mathbf{s})$ *is Lipschitz continuous with* $K = 2 + \Delta(G)$*.*

Proof. Using Eq. (10), the gradient of f with respect to $\mathbf{s}$ is

$$\nabla_{\mathbf{s}} f(\mathbf{s}) = (2\mathbf{I} - \mathbf{A})^{-\top}(2\mathbf{I} + \mathbf{D}^{\text{in}} - 2\mathbf{A})(2\mathbf{I} - \mathbf{A})^{-1}\mathbf{s}.$$

As shown in the proof of Lemma 7, the matrices $(2\mathbf{I} - \mathbf{A})^{-\top}$ and $(2\mathbf{I} - \mathbf{A})^{-1}$ are diagonally dominant, positive definite, and have spectral norm at most 1.

The middle term $2\mathbf{I} + \mathbf{D}^{\text{in}} - 2\mathbf{A}$ has spectral norm at most $2 + \Delta(G)$: the identity contributes 2, the diagonal matrix $\mathbf{D}^{\text{in}}$ is bounded by the maximum in-degree $\Delta(G)$, and $-2\mathbf{A}$ has norm at most 2 since $\mathbf{A}$ is row-stochastic.

Thus, by the submultiplicative property of matrix norms, the Lipschitz constant K is bounded by:

$$K \leq 1 \cdot (2 + \Delta(G)) \cdot 1 = 2 + \Delta(G).$$

$\square$

Lemma 9 (Polarization for Undirected graphs (P-Und))**.** *Given a continuous differentiable function* $f(\mathbf{s}) := \mathbf{s}^\top(\mathbf{I} + \mathbf{L})^{-2}\mathbf{s}$*, where* $\mathbf{A}$ *is a fixed row-stochastic matrix and* $\mathbf{s} \in \mathbb{R}^n$*, the function* $f(\mathbf{s})$ *is Lipschitz continuous with* $K = 2$*.*

Proof. Using Eq. (10), the gradient of f with respect to $\mathbf{s}$ is

$$\nabla_{\mathbf{s}} f(\mathbf{s}) = 2(\mathbf{I} + \mathbf{L})^{-2}\mathbf{s}.$$

The norm of this gradient ($\|2(\mathbf{I} + \mathbf{L})^{-2}\mathbf{s}\|_2$) is bounded, by sub-multiplicativity, by the eigenvalues of $2(\mathbf{I} + \mathbf{L})^{-1}(\mathbf{I} + \mathbf{L})^{-1}$. Therefore, by submultiplicativity the Lipschitz constant K for f is $2\lambda_{\text{max}}((\mathbf{I} + \mathbf{L})^{-1})^2$. Using the properties of eigenvalues we know that $\lambda_{\text{max}}(1 + \mathbf{X})^{-1} = \frac{1}{1+\lambda_{\text{min}}(\mathbf{X})}$; and since the minimum eigenvalue of a Laplacian matrix is 0, the maximum eigenvalue is 1. Thus, the Lipschitz constant K for f is bounded by 2. $\square$

Lemma 10 (Disagreement for undirected graphs (D-Undir))**.** *Given a continuous differentiable function* $f(\mathbf{L}, \mathbf{s}) = \mathbf{s}^\top(\mathbf{I} + \mathbf{L})^{-1}\mathbf{L}(\mathbf{I} + \mathbf{L})^{-1}\mathbf{s}$*, where* $\mathbf{L}$ *is a fixed symmetric Laplacian matrix and* $\mathbf{s} \in \mathbb{R}^n$*,* $f(\mathbf{s})$ *is Lipschitz continuous with a Lipschitz constant equal to* $2\Delta(G)$*.*

Proof. From Eq. (10), the gradient of f is $\nabla_{\mathbf{s}} f(\mathbf{L}, \mathbf{s}) = 2(\mathbf{I} + \mathbf{L})^{-1}\mathbf{s}$. The norm of this gradient ($\|2(\mathbf{I} + \mathbf{L})^{-1}\mathbf{s}\|_2$) is bounded, by the eigenvalues of $2(\mathbf{I}+\mathbf{L})^{-1}\mathbf{L}(\mathbf{I}+\mathbf{L})^{-1}$. Therefore, by submultiplicativity the Lipschitz constant K for f is $2\lambda_{\text{max}}((\mathbf{I} + \mathbf{L})^{-1})^2\lambda_{\text{max}}(\mathbf{L})$.

Using the properties of eigenvalues we know that $\lambda_{\text{max}}(1 + \mathbf{X})^{-1} = \frac{1}{1+\lambda_{\text{min}}(\mathbf{X})}$; and since the minimum eigenvalue of a Laplacian matrix is 0, the maximum eigenvalue is 1. Using the fact that the maximum eigenvalue of a graph Laplacian is bounded by the maximum degree in the graph, we get $K = 2\Delta(G)$. $\square$

Lemma 11 (Polarization plus Disagreement for undirected graphs (PD-Undir))**.** *Given a continuous differentiable function* $f(\mathbf{L}, \mathbf{s}) = \mathbf{s}^\top(\mathbf{I} + \mathbf{L})^{-1}\mathbf{s}$*, where* $\mathbf{L}$ *is a fixed symmetric Laplacian matrix and* $\mathbf{s} \in \mathbb{R}^n$*, the function* $f(\mathbf{s})$ *is Lipschitz continuous with a Lipschitz constant equal to* 2*.*

Proof. From Eq. (10), the gradient of f is $\nabla_{\mathbf{s}} f(\mathbf{L}, \mathbf{s}) = 2(\mathbf{I} + \mathbf{L})^{-1}\mathbf{s}$. The norm of this gradient ($\|2(\mathbf{I} + \mathbf{L})^{-1}\mathbf{s}\|_2$) is bounded, by sub-multiplicativity, by the eigenvalues of $(\mathbf{I} + \mathbf{L})^{-1}$. Therefore, the Lipschitz constant K for f is $2\lambda_{\text{max}}((\mathbf{I} + \mathbf{L})^{-1})$. Using the properties of eigenvalues we know that $\lambda_{\text{max}}(1 + \mathbf{A})^{-1} = \frac{1}{1+\lambda_{\text{min}}(\mathbf{A})}$; and since the minimum eigenvalue of a Laplacian matrix is 0, the maximum eigenvalue is 1. Hence, $K = 2\lambda_{\text{max}}((\mathbf{I} + \mathbf{L})^{-1}) = 2$. $\square$

B Detailed Methodological Choices

B.1 Node Selection Strategies

Selecting the best b-element subset of nodes to query that minimizes the statistical error of reconstruction is NP-hard in general since it can be cased as discrete optimal design problems [21, 33], such as A-optimal design for minimizing the mean squared error. To avoid computational bottlenecks, we devise three heuristic approaches that select the top-b nodes based on the following centrality measures:

Degree Centrality: This metric measures the total number of connections a node has, which includes both incoming and outgoing edges. Nodes with high degree centrality are considered influential due to their numerous direct connections. It is defined as:

$$\text{Degree Centrality of node } u = d_u^{\text{total}} = d_u^{\text{in}} + d_u^{\text{out}} = \sum_{v \in V} A_{vu} + \sum_{v \in V} A_{uv},$$

where A is the adjacency matrix of the graph.

Closeness Centrality [4]: This metric measures how close a node is to all other nodes in the graph. It captures the average shortest path distance from a node to all other nodes, indicating the node's ability to quickly interact with all parts of the network. The measure is inversely proportional to the sum of the shortest path distances between a node and all others.

$$\text{Closeness Centrality of node } u = \frac{1}{\sum_{v \in V} p(u, v)},$$

where $p(u, v)$ is the shortest path distance between nodes u and v.

PageRank [30]: This metric measures the importance of a node based on the structure of incoming links, assigning higher importance to nodes with many incoming links from other important nodes. The PageRank of a node u is defined by the recursive formula

$$u = PR(u) = \frac{1-d}{n} + d \sum_{v \in \text{In}(u)} \frac{PR(v)}{d_v},$$

where d is a damping factor (set to 0.85), $\text{In}(u)$ is the set of nodes that link to u, and d_v is the out-degree of v.

C Detailed Opinion Reconstruction Methods

Label Propagation (LP). Label Propagation (LP) is a semi-supervised learning algorithm originally designed for classification tasks, where labels are propagated through the network based on the labels of neighboring nodes [41]. We extended it to handle continuous values, making it suitable for reconstructing opinions in a network. This approach leverages the fact that neighboring nodes tend to have similar opinions.

The LP strategy for continuous values initializes the known opinions for the selected nodes and iteratively updates the values of the remaining nodes based on the average values of their neighbors. This process continues until convergence or a maximum number of iterations is reached.

Graph Neural Networks (GNN). We use a GCN (Graph Convolutional Network) [19] to propagate the known opinions of the selected nodes through the network to reconstruct the unknown opinions. The algorithm initializes the GCN with node features and trains it using the opinions of the selected nodes to define the squared error loss. After training, the GCN predicts the values for all nodes in the graph.

In our implementation, we use the queried opinions as the node features if available, setting to zero if not available. We also tested different structural features of the nodes, such as degree centrality, PageRank, and adjacency eigenvectors, but found no significant gains in accuracy, with only increased computational cost.

Graph Signal Processing (GSP). Graph Signal Processing (GSP) assumes that the graph signal (i.e., opinions in our case) can be described by a limited number $|F|$ of frequencies and $|X|$ nodes ($|X| = b$ in our problem) such that $|F|, |X| \leq n$. In GSP, frequencies correspond to eigenvectors of the Laplacian, which we denote with the column vectors in the matrix $\mathbf{U} \in \mathbb{R}^{|V| \times |V|}$. This signal can be written as

$$\mathbf{s}_X = \mathbf{P}_X^\top \mathbf{s} = \mathbf{P}_X^\top \mathbf{U}_F \tilde{\mathbf{s}}_F, \tag{11}$$

where $\mathbf{s}_X \in \mathbb{R}^{|X|}$ is the observation vector over the vertex set $X \subseteq V$, $\mathbf{P}_X \in \mathbb{R}^{|V| \times |X|}$ is a sampling matrix whose columns are indicator functions for nodes in X and satisfies $\mathbf{D}_X = \mathbf{P}_X \mathbf{P}_X^\top$, $\mathbf{U}_F \in \mathbb{R}^{|F| \times n}$ is the matrix of the F subset of Laplacian eigenvectors, and $\tilde{\mathbf{s}}_F$ is a sparse representation of the opinions in the frequency domain.

The problem of recovering a graph signal that is limited in the frequency domain (hence called *bandlimited*) from its samples is equivalent to the problem of properly selecting the sampling set X and then recovering $\mathbf{s}$ from $\mathbf{s}_X$ by inverting the system of equations in Eq. (11). Selecting an optimal sampling set X that maximizes a target cost function f is computationally challenging, as described in Section 3.1.

In our implementation, we use two results from Lorenzo *et al.* [21]. First, the necessary and sufficient conditions for perfect recovery, which requires $\|\mathbf{D}_{X^c} \mathbf{U}_F\|_2 < 1$ and $|X| \geq |F|$, where $\mathbf{D}_{X^c} = \mathbf{I} - \mathbf{D}_X$ projects onto the complement vertex set $X^c = V \backslash X$. Second, we assume that the signal can be affected by noise, where the observation model is given by

$$\mathbf{s}_X = \mathbf{P}_X^\top (\mathbf{s} + \mathbf{v}) = \mathbf{P}_X^\top \mathbf{U}_F \tilde{\mathbf{s}}_F + \mathbf{P}_X^\top \mathbf{v}, \tag{12}$$

where $\mathbf{v}$ is a zero-mean noise vector with covariance matrix $\mathbf{R}_v = \mathbb{E}\{\mathbf{v}\mathbf{v}^\top\}$. To design an interpolator in the presence of noise, we consider the best linear unbiased estimator (BLUE) [37]:

$$\hat{\mathbf{s}} = \mathbf{U}_F (\mathbf{U}_F^\top \mathbf{P}_X (\mathbf{P}_X^\top \mathbf{R}_v \mathbf{P}_X)^{-1} \mathbf{P}_X^\top \mathbf{U}_F)^{-1} \mathbf{U}_F^\top \mathbf{P}_S (\mathbf{P}_S^\top \mathbf{R}_v \mathbf{P}_X)^{-1} \mathbf{s}_X. \tag{13}$$

D Pseudocodes

D.1 Label Propagation (LP)

All node values are initialized to zero (line 1), and the values of the sampled nodes are set to their corresponding real values (lines 2-4). For a fixed number of iterations, the algorithm updates the value of each node (except the sampled ones) to the average value of its neighbors (lines 5-15). Specifically,

Algorithm 1: Label Propagation for Continuous Values

Data: Graph $G = (V, E)$, sampled indices S, real values $\mathbf{s}_{\mathrm{true}}$, maximum iterations max_iter
Result: Recovered node values $\mathbf{s}$

1 $\mathbf{s}[i] \leftarrow 0$ for all $i \in V$ `// Initialize node values`
2 **for** *each* $i \in S$ **do**
3 $\quad \mathbf{s}[i] \leftarrow \mathbf{s}_{\mathrm{true}}[i]$
4 **end**
5 **for** *each iteration t from* 1 *to* max_iter **do**
6 $\quad \mathbf{s}_{\text{new}} \leftarrow \mathbf{s}$
7 $\quad$ **for** *each* $i \in V$ **do**
8 $\quad\quad$ **if** $i \notin S$ **then**
9 $\quad\quad\quad N_i \leftarrow \{j \in V \mid (i, j) \in E\}$ `// Neighbors of` i **if** $N_i \neq \emptyset$ **then**
10 $\quad\quad\quad\quad \mathbf{s}_{\text{new}}[i] \leftarrow \frac{1}{|N_i|} \sum_{j \in N_i} \mathbf{s}[j]$
11 $\quad\quad\quad$ **end**
12 $\quad\quad$ **end**
13 $\quad$ **end**
14 $\quad \mathbf{s} \leftarrow \mathbf{s}_{\text{new}}$
15 **end**
16 **return** $\mathbf{s}$

for each iteration, it creates a copy of the current node values (line 6) and then, for each node that is not sampled (lines 7-8), it calculates the average value of its neighbors if it has any (lines 9-11). The process continues until the maximum number of iterations is reached (lines 5 and 15). The resulting node values represent the recovered opinions for the entire graph (line 16). Hence, the computational cost of Algorithm 1 is $O(\texttt{max\_iter} \times (|V| \times d_{\text{avg}}))$.

D.2 Graph Neural Networks (GNN)

In this method, all node features are initialized to zero (lines 2-3). The values of the sampled nodes are set to their corresponding real values (lines 4-5). A train mask is created to indicate which nodes have known values (lines 6-9). A GCN model is defined and initialized with the specified number of hidden channels. The model is trained for the specified number of epochs (lines 11-15), where it learns to predict node values by minimizing the MSE loss between the predicted and known values for the sampled nodes. After training, the model is used to predict the values for all nodes in the graph (line 16), and the predicted values are returned (line 17). Hence, the computational cost of Algorithm 2 is $O(\texttt{epochs} \times |E| \times \texttt{hidden\_channels})$.

E Time Complexity

The computational complexity of the overall algorithm depends on three main components: **node selection**, **innate opinion reconstruction**, and **opinion optimization**. Below, we provide the detailed time complexities for each step and the methods employed.

Node Selection. The node selection step involves computing centrality metrics that determine the importance of nodes

Algorithm 2: Training of GNN-based Recovery

Data: Graph G, sampled indices S, real values $\mathbf{s}_{\mathrm{true}}$, $hidden\_channels$, number of $epochs$, learning rate lr
Result: Recovered node values $\mathbf{s}$

1 **Input:** Graph G, node features X, targets Y
2 $X \leftarrow \mathbf{0}_{N \times 1}$ `// Initialize node features`
3 **for** *each* $i \in S$ **do**
4 $\quad X[i] \leftarrow \mathbf{s}_{\mathrm{true}}[i]$
5 **end**
6 train_mask $\leftarrow \mathbf{0}_N$ `// Initialize train mask`
7 **for** *each* $i \in S$ **do**
8 $\quad$ train_mask$[i] \leftarrow$ True
9 **end**
10 Define and initialize the GCN model with $hidden\_channels$
11 **for** *epoch* $i = 1$ *to epochs* **do**
12 $\quad \hat{\mathbf{s}} \leftarrow f_\theta(X, G)$ `// Forward`
13 $\quad$ loss $\leftarrow \frac{1}{|\text{train mask}|} \sum_{j \in \text{train mask}} (\hat{\mathbf{s}}[j] - \mathbf{s}_{\mathrm{true}}[j])^2$ `// Backward pass and optimization step`
14 $\quad \theta \leftarrow \mathrm{ADAM}(\theta, lr, \nabla_\theta \text{loss}, \beta_1, \beta_2)$ `// Backward`
15 **end**
16 $\mathbf{s} \leftarrow f_\theta(X, G)$ `// Predicted values for all nodes`
17 **Output:** $\mathbf{s}$

in the graph. The time complexities for the most common methods are as follows:

- *Closeness Centrality*: $\mathcal{O}(mn)$, where m is the number of edges and n is the number of nodes.
- *Degree Centrality*: $\mathcal{O}(n \log n)$, representing the sorting of node degrees.
- *PageRank*: $\mathcal{O}(m\tau)$, where τ is the number of iterations of the power method used to compute PageRank.

Innate Opinion Reconstruction. The reconstruction of innate opinions from observed graph data can be performed using various methods, each with distinct computational costs:

- *Graph Neural Networks (GNN)*: $\mathcal{O}(\text{epochs} \cdot m \cdot \text{hidden}_{\text{channels}})$, where *epochs* is the number of training iterations, m is the edge count, and $\text{hidden}_{\text{channels}}$ refers to the size of hidden layers in the GNN (see Appendix B.1 for details).
- *Graph Signal Processing (GSP)*: $\mathcal{O}(n^3)$, due to the matrix inversion and eigendecomposition required to reconstruct signals.
- *Label Propagation (LP)*: $\mathcal{O}(\max\{\text{iter} \cdot n \cdot d_{\text{avg}}\})$, where *iter* is the number of iterations, n is the node count, and d_{avg} is the average node degree.

Opinion Optimization. The opinion optimization step focuses on minimizing both *polarization* and *disagreement* in the graph. The time complexity depends heavily on the gradient-based methods and the stopping conditions chosen. In the directed case, the time complexity is given by:

$$\mathcal{O}(T(m+n)\tau), \tag{14}$$

where T is the number of iterations of the BiConjugate Gradient Stabilized (BiCGStab) solver in each optimization step, m is the number of edges, n is the number of nodes, and τ is the number of iterations of the gradient-based algorithm.

F Running Time Analysis

We report the running time of our method when applied to real-world networks. The objective function used in the experiments is **PD-Dir**, and the node selection strategy is based on degree centrality with $b = 0.2n$.

Table 7 summarizes the running times of three different methods—**GNN** (Graph Neural Networks), **GSP** (Graph Signal Processing), and **LP** (Label Propagation)—applied to three real-world networks: *brexit*, *referendum*, and *vaxNoVax*.

Table 7: Running times of methods on real-world networks.

Model	Network	Time (sec)
GNN	brexit	231.64
GNN	referendum	26.24
GNN	vaxNoVax	685.95
GSP	brexit	768.04
GSP	referendum	40.51
GSP	vaxNoVax	680.71
LP	brexit	1431.46
LP	referendum	174.62
LP	vaxNoVax	4213.62

The results highlight that the GNN-based approach is the most scalable method, demonstrating superior efficiency for large-scale networks such as *vaxNoVax* with 1.6M edges. GSP shows acceptable scalability but becomes computationally intensive on larger graphs. In contrast, LP struggles with scalability.

G Experimental Setup

Graph Instances. We consider 11 real-world directed graphs and 5 undirected graphs related to social media/networks, ranging from 30 to 12,000 nodes and up to 1,600,000 edges. Table 2 summarizes the statistics of the datasets. The first three datasets contain the follow network on $\mathbb{X}$ and are derived from Minici *et al.* [24]. The other networks are related to social activities and obtained from KONECT[2] [20].

We preprocess the graphs by removing disconnected nodes with zero outdegree, self-loops, and multiple edges. We ensure that all weights, if available, are positive. In the case of directed networks, we row-normalize the adjacency matrix as in Cinus *et al.* [9]. A directed edge (u, v) indicates that u has visibility of v's contents, and hence u can be influenced by v.

Opinions Instances. The `Referendum`, `Brexit`, and `VaxNoVax` datasets contain exposed opinions for each node, which are the average stances of tweets retweeted by the users; each tweet's stance was estimated with a supervised text-based classifier [24]. For graphs where such real proxies of opinions are not available, we generate opinions by sampling from either a uniform distribution between $[-0.5, +0.5]$ or from Gaussian distributions. To induce structural polarization, we use the Kernighan-Lin bisection method [18] to identify two communities in the graph. We then assign to each community average opinions symmetrically distant around 0. Node opinions are subsequently sampled from the Gaussian distribution centered around the average opinion assigned to their community. We preprocess opinions by inserting polarization in the statistical distribution with a parameter p using the function $f(x|p) = |x|^{\frac{1}{p}}$, using $p = 3$. We assume that the available opinions are those at equilibrium, hence we invert the FJ model dynamics to recover the innate opinions from the exposed one. Then we mean-centered the distribution according to Musco *et al.* [25] so that the average innate opinion is zero and we set its standard deviation is 1 (Table 2 shows the samples of these distributions).

Parameters. In the objective functions for directed graphs, to obtain fast convergence to a local minimum, the directed problems are solved using ADAM with $\beta_1 = 0.9$, $\beta_2 = 0.999$, and learning rate $\eta = 0.2$. We use early stopping when the relative change of the objective between different steps is less than 20%. For the problem with PD-UNDIR, to reach the global optimum, we used the Splitting Conic Solver (SCS) [28, 29] of CVXPY with 2,500 maximum number of iterations and a tolerance of 10^{-3}.

In opinion reconstruction, we set the querying budget b to 20% of nodes which enables fair comparison across all experiments, methods, and datasets. We also tested the effect of the number of budget b in reconstruction and optimization (See Figures 2 and 3). In the GSP strategy, we set the number of frequencies equal to 15% of nodes to satisfy the $b \geq |F|$. The dependency with respect to the number of frequencies and sampled nodes is reported in Figure 3. We also symmetrize the adjacency matrix to use a unique standard GSP method for undirected graphs. In the GNN strategy, we used GCN model [19] with 16 hidden channels, 200 epochs, and a learning rate of 0.01. In the LP strategy, we used 200 iterations. The above parameter settings were empirically evaluated to obtain a tradeoff between convergence and running time.

H Additional Experiments

H.1 Performance in Directed graphs with uniform opinion distribution

Results are presented in Table 8 are consistent with those in the main text. Here we consider the 8 real-world directed networks in Table 2 with a uniform opinion distribution.

In general, similarly to the polarized opinions configuration, our pipeline yields multiplicative errors usually below 2 compared to directly optimizing over error-free opinions. The LP reconstruction methodology consistently achieves multiplicative errors below 2; except for the "oz" network, which is consistently difficult to optimize for all methods. In "ciaodvd-trust" now all methods achieve comparable results,

[2] http://konect.cc/

meaning that the main difficulty for the methods was the presence of structural polarization.

Table 8: Average Multiplicative Error for the three objectives (P-DIR, D-DIR, PD-DIR) for 8 real-world directed graphs with different sizes (n). Opinions are uniformly distributed and reconstructed with $b = 0.20|V|$ nodes selected based on their degree.

Objective	Rec Method / Network	Multiplicative Error GNN	GSP	LP
P-DIR	highschool	2.02 ± 0.25	1.97 ± 0.24	$\mathbf{1.90 \pm 0.20}$
	wiki talk	1.27 ± 0.23	$\mathbf{1.21 \pm 0.13}$	1.33 ± 0.24
	innovation	1.89 ± 0.14	1.96 ± 0.20	$\mathbf{1.81 \pm 0.13}$
	oz	1.97 ± 0.10	1.93 ± 0.08	$\mathbf{1.88 \pm 0.09}$
	filmtrust	1.33 ± 0.05	1.35 ± 0.05	$\mathbf{1.27 \pm 0.04}$
	dnc-temporal	1.24 ± 0.05	$\mathbf{1.21 \pm 0.05}$	1.28 ± 0.11
	ciaodvd-trust	1.76 ± 0.05	1.86 ± 0.07	$\mathbf{1.64 \pm 0.04}$
	health	1.82 ± 0.03	1.83 ± 0.04	$\mathbf{1.72 \pm 0.03}$
D-DIR	highschool	1.67 ± 0.13	1.67 ± 0.10	$\mathbf{1.62 \pm 0.11}$
	wiki talk	1.20 ± 0.16	$\mathbf{1.18 \pm 0.09}$	1.28 ± 0.16
	innovation	1.82 ± 0.14	1.86 ± 0.11	$\mathbf{1.67 \pm 0.10}$
	oz	2.61 ± 0.12	2.52 ± 0.13	$\mathbf{2.46 \pm 0.11}$
	filmtrust	1.24 ± 0.03	1.24 ± 0.03	$\mathbf{1.23 \pm 0.03}$
	dnc-temporal	1.31 ± 0.05	$\mathbf{1.23 \pm 0.06}$	1.40 ± 0.08
	ciaodvd-trust	2.01 ± 0.06	1.99 ± 0.05	$\mathbf{1.75 \pm 0.03}$
	health	1.54 ± 0.03	1.54 ± 0.02	$\mathbf{1.45 \pm 0.02}$
PD-DIR	highschool	1.35 ± 0.08	1.31 ± 0.08	$\mathbf{1.27 \pm 0.08}$
	wiki talk	1.17 ± 0.18	$\mathbf{1.13 \pm 0.09}$	1.23 ± 0.19
	innovation	1.30 ± 0.06	1.31 ± 0.07	$\mathbf{1.24 \pm 0.05}$
	oz	1.37 ± 0.05	1.35 ± 0.03	$\mathbf{1.34 \pm 0.05}$
	filmtrust	1.14 ± 0.03	1.13 ± 0.02	$\mathbf{1.08 \pm 0.02}$
	dnc-temporal	1.11 ± 0.03	$\mathbf{1.09 \pm 0.03}$	$\mathbf{1.09 \pm 0.07}$
	ciaodvd-trust	1.30 ± 0.02	1.27 ± 0.02	$\mathbf{1.19 \pm 0.02}$
	health	1.29 ± 0.01	1.28 ± 0.01	$\mathbf{1.24 \pm 0.01}$

H.2 Performance in Undirected graphs with uniform opinion distribution

Results are presented in Table 9. We consider the 5 real-world undirected networks in Table 2 with uniform opinion distribution.

Reducing polarization and disagreement in a non-polarized configuration is more difficult if solved with reconstructed opinions compared to polarized configurations. In fact, multiplicative errors are higher compared to the ones in Table 5 in the main text. Again, errors show greater variability across different graphs in undirected settings compared to directed ones. This is because the global optimum corresponds to numerically low levels of polarization and disagreement. As a result, the denominator in the multiplicative error calculation is often smaller than 1, amplifying the observed errors. Given that opinions are uniform and so they reflect no structure, bounds are both sensitive to the difficulty of opinion reconstruction (which exploits the graph structure) and the numerical value of the global minimum of the objective. As a result, the bounds are larger than the actual error by one order of magnitude, indicating that, in practice, the problem is less challenging than theoretically predicted, and a tighter bound likely exists at this graph size scale.

Table 9: Average Multiplicative error in PD-UNDIR minimization in undirected graphs with uniform opinion distribution.

Rec Method / Network	Multiplicative Error GNN	GSP	LP
zachary	3.70 ± 0.78 (10)	$\mathbf{2.74 \pm 0.59}$ (11)	3.67 ± 0.92 (11)
beach	$\mathbf{6.67 \pm 4.12}$ (138)	10.93 ± 5.12 (166)	7.37 ± 7.12 (104)
train	$\mathbf{3.64 \pm 1.18}$ (11)	3.84 ± 1.07 (12)	4.24 ± 1.17 (11)
mit	$\mathbf{0.99 \pm 0.03}$ (10763)	1.00 ± 0.01 (12132)	1.00 ± 0.02 (7432)
football	6.31 ± 1.12 (14)	$\mathbf{6.25 \pm 1.20}$ (15)	7.00 ± 1.33 (13)

H.3 Performance of Node Selection Strategies in Directed graphs

Also in directed graphs with synthetic opinions, degree and PageRank are the best strategies for selecting nodes. Compared to real datasets, the increment with respect to random selection is smaller. Furthermore, in the "wiki talk" network random selection is comparable or superior to other methods. We hypothesize that this is due to the unique structural characteristics of the graph. Notably, the "wiki talk ht" and "dnc-temporalGraph" networks exhibit the lowest degree assortativity coefficient (-0.4), indicating that highly connected nodes tend to connect with low-degree peripheral ones. In such cases, labels propagating from hubs (selected by degree-based methods) are likely to remain confined within localized areas, which can lead to suboptimal performance. Additionally, we highlight that the Random Strategy is far from an arbitrary or ineffective approach. Selecting nodes uniformly at random often results in coverage of diverse regions of the network. Given the strong homophily typically observed in real-world innate opinions, this strategy can provide a reasonable basis for querying nodes to infer their innate opinions effectively.

Finally, results are consistent across the two different opinion distributions: polarized 10 and uniform 11. "ciaodvd-trust" results the most difficult network with polarized opinions, with multiplicative errors above 2 while reducing P-DIR. On the contrary, the non-polarized configuration in the "oz" dataset has the highest multiplicative errors in disagreement minimization.

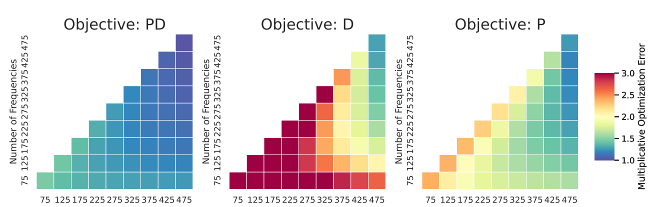

Figure 3: Average multiplicative error vs the number of selected nodes vs the number of frequencies. We used an Erdos Renyi graph model with $|V| = 500, p = 0.25$, and polarized opinions.

H.4 Sensitivity Analysis in Graph Signal Processing

In Figure 3 we quantify the multiplicative error induced by the selected nodes size and frequencies using the Graph Signal Processing reconstruction strategy. Recalling that the

Table 10: Multiplicative Error for the three objectives (P-DIR, D-DIR, PD-DIR) for different node selection strategies in directed networks with polarized opinions.

Objective	Sel Method / Network	Multiplicative Error Closeness centrality	Degree	PageRank	Random
P-DIR	highschool	2.01 ± 0.42	$\mathbf{1.68 \pm 0.28}$	1.71 ± 0.26	1.95 ± 0.33
	wiki talk	$\mathbf{1.31 \pm 0.24}$	1.34 ± 0.27	1.35 ± 0.27	1.32 ± 0.21
	innovation	1.77 ± 0.20	1.80 ± 0.21	$\mathbf{1.73 \pm 0.18}$	2.15 ± 0.32
	oz	1.80 ± 0.20	1.78 ± 0.17	$\mathbf{1.77 \pm 0.19}$	1.82 ± 0.17
	filmtrust-trust	1.32 ± 0.07	1.27 ± 0.06	$\mathbf{1.25 \pm 0.06}$	1.40 ± 0.10
	dnc-temporal	1.28 ± 0.16	1.25 ± 0.14	$\mathbf{1.24 \pm 0.13}$	1.33 ± 0.10
	ciaodvd-trust	2.53 ± 0.14	2.54 ± 0.14	2.43 ± 0.15	$\mathbf{2.01 \pm 0.10}$
	health	1.60 ± 0.08	$\mathbf{1.57 \pm 0.06}$	$\mathbf{1.57 \pm 0.07}$	1.68 ± 0.08
D-DIR	highschool	1.51 ± 0.23	$\mathbf{1.42 \pm 0.15}$	1.48 ± 0.18	1.51 ± 0.21
	wiki talk	1.36 ± 0.20	1.34 ± 0.17	1.35 ± 0.21	$\mathbf{1.28 \pm 0.16}$
	innovation	1.52 ± 0.20	$\mathbf{1.51 \pm 0.15}$	1.52 ± 0.19	1.71 ± 0.21
	oz	1.47 ± 0.16	1.47 ± 0.14	$\mathbf{1.45 \pm 0.14}$	1.57 ± 0.13
	filmtrust	$\mathbf{1.20 \pm 0.04}$	1.22 ± 0.05	1.22 ± 0.04	1.25 ± 0.07
	dnc-temporal	1.33 ± 0.10	1.35 ± 0.11	1.35 ± 0.11	$\mathbf{1.29 \pm 0.09}$
	ciaodvd-trust	1.51 ± 0.03	1.52 ± 0.03	1.50 ± 0.03	$\mathbf{1.48 \pm 0.04}$
	health	1.37 ± 0.05	$\mathbf{1.36 \pm 0.04}$	$\mathbf{1.36 \pm 0.05}$	1.40 ± 0.05
PD-DIR	highschool	1.47 ± 0.19	$\mathbf{1.32 \pm 0.09}$	1.36 ± 0.12	1.46 ± 0.13
	wiki talk	$\mathbf{1.26 \pm 0.22}$	1.28 ± 0.23	1.29 ± 0.23	$\mathbf{1.26 \pm 0.19}$
	innovation	1.36 ± 0.10	1.35 ± 0.10	$\mathbf{1.34 \pm 0.11}$	1.53 ± 0.15
	oz	$\mathbf{1.36 \pm 0.10}$	1.40 ± 0.09	1.38 ± 0.11	1.41 ± 0.11
	filmtrust	1.18 ± 0.04	1.15 ± 0.03	$\mathbf{1.13 \pm 0.04}$	1.23 ± 0.06
	dnc-temporal	1.20 ± 0.13	$\mathbf{1.16 \pm 0.12}$	$\mathbf{1.16 \pm 0.11}$	1.22 ± 0.10
	ciaodvd-trust	1.55 ± 0.05	1.56 ± 0.05	1.52 ± 0.05	$\mathbf{1.30 \pm 0.04}$
	health	$\mathbf{1.29 \pm 0.03}$	$\mathbf{1.29 \pm 0.02}$	$\mathbf{1.29 \pm 0.03}$	1.34 ± 0.04

Table 11: Multiplicative Error for the three objectives (P-DIR, D-DIR, PD-DIR) for different node selection strategies in directed networks with uniform opinion distribution.

Objective	Sel Method / Network	Multiplicative Error Closeness centrality	Degree	PageRank	Random
P-DIR	highschool	1.89 ± 0.26	1.90 ± 0.20	$\mathbf{1.84 \pm 0.21}$	1.91 ± 0.21
	wiki talk	1.34 ± 0.21	1.33 ± 0.24	1.36 ± 0.25	$\mathbf{1.31 \pm 0.21}$
	innovation	$\mathbf{1.80 \pm 0.13}$	1.81 ± 0.13	$\mathbf{1.80 \pm 0.13}$	1.86 ± 0.15
	oz	$\mathbf{1.84 \pm 0.09}$	1.88 ± 0.09	1.86 ± 0.08	1.93 ± 0.09
	filmtrust	1.30 ± 0.04	$\mathbf{1.27 \pm 0.04}$	1.31 ± 0.05	1.37 ± 0.05
	dnc-temporal	1.29 ± 0.12	1.28 ± 0.11	1.29 ± 0.11	$\mathbf{1.27 \pm 0.07}$
	ciaodvd-trust	$\mathbf{1.63 \pm 0.03}$	1.64 ± 0.04	1.64 ± 0.04	1.70 ± 0.04
	health	$\mathbf{1.70 \pm 0.03}$	1.72 ± 0.03	1.71 ± 0.03	1.79 ± 0.04
D-DIR	highschool	1.61 ± 0.11	1.62 ± 0.11	$\mathbf{1.58 \pm 0.10}$	1.66 ± 0.10
	wiki talk	1.31 ± 0.14	1.28 ± 0.16	1.28 ± 0.13	$\mathbf{1.22 \pm 0.14}$
	innovation	1.68 ± 0.11	$\mathbf{1.67 \pm 0.10}$	1.69 ± 0.10	1.83 ± 0.13
	oz	$\mathbf{2.44 \pm 0.11}$	2.46 ± 0.11	2.45 ± 0.13	2.56 ± 0.11
	filmtrust	1.24 ± 0.02	$\mathbf{1.23 \pm 0.03}$	1.25 ± 0.03	1.27 ± 0.03
	dnc-temporal	1.37 ± 0.08	1.40 ± 0.08	1.41 ± 0.08	$\mathbf{1.36 \pm 0.05}$
	ciaodvd-trust	$\mathbf{1.75 \pm 0.03}$	$\mathbf{1.75 \pm 0.03}$	$\mathbf{1.75 \pm 0.03}$	2.01 ± 0.06
	health	$\mathbf{1.45 \pm 0.02}$	$\mathbf{1.45 \pm 0.02}$	$\mathbf{1.45 \pm 0.02}$	1.54 ± 0.02
PD-DIR	highschool	$\mathbf{1.26 \pm 0.10}$	1.27 ± 0.08	1.27 ± 0.08	1.32 ± 0.09
	wiki talk	1.23 ± 0.20	1.23 ± 0.19	1.26 ± 0.24	$\mathbf{1.22 \pm 0.19}$
	innovation	1.25 ± 0.06	$\mathbf{1.24 \pm 0.05}$	1.25 ± 0.06	1.32 ± 0.10
	oz	$\mathbf{1.31 \pm 0.04}$	1.34 ± 0.05	1.34 ± 0.05	1.36 ± 0.05
	filmtrust	1.11 ± 0.02	$\mathbf{1.08 \pm 0.02}$	1.10 ± 0.02	1.16 ± 0.03
	dnc-temporal	1.10 ± 0.07	$\mathbf{1.09 \pm 0.07}$	$\mathbf{1.09 \pm 0.07}$	1.16 ± 0.05
	ciaodvd-trust	$\mathbf{1.19 \pm 0.01}$	$\mathbf{1.19 \pm 0.02}$	$\mathbf{1.19 \pm 0.01}$	1.29 ± 0.02
	health	$\mathbf{1.23 \pm 0.02}$	1.24 ± 0.01	1.24 ± 0.01	1.29 ± 0.01

Table 12: Multiplicative Error for the P-DIR objective for degree node selection strategy in directed networks with gaussian opinion distribution.

Sel Method Network	Multiplicative Error LP	Random
highschool	$\mathbf{1.90 \pm 0.20}$	2.26 ± 0.34
wiki talk	$\mathbf{1.33 \pm 0.24}$	1.36 ± 0.15
innovation	$\mathbf{1.81 \pm 0.13}$	2.46 ± 0.22
oz	$\mathbf{1.88 \pm 0.09}$	2.92 ± 0.14
filmtrust	$\mathbf{1.27 \pm 0.04}$	1.65 ± 0.09
temporal	$\mathbf{1.28 \pm 0.11}$	1.55 ± 0.10
ciaodvd	$\mathbf{1.64 \pm 0.04}$	2.73 ± 0.10
health	$\mathbf{1.72 \pm 0.03}$	2.08 ± 0.07

number of frequencies must be lower or equal to the number of selected nodes [21], we observe that multiplicative errors monotonically decrease along both dimensions as expected. Furthermore, the disagreement objective is the most difficult to optimize.

H.5 Additional Baseline

To further evaluate our approach, we tested an additional baseline where reconstructed innate opinions were assigned values drawn uniformly at random from the range of selected seed opinions.

We compared the performance of LP Opinion Reconstruction (LP) against this Random baseline under the condition where innate opinions were sampled from a Gaussian distribution, with the objective of minimizing P-DIR. The results indicate that the Random baseline consistently underperforms compared to LP Opinion Reconstruction.

Results are reported in Table 12.