# De-anonymizing Social Networks with Overlapping Community Structure

Luoyi Fu[1], Xinyu Wu[2], Shuaiqi Wang[2], Yichi Zhang[2], Zhongzhao Hu[2], Xinbing Wang[1,2], and Guihai Chen[1]

[1]Dept. of Computer Science, Shanghai Jiao Tong University, China.
[1]Email:{yiluofu, xwang8}@sjtu.edu.cn,gchen@cs.sjtu.edu.cn
[2]Dept. of Electrical Engineering, Shanghai Jiao Tong University, China.
[2] Email:{wuxinyu, wangshuaiqi, zyc1357038675, hzz5611577}@sjtu.edu.cn

***Abstract*—The advent of social networks poses severe threats on user privacy as adversaries can de-anonymize users' identities by mapping them to correlated cross-domain networks. Without ground-truth mapping, prior literature proposes various cost functions in hope of measuring the quality of mappings. However, there is generally a lacking of rationale behind the cost functions, whose minimizer also remains algorithmically unknown.**

**We jointly tackle above concerns under a more practical social network model parameterized by *overlapping communities*, which, neglected by prior art, can serve as side information for de-anonymization. Regarding the unavailability of ground-truth mapping to adversaries, by virtue of the Minimum Mean Square Error (MMSE), our first contribution is a well-justified cost function minimizing the expected number of mismatched users over all possible true mappings. While proving the NP-hardness of minimizing MMSE, we validly transform it into the weighted-edge matching problem (WEMP), which, as disclosed theoretically, resolves the tension between optimality and complexity: (i) WEMP asymptotically returns a negligible mapping error in large network size under mild conditions facilitated by higher overlapping strength; (ii) WEMP can be algorithmically characterized via the convex-concave based de-anonymization algorithm (CBDA), perfectly finding the optimum of WEMP. Extensive experiments further confirm the effectiveness of CBDA under overlapping communities: $90\%$ users are re-identified averagely in a series of networks when communities overlap densely, and the re-identification ratio is enhanced about $70\%$ compared to non-overlapping cases.**

## I. INTRODUCTION

With the mounting popularity of social networks, the privacy of users has been under great concern, as information of users in social networks is often released to public for wide usage in academy or advertisement [8]. Although users can be anonymized by removing personal identifiers such as names and family addresses, it is not sufficient for privacy protection since adversaries may re-identify these users by correlated side information.

Such user identification process in social networks resorting to auxiliary information is called *Social Network De-anonymization*. Initially proposed by Narayanan and Shimatikov [2], this fundamental issue has then gained increasing attention, leading to a large body of subsequent works [3], [4], [6], [7], [5], [8], [9]. Particularly, this family of works embarked on de-anonymization under a common framework, as will also be the framework of interest in our setting. To elaborate, in the framework there is an underlying network $G$ which characterizes the relationship among users. Then there are two networks observed in reality, named as published network $G_1$ and auxiliary network $G_2$, whose node sets are identical and edges are randomly and independently sampled from $G$ with probability $s_1$ and $s_2$ respectively. *The aim of de-anonymization is to discover the correct mapping between $V_1$ and $V_2$, the node set of $G_1$ and $G_2$, which corresponds the same user in two networks, with the network structure as the only side information available to the adversaries.*

Regardless of the considerable efforts paid to de-anonymization, there is still a severe lacking of a comprehensive understanding about the conditions under which the adversaries can perfectly de-anonymize user identities. It can be accounted for from three aspects. (i) Analytically, despite a variety of existing work [3], [4] that proposed several cost functions in measuring the quality of mappings, the theoretical devise of those costs functions lacks sufficient rationale behind. (ii) Algorithmically, previous works [3], [4] failed to provide any algorithm to demonstrate that the optimal solution of proposed cost functions can indeed be effectively obtained. (iii) Experimentally, due to the destitution of real cross-domain datasets, state-of-the-art research [6], [7] simply evaluated the performance of proposed algorithms on synthetic datasets or real cross-domain networks formed by artificial sampling, falling short of reproducing the genuine social networks.

**The above limitations motivate us to shed light on de-anonymization problem by jointly incorporating analytical, algorithmic and experimental aspects under the common framework noted earlier.** As far as we know, the only work that shares the closest correlation with us belongs to Fu et. al. [23], [24], who investigated this problem on social networks with non-overlapping communities and derived their cost function from the Maximum A Posterior (MAP) manner. However, we remark that the assumption of disjoint communities fails to reflect the real situation where a user belongs to multiple communities, as observed in massive real situations. For example, in social networks of scientific collaborators [9], actors and political blogospheres [18], users might belong to several research groups with different research topics, movies and political parties respectively. Furthermore,

*The early version of this paper is to appear in the Proceedings of IEEE INFOCOM 2017 [25].*

while MAP enables adversaries to find the correct mapping with the highest probability, it relies heavily on a prerequisite, i.e., a hypothetically true mapping between the given published and auxiliary networks. However, once the MAP estimation fails to exactly match this "true" mapping, then the mapping error becomes unpredictable, with the probability that the estimation deviates largely from the real ground-truth. For the first concern, by adopting the overlapping stochastic block model (OSBM), we allow the communities to overlap arbitrarily, which can well capture a majority of real social networks. For the second concern, we derive our cost function based on Minimum Mean Square Error (MMSE), which minimizes the expected number of mismatched users by incorporating all the possible true mappings between the given published and auxiliary networks. This incorporation, from an average perspective, keeps the estimation of MMSE from significant deviation from any possible hypothetic true mapping.

Hereinafter we unfold our main contributions in analytical, algorithmic and experimental aspects respectively as follows:

1. Analytically, we are the first to derive cost function based on MMSE, which justifiably ensures the minimum expected mapping error between our estimation and the ground-truth mapping. Then we demonstrate the NP-hardness of solving MMSE, whose intractability stems mainly from the calculation of all $n!$ possible mappings ($n$ is the total number of users). To cope with the hardness, we simplify MMSE by transforming it into a weighted-edge matching problem (WEMP), with mapping error negatively related to weights.

2. Algorithmically, in terms of solving WEMP, we theoretically reveal that WEMP alleviates the tension between optimality and complexity: Solving WEMP ensures optimality since its optimum, in large network size, negligibly deviates from the ground-truth mapping under mild conditions where on average a user belongs to asymptotically non-constant communities. Meanwhile it reduces complexity since perfectly deriving its optimum only entails a convex-concave based de-anonymization algorithm (CBDA) with polynomial time. The proposed CBDA serves as one of the very few attempts to address the algorithmic characterization, that has long remained open, of de-anonymization without pre-identification.

3. Experimentally, we validate our theoretical findings that minimizing WEMP indeed incurs negligible mapping error in large social networks based on real datasets. Interestingly, we also observe significant benefits that community overlapping effect brings to the performance of CBDA: (i) in notable true cross-domain co-author networks with dense overlapping communities, CBDA can correctly re-identify $90\%$ nodes on average; (ii) the overlapping communities bring about an enhancement of around $70\%$ re-identification ratio compared with non-overlapping cases.

Unlike de-anonymization with pre-identified seed nodes, to which a family of work pays endeavor, no prior knowledge of such seeds complicates this problem, thus leaving many aspects largely unexplored. Meanwhile, theoretical results on such seedless cases in prior art is short of experimental verification. Our work is, as far as we are concerned, the initial devotion to theoretically dissecting seedless cases with overlapping communities, under real cross-domain networks with more than $3000$ nodes. With novel exploitations of structural information, future design of more efficient mechanisms will be expected to further dilute the limitation of network size.

## II. Related Works

Social network de-anonymization problem has been in the dimelight in recent decades. Narayanan and Shimatikove [2] formulated this problem initially. They presented its framework and proposed a generic algorithm, which did not utilize any side information except the network structure and worked based on some pre-identified nodes, called seed nodes.

Predicated on this seminal paper, a large amount of work emerges focusing on different facets of de-anonymization. One major division is whether the anonymized network is seeded or seedless, i.e., whether pre-identified nodes exist. For seeded anonymized network, as the pioneering work [2], the common idea to solve the problem is to design algorithms based on *percolation*, which means that the re-identification process starts from the seed nodes and identify their neighbor nodes iteratively until all the nodes are de-anonymized [2], [12], [13], [14], [15]. Yartseva et al. [12], Kazemi et al. [13] and Fabiana et al. [14] studied seeded problem under Erdös-Rényi graph model, while Korula and Lattenzi [15] shed light on preferential attachment model.

However, in real situations it is often the case that adversaries are difficult to obtain seeded nodes before de-anonymizing [23], [24] due to the limited access to user profiles. For seedless networks, the major methodology is to propose cost functions and obtain an estimation of the correct mapping between two networks by optimizing these cost functions. Pedarsani and Grossglauer [3] are the precursors in de-anonymizing seedless networks. They studied this problem under Erdös-Rényi graph and their cost function was the number of mismatched edges. With the same cost function, Kazemi et al. [4] considered the situation where the nodes in two networks are overlapping partially, and Cullina and Kiyavash [5] further investigated the information-theoretic threshold for exact identification in [3]. However, the cost functions in [3], [4], [5] didn't have theoretical supports. One cost function based on Maximum A Posterior (MAP) has been justified by [8], [23], [24]. Onaran et al. [8] theoretically proved the validity of MAP and Fu et al. [23], [24] provided two approximation algorithms to solve this problem.

Another facet for de-anonymization problem is the amount of side information adversaries have. A large amount of work [2], [12], [13], [14], [15], [3], [4], [5], either in seeded or seedless situations, studied this problem without any side information except the topological structure of two networks, i.e., the edge sets in two networks. However, the clustering effect exists in real social networks, which has not been considered in work above. To incorporate clustering effect, Chiasserini et al. [16] studied clustering under seeded problem and drew the conclusion that the impact of clustering is double-edged, which may dramatically reduce the required seed nodes but make the algorithm more fragile to errors. Onaran et al. [8] and Fu et al. [23], [24] both studied clustering by modeling it as communities in two networks, and Fu et al.

[23], [24] showed that the side information of communities makes for higher accuracy of the algorithms intended for seedless problem. However, as far as we know, no existing work has ever focused on overlapping communities, which is omnipresent in real situations, especially the large-scale social networks nowadays.

## III. Models and Definitions

In this section, we will introduce the fundamental model and some related definitions. Before we start, we list some basic notations frequently used in our later analysis.

### A. Preliminary Notations

**Definition 1.** ***(Expectation Over Matrix)*** *Given a random matrix variable $\mathbf{A}$ and a function $f(\mathbf{A})$, the expectation of $f(\mathbf{A})$ over matrix $\mathbf{A}$ is denoted as $\boldsymbol{E}_{\mathbf{A}}(f(\mathbf{A}))$.*

**Definition 2.** ***(Frobenius Norm)*** *Given an $m \times n$ matrix $\mathbf{X}$, the Frobenius norm of $\mathbf{X}$ is $||\mathbf{X}||_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}(\mathbf{X}_{ij}^2)}$, where $\mathbf{X}_{ij}$ is the element at the $i_{th}$ row and $j_{th}$ column of $\mathbf{X}$.*

**Definition 3.** ***(Hadamard Product)*** *Given two $n \times n$ matrices $\mathbf{Y}$ and $\mathbf{Z}$, The Hadamard Product between $\mathbf{Y}$ and $\mathbf{Z}$ is defined as $\forall i, j \in \{1, 2, ..., n\}, (\mathbf{Y} \circ \mathbf{Z})_{ij} = \mathbf{Y}_{ij}\mathbf{Z}_{ij}$.*

### B. Social Network Models

The social network model considered in this paper is composed of three parts, i.e., the underlying network $G$, the published network $G_1$ and the auxiliary network $G_2$. $G_1$ and $G_2$ can be viewed as the incomplete observations of $G$, which represents the underneath relationship among all users. For instance, in reality $G$ may characterize the true underlying relationship among a group of people, while $G_1$ might represent the online network in Facebook of this group of people and $G_2$ might represent the communication records in the cell phones of them, both of which are observable.

*1) Underlying Social Network:* Let $G = (V, E, \mathbf{U})$, where $V$ is the node set, $E$ is the edge set and $\mathbf{U}$ is the adjacent matrix. We regard $G$ as undirected with $|V| = n$ nodes. To reflect the property of overlapping communities, we suppose $G$ is generated based on the overlapping stochastic block model (OSBM) [18], whose idea can be interpreted as follows:

Suppose there are $Q$ communities and each community $q$ contains a subset of nodes. For a generic node $i$, we introduce a latent $Q$-dimensional column vector $\boldsymbol{C_i}$, in which all elements are independent boolean variables $C_{iq} \in \{0, 1\}$, with $C_{iq}$ being the $q_{th}$ row in $\boldsymbol{C_i}$. $C_{iq} = 1$ means that node $i$ is in community $q$ and $C_{iq} = 0$ otherwise. We denote $p_q$ to be the probability of any node in $G$ belonging to community $q$, thus we have: $Pr(\boldsymbol{C_i} = \{C_{i1}, C_{i2}, \cdots, C_{iQ}\}^T) = \prod_{q=1}^{Q}(p_q)^{C_{iq}}(1-p_q)^{1-C_{iq}}$. We call $\boldsymbol{C_i}$ the *community representation* of node $i$, since $\boldsymbol{C_i}$ shows to which communities node $i$ belongs exactly.

In OSBM, the probability of edge existence between nodes $i$ and $j$ in $G$ relies on $\boldsymbol{C_i}$ and $\boldsymbol{C_j}$. Hence we denote $Pr\{(i,j) \in E\} = p_{\boldsymbol{C_i}\boldsymbol{C_j}}$, where $p_{\boldsymbol{C_i}\boldsymbol{C_j}}$ is pre-defined depending on the number of communities nodes $i$ and $j$ co-exist in, which is easy to obtain in real de-anonymization.

*2) Published Network and Auxiliary Network:* We let $G_1(V_1, E_1, \mathbf{A})$ denote the published network, whose node labeling is identical with the underlying graph $G$ and edges are independently sampled from $G$ with probability $s_1$. In contrast, an auxiliary network, denoted by $G_2(V_2, E_2, \mathbf{B})$, does not necessarily share the same node labeling as $G$, and the edges are independently sampled from $G$ with probability $s_2$. $\mathbf{A}$ and $\mathbf{B}$ respectively represent the adjacency matrix of $G_1$ and $G_2$. In correspondence to real situations, $G_1$ characterizes the anonymized network where users' identities are unavailable for privacy concern. On the contrary, $G_2$ characterizes an un-anonymized network where users' identities are all available.

Adversaries can leverage $G_2$ to identify nodes in $G_1$ based on the edge relationship and community information: (i) For edge relationship, adversaries can harness the *degree similarity* that a node of high degree in $G_1$ should be inclined to match a node of high degree in $G_2$; (ii) For community information, adversaries can exploit the *community representation similarity* that nodes in $G_1$ and $G_2$ with the same community representation should be matched with higher probability.

For the edge set $E_k$ $(k \in \{1, 2\})$ of either network, $Pr\{(i,j) \in E_k\} = s_k$ if $(i,j) \in E_k$ and $Pr\{(i,j) \in E_k\} = 0$ otherwise. For the node sets $V_1$ and $V_2$, we assume same number of nodes in $G$, $G_1$ and $G_2$, i.e., $|V| = |V_1| = |V_2| = n$.

Furthermore, we should clarify that we render each node pair $(i,j)$ a weight $w_{ij}$, which, quantified in Section III-C, is the cost of mistakenly matching the node pair $(i,j)$ and is contingent on $p_{\boldsymbol{C_i}\boldsymbol{C_j}}$, $s_1$ and $s_2$. As we will show in Section IV-A, $w_{ij}$ is negatively proportional to the number of communities nodes $i$ and $j$ co-exist in, evincing the cost reduction arose from higher *overlapping strength* of communities.

**Remark:** In fact $G$, $G_1$ and $G_2$ are all random variables. We directly use $G$, $G_1$, $G_2$ as notations for the realizations of these random variables with no loss of clearance. Moreover, we set $\boldsymbol{\theta} = \{\{p_{\boldsymbol{C_i}\boldsymbol{C_j}} | 1 \le i, j \le n\}, s_1, s_2\}$ as the parameter set incorporating all pre-defined parameters in the model.

### C. Social Network De-anonymization

The goal of social network de-anonymization problem is to find a mapping $\pi : V_1 \mapsto V_2$, which corresponds nodes on behalf of the same user in $G_1$ and $G_2$. We can equivalently express this mapping by forming a permutation matrix $\mathbf{\Pi} \in \{0,1\}^{n \times n}$, where $\mathbf{\Pi}(i,j) = 1$ if $\pi(i) = j$ and $\mathbf{\Pi}(i,j) = 0$ otherwise (If $|V_1| \neq |V_2|$, then $\mathbf{\Pi}$ is a non-square matrix which does not affect our analysis and algorithm design). We denote $\mathbf{\Pi_0}(\pi_0)$ as the true permutation matrix (mapping) between $G_1$ and $G_2$. We do not have any prior knowledge of $\mathbf{\Pi_0}$ and access to the underlying graph $G$. We formally define the social network de-anonymization problem in Definition 4 along with an illustrative instance in Fig. 1.

**Definition 4.** ***(Social Network De-anonymization Problem)*** *Given the published network $G_1$, the auxiliary network $G_2$, parameter set $\boldsymbol{\theta}$, social network de-anonymization problem aims to construct the true mapping $\pi_0$ between $V_1$ and $V_2$.*

However, our estimated permutation, $\hat{\mathbf{\Pi}}$, may deviate from the ground-truth $\mathbf{\Pi_0}$. To quantify this difference, we introduce a metric called "node mapping error (NME)" as follows.

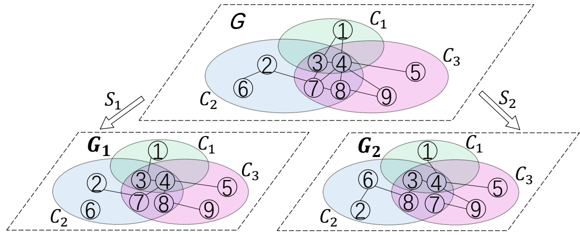



Fig. 1: An example of $G$, $G_1$ and $G_2$. The edges of $G_{1(2)}$ are sampled independently from $G$ with probability $s_{1(2)}$. $C_1, C_2, C_3$ denote 3 different communities in OSBM. The true mapping $\pi_0 = \{(1,1),(2,6),(3,3),(4,4),(5,5),(6,2),(7,8),(8,7),(9,9)\}$.

**Definition 5.** ***(Node Mapping Error)*** *Given the estimated* $\tilde{\mathbf{\Pi}}$ *and ground-truth* $\mathbf{\Pi_0}$*, the node mapping error (NME) between* $\hat{\mathbf{\Pi}}$ *and* $\mathbf{\Pi_0}$ *is defined as* $d(\hat{\mathbf{\Pi}}, \mathbf{\Pi_0}) = \frac{1}{2}||\hat{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2$.

Obviously $d(\hat{\mathbf{\Pi}}, \mathbf{\Pi_0})$ equals to 0 if and only if two permutations are identical, and if $k$ nodes are mapped mistakenly, then NME equals to $k$, showing that NME is well-defined. Thus the goal of de-anonymization is to minimize NME.

Moreover, since adversaries is uncertain about the true mapping between the given $G_1$ and $G_2$, $\mathbf{\Pi_0}$ can be viewed as a random variable whose probability distribution is conditioned on $G_1$ and $G_2$ in adversaries' perspectives. Naturally adversaries prefer an estimation of $\mathbf{\Pi_0}$ keeping from severe NME on average. To this end, we consider selecting $\hat{\mathbf{\Pi}}$ in the light of "Minimum Mean Square Error (MMSE)" criterion, which, formally presented in Definition 6, is the minimizer of the expected NME in the form of mean square.

**Definition 6.** ***(The MMSE Estimator)*** *Given* $G_1$*,* $G_2$ *and* $\boldsymbol{\theta}$*, the MMSE estimator is an estimation of* $\mathbf{\Pi_0}$ *minimizing the number of mistakenly matched nodes in expectation, which is*

$$\begin{aligned}\hat{\mathbf{\Pi}} &= \arg\min_{\mathbf{\Pi}\in\mathbf{\Pi^n}} \mathbf{E}_{\mathbf{\Pi_0}}\{d(\mathbf{\Pi}, \mathbf{\Pi_0})\} \\ &= \arg\min_{\mathbf{\Pi}\in\mathbf{\Pi^n}} \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi} - \mathbf{\Pi_0}||_F^2 Pr(\mathbf{\Pi_0}|G_1, G_2, \boldsymbol{\theta}),\end{aligned} \tag{1}$$

*where* $\mathbf{\Pi^n}$ *is the set of* $n \times n$ *permutation matrices.*

**Remark:** Recall that prior effort [8] has leveraged Maximum A Posterior (MAP), which provides the solution with the highest probability being exactly identical to the true permutation. MMSE and MAP characterize different aspects of minimizing NME. As far as we know, no previous work has learned de-anonymization under MMSE, which, however, is also of great significance as MAP in reducing NME.

The main notations in our work are summarized in Table 1.

## IV. Analytical Aspect of De-anonymization Problem

In this section, we start to provide analysis of the social network de-anonymization problem that we have defined earlier. In doing so, we firstly prove that this problem is NP-hard. To facilitate the problem analysis, we then give an approximation to the original MMSE estimator and verify it under the expectation of different possible network structures. Furthermore, we validate this approximation by proving that the approximation ratio is not small for a single possible network structure.

TABLE I: **Notions and Definitions**

| Notation | Definition |
|---|---|
| $G$ | Underlying social network |
| $G_1, G_2$ | Published and auxiliary networks |
| $V, V_1, V_2$ | Vertex sets of graphs $G$, $G_1$ and $G_2$ |
| $E, E_1, E_2$ | Edge sets of graphs $G$, $G_1$, $G_2$ |
| $s_1, s_2$ | Edge sampling probabilities of graphs $G_1$, $G_2$ |
| $n$ | Total number of nodes |
| $w_{ij}$ | The weight of node pair $(i, j)$ |
| $\boldsymbol{C_i}$ | Community representation of node $i$ |
| $p_{\boldsymbol{C_i}\boldsymbol{C_j}}$ | Probability of edge existence between node $i$ and $j$ in $G$ |
| $\boldsymbol{\theta}$ | Parameter set |
| $\mathbf{W}$ | The weight matrix |
| $\mathbf{U}, \mathbf{A}, \mathbf{B}$ | Adjacency matrices of $G$, $G_1$, $G_2$ |
| $\mathbf{\Pi_0}(\pi_0)$ | True permutation matrix (True mapping) between $V_1$ and $V_2$ |
| $\mathbf{\Pi}(\pi)$ | A permutation matrix (A mapping) between $V_1$ and $V_2$ |
| $\hat{\mathbf{\Pi}}(\hat{\pi})$ | The MMSE estimator (the corresponding mapping) |
| $\tilde{\mathbf{\Pi}}(\tilde{\pi})$ | The minimizer of WEMP (the corresponding mapping) |
| $\mathbf{\Pi^n}$ | The set of $n \times n$ permutation matrices. |

### A. Transformation of MMSE Estimator

As can be seen from the definition of MMSE (Eqn. (1) in Section III-C), the posterior probability $Pr(\mathbf{\Pi_0}|G_1, G_2, \boldsymbol{\theta})$ still needs to be expressed more explicitly. Inspired by the derivation in [8], we have the following theorem about the transformation of MMSE estimator.

**Theorem 1.** *Given the published graph* $G_1$*, the auxiliary graph* $G_2$ *and the parameter set* $\boldsymbol{\theta}$*, the MMSE estimator can be equivalently transformed into*

$$\hat{\mathbf{\Pi}} = \arg\max_{\mathbf{\Pi}\in\mathbf{\Pi^n}} \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi} - \mathbf{\Pi_0}||_F^2 ||\mathbf{W} \circ (\mathbf{\Pi_0}\mathbf{A} - \mathbf{B}\mathbf{\Pi_0})||_F^2 \tag{2}$$

*where in the metric* $\mathbf{W}$*,* $W(i,j) = W(j,i) = \sqrt{w_{ij}}$*,* $w_{ij} = \log\left(\frac{1-p_{\boldsymbol{C_i}\boldsymbol{C_j}}(s_1+s_2-s_1s_2)}{p_{\boldsymbol{C_i}\boldsymbol{C_j}}(1-s_1)(1-s_2)}\right)$ *is weight between nodes* $i$ *and* $j$*, and "*$\circ$*" denotes the Hadamard product.*

We shall provide here a sketch of the proof of Theorems 1. The complete proof, including all mathematical details, can be found in Appendix A.

*Proof.* First, we define $\mathcal{G}_{\mathbf{\Pi}}$ as the set of all realizations of the underlying network which is in consistency with the given $G_1$, $G_2$ and $\mathbf{\Pi_0}$. Then the MMSE estimator can be written as

$$\hat{\mathbf{\Pi}} = \arg\min_{\mathbf{\Pi}\in\mathbf{\Pi^n}} \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi} - \mathbf{\Pi_0}||_F^2 \sum_{G\in\mathcal{G}_{\mathbf{\Pi}}} Pr(G, \mathbf{\Pi_0}|G_1, G_2, \boldsymbol{\theta}). \tag{3}$$

Focusing on $Pr(G, \mathbf{\Pi_0}|G_1, G_2, \boldsymbol{\theta})$ in Eqn. 3, we derive that

$$Pr(G, \mathbf{\Pi_0}|G_1, G_2, \boldsymbol{\theta}) \sim Pr(G)Pr(G_1|G)Pr(G_2|G, \mathbf{\Pi_0}) \tag{4}$$

by Bayes formula as well as the independency of the sampling process of $G_1$ and $G_2$. What's more, $a \sim b$ means that $a$ and $b$ are different only in parameters unrelated to $\mathbf{\Pi_0}$.

Then, to further analyze $Pr(G)Pr(G_1|G)Pr(G_2|G, \mathbf{\Pi_0})$, we define $G^*_{\pi_0}$ as the graph which has the smallest number of edges in $\mathcal{G}_{\mathbf{\Pi}}$, i.e., $G^*_{\pi_0} = (V, E_1 \cup \pi_0(E_1))$, where $\pi_0(E_1) = \{(\pi_0(i), \pi_0(j))|(i,j) \in E_1\}$ and set $E^*_{\pi_0}$ as the edge set of $G^*_{\pi_0}$, and $E^{*ij}_{\pi_0}$ as the indicator variable between nodes $i$ and $j$. By these definitions, we explicitly express that

$$\begin{aligned}&\sum_{G\in\mathcal{G}_{\mathbf{\Pi}}} Pr(G)Pr(G_1|G)Pr(G_2|G, \mathbf{\Pi_0}) \\ &\qquad \sim \sum_{i<j}^{n} E^{*ij}_{\pi_0} \log\left(\frac{p_{\boldsymbol{C_i}\boldsymbol{C_j}}(1-s_1)(1-s_2)}{1-p_{\boldsymbol{C_i}\boldsymbol{C_j}}(s_1+s_2-s_1s_2)}\right)\end{aligned} \tag{5}$$

Ultimately, by analyzing the parameter $E^{*ij}_{\pi_0}$, we derive that

$$\begin{aligned}\hat{\mathbf{\Pi}} &= \arg\min_{\mathbf{\Pi}\in\mathbf{\Pi^n}} \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi}-\mathbf{\Pi_0}||_F^2 \sum_{G\in\mathcal{G}_{\mathbf{\Pi}}} Pr(G,\mathbf{\Pi_0}|G_1,G_2,\boldsymbol{\theta}) \\ &\sim \arg\min_{\mathbf{\Pi}\in\mathbf{\Pi^n}} \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi}-\mathbf{\Pi_0}||_F^2 \\ &\qquad \sum_{i<j}^{n} E^{*ij}_{\pi_0} \log\left(\frac{p_{\boldsymbol{C_i}\boldsymbol{C_j}}(1-s_1)(1-s_2)}{1-p_{\boldsymbol{C_i}\boldsymbol{C_j}}(s_1+s_2-s_1s_2)}\right) \\ &= \arg\max_{\mathbf{\Pi}\in\mathbf{\Pi^n}} \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi}-\mathbf{\Pi_0}||_F^2 ||\mathbf{W}\circ(\mathbf{\Pi_0}\mathbf{A}-\mathbf{B}\mathbf{\Pi_0})||_F^2. \end{aligned} \tag{6}$$

$\square$

**Remark:** Additionally, to simplify the form of $||\mathbf{W}\circ(\mathbf{\Pi_0}\mathbf{A}-\mathbf{B}\mathbf{\Pi_0})||_F^2$, we use $\mathbf{\Pi_0}\hat{\mathbf{A}}$ to represent $\mathbf{W}\circ\mathbf{\Pi_0}\mathbf{A}$, and $\hat{\mathbf{B}}\mathbf{\Pi_0}$ to represent $\mathbf{W}\circ\mathbf{B}\mathbf{\Pi_0}$[1]. Therefore we can rewrite the MMSE estimator in Eqn. (37) as

$$\hat{\mathbf{\Pi}} = \arg\max_{\mathbf{\Pi}\in\mathbf{\Pi^n}} \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi}-\mathbf{\Pi_0}||_F^2 ||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2, \tag{7}$$

and $g(\mathbf{\Pi}) = \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi}-\mathbf{\Pi_0}||_F^2 ||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$. In the following analysis, we use the form in Eqn. (7). In Section V-A, we will discuss the condition under which $\mathbf{W}\circ\mathbf{A}=\hat{\mathbf{A}}$ and $\mathbf{W}\circ\mathbf{B}=\hat{\mathbf{B}}$.

### *B. NP-hardness of Solving the MMSE Estimator*

Since we have derived a more explicit form of MMSE estimator, we are interested in whether there exists a polynomial-time algorithm that can solve the MMSE problem. However, as we will prove in the sequel, this problem is NP-hard, meaning that no polynomial time (pseudo-polynomial time) approximation algorithm exists for solving the MMSE estimator.

**Proposition 1.** *Solving the MMSE estimator is an NP-hard problem. There is no polynomial time or pseudo-polynomial time approximation algorithm for this problem with any multiplicative approximation guarantee unless P=NP.*

*Proof.* Hereinafter, we provide the outline of our proof divided into two steps (the full version of the proof is available in Appendix B):

**i)** reducing the 1-median problem to our MMSE problem [2]

**ii)** finding the lower bound for 1-median problem [3].

For step i,**Reduction from 1-median problem:** To equivalently transform our problem into the form of 1-median problem, we construct a clique with $n!$ nodes, each node $i$ representing a possible $\mathbf{\Pi_0}(i)\in\mathbf{\Pi^n}$. We modify Eqn. (7) equivalently into

$$\hat{\mathbf{\Pi}} = \arg\min_{\mathbf{\Pi}\in\mathbf{\Pi}^n} \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} (4n - ||\mathbf{\Pi}-\mathbf{\Pi_0}||_F^2) ||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2, \tag{8}$$

where we set $D(i,j) = 4n - ||\mathbf{\Pi_0}(i)-\mathbf{\Pi_0}(j)||_F^2$ and $\omega(i) = ||\mathbf{\Pi_0}(i)\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}(i)||_F^2$ (Note that $\mathbf{\Pi_0}(i)$ is a node in the clique). Both $D(i,j)$ and $\omega(i)$ meet the requirements in 1-median problem. Thus, we complete the reduction.

For step ii),**The Lower Bound for 1-Median Problem:** For a 1-median problem with $n!$ nodes, obviously we need to calculate at least $\lceil \frac{n!}{2} \rceil$ edges to form an edge set such that the endpoints of all edges inside cover all nodes in the graph, or else at least one node will not be calculated for any edge connecting it, which we can not judge if it is the node we intend to find. Since the size of our input, a matrix, is $n^2$, the complexity turns out to be $\Omega((\sqrt{n}/2)!) = \Omega(\sqrt{n}!)$, exceeding polynomial. $\square$

The NP-hardness of MMSE estimator shows the impossibility to pursue an exact algorithm or any approximation algorithm with multiplicative guarantee. Thus we need to simplify this problem by conducting reasonable approximation to make it possible to solve this problem, with certain tolerance of mapping error. In the following we propose one way to approximate this problem, the analysis of which will indicate that the error arose by this approximation can be bounded.

### *C. Approximation of the MMSE estimator*

As we have just stated above, the NP-hardness of MMSE problem urges us to find proper approximation for the original problem. Recall that MMSE involves all the possible true mappings, the number of which is $n!$, thus leading to fairly prohibitive computational cost. To tackle the difficulty, we firstly transform the original MMSE problem into a weighted-edge matching problem (WEMP), which, as we will define and present more details later, simplifies the form of objective function of the original MMSE problem and makes it tractable. Then we demonstrate that this transformation is valid, meaning that the solution of WEMP will not deviate much from the solution of the original MMSE problem by proving its high approximation ratio. Definition 7 provides the formal statement of WEMP.

**Definition 7.** ***(Weighted-Edge Matching Problem)*** *Given the adjacent matrices of $G_1$ and $G_2$, denoted as $\mathbf{A}$ and $\mathbf{B}$ respectively, set $\mathbf{W}\circ\mathbf{A}=\hat{\mathbf{A}}$ and $\mathbf{W}\circ\mathbf{B}=\hat{\mathbf{B}}$ where $\mathbf{W}$ is the weight matrix, the weight-edge matching problem is to find*

$$\tilde{\mathbf{\Pi}} = \arg\min_{\mathbf{\Pi}\in\mathbf{\Pi^n}} ||\mathbf{\Pi}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi}||_F^2.$$

Hereinafter we discuss the following two aspects of WEMP:

- How do we transform from the original MMSE problem into WEMP?
- How is the validity of this transformation?

*1) **The Idea of Transformation**:* We intend to transform the original problem of solving the MMSE estimator into WEMP.

[1] We should clarify that we only provide a simpler form to represent $\mathbf{W}\circ\mathbf{\Pi_0}\mathbf{A}$ and $\mathbf{W}\circ\mathbf{B}\mathbf{\Pi_0}$, and it does NOT imply that $\mathbf{W}\circ\mathbf{A}=\hat{\mathbf{A}}$ and $\mathbf{W}\circ\mathbf{B}=\hat{\mathbf{B}}$. But some operations under this new notation still hold, for example, multiplying a permutation matrix does not change the value of the Frobenius norm, i.e., $||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 = ||\mathbf{W}\circ(\mathbf{\Pi_0}\mathbf{A}-\mathbf{B}\mathbf{\Pi_0})||_F^2 = ||\mathbf{W}\circ\mathbf{\Pi_0^T}(\mathbf{\Pi_0}\mathbf{A}-\mathbf{B}\mathbf{\Pi_0}))||_F^2 = ||\mathbf{W}\circ(\mathbf{A}-\mathbf{\Pi_0^T}\mathbf{B}\mathbf{\Pi_0})||_F^2$ and $||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 = ||\mathbf{\Pi_0}\hat{\mathbf{A}}\mathbf{\Pi_0^T}-\hat{\mathbf{B}}||_F^2$.

[2] Here the 1-median problem [21] refers to that: Given a connected undirected graph $G=(V,E)$ in which no isolated vertices exist and each node $v$ is endowed with a nonnegative weight $\omega(v)$, find the vertex $v^*$ which minimizes weighted sum: $H(v^*) = \sum_{v\in V}\omega(v)\cdot D(v,v^*)$, where $D(v,v^*)$ means the shortest path length (also nonnegative) between nodes $v$ and $v^*$.

[3] Note that 1-median itself is not NP-hard if the problem size is O(n), but we demonstrate that when applied in our problem it becomes larger than polynomial.

The idea of this transformation can be interpreted in the following sense: for any fixed $\mathbf{\Pi}$, define a set $S_k(\mathbf{\Pi}), 0 \le k \le n$, any element of which is an $n \times n$ permutation matrix $\mathbf{\Pi_0}$ such that $d(\mathbf{\Pi}, \mathbf{\Pi_0}) = k$. When $k = 0$, $d(\mathbf{\Pi}, \mathbf{\Pi_0}) = 0$. Thus $\mathbf{\Pi} = \mathbf{\Pi_0}$ and $S_0(\mathbf{\Pi}) = \{\mathbf{\Pi}\}$. What's more, if $\mathbf{\Pi} \neq \mathbf{\Pi_0}$, at least one node pair will be mismatched, thus $k \ge 2$ and $S_1(\mathbf{\Pi}) = \emptyset$. Then we can transform the original problem as

$$\hat{\mathbf{\Pi}} = \arg\max_{\mathbf{\Pi}\in\mathbf{\Pi^n}} \sum_{k=0}^{n} k \left( \sum_{\mathbf{\Pi_0}\in S_k(\mathbf{\Pi})} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 \right). \quad (9)$$

Based on Eqn. (9), we propose our idea of transforming it into WEMP. To present our idea clearly, we divide our analysis into three parts; First we analyze a single term, $||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$, where $\mathbf{\Pi_0} \in S_2(\mathbf{\Pi})$; Then we analyze $\mathbf{\Pi_0} \in S_k(\mathbf{\Pi})$ based on the analysis of $\mathbf{\Pi_0} \in S_2(\mathbf{\Pi})$; Finally we analyze the R.H.S of Eqn. (9) based on Sequence Inequality. The detailed analysis of these three parts are unfolded in Appendix C and in the sequel we provide the sketch of three parts:

i. Analysis of $||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$ where $\mathbf{\Pi_0} \in S_2(\mathbf{\Pi})$

Note that any permutation in $S_2(\mathbf{\Pi})$ only causes matching error on one pair of nodes. Thus we set one specific $\mathbf{\Pi_0} \in S_2(\tilde{\mathbf{\Pi}})$, which differs from $\tilde{\mathbf{\Pi}}$ only in the $i_{th}$ and $j_{th}$ row. By using the condition that $\mathbf{A}$ and $\mathbf{B}$ are conditionally independent[4], we can derive that

$$\mathbf{E_{A,B}}(||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 - ||\tilde{\mathbf{\Pi}}\hat{\mathbf{A}} - \hat{\mathbf{B}}\tilde{\mathbf{\Pi}}||_F^2) = 2 \sum_{k\neq i,j}^{n} \Delta_{i,j,k,\pi_0},$$

where

$$\begin{aligned}\Delta_{i,j,k,\pi_0} = &(w_{ik}(1 - 2p_{\boldsymbol{C_i C_k}} s_2) - w_{jk}(1 - 2p_{\boldsymbol{C_j C_k}} s_2)) \\ &\cdot (p_{\boldsymbol{C_{\pi_0(i)} C_{\pi_0(k)}}} - p_{\boldsymbol{C_{\pi_0(j)} C_{\pi_0(k)}}}) s_1.\end{aligned}$$

and it reflects a part of the difference $||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 - ||\tilde{\mathbf{\Pi}}\hat{\mathbf{A}} - \hat{\mathbf{B}}\tilde{\mathbf{\Pi}}||_F^2$ caused by the difference of a single element in matrices $\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}$ and $\tilde{\mathbf{\Pi}}\hat{\mathbf{A}} - \hat{\mathbf{B}}\tilde{\mathbf{\Pi}}$. We set $\hat{\Delta} = \mathbf{E}_{i,j,\pi_0}(\Delta_{i,j,\pi_0})$ and note that $\mathbf{E_{A,B}}(||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 - ||\tilde{\mathbf{\Pi}}\hat{\mathbf{A}} - \hat{\mathbf{B}}\tilde{\mathbf{\Pi}}||_F^2) > 0$ since $\tilde{\mathbf{\Pi}}$ is the minimizer of $||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$. Therefore $\hat{\Delta} = \mathbf{E}_{i,j,\pi_0}(\Delta_{i,j,\pi_0}) > 0$.

ii. Analysis of $\sum_{\mathbf{\Pi_0}\in S_k(\mathbf{\Pi})} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$

We first focus on $S_k(\mathbf{\Pi_0})$, and count the number of elements in $S_k(\mathbf{\Pi_0})$, denoted as $|S_k|$. We can obtain the relationship between $|S_k|$ and $|S_{k-1}|$ when $k \ge 2$ as

$$|S_k| = C_n^k |T_k| \ge (k-1)\frac{C_n^k}{C_n^{k-1}} |S_{k-1}| = (1-\frac{1}{k})(n-k+1)|S_{k-1}|, \quad (10)$$

Then we consider $\mathbf{\Pi_0} \in S_k(\mathbf{\Pi})$. Note that for any $\mathbf{\Pi_0} \in S_k(\mathbf{\Pi})$, there are $k$ rows and columns that may cause the difference between $||\mathbf{W}\circ(\mathbf{\Pi_0}\mathbf{A}\mathbf{\Pi_0^T} - \mathbf{B})||_F^2$ and $||\mathbf{W}\circ(\tilde{\mathbf{\Pi}}\mathbf{A}\tilde{\mathbf{\Pi}}^\mathbf{T} - \mathbf{B})||_F^2$. We can discover that for any $\mathbf{\Pi_0} \in S_k(\mathbf{\Pi})$, the number of node pairs $(i, j)$ which may influence the difference between $||\mathbf{W}\circ(\mathbf{\Pi_0}\mathbf{A}\mathbf{\Pi_0^T} - \mathbf{B})||_F^2$ and $||\mathbf{W}\circ(\tilde{\mathbf{\Pi}}\mathbf{A}\tilde{\mathbf{\Pi}}^\mathbf{T} - \mathbf{B})||_F^2$ is approximately $\sum_{i=1}(n - i) = \frac{(2n-k-1)k}{2}$. Thus, denoting $N_k$ as the number of such node pairs for all $\mathbf{\Pi_0} \in S_k(\mathbf{\Pi})$, we can obtain

$$N_k = \frac{(2n-k-1)k}{2}|S_k| \ge (n-k+1)\frac{2n-k-1}{2n-k} N_{k-1}.$$

Therefore on average, we have

$$\begin{aligned}\sum_{\mathbf{\Pi_0}\in S_k} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 &= N_k\hat{\Delta} \\ &\ge (n-k+1)\sum_{\mathbf{\Pi_0}\in S_{k-1}} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2,\end{aligned} \quad (11)$$

where $k \le n$ and $\hat{\Delta}$ is denoted in part i.

Thus, we can claim that in average case, if $k_1 > k_2$, then

$$\sum_{\mathbf{\Pi_0}\in S_{k_1}} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 > \sum_{\mathbf{\Pi_0}\in S_{k_2}} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2. \quad (12)$$

iii. Maximum Value Under Sequence Inequality

Based on the analysis above, we set $\mathbf{\Pi} = \tilde{\mathbf{\Pi}}$, if $\mathbf{\Pi_0} \in S_0(\tilde{\mathbf{\Pi}})$, $||\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2$ is the minimum value. Moreover, $\sum_{\mathbf{\Pi_0}\in S_0(\tilde{\mathbf{\Pi}})} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 = ||\tilde{\mathbf{\Pi}}\hat{\mathbf{A}} - \hat{\mathbf{B}}\tilde{\mathbf{\Pi}}||_F^2$ is also the minimum value in the set

$$\left\{ \begin{aligned} &\sum_{\mathbf{\Pi_0}\in S_0(\tilde{\mathbf{\Pi}})} ||(\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0})||_F^2, \sum_{\mathbf{\Pi_0}\in S_2(\tilde{\mathbf{\Pi}})} ||(\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0})||_F^2, \\ &\sum_{\mathbf{\Pi_0}\in S_3(\tilde{\mathbf{\Pi}})} ||(\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0})||_F^2, ..., \sum_{\mathbf{\Pi_0}\in S_n(\tilde{\mathbf{\Pi}})} ||(\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0})||_F^2 \end{aligned} \right\}.$$

Thus according to sequence inequality, we know that in average case, by setting $\mathbf{\Pi}$ in the original MMSE objective function

$$\sum_{\mathbf{\Pi_0}\in\Pi^n} ||\mathbf{\Pi} - \mathbf{\Pi_0}||_F^2 ||\mathbf{W}\circ(\mathbf{\Pi_0}\mathbf{A} - \mathbf{B}\mathbf{\Pi_0})||_F^2$$

equal to $\tilde{\mathbf{\Pi}}$, the minimizer of WEMP, then this original MMSE objective function reaches its largest value [5] under Sequence Inequality.

However, as we noted earlier, we can only transform the original MMSE problem into WEMP in an average case of network structures. This implies that the transformation is not necessarily the best approximation of a single network structure. In the following we further analyze the validity of this transformation in a possible network structure by showing the approximation ratio of our transformation is large (at least larger than 0.5).

*2) The Validity of Transformation:* As we have stated above, $\tilde{\mathbf{\Pi}}$ does not necessarily achieve the maximum of the original MMSE problem for a specific network structure. That is to say there may exist error in $g(\tilde{\mathbf{\Pi}})$ and $g(\hat{\mathbf{\Pi}})$, where $g(\hat{\mathbf{\Pi}})$ is the maximum value of the original MMSE objective function and $g(\tilde{\mathbf{\Pi}})$ is the value of MMSE objective function when $\mathbf{\Pi}$ equals to the minimizer of WEMP. Theorem 2 shows that under the mild condition indicated by Inequality (11), we can get approximation ratio $g(\tilde{\mathbf{\Pi}})/g(\hat{\mathbf{\Pi}})$ larger than 0.5, which, to some extent, makes our estimation reasonable.

[4]Since $G_1$ and $G_2$ are independently sampled from $G$.

[5]We set $k_1 > k_2$, then $\sum_{\mathbf{\Pi_0}\in S_{k_1}} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 > \sum_{\mathbf{\Pi_0}\in S_{k_2}} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$, and $||\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2$ where $\mathbf{\Pi_0} \in S_{k_1}(\tilde{\mathbf{\Pi}})$ is larger than $||\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2$ where $\mathbf{\Pi_0} \in S_{k_2}(\tilde{\mathbf{\Pi}})$. Thus, $\sum_{\mathbf{\Pi_0}\in\Pi^n} ||\mathbf{\Pi} - \mathbf{\Pi_0}||_F^2 ||\mathbf{W}\circ(\mathbf{\Pi_0}\mathbf{A} - \mathbf{B}\mathbf{\Pi_0})||_F^2 = \sum_{k=0}^{n} 2k \sum_{\mathbf{\Pi_0}\in S_k(\tilde{\mathbf{\Pi}})} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$ reaches its largest value.

**Theorem 2.** *Given the published graph $G_1$, the auxiliary graph $G_2$, the parameter set $\boldsymbol{\theta}$ and the weight matrix $\mathbf{W}$, in average case we have the approximation ratio $g(\tilde{\mathbf{\Pi}})/g(\hat{\mathbf{\Pi}})$ larger than* $0.5$.

*Proof.* We present the sketch of the proof here and defer the complete version to the Appendix D.

First we derive the upperbound of $\frac{g(\hat{\mathbf{\Pi}})-g(\tilde{\mathbf{\Pi}})}{g(\tilde{\mathbf{\Pi}})}$:

$$\frac{g(\hat{\mathbf{\Pi}})-g(\tilde{\mathbf{\Pi}})}{g(\tilde{\mathbf{\Pi}})} \leq \frac{2\beta n \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi_0}\tilde{\mathbf{A}}-\tilde{\mathbf{B}}\mathbf{\Pi_0}||_F^2}{\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0}||_F^2||\mathbf{\Pi_0}\tilde{\mathbf{A}}-\tilde{\mathbf{B}}\mathbf{\Pi_0}||_F^2}. \tag{13}$$

where $||\tilde{\mathbf{\Pi}}-\hat{\mathbf{\Pi}}||_F^2 = 2\beta n$ and $\beta \in [0,1]$ is the ratio between the number of mistakenly matched nodes and that of all the nodes.

Then we divide $\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$ into two parts:

$$D_1 = \sum_{k\leq \rho n}\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2;$$

$$D_2 = \sum_{\rho n< k\leq n}\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2.$$

where $\rho$ is any real number in $[0,1]$ and we assume that $\rho n$ is an integer, since if it is not an integer, we can easily modify it by rounding.

For $D_1$ and $D_2$, in average case we can obtain $D_1 \leq (2n)^{\rho n}$ and $D_2 \geq (1-\rho)n\frac{n!}{((1-\rho)n)!}$.

Note that if we set $\rho = \Omega(1) = c_0$, where $c_0 \to 1$, then $(1-\rho)n$ can be upper bounded by a constant $c_1$ and

$$D_2 \geq c_1\frac{n!}{c_1!} = cn! \sim c\sqrt{2\pi n}(\frac{n}{e})^n,$$

where $c$ is a constant and the last step holds due to the Stirling's formula. Therefore we can upper bound $\frac{D_2}{D_1}$ as

$$\frac{D_2}{D_1} \geq c\frac{\sqrt{2\pi n}(\frac{n}{e})^n}{(2n)^{\rho n}} = c\sqrt{2\pi n}\left(\frac{n^{1-\rho}}{2^\rho e}\right)^n.$$

Then if $\rho$ is a constant and $\rho \to 1$, we can find that when $n \to \infty$, $D_2$ is of higher order of $n$ than $D_1$. Therefore we can easily verify that $\sum_{\rho n<k\leq n}\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0}||_F^2||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$ is of higher order of $n$ than $\sum_{k\leq\rho n}\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0}||_F^2||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$, since for $k_1 > \rho n$ and $k_2 < \rho n$, $\mathbf{\Pi_1'} \in S_{k_1}(\tilde{\mathbf{\Pi}})$ and $\mathbf{\Pi_2'} \in S_{k_2}(\tilde{\mathbf{\Pi}})$, we have $||\mathbf{\Pi_1'}-\tilde{\mathbf{\Pi}}||_F^2 \geq ||\mathbf{\Pi_2'}-\tilde{\mathbf{\Pi}}||_F^2$. Thus we can obtain

$$\begin{aligned}&\frac{2\beta n \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2}{\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0}||_F^2||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2}\\ &\approx \frac{2\beta n\sum_{\rho n<k\leq n}\sum_{\mathbf{\Pi_0}\in S_k(\tilde{\mathbf{\Pi}})} ||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2}{\sum_{\rho n<k\leq n}\sum_{\mathbf{\Pi_0}\in S_k(\tilde{\mathbf{\Pi}})} ||\mathbf{\Pi_0}-\tilde{\mathbf{\Pi}}||_F^2||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2}\\ &\leq \frac{2\beta n}{2\rho n}\frac{\sum_{\rho n<k\leq n}\sum_{\mathbf{\Pi_0}\in S_k(\tilde{\mathbf{\Pi}})} ||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2}{\sum_{\rho n<k\leq n}\sum_{\mathbf{\Pi_0}\in S_k(\tilde{\mathbf{\Pi}})} ||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2} = \frac{\beta}{\rho}.\end{aligned}$$

Then we have the approximation ratio

$$\frac{g(\tilde{\mathbf{\Pi}})}{g(\hat{\mathbf{\Pi}})} \geq \frac{1}{1+\frac{\beta}{\rho}} \approx \frac{1}{1+\beta} \geq \frac{1}{2}.$$

□

## V. Algorithmic Aspect of De-anonymization Problem

In this section, we show that WEMP is of significant advantages in seedless de-anonymization since it resolves the tension between *optimality* and *complexity*. For optimality, We prove the good performance of solving WEMP that the result makes the node mapping error (NME) negligible in large social networks under mild conditions, facilitated by higher overlapping strength; For complexity, the optimal mapping of WEMP, $\tilde{\mathbf{\Pi}}$, can be perfectly sought algorithmically by our convex-concave based de-anonymization algorithm (CBDA).

### A. Optimality: WEMP Returns Negligible NME

Recall that our aim is to minimize NME in expectation, thus a natural question arises: *how much NME $\tilde{\mathbf{\Pi}}$ may cause for any probable real permutation matrix $\mathbf{\Pi_0}$?* The answer reflects the ability of solving WEMP in enhancing mapping accuracy. To answer it, we demonstrate that under mild conditions, the *relative NME*, defined as $\frac{||\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0}||_F^2}{||\mathbf{\Pi_0}||_F^2}$, vanishes to 0 as $n \to \infty$. This implies that under large network size, NME caused by $\tilde{\mathbf{\Pi}}$ is negligible compared with $|V| = n$. Furthermore, we surprisingly find that the conditions are facilitated under higher overlapping strength, explicitly delineating benefits brought by overlapping communities in NME reduction. Theorem 3 formally presents our result mentioned above. Before that, we give Lemma 1, a prerequisite in proving Theorem 3, with its proof in Appendix E.

**Lemma 1.** *Suppose the permutation matrix $\mathbf{\Pi}$ keeps invariant of the community representation of all the nodes, i.e., $\forall\mathbf{\Pi} \in \mathbf{\Pi^n}$ such that $\mathbf{\Pi}(i,j) = 1$, $\boldsymbol{C_i} = \boldsymbol{C_j}$, then $\hat{\mathbf{A}} = \mathbf{W}\circ\mathbf{A}$, $\hat{\mathbf{B}} = \mathbf{W}\circ\mathbf{B}$ and $||\mathbf{W}\circ(\mathbf{\Pi}\mathbf{A}\mathbf{\Pi^T}-\mathbf{B})||_F = ||\mathbf{\Pi}\hat{\mathbf{A}}\mathbf{\Pi^T}-\hat{\mathbf{B}}||_F$.*

**Remark:** Note that there are no differences in form between $||\mathbf{\Pi_1}\hat{\mathbf{A}}\mathbf{\Pi_1^T}-\hat{\mathbf{B}}||_F$ and $||\hat{\mathbf{A}}-\mathbf{\Pi_2}\hat{\mathbf{B}}\mathbf{\Pi_2^T}||_F$ since we can simply set $\mathbf{\Pi_2} = \mathbf{\Pi_1^T}$. Therefore, we do not distinguish the forms $||\mathbf{\Pi}\hat{\mathbf{A}}\mathbf{\Pi^T}-\hat{\mathbf{B}}||_F$ and $||\hat{\mathbf{A}}-\mathbf{\Pi}\hat{\mathbf{B}}\mathbf{\Pi^T}||_F$ anymore.

**Theorem 3.** *Given $G_1(V_1,E_1,\mathbf{A})$, $G_2(V_2,E_2,\mathbf{B})$, $\boldsymbol{\theta}$ and $\mathbf{W}$. Let $K = \min_{s,t,j}\{(p_{\boldsymbol{C_s}\boldsymbol{C_j}}+p_{\boldsymbol{C_t}\boldsymbol{C_j}})\min\{s_1,s_2\}\}, L = \max_{s,t,j}\{[(p_{\boldsymbol{C_s}\boldsymbol{C_j}}+p_{\boldsymbol{C_t}\boldsymbol{C_j}})\max\{s_1,s_2\}]^2\}$. If (i) $\frac{L}{K} = o(1)$; (ii) $\frac{||\hat{\mathbf{A}}-\mathbf{\Pi_0}\hat{\mathbf{B}}\mathbf{\Pi_0^T}||_F^2}{||\hat{\mathbf{A}}-\tilde{\mathbf{\Pi}}\hat{\mathbf{B}}\tilde{\mathbf{\Pi}}^\mathbf{T}||_F^2} = \Omega(1)$; (iii) $||\hat{\mathbf{A}}-\mathbf{\Pi_0}\hat{\mathbf{B}}\mathbf{\Pi_0^T}||_F^2 = o(Kn^2)$; (iv) $\mathbf{\Pi_0}$ and $\tilde{\mathbf{\Pi}}$ keep invariant of community representations, then as $n \to \infty$, $\frac{||\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0}||_F^2}{||\mathbf{\Pi_0}||_F^2} \to 0$.*

*Proof.* The proof has four steps:

**i)** Upper bounding $||\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0}||_F$ by $||(\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0})\hat{\mathbf{B}}||_F$;

**ii)** Finding the relationship between $||(\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0})\hat{\mathbf{B}}||_F$ and $\mathbf{tr}((\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0})\hat{\mathbf{B}}(\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0})^\mathbf{T}\hat{\mathbf{A}})$;

**iii)** Upper bounding $\mathbf{tr}((\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0})\hat{\mathbf{B}}(\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0})^\mathbf{T}\hat{\mathbf{A}})$;

**iv)** Upper bounding $\frac{||\mathbf{\Pi_0}-\tilde{\mathbf{\Pi}}||_F^2}{||\mathbf{\Pi_0}||_F^2}$.

i. Upper bounding $||\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0}||_F$ by $||(\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0})\hat{\mathbf{B}}||_F$:

For the $i_{th}$ row of $(\mathbf{\Pi_0}-\tilde{\mathbf{\Pi}})$, we set $\pi_0(i) = s$ and $\tilde{\pi}(i) = t$. If $s = t$, then the $i_{th}$ row of $(\mathbf{\Pi_0}-\tilde{\mathbf{\Pi}})\hat{\mathbf{B}}$ is a zero vector; else the $i_{th}$ row of $(\mathbf{\Pi_0}-\tilde{\mathbf{\Pi}})\hat{\mathbf{B}}$ is $(\mathbf{B}_{s1}-\mathbf{B}_{t1},\mathbf{B}_{s2}-\mathbf{B}_{t2},\cdots,\mathbf{B}_{sn}-$

$\mathbf{B}_{tn}$). For an element, $([(\mathbf{\Pi_0} - \tilde{\mathbf{\Pi}})\hat{\mathbf{B}}]_{ij})^2 = (\mathbf{B}_{sj} - \mathbf{B}_{tj})^2$. Taking the expectation on both sides, we can derive that

$$\mathbf{E_B}[(\mathbf{\Pi_0} - \tilde{\mathbf{\Pi}})\hat{\mathbf{B}}]^2_{ij} = (p_{\boldsymbol{C_s C_j}} + p_{\boldsymbol{C_t C_j}} - 2p_{\boldsymbol{C_s C_j}} p_{\boldsymbol{C_t C_j}} s_2)s_2,$$

where $\mathbf{E_B}$ means taking expectation on every element in $\mathbf{B}$. By summing up all the rows and columns,

$$\begin{aligned}
&||(\mathbf{\Pi_0} - \tilde{\mathbf{\Pi}})\hat{\mathbf{B}}||^2_F = \mathbf{E}\sum_{i=1}^{n}\sum_{j=1}^{n}[(\mathbf{\Pi_0} - \tilde{\mathbf{\Pi}})\hat{\mathbf{B}}]^2_{ij} \\
&= \sum_{i=1}^{n} \mathbb{1}\{\pi_0(i) \neq \tilde{\pi}(i)\} \sum_{j=1}^{n} (p_{\boldsymbol{C_s C_j}} + p_{\boldsymbol{C_t C_j}} - 2p_{\boldsymbol{C_s C_j}} p_{\boldsymbol{C_t C_j}} s_2)s_2 \\
&\geq \sum_{i=1}^{n} n\mathbb{1}\{\pi_0(i) \neq \tilde{\pi}(i)\} \min_j (p_{\boldsymbol{C_s C_j}} + p_{\boldsymbol{C_t C_j}} - 2p_{\boldsymbol{C_s C_j}} p_{\boldsymbol{C_t C_j}} s_2)s_2,
\end{aligned}$$

Setting $K = \min_{s,t,j}(p_{\boldsymbol{C_s C_j}} + p_{\boldsymbol{C_t C_j}} - 2p_{\boldsymbol{C_s C_j}} p_{\boldsymbol{C_t C_j}} s_2)s_2$. Note that $||(\mathbf{\Pi_0} - \tilde{\mathbf{\Pi}})||^2_F = 2\sum_{i=1}^n \mathbf{1}\{\pi_0(i) \neq \tilde{\pi}(i)\}$, we have

$$||\mathbf{\Pi_0} - \tilde{\mathbf{\Pi}}||^2_F \leq \frac{2}{nK}||(\mathbf{\Pi_0} - \tilde{\mathbf{\Pi}})\hat{\mathbf{B}}||^2_F. \tag{14}$$

Similarly we can replace $\hat{\mathbf{B}}$ by $\hat{\mathbf{A}}$ and change $s_2$ to $s_1$ in $K$.

ii. $||(\mathbf{\Pi_0} - \tilde{\mathbf{\Pi}})\hat{\mathbf{B}}||_F$ and $\mathbf{tr}((\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0})\hat{\mathbf{B}}((\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0})^\mathbf{T})\hat{\mathbf{A}})$: Note that

$$\begin{aligned}
&||(\mathbf{\Pi_0} - \tilde{\mathbf{\Pi}})\hat{\mathbf{B}}||_F = ||\hat{\mathbf{B}}(\mathbf{\Pi_0} - \tilde{\mathbf{\Pi}})^\mathbf{T}||_F = ||\tilde{\mathbf{\Pi}}\hat{\mathbf{B}}(\mathbf{\Pi_0} - \tilde{\mathbf{\Pi}})^\mathbf{T}||_F \\
&\quad \leq ||(\tilde{\mathbf{\Pi}}\hat{\mathbf{B}}\mathbf{\Pi_0} - \hat{\mathbf{A}}) - (\tilde{\mathbf{\Pi}}\hat{\mathbf{B}}\tilde{\mathbf{\Pi}} - \hat{\mathbf{A}})||_F \\
&\quad \leq ||\tilde{\mathbf{\Pi}}\hat{\mathbf{B}}\tilde{\mathbf{\Pi}}^\mathbf{T} - \hat{\mathbf{A}}||_F + ||\tilde{\mathbf{\Pi}}\hat{\mathbf{B}}\mathbf{\Pi_0^T} - \hat{\mathbf{A}}||_F,
\end{aligned}$$

where the second equation holds since the permutation matrix $\tilde{\mathbf{\Pi}}$ keeps invariant of Frobenius norm, and the second inequality holds due to the triangular inequality of Frobenius norm. Then we obtain

$$||(\mathbf{\Pi_0} - \tilde{\mathbf{\Pi}})\hat{\mathbf{B}}||^2_F \leq 2(||\tilde{\mathbf{\Pi}}\hat{\mathbf{B}}\tilde{\mathbf{\Pi}}^\mathbf{T} - \hat{\mathbf{A}}||^2_F + ||\tilde{\mathbf{\Pi}}\hat{\mathbf{B}}\mathbf{\Pi_0^T} - \hat{\mathbf{A}}||^2_F).$$

For the term $||\tilde{\mathbf{\Pi}}\hat{\mathbf{B}}\mathbf{\Pi_0^T} - \hat{\mathbf{A}}||^2_F$,

$$\begin{aligned}
&||\tilde{\mathbf{\Pi}}\hat{\mathbf{B}}\mathbf{\Pi_0^T} - \hat{\mathbf{A}}||^2_F = \mathbf{tr}((\tilde{\mathbf{\Pi}}\hat{\mathbf{B}}\mathbf{\Pi_0^T} - \hat{\mathbf{A}})^\mathbf{T}(\tilde{\mathbf{\Pi}}\hat{\mathbf{B}}\mathbf{\Pi_0^T} - \hat{\mathbf{A}})) \\
&= \mathbf{tr}(\hat{\mathbf{A}}^\mathbf{T}\hat{\mathbf{A}}) + \mathbf{tr}(\hat{\mathbf{B}}^\mathbf{T}\hat{\mathbf{B}}) - 2\mathbf{tr}(\mathbf{\Pi_0}\hat{\mathbf{B}}\tilde{\mathbf{\Pi}}^\mathbf{T}\hat{\mathbf{A}}) \\
&= ||\hat{\mathbf{A}}||^2_F + ||\hat{\mathbf{B}}||^2_F - 2\mathbf{tr}(\mathbf{\Pi_0}\hat{\mathbf{B}}\tilde{\mathbf{\Pi}}^\mathbf{T}\hat{\mathbf{A}}) \\
&= \frac{1}{2}(||\tilde{\mathbf{\Pi}}\hat{\mathbf{B}}\tilde{\mathbf{\Pi}}^\mathbf{T} - \hat{\mathbf{A}}||^2_F + ||\mathbf{\Pi_0}\hat{\mathbf{B}}\mathbf{\Pi_0^T} - \hat{\mathbf{A}}||^2_F) \\
&\quad + \mathbf{tr}(\mathbf{\Pi_0}\hat{\mathbf{B}}\mathbf{\Pi_0^T}\hat{\mathbf{A}}) + \mathbf{tr}(\tilde{\mathbf{\Pi}}\hat{\mathbf{B}}\tilde{\mathbf{\Pi}}^\mathbf{T}\hat{\mathbf{A}}) - 2\mathbf{tr}(\mathbf{\Pi_0}\hat{\mathbf{B}}\tilde{\mathbf{\Pi}}^\mathbf{T}\hat{\mathbf{A}}) \\
&\leq ||\mathbf{\Pi_0}\hat{\mathbf{B}}\mathbf{\Pi_0}^T - \hat{\mathbf{A}}||^2_F + \mathbf{tr}((\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0})\hat{\mathbf{B}}((\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0})^\mathbf{T})\hat{\mathbf{A}}),
\end{aligned} \tag{15}$$

the last inequality holds since $||\tilde{\mathbf{\Pi}}\hat{\mathbf{B}}\tilde{\mathbf{\Pi}}^\mathbf{T} - \hat{\mathbf{A}}||^2_F \leq ||\mathbf{\Pi_0}\hat{\mathbf{B}}\mathbf{\Pi_0^T} - \hat{\mathbf{A}}||^2_F$.

iii. Upper Bound of $\mathbf{tr}((\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0})\hat{\mathbf{B}}(\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0})^\mathbf{T}\hat{\mathbf{A}})$:

Set $\mathbf{Z} = (\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0})\hat{\mathbf{B}}(\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0})^\mathbf{T}\hat{\mathbf{A}}$. For simplicity, we define $\mathbf{Y} = (\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0})\hat{\mathbf{B}}$ and $\mathbf{X} = (\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0})^\mathbf{T}\hat{\mathbf{A}}$, thus $\mathbf{Z} = \mathbf{YX}$. We focus on $\mathbf{tr}(\mathbf{Z})$. It is easy to verify that for any node $i$, when $\tilde{\mathbf{\Pi}}$ and $\mathbf{\Pi_0}$ map it to the same node, then $\mathbf{Z}_{ii} = 0$. If not, for node $i$ we assume that $\tilde{\mathbf{\Pi}}$ maps it to $s$ and $\mathbf{\Pi_0}$ maps it to $t$, where $s \neq t$. We can obtain the $i_{th}$ row of $\mathbf{Y}$ as $\mathbf{Y_{i\cdot}} = (\hat{\mathbf{B}}_{s1} - \hat{\mathbf{B}}_{t1}, \hat{\mathbf{B}}_{s2} - \hat{\mathbf{B}}_{t2}, \cdots, \hat{\mathbf{B}}_{sn} - \hat{\mathbf{B}}_{tn})$. Similarly, we can obtain the $i_{th}$ column of $\mathbf{X}$ as $\mathbf{X_{\cdot i}} = (\hat{\mathbf{A}}_{p_1 1} - \hat{\mathbf{A}}_{q_1 1}, \hat{\mathbf{A}}_{p_2 2} - \hat{\mathbf{A}}_{q_2 2}, \cdots, \hat{\mathbf{A}}_{p_n n} - \hat{\mathbf{A}}_{q_n n})^\mathbf{T}$, where $p_i(q_i)$ means the row index of the $1(-1)$ in the $i_{th}$ column of $\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0}$. If $\pi_0(j) = \tilde{\pi}(j)$, we simply set $\mathbf{X}_{ji} = 0$. Therefore $\mathbf{Z}_{ii}$, an element on diagonal of $\mathbf{Z}$, satisfies

$$\begin{aligned}
|\mathbf{Z}_{ii}| &= |\langle \mathbf{Y_{i\cdot}} \mathbf{X_{\cdot i}} \rangle| \leq ||\mathbf{Y_{i\cdot}}||_F ||\mathbf{X_{\cdot i}}||_F \\
&\leq n \max_k |\hat{\mathbf{B}}_{sk} - \hat{\mathbf{B}}_{tk}| \max_\ell |\hat{\mathbf{A}}_{p_\ell \ell} - \hat{\mathbf{A}}_{q_\ell \ell}|.
\end{aligned} \tag{16}$$

Note that if we normalize $w_{ij}$ to $[0,1]$ by dividing $||\mathbf{W}||_F$, with no impact on $\tilde{\mathbf{\Pi}}$ since it is irrelevant with $||\mathbf{W}||_F$, then $|\mathbf{Z}_{ii}| \leq n$. Taking the expectation of $\mathbf{A}$ and $\mathbf{B}$ on both sides of Inequality (16), we can obtain that

$$\begin{aligned}
\mathbf{E_{A,B}}|\mathbf{Z}_{ii}| &= \mathbf{E_{A,B}}(\max_{s,t,k} |\hat{\mathbf{B}}_{sk} - \hat{\mathbf{B}}_{tk}| \max_{p,q,\ell} |\hat{\mathbf{A}}_{p_\ell \ell} - \hat{\mathbf{A}}_{q_\ell \ell}|) \\
&\leq \mathbf{E_{A,B}}(\max_{s,t,k} |\mathbf{B}_{sk} - \mathbf{B}_{tk}| \max_{p,q,\ell} |\mathbf{A}_{p_\ell \ell} - \mathbf{A}_{q_\ell \ell}|) \\
&\leq \max_{p,q,\ell}\{[(p_{\boldsymbol{C_s C_j}} + p_{\boldsymbol{C_t C_j}}) \max\{s_1, s_2\}]^2\} = L,
\end{aligned}$$

where the first inequality holds since for any $s, t, k$ and the normalized weights $w_{sk}, w_{tk} \leq 1$, $|\hat{\mathbf{B}}_{sk} - \hat{\mathbf{B}}_{tk}| = |\sqrt{w_{sk}}\mathbf{B}_{sk} - \sqrt{w_{tk}}\mathbf{B}_{tk}| \leq |\mathbf{B}_{sk} - \mathbf{B}_{tk}|$, and $|\hat{\mathbf{A}}_{p_\ell \ell} - \hat{\mathbf{A}}_{q_\ell \ell}|$ is similar. Hence

$$|\mathbf{tr}((\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0})\hat{\mathbf{B}}((\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0})^\mathbf{T})\hat{\mathbf{A}})| \leq n \max_i |\langle \mathbf{Y_{i\cdot}} \mathbf{X_{\cdot i}} \rangle| \leq n^2 L. \tag{17}$$

iv. Upper Bound of $\frac{||\mathbf{\Pi_0} - \tilde{\mathbf{\Pi}}||^2_F}{||\mathbf{\Pi_0}||^2_F}$:

From Inequalities (14), (15) and (17), we can obtain

$$\begin{aligned}
&||(\mathbf{\Pi_0} - \tilde{\mathbf{\Pi}})||^2_F \leq \frac{2}{nK}||(\mathbf{\Pi_0} - \tilde{\mathbf{\Pi}})\hat{\mathbf{B}}||^2_F \\
&\leq \frac{8}{nK}||\mathbf{\Pi_0}\hat{\mathbf{B}}\mathbf{\Pi_0^T} - \hat{\mathbf{A}}||^2_F + 2\mathbf{tr}((\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0})\hat{\mathbf{B}}((\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0})^\mathbf{T})\hat{\mathbf{A}}) \\
&\leq \frac{8}{nK}||\mathbf{\Pi_0}\hat{\mathbf{B}}\mathbf{\Pi_0^T} - \hat{\mathbf{A}}||^2_F + \frac{4nL}{K}.
\end{aligned}$$

Since condition 2 holds, there exists a constant $\tilde{c} \geq 1$ such that $||\hat{\mathbf{A}} - \mathbf{\Pi_0}\hat{\mathbf{B}}\mathbf{\Pi_0}^\mathbf{T}||_F \leq \tilde{c}||\hat{\mathbf{A}} - \tilde{\mathbf{\Pi}}\hat{\mathbf{B}}\tilde{\mathbf{\Pi}}^\mathbf{T}||_F$. Therefore since $||\mathbf{\Pi_0}||^2_F = 2n$ and the first and third condition, we can bound the relative NME when $n \to \infty$ as:

$$\begin{aligned}
\frac{||(\mathbf{\Pi_0} - \tilde{\mathbf{\Pi}})||^2_F}{||\mathbf{\Pi_0}||^2_F} &\leq \frac{4}{n^2 K}||\mathbf{\Pi_0}\hat{\mathbf{B}}\mathbf{\Pi_0^T} - \hat{\mathbf{A}}||^2_F + \frac{2L}{K} \\
&= \frac{4\tilde{c}}{n^2 K}||\tilde{\mathbf{\Pi}}\hat{\mathbf{B}}\tilde{\mathbf{\Pi}}^\mathbf{T} - \hat{\mathbf{A}}||^2_F + \frac{2L}{K} \to 0.
\end{aligned}$$

This completes our proof. □

**Remark:** Although Theorem 3 does not ensure NME=0 exactly, it makes sense in de-anonymization since NME can be neglected when the size of network is very large and we can map asymptotically all nodes correctly under mild conditions. We show the mildness of these conditions under a particularly network structure: the whole networks connected with high probability, which must follow $p_{\boldsymbol{C_i C_j}} = \Omega(\frac{\log n}{n}), \forall i, j \in \{1, 2, ..., n\}$ [10].

Meanwhile, we take $s = s_1 = s_2 = o(1)$ denoting sparse sampling from $G$. For condition (i), $\frac{L}{K} = O(\frac{p^2_{\boldsymbol{C_i C_j}} s^2}{p_{\boldsymbol{C_i C_j}} s}) = o(1)$; For conditions (ii) and (iii), $|\mathbf{E}[(\mathbf{A} - \mathbf{\Pi B \Pi^T})_{ij}]| = p_{\boldsymbol{C_i C_j}} s + p_{\boldsymbol{C_{\pi(i)} C_{\pi(j)}}} s = O(\frac{s \log n}{n})$ satisfies both; For condition (iv), we show in Section V-B that it holds if $\tilde{\mathbf{\Pi}}$ *keeps invariant of community representations* (Recall Lemma 1) as $\mathbf{\Pi_0}$, which is easily realizable.

**Overlapping Communities Benefits De-anonymization:** Now we show that overlapping communities positively impact on reducing relative NME through facilitating conditions in Theorem 3, specifically condition 3.

For convenience, we assume $s = s_1 = s_2$. When $\pi_0$ keeps invariant of community representations, then on average condition 3 can be written as

$$2\sum_{1\le i<j\le n} p_{\boldsymbol{C_i C_j}} s\log\left(\frac{1-p_{\boldsymbol{C_i C_j}}(2s-s^2)}{p_{\boldsymbol{C_i C_j}}(1-s)^2}\right) = o(Kn^2). \quad (18)$$

To characterize the global situation in the networks, we define an *average probability* $\hat{p}$ such that

$$\begin{aligned}&\sum_{1\le i<j\le n} p_{\boldsymbol{C_i C_j}} s\log\left(\frac{1-p_{\boldsymbol{C_i C_j}}(2s-s^2)}{p_{\boldsymbol{C_i C_j}}(1-s)^2}\right)\\ &\quad=\frac{n(n-1)}{2}\log\left(\frac{1-\hat{p}(2s-s^2)}{\hat{p}(1-s)^2}\right)\hat{p}s,\end{aligned} \quad (19)$$

where $\hat{p}$ is positively correlated to the overlapping strength of the whole networks. Taking the derivative of $\hat{p}$ over $\left(\frac{1-\hat{p}(2s-s^2)}{\hat{p}(1-s)^2}\right)e^{\hat{p}s}$, we find that

$$d\left(\left(\frac{1-\hat{p}(2s-s^2)}{\hat{p}(1-s)^2}\right)e^{\hat{p}s}\right)/d\hat{p} = \frac{e^{\hat{p}s}}{(1-s)^2}\left(\frac{\hat{p}-1}{\hat{p}^2}-(2s-s^2)\right)\le 0, \quad (20)$$

indicating that $\left(\frac{1-\hat{p}(2s-s^2)}{\hat{p}(1-s)^2}\right)e^{\hat{p}s}$ is a decreasing function becoming smaller as overlapping strength increases. Therefore if the order of $\hat{p}$ rises, then the order of $||\mathbf{\hat{A}}-\mathbf{\Pi_0}\mathbf{\hat{B}}\mathbf{\Pi_0^T}||_F^2$ turns smaller, facilitating $||\mathbf{\hat{A}}-\mathbf{\Pi_0}\mathbf{\hat{B}}\mathbf{\Pi_0^T}||_F^2 = o(Kn^2)$.

Taking a vivid example of the proposed OSBM [18] in which

$$p_{\boldsymbol{C_i C_j}} = \frac{1}{1+ae^{-x}}, \quad (21)$$

where $a$ is an adjustable parameter and $x$ is the number of overlapping communities. We find that $\min_{i,j} p_{\boldsymbol{C_i C_j}} = \frac{1}{1+a}$ is a constant if $a = \Omega(1)$, and can be arbitrarily close to 1 when $x$ is large enough. So if $s = o(1)$ and $\hat{p} = 1-o(1)$, which means that the overlapping strength is very large, then

$$\begin{aligned}\hat{p}\log\left(\frac{1-\hat{p}(2s-s^2)}{\hat{p}(1-s)^2}\right) &= \hat{p}\log\left(1+\frac{1-\hat{p}}{\hat{p}(1-s)^2}\right)\\ &\approx \frac{1-\hat{p}}{(1-s)^2} = o(1) = o(\min_{i,j} p_{\boldsymbol{C_i C_j}}),\end{aligned} \quad (22)$$

thus condition (iii) holds. Meanwhile $s = o(1)$ makes condition (i) hold as well. Therefore all the four conditions in Theorem 3 hold, thus the relative NME vanishes to 0.

### *B. Complexity: WEMP can be Algorithmically Solved*

Upon proving the good performance of solving WEMP in large-scale networks, now we algorithmically demonstrate that WEMP reduces the complexity of the MMSE problem since the optimal mapping of WEMP can be perfectly found by the convex-concave based de-anonymization algorithm (CBDA).

*1) Formulation of WEMP in Constrained Optimization Form:* We first restate WEMP in the form of the following constrained optimization problem:

$$\begin{aligned}&\text{minimize}||(\mathbf{\hat{A}}-\mathbf{\Pi}\mathbf{\hat{B}}\mathbf{\Pi^T})||_F^2\\ \textbf{s.t. } &\forall i\in V_1,\ \textstyle\sum_i \mathbf{\Pi}_{ij}=1 && (23)\\ &\forall j\in V_2,\ \textstyle\sum_j \mathbf{\Pi}_{ij}=1 && (24)\\ &\forall i,j,\ \mathbf{\Pi}_{ij}\in\{0,1\}, && (25)\\ &\forall i\in V_1, \boldsymbol{C_i}=\boldsymbol{C_{\pi(i)}} && (26)\end{aligned}$$

Constraints 23, 24 and 25 are the attributes of permutation matrices. What's more, we append constraint 26, which means that our estimated mapping $\pi$ should keep the community representation of all the nodes in $V_1$ unchanged before and after mapping. Note that it is hard to implement this constraint directly in the optimization problem since it is not in the form of permutation matrix. However, we can easily convert it into a suitable one by defining a new matrix to characterize the community representation of all the nodes, which we call as "Community Representation Matrix", denoted as $\mathbf{M}$. Its formal definition is as follows.

Take Fig. 1 as an instance again. The community representation matrix of $G$, denoted as $\mathbf{M_G}$, satisfies

$$\mathbf{M_G^T} = \begin{bmatrix}1&0&1&1&0&0&0&0&0\\0&1&1&1&0&1&1&1&0\\0&0&1&1&1&0&1&1&1\end{bmatrix}.$$

Note that the community representation matrices for $G$, $G_1$ and $G_2$ are identical. So we set all of them to be $\mathbf{M}$. Hence the constraint (26) can be rewritten as $||\mathbf{\Pi M}-\mathbf{M}||_F^2 = 0$. According to optimization theory, we can form this constraint into the objective function by regarding it as the penalty term and obtain a new objective function

$$F_0(\mathbf{\Pi}) = ||\mathbf{\hat{A}}-\mathbf{\Pi}\mathbf{\hat{B}}\mathbf{\Pi^T}||_F^2 + \mu||\mathbf{\Pi M}-\mathbf{M}||_F^2,$$

where $\mu$ is an adjustable penalty parameter, which is large enough such that when the objective function reaches its minimum value, $||\mathbf{\Pi M}-\mathbf{M}||_F^2$ is exactly. Note that this transformation of objective function does not affect the previous analytical results of WEMP since we have the assumption that the true mapping keeps invariant of the community representation of every single node before and after mapping.

*2) Problem Relaxation and Idea of Algorithm Design:* Hereinafter, we focus on how we design our algorithm targeting the WEMP.

**Problem Relaxation:** WEMP is an integer program problem which cannot be solved efficiently. We relax the original feasible region of WEMP $\Omega_0$ into $\Omega$, which are respectively

$$\Omega_0 = \{\mathbf{\Pi}_{ij}\in\{0,1\}|\forall i,j, \textstyle\sum_i \mathbf{\Pi}_{ij}=1\ , \sum_j \mathbf{\Pi}_{ij}=1\};$$
$$\Omega = \{\mathbf{\Pi}_{ij}\in[0,1]|\forall i,j, \textstyle\sum_i \mathbf{\Pi}_{ij}=1\ , \sum_j \mathbf{\Pi}_{ij}=1\}.$$

After this relaxation the problem becomes tractable. However, a natural question arises: *How to obtain the solution of the original unrelaxed problem from that of the relaxed problem*?

**Idea of Convex-Concave Relaxation Method:** Note that the minimizer of a concave function must be at the boundary of the feasible region, coinciding that $\Omega_0$, the original feasible set, is just the boundary of $\Omega$. Therefore, a natural idea emerges: *We can modify the convex relaxed problem into a concave problem gradually.* Thus we apply the convex-concave optimization method (CCOM), whose concept is pioneeringly proposed in [22] to solve pattern matching problems: For $F_0(\mathbf{\Pi})$, we find its convex and concave relaxed version respectively $F_1(\mathbf{\Pi})$ and $F_2(\mathbf{\Pi})$. Then we obtain a new objective function as $F(\mathbf{\Pi}) = (1-\alpha)F_1(\mathbf{\Pi})+\alpha F_2(\mathbf{\Pi})$. We modify $\alpha$ gradually from 0 to 1 with interval $\Delta\alpha$, each time solving the new $F(\mathbf{\Pi})$ initialized by the optimizer last time. $F(\mathbf{\Pi})$ becomes more concave, with its optimum closer to $\Omega_0$ where $\tilde{\mathbf{\Pi}}$ lies.

*3) **Implementation of CCOM and Algorithm Design**:* Although [22] has proposed the general framework of CCOM, the way it presents to obtain $F_1(\mathbf{\Pi})$ and $F_2(\mathbf{\Pi})$ is rather complex, as it involves Kronecker product and the Laplacian matrix of graphs. Here we provide a simple way, as defined in Lemma 2, to get the convex relaxation and concave relaxation.

**Lemma 2.** *A proper way to get the convex relaxation and concave relaxation is*

$$F_1(\mathbf{\Pi}) = F_0(\mathbf{\Pi}) + \frac{\lambda_{min}}{2}(n - ||\mathbf{\Pi}||_F^2);$$

$$F_2(\mathbf{\Pi}) = F_0(\mathbf{\Pi}) + \frac{\lambda_{max}}{2}(n - ||\mathbf{\Pi}||_F^2).$$

*where $\lambda_{min}$ ($\lambda_{max}$) is the smallest (largest) eigenvalue of the Hessian matrix of $F_0(\mathbf{\Pi})$. Therefore we form our new objective function in CCOM as*

$$F(\mathbf{\Pi}) = (1-\alpha)F_1(\mathbf{\Pi}) + \alpha F_2(\mathbf{\Pi}) = F_0(\mathbf{\Pi}) + 2\xi(n - ||\mathbf{\Pi}||_F^2),$$

*where $\xi = (1-\alpha)\lambda_{min} + \alpha\lambda_{max}$, $\xi \in [\lambda_{min}, \lambda_{max}]$.*

The proof of Lemma 2, which is left in Appendix F, uses the sufcient and necessary condition that for a function whose variable is matrix is convex (concave) is that the Hessian matrix of this function is positive (negative) semi-denite.

Lemma 2 presents a simple way to implement CCOM algorithmically, since $F_0(\mathbf{\Pi})$ is just our objective function in Section V-B1 and $||\mathbf{\Pi}||_F^2$ can be computed easily. We can modify $F(\mathbf{\Pi})$ step by step from a convex function to a concave function by modifying the value of $\xi$ or $\alpha$. In the following analysis, we set $F_\xi(\mathbf{\Pi})$ equivalent to $F(\mathbf{\Pi})$ since $\xi$ is an adjustable parameter in $F(\mathbf{\Pi})$.

A vivid example of the CCOM under the formulation of $F_\xi(\mathbf{\Pi})$ by Lemma 2 is illustrated in Fig. 2. As can be seen in the figure, when $\xi$ starts at $\lambda_{min}$, $F_\xi(\mathbf{\Pi})$ is a convex function, thus we can obtain the minimizer of this objective function. After we find the minimizer, we increase $\alpha$ which makes the objective function become less convex. To obtain the minimizer of this new objective function, we have the prior knowledge of the previous minimizer, and since we only slightly modify the objective function, the optimal solution of new objective function should not deviate much from the previous one intuitively. Therefore we can start from the previous minimizer to find the new minimizer. Gradually, as $\alpha$ becomes increasingly larger, the objective function tends to be concave while the minimizer of it tends to get close to the boundary, on which the optimal solution of the original WEMP exists. The trail for the minimizer can be referred to the red line with arrows in Fig. 2.

Based on the above analysis, we propose Algorithm 1, called *Convex-concave Based De-anonymization Algorithm (CBDA)*, as our main algorithm for the weighted-edge matching problem (WEMP) under CCOM. Note that $F_0(\mathbf{\Pi})$ itself is convex in our problem, thus we can set $\xi$ from $0$ to an arbitrarily large number, which obviates the great complexity to calculate eigenvalues of Hessian matrices.

CBDA consists of an outer loop (lines 3 to 10) and an inner loop (lines 4 to 8). The outer loop modifies $\xi$ in CCOM.

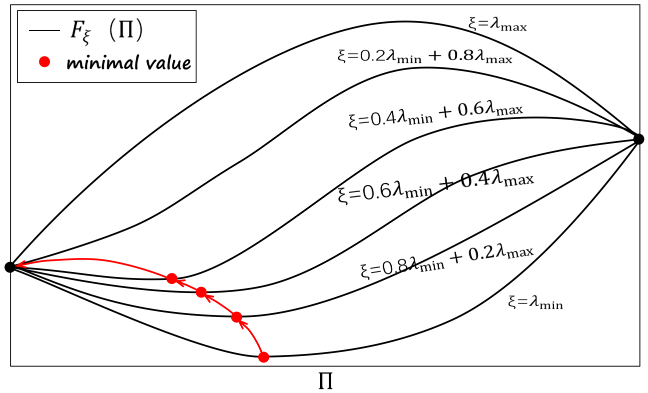



Fig. 2: An Illustration of the Implementation of CCOM

The inner loop finds the minimizer of $F(\mathbf{\Pi})$, whose main idea resembles descending algorithms: In line 5, we obtain descending direction by minimizing $\mathbf{tr}(\nabla_{\mathbf{\Pi_k}} F(\mathbf{\Pi_k})^T \mathbf{X}^\perp)$, dangling the highest probability to find a descending direction characterized by $\mathbf{tr}(\nabla_{\mathbf{\Pi_k}} F(\mathbf{\Pi_k})^T \mathbf{X}^\perp) < 0$. In line 6 we search for step length $\gamma_k$ contributing most to lowering $F(\mathbf{\Pi})$ on this descending direction. Line 7 is the update of estimation.

---

**Algorithm 1** Convex-concave Based De-anonymization Algorithm (CBDA)

---

**Input:** Adjacent matrices $\mathbf{A}$ and $\mathbf{B}$; Community assignment matrix $\mathbf{M}$;
Weight controlling parameter $\mu$; Adjustable parameters $\delta$, $\Delta\xi$.

**Output:** Estimated permutation matrix $\tilde{\mathbf{\Pi}}$.

1: Form the objective function $F_0(\mathbf{\Pi})$ and $F(\mathbf{\Pi})$.
2: $\xi \leftarrow 0$, $k \leftarrow 1$. $\mathbf{\Pi_1} \leftarrow \mathbf{1}_{n\times n}./n$. Set $\xi_m$, the upper limit of $\xi$.
3: **while** $\xi < \xi_m$ and $\mathbf{\Pi_k} \notin \Omega_0$ **do**
4:     **while** $k = 1$ or $|F(\mathbf{\Pi_{k+1}}) - F(\mathbf{\Pi_k})| \geq \delta$ **do**
5:         $\mathbf{X}^\perp \leftarrow \arg\min_{\mathbf{X}^\perp} \mathbf{tr}(\nabla_{\mathbf{\Pi_k}} F(\mathbf{\Pi_k})^T \mathbf{X}^\perp)$, where $\mathbf{X}^\perp \in \Omega$.
6:         $\gamma_k \leftarrow \arg\min_\gamma F(\mathbf{\Pi_k} + \gamma(\mathbf{X}^\perp - \mathbf{\Pi_k}))$, where $\gamma_k \in [0,1]$.
7:         $\mathbf{\Pi_{k+1}} \leftarrow \mathbf{\Pi_k} + \gamma_k(\mathbf{X}^\perp - \mathbf{\Pi_k})$, $k \leftarrow k+1$.
8:     **end while**
9:     $\xi \leftarrow \xi + \Delta\xi$.
10: **end while**

---

*4) Time Complexity and Convergence Analysis:* **Time Complexity:** The inner loop is similar to the Frank-Wolfe algorithm, with $O(n^6)$ in a round (since the input is an $n \times n$ matrix). If the maximum number of inner loops as $T$, thus the whole algorithm has a complexity of $O\left(\frac{n^6 T\xi}{\Delta\xi}\right)$ . As far as we know, a dearth of algorithmic analysis of seedless de-anonymization exists except for [23], [24], with their proposed algorithm sharing identical complexity of $O(n^6)$ with ours.

**Convergence:** Before the convergence analysis, we first clarify that:

- We set $\mathbf{\Pi_k}$ as the estimation after $k$ rounds in the inner loop, thus we have $\mathbf{\Pi_{k+1}} = \mathbf{\Pi_k} + \gamma_k(\mathbf{X}_\perp - \mathbf{\Pi_k})$.

- We set $F_\xi(\mathbf{\Pi}) = F_0(\mathbf{\Pi}) + \xi(n - ||\mathbf{\Pi}||_F^2)$ and $\mathbf{\Pi}^\xi$ as the minimizer of $F_\xi(\mathbf{\Pi})$.

Then we analyze the convergence of CBDA and propose Lemma 3.

**Lemma 3.** *CBDA converges and the final output is a permutation matrix in the original feasible region* $\Omega_0$.

*Proof.* There are inner and outer loops in CBDA and we show the convergence of them respectively.

**1. Inner Loop:** We provide the outline of the proof here and put the detailed one in Appendix G.

We focus on $F_\xi(\mathbf{\Pi_{k+1}})$ and $F_\xi(\mathbf{\Pi}^\xi)$. According to Taylor's Theorem, we can derive that

$$\begin{aligned} F_\xi(\mathbf{\Pi_{k+1}}) &= F_\xi(\mathbf{\Pi_k} + \gamma_k(\mathbf{X}_\perp - \mathbf{\Pi_k})) \\ &\leq F_\xi(\mathbf{\Pi_k}) + \gamma_k \mathbf{tr}(\nabla F_\xi^T(\mathbf{\Pi_k})(\mathbf{\Pi}^\xi - \mathbf{\Pi_k})) + \gamma_k \mathbf{R_k}, \end{aligned} \tag{27}$$

$$\begin{aligned} F_\xi(\mathbf{\Pi}^\xi) &= F_\xi(\mathbf{\Pi_k} + \mathbf{\Pi}^\xi - \mathbf{\Pi_k}) \\ &= F_\xi(\mathbf{\Pi_k}) + \mathbf{tr}(\nabla F_\xi^T(\mathbf{\Pi_k})(\mathbf{\Pi}^\xi - \mathbf{\Pi_k})) + \mathbf{R}'_\mathbf{k}, \end{aligned} \tag{28}$$

where $\gamma_k \mathbf{R_k}$ and $\mathbf{R}'_\mathbf{k}$ is the remainder of this Taylor series.

Combining Eqn. (27) and (28), we can obtain

$$\begin{aligned} & F_\xi(\mathbf{\Pi_{k+1}}) - F_\xi(\mathbf{\Pi}^\xi) \\ & \leq (1-\gamma_k)(F_\xi(\mathbf{\Pi_k}) - F_\xi(\mathbf{\Pi}^\xi)) + \gamma_k \Delta \mathbf{R_k} \\ & \leq \prod_{i=1}^{k}(1-\gamma_i)(F_\xi(\mathbf{\Pi_1}) - F_\xi(\mathbf{\Pi}^\xi)) + \sum_{i=1}^{k}\gamma_i \prod_{j=1}^{k-i}(1-\gamma_j)\Delta \mathbf{R}_i. \end{aligned} \tag{29}$$

For $F_\xi(\mathbf{\Pi_1}) - F_\xi(\mathbf{\Pi}^\xi)$, note that $\mathbf{\Pi_1} = \mathbf{\Pi}^{\xi - \mathbf{\Delta}\xi}$, then we can derive that

$$F_\xi(\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}) - F_\xi(\mathbf{\Pi}^\xi) \leq \Delta\xi(||\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}||_F^2 - ||\mathbf{\Pi}^\xi||_F^2). \tag{30}$$

Therefore by combining Inequalities (29) and (30), we can obtain that both $\prod_{i=1}^{k}(1-\gamma_i)(F_\xi(\mathbf{\Pi_1}) - F_\xi(\mathbf{\Pi}^\xi))$ and $\sum_{i=1}^{k}\gamma_i \prod_{j=1}^{k-i}(1-\gamma_j)\Delta \mathbf{R}_i$ approach to 0.

**2. Outer Loop:** Note that from Eqn. (30), we know $(||\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}||_F^2 - ||\mathbf{\Pi}^\xi||_F^2)$ is nonnegative since $\Delta\xi > 0$ and $\mathbf{\Pi}^\xi$ is the minimizer of $F_\xi(\mathbf{\Pi})$. Thus $||\mathbf{\Pi}^\xi||_F^2 \leq ||\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}||_F^2$. Note that $||\mathbf{\Pi}||_F^2 \leq n$. From Inequality (30), we find that

$$\begin{aligned} F_\xi(\mathbf{\Pi}^\xi) &\geq F_0(\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}) + (\xi - \Delta\xi)(n - ||\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}||_F^2) \\ &\quad - \Delta\xi(n - ||\mathbf{\Pi}^\xi||_F^2) \\ &= F_{\xi-\Delta\xi}(\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}) - \Delta\xi \mathbf{tr}(||\mathbf{\Pi}^\xi||_F^2 - n). \end{aligned}$$

Therefore

$$\begin{aligned} & |F_\xi(\mathbf{\Pi}^\xi) - F_{\xi-\Delta\xi}(\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi})| \\ & \leq \Delta\xi|||(\mathbf{\Pi}^\xi)||_F^2 - n| \leq \Delta\xi|||\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}||_F^2 - n| \\ & \leq \Delta\xi|||\mathbf{\Pi}^{\xi_\mathbf{0}}||_F^2 - n| \leq \Delta\xi(n-1), \end{aligned}$$

where the third inequality holds since $\mathbf{\Pi}^{\xi_\mathbf{0}}$ is the minimizer of $F_{\lambda_{min}}(\mathbf{\Pi})$, i.e., the convex relaxation of $F_0(\mathbf{\Pi})$, and the fourth inequality holds since $\min_{\mathbf{\Pi}\in\Omega} ||\mathbf{\Pi}||_F^2 = 1$ and $\mathbf{\Pi} = \mathbf{1}_{n\times n}./n$ is the minimizer. Therefore, if $\Delta\xi = o(\frac{1}{n})$, the outer loop converges.

Combining the convergence analysis of both inner and outer loops above, we complete the proof of the convergence of CBDA. □

TABLE II: **Main Experimental Parameters**

| Notation | Definition | Range |
|---|---|---|
| $N$ | Number of Nodes | {500, 1000, 1500, 2000} |
| $s$ | Sampling Probability ($s_1 = s_2 = s$) | 0.3-0.9 |
| $a$ | OSBM Parameter | {3, 5, 7, 9} |
| $\eta$ | Community Ratio | {0.05, 0.1} |
| $OL/NOL$ | Overlapping or Non-Overlapping | {OL, NOL} |

Lemma 3 ensures that CBDA can perfectly solve WEMP, which vanishes the relative NME. Therefore CBDA is an algorithmic approach for seedless de-anonymization with high feasibility and good performance.

## VI. Experimental Aspect of Social Network De-anonymization Problem

In this section, we utilize three datasets: synthetic networks, sampled real social networks and true cross-domain networks, to validate our theoretical results and the performance of CBDA.Before presenting empirical results, we first introduce our experimental setup.

### A. Experiment Setup

*1) Main Parameters:* We list our adjustable parameters involved in our experiments in Table II. Three parameters are in need of further explanations:

i) $a$. This is a parameter in the overlapping stochastic block model (OSBM) which determines the $p_{\boldsymbol{C_i C_j}}$, the probability of edge existence between nodes $i$ and $j$ in underlying graph. Specifically, $p_{\boldsymbol{C_i C_j}}$ can be expressed as

$$p_{\boldsymbol{C_i C_j}} = \frac{1}{1 + ae^{-x}},$$

where $x$ is the number of communities that both nodes $i$ and $j$ belong to. Note that if $a$ becomes larger (smaller), then $p_{\boldsymbol{C_i C_j}}$ is smaller (larger) so that the graph becomes sparser (denser).

ii) $\eta$. This is the community ratio. It means the ratio between the number of communities and nodes. This ratio reflects the density of community numbers regardless of the size of network. In performance validation of CBDA we set $\eta = 0.05$ or $0.1$, while when studying the influence of $\eta$ on de-anonymization accuracy, it will be endowed with more values.

iii) $OL/NOL$. $OL$ means that communities are overlapping while $NOL$ means not. This makes for illustrating the impact of the overlapping property of communities on the mapping accuracy.

*2) Experimental Datasets:* We discuss three adopted datasets, which is shown is Table III, in an order from model-based to real social networks.

i) Synthetic Networks: We generate networks by setting the community representation of every node independently and randomly deciding the edge existence in node pair $(i, j)$ based on OSBM [18]

ii) Sampled Real Social Networks: The underlying social network $G$ is extracted from LiveJournal [17], while $G_1$ and $G_2$ are sampled from $G$ with the same probability $s$.

iii) Cross-Domain Co-author Networks: The co-author networks are from the Microsoft Academic Graph (MAG) [11]. We extract 4 networks belonging to different sub-areas in the

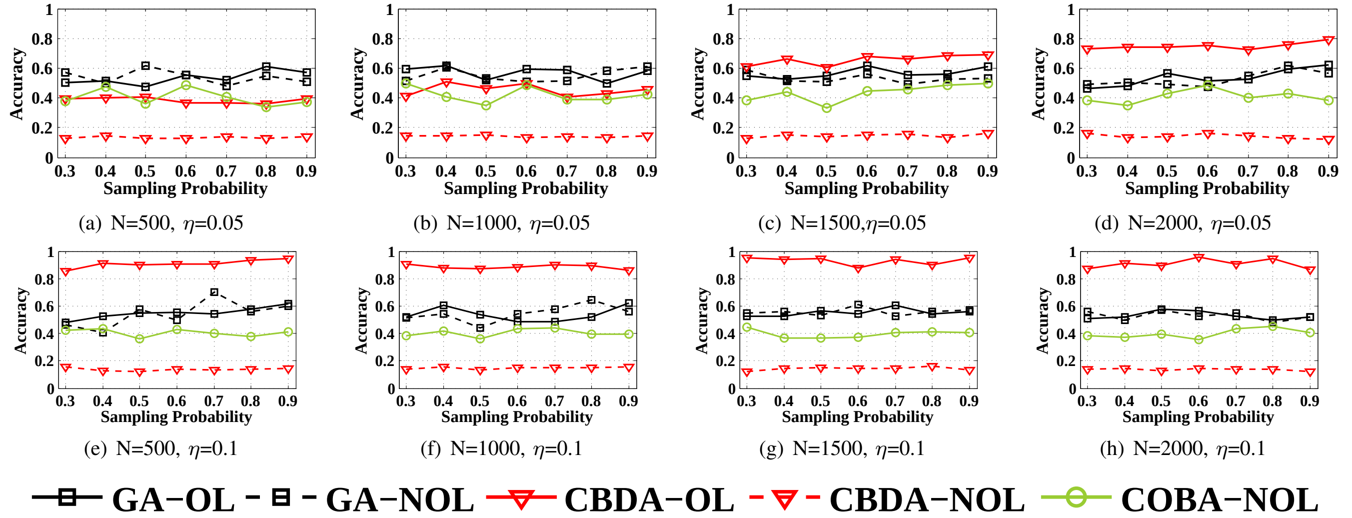



Fig. 3: Experiments on Synthetic Networks with $a = 5$.

TABLE III: **Datasets in Basic Experiments**

| **Dataset** | Synthetic | Sampled Real Social | Cross-Domain Co-author |
|---|---|---|---|
| **Source** | OSBM | LiveJournal [1] | MAG [11] |
| **Num. of Nodes** | $500 \sim 2000$ | $500 \sim 2000$ | 3176 |
| **Num. of Communities** | $25 \sim 1000$ | $25 \sim 1000$ | 89 |

field of computer science, with the same group of authors, each of whom has a unique 8-bit hexadecimal ID enabling us to construct the true mapping between two networks as the one mapping nodes with same ID. Each network can be viewed as $G_1$ or $G_2$, thus there are $C_4^2 = 6$ combinations. Note that we can assign $w_{ij}$ on all these 3 datasets since the prior knowledge is just the community assignment matrix $M$, which can be generated or known from the real networks.

*3) Algorithms for Comparison and Performance Metric:* We exclude algorithms for seeded de-anonymization and select algorithms suitable for seedless cases related to our main point: showing the impact of overlapping communities on reducing NME, though other algorithms might outperform ours. We select two algorithms for comparison: (i) the Genetic Algorithm (GA), an epitome of heuristic algorithms; (ii) the Convex Optimization-Based Algorithm (COBA) in [23], [24], assigning a node to a unique community, which primarily suits non-overlapping cases. The performance metric is *accuracy*, the proportion of correctly mapped nodes.

*4) Supplementary Experiments:* To make our experimental validation more comprehensive and convincing, we study (i) the effect of different community ratios ($\eta$), which is modified from $0.025$ to $0.2$ with interval $0.025$, on the accuracy based on sampled real social networks; (ii) the priority of our cost function with $\mathbf{W}$ derived from MMSE makes for higher accuracy, comparing with the cost function without $\mathbf{W}$ in [3]; (iii) the instability of GA revealing its practical limitation *and thus in validation on* 3 *datasets we take the average preformance of GA* 10 *times as its accuracy.*

## *B. Experiment Results*

*1) Synthetic Networks:* Fig. 3 and 4 illustrate our experimental results on synthetic networks, where the community ratio $\eta \in \{0.05, 0.1\}$, the network size $N$ range from 500 to 2000 in Fig 3 and the OSBM parameter $a$ range from 3 to 9 in Fig 4. More exhaustive results under variable $N$, $a$ and $\eta$ are shown in Appendix H in two figures, Fig. 5 and Fig. 6, whose $\eta$ is 0.05 and 0.1 respectively.

From Fig. 3, we observe that: (i) The *average* accuracy of genetic algorithm (GA) under different settings keeps at levels around $40\% - 60\%$, which illustrates that different sizes, densities and whether the communities overlap or not do not make a difference on the performance of GA averagely. This is because GA examines the edges one by one to make the cost function as small as possible, therefore it is not seriously affected by the global setting of the networks. (ii) The accuracy of COBA also keeps at a stable level in different situations. However, COBA can only cope with non-overlapping situations, and generally its performance is inferior to GA when communities are not overlapped, which is in line with the results in [23], [24]. (iii) The accuracy of CBDA, our algorithm, rises with the increasing network size $N$ when $\eta = 0.05$. Specifically, when $N$ goes from 500 to 2000, the accuracy rises from approximately $40\%$ to $80\%$. This corresponds to our Theorem 3 that as the size of networks becomes larger, the relative NME becomes smaller. When $\eta = 0.1$, which indicates denser communities, the accuracy of CBDA is at a high level (around $90\%$) even if the network size is small. On the other hand, however, when dealing with non-overlapping situations, our CBDA works stably but not as efficiently as GA or COBA, with the accuracy only around $20\%$.

From Fig. 4, we can observe that when communities are non-overlapping, for both COBA and our CBDA, curves at $\eta = 0.05$ have the same trends as curves at $\eta = 0.1$, showing that the community density under non-overlapping situations does not affect the performance of all these algorithms. However, our CBDA always performs better than other algorithms when the communities are overlapping each other. With a certain community size, the accuracy of CBDA keeps at a stable level when $\eta = 0.05$ and varies with edge density

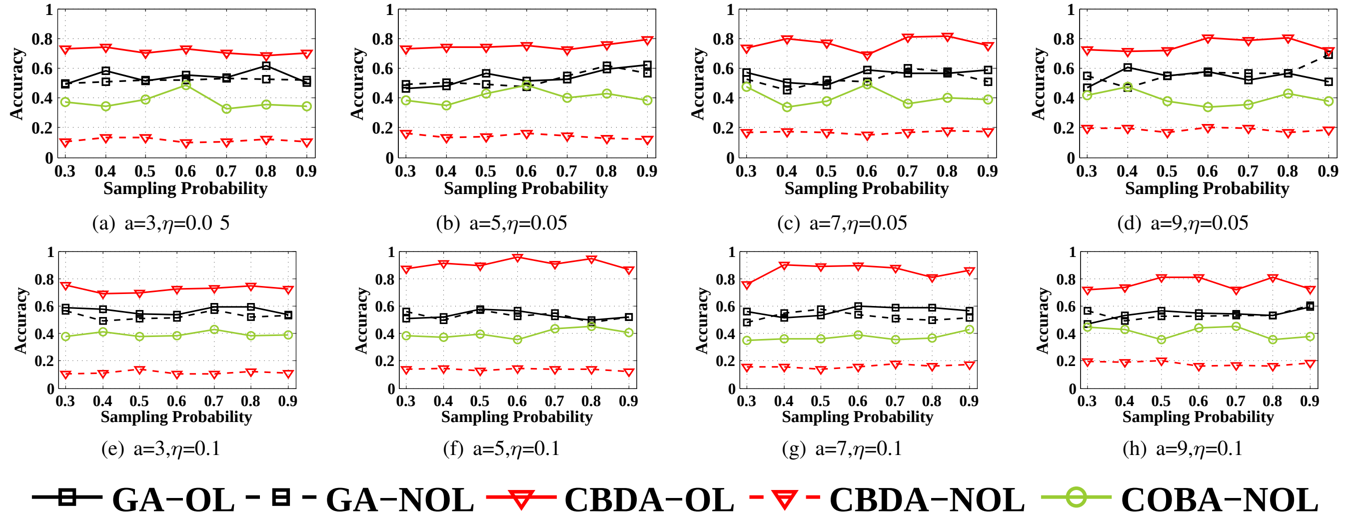



Fig. 4: Experiments on Synthetic Networks with $N = 2000$.

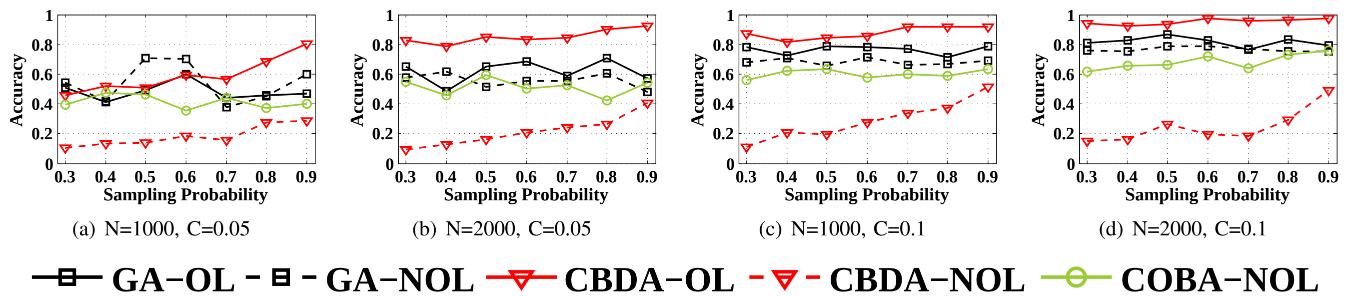



Fig. 5: Experiments on Sampled Real Social Networks.

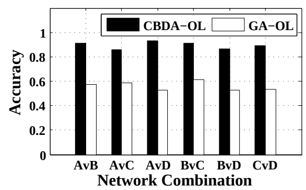
$a$ when $\eta = 0.1$. Thus, when the community density is large, the performance of CBDA is mainly decided by the edge density ($a$), positively correlated to community density; when the community density is small, then the performance of CBDA is mainly decided by the size of the networks ($N$). This shows that the community ratio (density) determines the dominant factor ($a$ or $N$) in de-anonymization accuracy in networks with overlapping communities.

*2) Sampled Real Social Networks:* In sampled real social networks, we utilize the real underlying network, thus no modifications on $a$ exist. We select representative results in Fig. 5 and leave more exhaustive results in Fig. 7 in Appendix H. We can observe: (i) GA performs better in larger networks and under denser communities, either overlapping or non-overlapping; (ii) The performance of COBA is also enhanced when the size of networks become larger and the community becomes denser; (iii) The performance of CBDA under non-overlapping situations does not outperform other algorithms, but a rising tendency exists as the sampling probability $s$ becomes larger; (iv) The performance of CBDA under overlapping situations still performs well under denser communities and larger network size, with the highest point $95\%$ and the highest average level around $90\%$ when $N = 2000$ and $\eta = 0.1$, the largest size and densest communities in Table III.

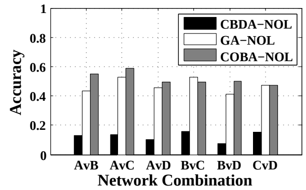
Synthesizing the above four observations, we can learn that the OSBM does not reflect the real social networks very precisely, since the performance of all three algorithms under non-overlapping or overlapping communities differs in two datasets. Moreover, with the same experimental setting, we discover that the performance of our CBDA is better in sampled real social networks than in OSBM-based synthetic networks, which further undergirds the high performance of our algorithm in practical use. Additionally, the results in Fig. 5 also meet Theorem 3 that as the network size becomes larger, the relative NME is much smaller and close to $0$, indicating that Theorem 3 also works in real social networks.

*3) Cross-Domain Co-author Networks:* In cross-domain co-author networks, we pick up four networks with the same set of $3176$ users. Fig. 6 illustrates our results. We find that in non-overlapping situation, the results correspond to those in previous datasets that our CBDA does not perform well, while GA and COBA work well. On the other hand, in overlapping situation, we find our CBDA reaches accuracy around $90\%$, outstripping GA whose accuracy is averagely $60\%$. This phenomenon upgrades the significance of our CBDA in de-anonymization with overlapping communities since the dataset is entirely realistic. Moreover, since overlapping cases are much more quotidian in real social networks, CBDA has wider usage than GA and COBA.

*4) The Effect of Community Density:* We apply the sampled real social networks under which we can adjust the community ratio $\eta$. We modify $\eta$ from $0.025$ to $0.2$, with interval $0.025$.

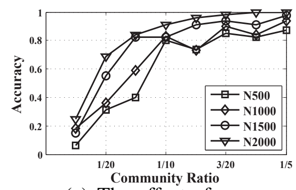



(a) Overlapping

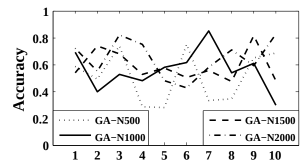



(b) Non-Overlapping

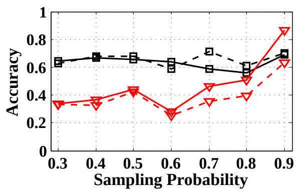
Fig. 6: Experiments on Cross-Domain Co-author Networks.



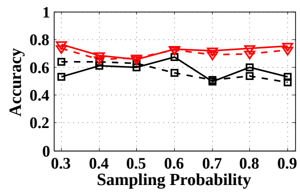

(a) The effect of $\eta$




(b) The Instability of GA




(c) The effect of $\mathbf{W}$, N=1000




(d) The effect of $\mathbf{W}$, N=2000




Fig. 7: Supplementary Experiments

The results are shown in Fig. 7(a). We can observe that our CBDA performs better when the network size is larger, which again echoes the conclusion in Theorem 3. Moreover, with the larger community ratio, the accuracy of CBDA rises up, showing that CBDA is suitable for social networks with highly overlapping communities. What's more, a huge gap occurs between the accuracy of $\eta = 0.025$ and $\eta = 0.075$, and when $\eta \geq 0.1$, the accuracy of CBDA under all the network sizes involved keeps at high levels, around $80\%$ or higher. The results further illustrate that the higher community ratio $\eta$, the better de-anonymizing result will be.

*5) The Effect of Weight Matrix* $\mathbf{W}$*:* We put representative results in Fig. 7(c) and 7(d) and leave more exhaustive results in Fig. 4 in Appendix H. As these two figures show, CBDA works better appending $\mathbf{W}$ derived by MMSE, since the non-weighted cost function, adopted in [3], fails to distinguish nodes belonging to different number of communities. It shows the superiority of cost functions derived with rationale, as we claim in Section IV. Under larger network size, however, the difference becomes fainter since the impact of distinguishing a single node by $w_{ij}$ is weaker than the benefits brought by large size shown in Theorem 3.

*6) The Instability of Genetic Algorithm:* We disclose the instability of GA in Fig. 7(b). We run GA 10 times under sampled real social networks with different sizes. The performance of GA fluctuates violently, bewildering adversaries in the quality of a specific estimation, which inhibits the usage of GA in practice.

## VII. Conclusion

We tackle seedless de-anonymization under a more practical social network model parameterized by *overlapping communities* than existing work. By MMSE, we derive a well-justified cost function minimizing the expected number of mismatched users. While showing the NP-hardness of minimizing MMSE, we validly transform it into WEMP which resolves the tension between optimality and complexity: (i) WEMP asymptotically returns a negligible mapping error under mild conditions facilitated by higher overlapping strength; (ii) WEMP can be algorithmically solved via CBDA, which exactly finds the optimum of WEMP. Extensive experiments further confirm the effectiveness of CBDA under overlapping communities.

## References


[1] J. Leskovec and A. Krevl, "SNAP Datasets: Stanford Large Network Dataset Collection", http://snap.stanford.edu/data, 2014.

[2] A. Narayanan and V. Shmatikov, "De-anonymizing social networks", in *IEEE Symposium on Security and Privacy*, pp. 173-187, 2009.

[3] P. Pedarsani and M. Grossglauser, "On the privacy of anonymized networks" in *Proc. ACM SIGKDD*, pp. 1235-1243, 2011.

[4] E. Kazemi, L. Yartseva and M. Grossglauser, "When can two unlabeled networks be aligned under partial overlap?", in *IEEE 53rd Annual Allerton Conference on Communication, Control, and Computing*, pp. 33-42, 2015.

[5] D. Cullina and N. Kiyavash, "Improved achievability and converse bounds for Erdős-Rényi graph matching", in *Proc. ACM SIGMETRICS*, pp. 63-72, 2016.

[6] S. Ji, W. Li, M. Srivatsa and R. Beyah, "Structural data de-anonymization: Quantification, practice, and implications", in *Proc. ACM CCS*, pp. 1040-1053, 2014.

[7] S. Ji, W. Li, N. Z. Gong, P. Mittal and R. Beyah, "On your social network de-anonymizablity: Quantification and large scale evaluation with seed knowledge" in *NDSS* 2015.

[8] E. Onaran, G. Siddharth and E. Erkip, "Optimal de-anonymization in random graphs with community structure", arXiv preprint arXiv:1602.01409, 2016.

[9] G. Palla, I. Derenyi, L. J. Farkas and T. Vicsek, "Uncovering the overlapping community structure of complex networks in nature and society", in *Nature*, No. 7043, Vol. 435, pp. 814-818, 2005.

[10] P. Erdős and A. Rényi, "On random graphs", in *Publicationes Mathematicae*, pp. 290-297, 1959.

[11] https://www.microsoft.com/en-us/research/project/microsoft-academic-graph/

[12] L. Yartseva and M. Grossglauser, "On the performance of percolation graph matching", in *Proc. ACM COSN*, pp. 119-130, 2013.

[13] E. Kazemi, S. H. Hassani and M. Grossglauser, "Growing a graph matching from a handful of seeds", in *Proc. the VLDB Endowment*, pp. 1010-1021, 2015.

[14] C. F. Chiasserini, M. Garetto and E. Leonardi, "Social network de-anonymization under scale-free user relations", in *IEEE/ACM Trans. on Networking*, Vol. 24, No. 6, pp. 3756-3769, 2016.

[15] N. Korula and S. Lattanzi, "An efficient reconciliation algorithm for social networks", in *Proc. the VLDB Endowment*, pp. 377-388, 2014.

[16] C. F. Chiasserini, M. Garetto and E. Leonardi, "Impact of clustering on the performance of network de-anonymization", in *Proc. ACM COSN*, pp. 83-94, 2015.

[17] J. Yang and J. Leskovec, "Defining and evaluating network communities based on ground-truth", in *Knowledge and Information Systems*, No. 42, Vol. 1, pp. 181-213, 2015.

[18] P. Latouche, E. Birmele and C. Ambroise, "Overlapping stochastic block models with application to the french political blogosphere", in *The Annals of Applied Statistics* pp.309–336, 2011.

[19] A. Decelle, F. Krzakala, C. Moore and L. Zdeborova, "Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications" in *Physical Review E*, No. 84, Vol. 6, pp. 066106, 2011.

[20] G. H. Hardy, J. E. Littlewood and G. Pólya, "Inequalities. Reprint of the 1952 edition." in *Cambridge Mathematical Library*, 1988

[21] O. Kariv and S. L. Hakimi, "Algorithm approach to network location problems - 2. the p-medians", in *Siam Journal on Applied Mathematics*, No. 3, Vol. 37, pp. 539-560, 1979.

[22] M. Zaslavskiy, F. Bach and J. P. Vert, "A path following algorithm for the graph matching problem" , in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, No. 12, Vol. 31, pp. 2227-2242, 2009.

[23] L. Fu, X. Fu, Z. Hu, Z. Xu and X. Wang, "De-anonymization of Social Networks with Communities: When Quantifications Meet Algorithms", arXiv preprint arXiv:1703.09028, 2017.

[24] X. Fu, Z. Hu, Z. Xu, L. Fu and X. Wang, "De-anonymization of Networks with Communities: When Quantifications Meet Algorithms", in *IEEE Globecom*, 2017.

[25] X. Wu, Z. Hu, X. Fu, L. Fu, X. Wang and S. Lu, "Social Network De-anonymization with Overlapping Communities: Analysis, Algorithm and Experiments", in *IEEE INFOCOM 2018*, Hawaii, USA, 2018.

## Appendix

### A. PROOF OF THEOREM 1

*Proof.* Define $\mathcal{G}_{\mathbf{\Pi}}$ as the set of all realizations of the underlying network which is in consistency with the given $G_1$, $G_2$ and $\mathbf{\Pi}$. Then the MMSE estimator can be written as

$$\hat{\mathbf{\Pi}} = \arg\min_{\mathbf{\Pi}\in\mathbf{\Pi^n}} \sum_{\mathbf{\Pi_0}\in\Pi^n} ||\mathbf{\Pi}-\mathbf{\Pi_0}||_F^2 \sum_{G\in\mathcal{G}_{\mathbf{\Pi}}} Pr(G,\mathbf{\Pi_0}|G_1,G_2,\boldsymbol{\theta}).$$

Let us focus on the conditional probability $Pr(G,\mathbf{\Pi_0}|G_1,G_2,\boldsymbol{\theta})$ in Eqn. (2). According to Bayesian's formula, along with the fact that $G_1$ and $G_2$ are sampled independently from each other, we obtain

$$\begin{aligned} Pr(G,\mathbf{\Pi_0}|G_1,G_2,\boldsymbol{\theta}) &= \frac{Pr(G,G_1,G_2,\mathbf{\Pi_0})}{Pr(G_1,G_2)} \\ &\sim Pr(G)Pr(G_1|G)Pr(G_2|G,\mathbf{\Pi_0}), \end{aligned} \tag{31}$$

where $a \sim b$ means that $a$ and $b$ are different only in parameters unrelated to $\mathbf{\Pi_0}$, which will not change the value of $\arg\max$ or $\arg\min$.[6] Note that the parameter set $\boldsymbol{\theta}$ remains invariant, so we need not add $\boldsymbol{C_i}$ and $\boldsymbol{\theta}$ into further consideration.

Set $E^{ij}$ as the indicator variable about whether an edge exists between nodes $i$ and $j$ in the edge set $E$. If an edge exists then $E^{ij}=1$, otherwise $E^{ij}=0$. The same rule also holds for indicators $E_1^{ij}$ and $E_2^{ij}$. Therefore Eqn. (31) can be further written as

$$\begin{aligned} &\sum_{G\in\mathcal{G}_{\mathbf{\Pi}}} Pr(G)Pr(G_1|G)Pr(G_2|G,\Pi_0) \\ &= \sum_{G\in\mathcal{G}_{\mathbf{\Pi}}} \prod_{i<j}^{n} s_1^{E_1^{ij}}(1-s_1)^{E^{ij}-E_1^{ij}} s_2^{E_2^{\pi_0(i)\pi_0(j)}} \\ &\quad\cdot (1-s_2)^{E^{ij}-E_2^{\pi_0(i)\pi_0(j)}} p_{\boldsymbol{C_i}\boldsymbol{C_j}}^{E^{ij}}(1-p_{\boldsymbol{C_i}\boldsymbol{C_j}})^{1-E^{ij}} \\ &= \prod_{i<j}\left(\frac{s_1}{1-s_1}\right)^{E_1^{ij}}\left(\frac{s_2}{1-s_2}\right)^{E_2^{\pi_0(i)\pi_0(j)}} \\ &\quad\cdot \sum_{G\in\mathcal{G}_{\mathbf{\Pi}}}\left((1-s_1)(1-s_2)\frac{p_{\boldsymbol{C_i}\boldsymbol{C_j}}}{1-p_{\boldsymbol{C_i}\boldsymbol{C_j}}}\right)^{E^{ij}} \\ &\sim \sum_{G\in\mathcal{G}_{\mathbf{\Pi}}}\left((1-s_1)(1-s_2)\frac{p_{\boldsymbol{C_i}\boldsymbol{C_j}}}{1-p_{\boldsymbol{C_i}\boldsymbol{C_j}}}\right)^{E^{ij}}. \end{aligned} \tag{32}$$

Note that the last equivalence in Eqn. (32) holds since the term $\left(\frac{s_1}{1-s_1}\right)^{E_1^{ij}}$ does not depend on $\pi_0$ and the product $\prod_{i<j}\left(\frac{s_2}{1-s_2}\right)^{E_2^{\pi_0(i)\pi_0(j)}}$ is independent of $\pi_0$ due to the bijective property of $\pi_0$.

Then we define $G^*_{\pi_0}$ as the graph which has the smallest number of edges in $\mathcal{G}_{\mathbf{\Pi}}$. Equivalently $G^*_{\pi_0}=(V,E_1\cup\pi_0(E_1))$, where $\pi_0(E_1)=\{(\pi_0(i),\pi_0(j))|(i,j)\in E_1\}$. Now we set $E^*_{\pi_0}$ as the edge set of $G^*_{\pi_0}$, and $E^{*ij}_{\pi_0}$ as the indicator variable between nodes $i$ and $j$, i.e., $E^{*ij}_{\pi_0}=1$ if $(i,j)\in E^*_{\pi_0}$ and $E^{*ij}_{\pi_0}=0$ otherwise. Then we sum up all the graphs in $\mathcal{G}_{\mathbf{\Pi}}$

$$\begin{aligned} &\sum_{G\in\mathcal{G}_{\mathbf{\Pi}}}\left((1-s_1)(1-s_2)\frac{p_{\boldsymbol{C_i}\boldsymbol{C_j}}}{1-p_{\boldsymbol{C_i}\boldsymbol{C_j}}}\right)^{E^{ij}} \\ &= \prod_{i<j}^{n}\left((1-s_1)(1-s_2)\frac{p_{\boldsymbol{C_i}\boldsymbol{C_j}}}{1-p_{\boldsymbol{C_i}\boldsymbol{C_j}}}\right)^{E^{*ij}_{\pi_0}} \\ &\quad\cdot \sum_{k=0}^{E_{ij}-E^{*ij}_{\pi_0}} C^k_{E_{ij}-E^{*ij}_{\pi_0}}\left((1-s_1)(1-s_2)\frac{p_{\boldsymbol{C_i}\boldsymbol{C_j}}}{1-p_{\boldsymbol{C_i}\boldsymbol{C_j}}}\right)^{k}. \end{aligned} \tag{33}$$

Note that in Eqn. (33) last multiplicative factor ,

$$\sum_{k=0}^{E_{ij}-E^{*ij}_{\pi_0}} C^k_{E_{ij}-E^{*ij}_{\pi_0}}\left((1-s_1)(1-s_2)\frac{p_{\boldsymbol{C_i}\boldsymbol{C_j}}}{1-p_{\boldsymbol{C_i}\boldsymbol{C_j}}}\right)^{k},$$

yields as a Bernoulli sum, therefore Eqn. (33) can be further written as

$$\begin{aligned} &\sum_{G\in\mathcal{G}_{\mathbf{\Pi}}}\left((1-s_1)(1-s_2)\frac{p_{\boldsymbol{C_i}\boldsymbol{C_j}}}{1-p_{\boldsymbol{C_i}\boldsymbol{C_j}}}\right)^{E^{ij}} \\ &= \prod_{i<j}^{n}\left((1-s_1)(1-s_2)\frac{p_{\boldsymbol{C_i}\boldsymbol{C_j}}}{1-p_{\boldsymbol{C_i}\boldsymbol{C_j}}}\right)^{E^{*ij}_{\pi_0}} \\ &\quad\cdot\left(1+(1-s_1)(1-s_2)\frac{p_{\boldsymbol{C_i}\boldsymbol{C_j}}}{1-p_{\boldsymbol{C_i}\boldsymbol{C_j}}}\right)^{1-E^{*ij}_{\pi_0}} \\ &\sim \prod_{i<j}^{n}\left(\frac{p_{\boldsymbol{C_i}\boldsymbol{C_j}}(1-s_1)(1-s_2)}{1-p_{\boldsymbol{C_i}\boldsymbol{C_j}}(s_1+s_2-s_1s_2)}\right)^{E^{*ij}_{\pi_0}} \\ &\sim \sum_{i<j}^{n} E^{*ij}_{\pi_0}\log\left(\frac{p_{\boldsymbol{C_i}\boldsymbol{C_j}}(1-s_1)(1-s_2)}{1-p_{\boldsymbol{C_i}\boldsymbol{C_j}}(s_1+s_2-s_1s_2)}\right). \end{aligned} \tag{34}$$

Here the last line in Eqn. (34) holds since the log operator keeps the minimum $\mathbf{\Pi_0}$ invariant. Note that $G^*_{\pi_0}=(V,E_1\cup\pi_0(E_1))$. Then we can find that $E^{*ij}_{\mathbf{\Pi_0}}=0$ if and only if both $E_1^{ij}$ and $E_2^{ij}$ are equal to 0, and $E^{*ij}_{\mathbf{\Pi_0}}=1$ occurs in the following three conditions:

- $(i,j)\in E_1$ but $(i,j)\notin E_2$. Note that this condition also ensures that $(\pi_0(i),\pi_0(j))\in E_2$.
- $(i,j)\in E_2$ but $(i,j)\notin E_1$. Note that this condition also ensures that $(\pi_0(i),\pi_0(j))\notin E_2$.
- $(i,j)\in E_1$ and $(i,j)\in E_2$. Note that this condition also ensures that $(\pi_0(i),\pi_0(j))\in E_2$.

Synthesizing all the above conditions, we can express $E^{*ij}_{\pi_0}$ as

$$E^{*ij}_{\pi_0}=\frac{1}{2}(E_1^{ij}+E_2^{ij}+|\mathbb{1}\{(i,j)\in E_1\}-\mathbb{1}\{(\pi_0(i),\pi_0(j))\in E_2\}|), \tag{35}$$

where $\mathbb{1}\{P\}=1$ if the random event $P$ happens and $\mathbb{1}\{P\}=0$ otherwise. Substituting Eqn. (35) into the last line in Eqn. (34), we get

$$\begin{aligned} &\arg\min_{\mathbf{\Pi}\in\mathbf{\Pi^n}}\sum_{i<j}^{n}E^{*ij}_{\pi_0}\log\left(\frac{p_{\boldsymbol{C_i}\boldsymbol{C_j}}(1-s_1)(1-s_2)}{1-p_{\boldsymbol{C_i}\boldsymbol{C_j}}(s_1+s_2-s_1s_2)}\right) \\ &= \arg\max_{\mathbf{\Pi}\in\mathbf{\Pi^n}}\sum_{i<j}^{n}w_{ij}|\mathbb{1}\{(i,j)\in E_1\}-\mathbb{1}\{(\pi_0(i),\pi_0(j))\in E_2\}| \\ &= \arg\max_{\mathbf{\Pi}\in\mathbf{\Pi^n}}||\mathbf{W}\circ(\mathbf{\Pi_0}\mathbf{A}-\mathbf{B}\mathbf{\Pi_0})||_F^2, \end{aligned} \tag{36}$$

[6] There is a notation abuse for $\sim$ between the one in Eqn. (**??**) and here.

where $w_{ij} = \log\left(\frac{1-p_{\boldsymbol{C_i C_j}}(s_1+s_2-s_1s_2)}{p_{\boldsymbol{C_i C_j}}(1-s_1)(1-s_2)}\right)$ is weight between nodes $i$ and $j$, $\mathbf{W}$ is the symmetric weight matrix where $\mathbf{W}(i,j) = \sqrt{w_{ij}} = \mathbf{W}(j,i)$, and "$\circ$" denotes the Hadamard product.

Substituting Eqn. (36) into Eqn. (31), now we can formulate the MMSE estimator as

$$\hat{\mathbf{\Pi}} = \arg\max_{\mathbf{\Pi}\in\mathbf{\Pi^n}} \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi}-\mathbf{\Pi_0}||_F^2 ||\mathbf{W}\circ(\mathbf{\Pi_0 A}-\mathbf{B\Pi_0})||_F^2. \tag{37}$$

□

### B. PROOF OF PROPOSITION 1

*Proof.* **Reduction from 1-median problem:** Our construction of the clique works as follows: Suppose there are $n$ nodes in $G_1$ and $G_2$. Then for any permutation matrix $\mathbf{\Pi}\in\mathbf{\Pi^n}$, we have

$$\begin{aligned}\hat{\mathbf{\Pi}} &= \arg\max_{\mathbf{\Pi}\in\Pi^n} \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi}-\mathbf{\Pi_0}||_F^2 ||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 \\ &= \arg\min_{\mathbf{\Pi}\in\Pi^n} \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} (4n-||\mathbf{\Pi}-\mathbf{\Pi_0}||_F^2)||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2,\end{aligned}$$

in which all the multiplicative factors are nonnegative. Since the number of elements in $\mathbf{\Pi^n}$ is $n!$, then we construct a clique with $n!$ nodes, with every node representing an $n\times n$ permutation matrix. We set the distance between two nodes $i$ and $j$ as $D(i,j) = 4n - ||\mathbf{\Pi}(i)-\mathbf{\Pi}(j)||_F^2$. Note that this distance satisfies the triangular equality $D(i,k)+D(k,j)\geq D(i,j)$, which assures that the edge directly connecting nodes $i$ and $j$ has the minimum distance among all possible paths between them. So the shortest path length between nodes $i$ and $j$ is just the distance $D(i,j)$. We define the weight of node $i$ as $\omega(i) = ||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$ (Note that each $\mathbf{\Pi_0}$ is a node in the graph with $n!$ nodes). For ease of understanding, Fig. 8 illustrates the constructed clique with 5 nodes.

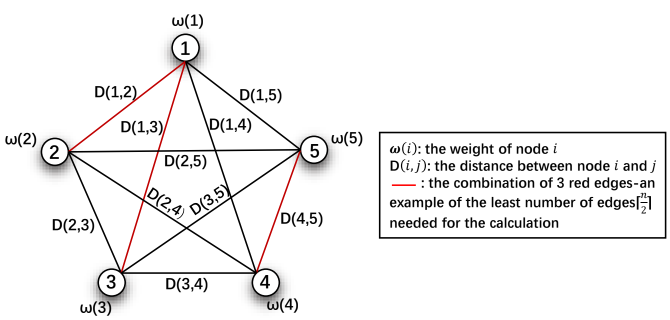



Fig. 8: An Illustration of the Constructed Clique with 5 Nodes

**The Lower Bound for 1-Median Problem:** Based on the above construction, we equivalently transform our problem into the form of a 1-median problem:

$$i_0 = \arg\min_{i^*\in V} \sum_{i\in V} \omega(i)\cdot D(i,i^*).$$

For a 1-median problem with $n$ nodes, it is easy to discover that we need to calculate at least $\lceil n/2\rceil$ times, since we need at least $\lceil n/2\rceil$ edges to form an edge set such that the endpoints of all edges in this edge set cover all the vertices in the graph. Or else one node will not be calculated for any edge connecting it, thus no information about this node is revealed, and then we can not judge whether this node is the one we intend to find. The red lines in Fig. 8 illustrates an example that when there are 5 nodes, the least number of edges needed to be calculated is $\lceil 5/2\rceil = 3$. For our MMSE estimator problem we have $n!$ nodes, thus the calculation times is at least $(n/2)!$, which means that we need to calculate $(n/2)!$ permutation matrices. Compared with the size of the problem, $n^2$, the complexity turns out to be $\Omega(((\sqrt{n})/2)!) = \Omega(\sqrt{n}!)$, which exceeds polynomial.

□

### C. ANALYSIS OF TRANSFORMATION FROM MMSE TO WEMP

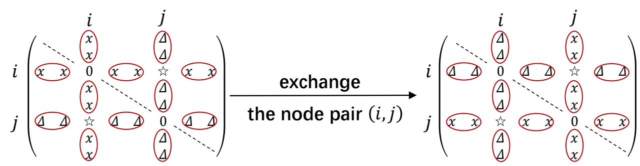



Fig. 9: An example of the effect of $\mathbf{\Pi_0}$ which differs from $\tilde{\mathbf{\Pi}}_\mathbf{0}$ only in the $i_{th}$ and $j_{th}$ row. The triangles denote the $j_{th}$ row and column the "x"es denote the $i_{th}$ row and column of $\mathbf{W}\circ(\mathbf{\Pi_0 A\Pi_0^T}-\mathbf{B})$. And the triangles denote the $i_{th}$ row and column the "x"es denote the $j_{th}$ row and column of $\mathbf{W}\circ(\tilde{\mathbf{\Pi}}\mathbf{A}\tilde{\mathbf{\Pi}}^\mathbf{T}-\mathbf{B})$. Note that the difference between $\mathbf{W}\circ(\mathbf{\Pi_0 A\Pi_0^T}-\mathbf{B})$ and $\mathbf{W}\circ(\tilde{\mathbf{\Pi}}\mathbf{A}\tilde{\mathbf{\Pi}}^\mathbf{T}-\mathbf{B})$ exists in the $i_{th}$ and $j_{th}$ row and column except the intersections (those 0s and stars).

1. Analysis of $||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$ where $\mathbf{\Pi_0}\in S_2(\mathbf{\Pi})$

Now we focus on the value of $||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$, where $\mathbf{\Pi_0}\in S_2(\mathbf{\Pi})$. Note that any permutation in $S_2(\mathbf{\Pi})$ only causes matching error on one pair of nodes. Thus if we consider $\mathbf{\Pi} = \tilde{\mathbf{\Pi}}$ and set one specific $\mathbf{\Pi_0}\in S_2(\tilde{\mathbf{\Pi}})$, which differs from $\tilde{\mathbf{\Pi}}$ only in the $i_{th}$ and $j_{th}$ row, we can derive that

$$\begin{aligned}&||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 - ||\tilde{\mathbf{\Pi}}\hat{\mathbf{A}}-\hat{\mathbf{B}}\tilde{\mathbf{\Pi}}||_F^2 \\ &= ||\mathbf{W}\circ(\mathbf{\Pi_0 A\Pi_0^T}-\mathbf{B})||_F^2 - ||\mathbf{W}\circ(\tilde{\mathbf{\Pi}}\mathbf{A}\tilde{\mathbf{\Pi}}^\mathbf{T}-\mathbf{B})||_F^2 \\ &= 2\Bigg(\sum_{k\neq i,j}^{n}[(\mathbf{W}\circ(\mathbf{\Pi_0 A\Pi_0^T}-\mathbf{B}))_{ik}^2 - (\mathbf{W}\circ(\tilde{\mathbf{\Pi}}\mathbf{A}\tilde{\mathbf{\Pi}}^\mathbf{T}-\mathbf{B}))_{ik}^2] \\ &\quad + \sum_{k\neq i,j}^{n}[(\mathbf{W}\circ(\mathbf{\Pi_0 A\Pi_0^T}-\mathbf{B}))_{jk}^2 - (\mathbf{W}\circ(\tilde{\mathbf{\Pi}}\mathbf{A}\tilde{\mathbf{\Pi}}^\mathbf{T}-\mathbf{B}))_{jk}^2]\Bigg) \\ &= 2\Bigg(\sum_{k\neq i,j}^{n} w_{ik}[(\mathbf{\Pi_0 A\Pi_0^T}-\mathbf{B})_{ik}^2 - (\tilde{\mathbf{\Pi}}\mathbf{A}\tilde{\mathbf{\Pi}}^\mathbf{T}-\mathbf{B})_{ik}^2] \\ &\quad + \sum_{k\neq i,j}^{n} w_{jk}[(\mathbf{\Pi_0 A\Pi_0^T}-\mathbf{B})_{jk}^2 - (\tilde{\mathbf{\Pi}}\mathbf{A}\tilde{\mathbf{\Pi}}^\mathbf{T}-\mathbf{B})_{jk}^2]\Bigg) \\ &= 2\Bigg(\sum_{k\neq i,j}^{n} w_{ik}[\mathbf{\Pi_0 A\Pi_0^T}-\tilde{\mathbf{\Pi}}\mathbf{A}\tilde{\mathbf{\Pi}}^\mathbf{T}]_{ik}\psi(\mathbf{B}_{ik}) \\ &\quad + \sum_{k\neq i,j}^{n} w_{jk}[\mathbf{\Pi_0 A\Pi_0^T}-\tilde{\mathbf{\Pi}}\mathbf{A}\tilde{\mathbf{\Pi}}^\mathbf{T}]_{jk}\psi(\mathbf{B}_{jk})\Bigg),\end{aligned} \tag{38}$$

where $\psi(x) = -1$ if $x = 1$ and $\psi(x) = 1$ if $x = 0$. Fig. 9 illustrates how Eqn. (38) can be derived intuitively. Note that if $\mathbf{\Pi_0}$ and $\tilde{\mathbf{\Pi}}$ are different only in the $i_{th}$ and $j_{th}$ rows, then the difference between $||\mathbf{W} \circ (\mathbf{\Pi_0 A \Pi_0^T} - \mathbf{B})||_F^2$ and $||\mathbf{W} \circ (\tilde{\mathbf{\Pi}} \mathbf{A} \tilde{\mathbf{\Pi}}^\mathbf{T} - \mathbf{B})||_F^2$ exists in the red circles in Fig. 9, which corresponds to the third line in Eqn. (38). Note that the intersection part, i.e., the stars in Fig. 9, does not contribute to the $||\mathbf{W} \circ (\mathbf{\Pi_0 A \Pi_0^T} - \mathbf{B})||_F^2$ and $||\mathbf{W} \circ (\tilde{\mathbf{\Pi}} \mathbf{A} \tilde{\mathbf{\Pi}}^\mathbf{T} - \mathbf{B})||_F^2$.

Note that since $\mathbf{\Pi_0}$ and $\tilde{\mathbf{\Pi}}$ are different in the $i_{th}$ and $j_{th}$ rows, then $(\tilde{\mathbf{\Pi}} \mathbf{A} \tilde{\mathbf{\Pi}}^\mathbf{T})_{ik} = (\mathbf{\Pi_0 A \Pi_0^T})_{jk}$. Therefore

$$
\begin{aligned}
&||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 - ||\tilde{\mathbf{\Pi}}\hat{\mathbf{A}} - \hat{\mathbf{B}}\tilde{\mathbf{\Pi}}||_F^2 \\
&= 2\Big(\sum_{k\neq i,j}^{n} w_{ik}\psi(\mathbf{B}_{ik})([\mathbf{\Pi_0 A \Pi_0^T}]_{ik} - [\mathbf{\Pi_0 A \Pi_0^T}]_{jk}) \\
&\quad + \sum_{k\neq i,j}^{n} w_{jk}\psi(\mathbf{B}_{jk})([\mathbf{\Pi_0 A \Pi_0^T}]_{jk} - [\mathbf{\Pi_0 A \Pi_0^T}]_{ik})\Big) \\
&= 2\Big(\sum_{k\neq i,j}^{n} (w_{ik}\psi(\mathbf{B}_{ik}) - w_{jk}\psi(\mathbf{B}_{jk})) \\
&\quad \cdot [(\mathbf{\Pi_0 A \Pi_0^T})_{ki} - (\mathbf{\Pi_0 A \Pi_0^T})_{kj}]\Big),
\end{aligned} \tag{39}
$$

Since $G_1$ and $G_2$ are independently sampled from $G$, then $\mathbf{A}$ and $\mathbf{B}$ are independent. Thus we can first take the expectation of $\mathbf{B}$ on both sides of Eqn. (39). Note that the probability for the edge existence between nodes $i$ and $j$ in $\mathbf{B}$ is $p_{\boldsymbol{C_i C_j}} s_2$, therefore $\mathbf{E}[\psi(B_{ij})] = (-1)p_{\boldsymbol{C_i C_j}} s_2 + (1 - p_{\boldsymbol{C_i C_j}} s_2) = 1 - 2p_{\boldsymbol{C_i C_j}} s_2$. Hence, taking the expectation of $\mathbf{B}$ on both sides of Eqn. (39) and we obtain

$$
\begin{aligned}
&\mathbf{E_B}(||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 - ||\tilde{\mathbf{\Pi}}\hat{\mathbf{A}} - \hat{\mathbf{B}}\tilde{\mathbf{\Pi}}||_F^2) \\
&= 2\sum_{k\neq i,j}^{n} [w_{ik}[(1 - 2p_{\boldsymbol{C_i C_k}} s_2) - w_{jk}(1 - 2p_{\boldsymbol{C_j C_k}} s_2)) \\
&\quad \cdot ((\mathbf{\Pi_0 A \Pi_0^T})_{ki} - (\mathbf{\Pi_0 A \Pi_0^T})_{kj})].
\end{aligned}
$$

Similarly, taking the expectation of $\mathbf{A}$ on both sides, we have

$$
\begin{aligned}
&\mathbf{E_{A,B}}(||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 - ||\tilde{\mathbf{\Pi}}\hat{\mathbf{A}} - \hat{\mathbf{B}}\tilde{\mathbf{\Pi}}||_F^2) \\
&= 2\sum_{k\neq i,j}^{n} (w_{ik}(1 - 2p_{\boldsymbol{C_i C_k}} s_2) - w_{jk}(1 - 2p_{\boldsymbol{C_j C_k}} s_2)) \\
&\quad \cdot (p_{\boldsymbol{C_{\pi_0(i)} C_{\pi_0(k)}}} - p_{\boldsymbol{C_{\pi_0(j)} C_{\pi_0(k)}}}) s_1 \\
&= 2\sum_{k\neq i,j}^{n} \Delta_{i,j,k,\pi_0},
\end{aligned}
$$

where

$$
\begin{aligned}
\Delta_{i,j,k,\pi_0} = &(w_{ik}(1 - 2p_{\boldsymbol{C_i C_k}} s_2) - w_{jk}(1 - 2p_{\boldsymbol{C_j C_k}} s_2)) \\
&\cdot (p_{\boldsymbol{C_{\pi_0(i)} C_{\pi_0(k)}}} - p_{\boldsymbol{C_{\pi_0(j)} C_{\pi_0(k)}}}) s_1.
\end{aligned}
$$

$\Delta_{i,j,k,\pi_0}$ reflects a part of the difference $||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 - ||\tilde{\mathbf{\Pi}}\hat{\mathbf{A}} - \hat{\mathbf{B}}\tilde{\mathbf{\Pi}}||_F^2$ caused by the difference of a single element in matrices $\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}$ and $\tilde{\mathbf{\Pi}}\hat{\mathbf{A}} - \hat{\mathbf{B}}\tilde{\mathbf{\Pi}}$.[7] Since we consider the average case of all possible $\mathbf{\Pi_0}$, we also consider the average value of $\Delta_{i,j,\pi_0}$, which we set as $\hat{\Delta} = \mathbf{E}_{i,j,\pi_0}(\Delta_{i,j,\pi_0})$. Note that $\mathbf{E_{A,B}}(||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 - ||\tilde{\mathbf{\Pi}}\hat{\mathbf{A}} - \hat{\mathbf{B}}\tilde{\mathbf{\Pi}}||_F^2) > 0$ since $\tilde{\mathbf{\Pi}}$ is the minimizer of $||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$. Therefore $\hat{\Delta} = \mathbf{E}_{i,j,\pi_0}(\Delta_{i,j,\pi_0}) > 0$.

[7]For example, the difference of the corresponding element (with the same notation, e.g., (i,k) in the left matrix and (j,k) in the right matrix, both of which are triangles.) in two matrices in Fig. 9 inside one of the red circles

2. Analysis of $\sum_{\mathbf{\Pi_0} \in S_k(\mathbf{\Pi})} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$

Now we move to the second part involved in our idea. We first focus on $S_k(\mathbf{\Pi_0})$, and count the number of elements in $S_k(\mathbf{\Pi_0})$, denoted as $|S_k|$. Note that if there are $k$ mismatched nodes in a graph with $n$ nodes, there are $C_n^k$ possible sets of mismatched nodes. We define $|T_k|$ as the number of elements in each possible set, and can get $|S_k| = C_n^k |T_k|$. For $|T_k|$, we can find that it satisfies

$$
\begin{aligned}
|T_k| &= (k-1)(|T_{k-2}| + (k-2)(|T_{k-3}| + (k-3)(|T_{k-4}| + ...))) \\
&= \sum_{t=1}^{k-1}(\prod_{i=1}^{t}(k-i))|T_{k-t-1}|.
\end{aligned} \tag{40}
$$

Consider $|T_k|$ and $|T_{k-1}|$ in Eqn. (40), we can discover that

$$
|T_k| = (k-1)(|T_{k-2}| + |T_{k-1}|) \geq (k-1)|T_{k-1}|, k \geq 2.
$$

Therefore we obtain the relationship between $|S_k|$ and $|S_{k-1}|$ as

$$
|S_k| = C_n^k |T_k| \geq (k-1)\frac{C_n^k}{C_n^{k-1}}|S_{k-1}| = (1-\frac{1}{k})(n-k+1)|S_{k-1}|, \tag{41}
$$

where $k \geq 2$. Eqn. (41) shows that when $k$ is much smaller than $n$, then $\frac{|S_k|}{|S_{k-1}|} = (1-\frac{1}{k})(n-k+1)$ is large; when $k$ gets close to $n$, then $\frac{|S_k|}{|S_{k-1}|}$ approaches $1$, which means that $|S_k|$ and $|S_{k-1}|$ are almost the same.

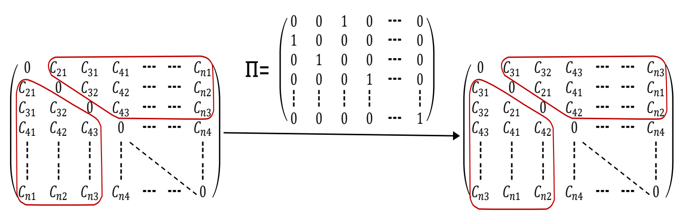


Fig. 10: An example of the effect of $\mathbf{\Pi} \in S_3(\tilde{\mathbf{\Pi}})$, where we set $\tilde{\mathbf{\Pi}} = \mathbf{I}$. $\mathbf{I}$ is the identity matrix. Note that under the $\mathbf{\Pi}$ above the arrow, which differs from $\mathbf{I}$ only in the first three rows (columns). Thus the possible difference between two matrices only exists in the red circles, with $6n - 6$ elements in the matrix involved.

Now we consider $\mathbf{\Pi_0} \in S_k(\mathbf{\Pi})$. Note that for any $\mathbf{\Pi_0} \in S_k(\mathbf{\Pi})$, there are $k$ rows and columns that may cause the difference between $||\mathbf{W} \circ (\mathbf{\Pi_0 A \Pi_0^T} - \mathbf{B})||_F^2$ and $||\mathbf{W} \circ (\tilde{\mathbf{\Pi}} \mathbf{A} \tilde{\mathbf{\Pi}}^\mathbf{T} - \mathbf{B})||_F^2$. Fig. 10 illustrates an example of $\mathbf{\Pi_0} \in S_3(\mathbf{\Pi})$. Therefore we can discover for any $\mathbf{\Pi_0} \in S_k(\mathbf{\Pi})$, the number of node pairs $(i,j)$ which may influence the difference between $||\mathbf{W} \circ (\mathbf{\Pi_0 A \Pi_0^T} - \mathbf{B})||_F^2$ and $||\mathbf{W} \circ (\tilde{\mathbf{\Pi}} \mathbf{A} \tilde{\mathbf{\Pi}}^\mathbf{T} - \mathbf{B})||_F^2$ is approximately $\sum_{i=1}(n-i) = \frac{(2n-k-1)k}{2}$.[8] Thus, denoting

[8]For example, in Fig. 10 when $k = 3$ the number is $6n - 6$. Although there may be some elements which do not cause error, such as the two stars in Fig. 9, the number of this kinds of node pairs can be neglected when $n$ is large enough.

$N_k$ as this number of node pair, we can obtain

$$\begin{aligned}N_k &= \frac{(2n-k-1)k}{2}|S_k| \\ &\geq \frac{(2n-k-1)k}{2}(1-\frac{1}{k})(n-k+1)|S_{k-1}| \\ &= (1-\frac{1}{k})(n-k+1)\frac{(2n-k-1)k}{(2n-k)(k-1)}N_{k-1} \\ &= (n-k+1)\frac{2n-k-1}{2n-k}N_{k-1}.\end{aligned}$$

Therefore in average, we have

$$\begin{aligned}\sum_{\mathbf{\Pi_0}\in S_k} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 &= N_k\hat{\Delta} \\ &\geq (n-k+1)\frac{2n-k-1}{2n-k}N_{k-1}\hat{\Delta} \\ &\geq (n-k+1)\frac{2n-k-1}{2n-k}\sum_{\mathbf{\Pi_0}\in S_{k-1}} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 \\ &\approx (n-k+1)\sum_{\mathbf{\Pi_0}\in S_{k-1}} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2,\end{aligned} \tag{42}$$

where the last approximation holds since $k \leq n$ and when $n \to \infty$, $\frac{2n-k-1}{2n-k} \to 1$.

Therefore, we can claim that in average, if $k_1 > k_2$, then

$$\sum_{\mathbf{\Pi_0}\in S_{k_1}} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 > \sum_{\mathbf{\Pi_0}\in S_{k_2}} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2. \tag{43}$$

3. Maximum Value Under Sequence Inequality

In average case, by setting $\mathbf{\Pi}$ in the original MMSE objective function

$$\sum_{\mathbf{\Pi_0}\in \Pi^n} ||\mathbf{\Pi} - \mathbf{\Pi_0}||_F^2 ||\mathbf{W} \circ (\mathbf{\Pi_0 A} - \mathbf{B\Pi_0})||_F^2$$

equal to $\tilde{\mathbf{\Pi}}$, the minimizer of WEMP, then this original MMSE objective function reaches its largest value under Sequence Inequality.

Moreover, note that if we do not set $\tilde{\mathbf{\Pi}} = \hat{\mathbf{\Pi}}$, for example set $\tilde{\mathbf{\Pi}} = \mathbf{\Pi} \in S_k(\mathbf{\Pi})$, we can verify that $\mathbf{\Pi}$ does not make the objective function in Eqn. (9) larger than $\mathbf{\Pi_0}$ since

$$\begin{aligned}&0||\mathbf{\Pi}\hat{\mathbf{A}}\mathbf{\Pi^T} - \hat{\mathbf{B}}||_{\mathbf{F}}^{\mathbf{2}} + 2k||\mathbf{\Pi}\hat{\mathbf{A}}\mathbf{\Pi^T} - \hat{\mathbf{B}}||_{\mathbf{F}}^{\mathbf{2}} \\ &\geq 2k||\hat{\mathbf{\Pi}}\hat{\mathbf{A}}\hat{\mathbf{\Pi}}^{\mathbf{T}} - \hat{\mathbf{B}}||_{\mathbf{F}}^{\mathbf{2}} + 0||\mathbf{\Pi}\hat{\mathbf{A}}\mathbf{\Pi^T} - \hat{\mathbf{B}}||_{\mathbf{F}}^{\mathbf{2}},\end{aligned}$$

which means that the Sequency Inequality preserves that when $||\mathbf{\Pi_0}\hat{\mathbf{A}}\mathbf{\Pi_0^T} - \hat{\mathbf{B}}||_F^2$ achieves its minimum, then $||\Pi - \Pi_0||_F^2$ also achieves its minimum. Therefore by setting $\tilde{\mathbf{\Pi}} = \hat{\mathbf{\Pi}}$ we can achieve the largest value of the original MMSE problem under this sequence inequality.

## D. PROOF OF THEOREM 2

*Proof.*

$$\begin{aligned}&g(\hat{\mathbf{\Pi}}) - g(\tilde{\mathbf{\Pi}}) \\ &= \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} (||\hat{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2 - ||\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2)||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2.\end{aligned} \tag{44}$$

Then we divide the set $\mathbf{\Pi^n}$ into two subsets:

$$\mathbf{\Pi_1^n} = \{\mathbf{\Pi} \in \mathbf{\Pi^n} | ||\hat{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2 > ||\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2\};$$
$$\mathbf{\Pi_2^n} = \{\mathbf{\Pi} \in \mathbf{\Pi^n} | ||\hat{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2 < ||\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2\}.$$

Following that we divide the Eqn. (44) into two sets, $\mathbf{\Pi_1^n}$ and $\mathbf{\Pi_2^n}$:

$$\begin{aligned}g(\hat{\mathbf{\Pi}}) - g(\tilde{\mathbf{\Pi}}) &= \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi_1^n}} (||\hat{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2 - ||\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2)||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 \\ &- \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi_2^n}} (||\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2 - ||\hat{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2)||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 \\ &\leq ||\tilde{\mathbf{\Pi}} - \hat{\mathbf{\Pi}}||_F^2 \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi_1^n}} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2.\end{aligned} \tag{45}$$

where the last inequality holds due to the triangular inequality $||\hat{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2 - ||\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2 \leq ||\tilde{\mathbf{\Pi}} - \hat{\mathbf{\Pi}}||_F^2$ and the term $\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi_2^n}}(||\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2 - ||\hat{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2)||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$ is positive. Then we have

$$\begin{aligned}\frac{g(\hat{\mathbf{\Pi}}) - g(\tilde{\mathbf{\Pi}})}{g(\tilde{\mathbf{\Pi}})} &= \frac{(||\tilde{\mathbf{\Pi}} - \hat{\mathbf{\Pi}}||_F^2)\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi_1^n}} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2}{\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2 ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2} \\ &\leq \frac{2\beta n \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi_0}\tilde{\mathbf{A}} - \tilde{\mathbf{B}}\mathbf{\Pi_0}||_F^2}{\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\tilde{\mathbf{\Pi}} - \mathbf{\Pi_0}||_F^2 ||\mathbf{\Pi_0}\tilde{\mathbf{A}} - \tilde{\mathbf{B}}\mathbf{\Pi_0}||_F^2}.\end{aligned} \tag{46}$$

where $||\tilde{\mathbf{\Pi}} - \hat{\mathbf{\Pi}}||_F^2 = 2\beta n$ and $\beta \in [0,1]$ is the ratio between the number of mistakenly matched nodes and that of all the nodes. The last inequality in (46) holds because $\mathbf{\Pi_1^n} \subset \mathbf{\Pi^n}$.

Now we divide the sum $\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$ into two parts:

$$D_1 = \sum_{k\leq \rho n}\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2;$$
$$D_2 = \sum_{\rho n < k \leq n}\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2.$$

where $\rho$ is any real number in $[0,1]$ and we assume that $\rho n$ is an integer[9].

For $D_1$, in average case we can obtain

$$\begin{aligned}D_1 &\leq \sum_{i=1}^{\rho n}\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}} ||\mathbf{\Pi_0}\hat{\mathbf{A}} - \hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 \leq \sum_{i=1}^{\rho n}\prod_{j=1}^{i} 2(n-j+1) \\ &\leq \sum_{i=1}^{\rho n}(2n)^i = 2n\frac{(2n)^{\rho n}-1}{2n-1} \approx (2n)^{\rho n}.\end{aligned}$$

For $D_2$, according to Inequality (42), in average case we can get

$$\begin{aligned}D_2 &\geq \sum_{k=\rho n+1}^{n}\prod_{j=1}^{k}(n-j+1) = \sum_{k=\rho n+1}^{n}\frac{n!}{(n-k)!} \\ &\geq \sum_{k=\rho n+1}^{n}\frac{n!}{((1-\rho)n)!} = (1-\rho)n\frac{n!}{((1-\rho)n)!}.\end{aligned}$$

Note that if we set $\rho = \Omega(1) = c_0$, where $c_0 \to 1$, then $\rho \to 1$ and

$$D_2 \geq c_0\frac{n!}{c_0!} = cn! \sim c\sqrt{2\pi n}(\frac{n}{e})^n,$$

[9] If it is not an integer, we can easily modify it by rounding.

where $c$ is a constant and the last step holds due to the Stirling's formula. Therefore we can upper bound $\frac{D_2}{D_1}$ as

$$\frac{D_2}{D_1} \geq c\frac{\sqrt{2\pi n}(\frac{n}{e})^n}{(2n)^{\rho n}} = c\sqrt{2\pi n}\left(\frac{n^{1-\rho}}{2^\rho e}\right)^n.$$

Then if $\rho$ is a constant which approaches 1 but does not equal to 1, then we find that when $n \to \infty$, $D_2$ is of higher order of $n$ than $D_1$. Therefore we can easily verify that in the denominator of the last term in Inequality (46), $\sum_{\rho n<k\leq n}\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}}||\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0}||_F^2||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$ is of higher order of $n$ than $\sum_{k\leq\rho n}\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}}||\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0}||_F^2||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$, since for $k_1 > \rho n$ and $k_2 < \rho n$, $\mathbf{\Pi_1'} \in S_{k_1}(\tilde{\mathbf{\Pi}})$ and $\mathbf{\Pi_2'} \in S_{k_2}(\tilde{\mathbf{\Pi}})$, we have $||\mathbf{\Pi_1'}-\tilde{\mathbf{\Pi}}||_F^2 \geq ||\mathbf{\Pi_2'}-\tilde{\mathbf{\Pi}}||_F^2$. Therefore according to Lemma **??**, we can leave the term with highest order of $n$ in the denominator and numerator in the last term in Inequality (46) when $n \to \infty$ and thus we can obtain

$$\begin{aligned}
&\frac{2\beta n\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}}||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2}{\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}}||\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0}||_F^2||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2} \\
&\approx \frac{2\beta n\sum_{\rho n<k\leq n}\sum_{\mathbf{\Pi_0}\in S_k(\tilde{\mathbf{\Pi}})}||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2}{\sum_{\rho n<k\leq n}\sum_{\mathbf{\Pi_0}\in S_k(\tilde{\mathbf{\Pi}})}||\mathbf{\Pi_0}-\tilde{\mathbf{\Pi}}||_F^2||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2} \\
&\leq \frac{2\beta n}{2\rho n}\frac{\sum_{\rho n<k\leq n}\sum_{\mathbf{\Pi_0}\in S_k(\tilde{\mathbf{\Pi}})}||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2}{\sum_{\rho n<k\leq n}\sum_{\mathbf{\Pi_0}\in S_k(\tilde{\mathbf{\Pi}})}||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2} = \frac{\beta}{\rho}.
\end{aligned}$$

Thus we have the approximation ratio

$$\frac{g(\tilde{\mathbf{\Pi}})}{g(\hat{\mathbf{\Pi}})} \geq \frac{1}{1+\frac{\beta}{\rho}} \approx \frac{1}{1+\beta} \geq \frac{1}{2}.$$

□

Note that in the proof of Theorem 2, we use several times of inequality scaling method to derive the lower bound of approximation ratio, which is 0.5. These inequality scaling may cause this lower bound to be smaller than the real approximation ratio. That is to say, the approximation ratio 0.5 may be even worse than the approximation ratio in the worst case in real situations. For example, in Inequality (46) we directly use

$$\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi_1^n}}||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 \leq \sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}}||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2,$$

which may cause a big gap. Therefore, for a more general situation we have the following corollary.

**Corollary 1.** *Given the published graph $G_1$, the auxiliary graph $G_2$, the parameter set $\boldsymbol{\theta}$ and the weight matrix $\mathbf{W}$, and we let*

$$\chi = \left(\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi_1^n}}||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2\right)\Big/\left(\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi^n}}||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2\right),$$

*then in average case, the approximation $g(\tilde{\mathbf{\Pi}})/g(\hat{\mathbf{\Pi}})$ ratio is larger than $\frac{1}{1+\beta\chi}$.*

This corollary can be easily proved by slightly changing the form of Eqn. (46). To take an example to illustrate the gap of approximation ratio caused by $\chi$ more intuitively, we assume that $\sum_{\mathbf{\Pi}\in\mathbf{\Pi_1^n}}||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2 = \sum_{\mathbf{\Pi}\in\mathbf{\Pi_2^n}}||\mathbf{\Pi_0}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi_0}||_F^2$ [10]. Then $\chi = \frac{1}{2}$ and $(\frac{1}{1+\beta\chi}) > \frac{2}{3}$, which causes the gap of the lower bound of approximation ratio to be $\frac{2}{3}-\frac{1}{2}=\frac{1}{6}$.

Note that we still claim that the approximation ratio is *larger* than $(\frac{1}{1+\beta\chi})$. This is because we eliminate the sum $\sum_{\mathbf{\Pi_0}\in\mathbf{\Pi_2^n}}(||\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0}||_F^2 - ||\tilde{\mathbf{\Pi}}-\mathbf{\Pi_0}||_F^2)||\mathbf{\Pi}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi}||_F^2$ in Eqn. (45), which also generates a gap between the lower bound $\frac{1}{1+\beta\chi}$ and the real approximation ratio. We leave it a future direction to find a proper estimation of this gap. However, the current gap still ensures the real approximation ratio strictly larger than $\frac{1}{1+\beta\chi}$, which further strengthens our claim at the beginning of Section IV-C that the transformation of the original MMSE problem is valid.

### E. PROOF OF LEMMA 1

*Proof.* We know $||\mathbf{\Pi}\hat{\mathbf{A}}-\hat{\mathbf{B}}\mathbf{\Pi}||_F = ||\mathbf{W}\circ(\mathbf{\Pi}\mathbf{A}-\mathbf{B}\mathbf{\Pi})||_F$, thus we only need to prove that

$$\mathbf{W}\circ\mathbf{\Pi}\mathbf{A} = \mathbf{\Pi}\mathbf{W}\circ\mathbf{A}$$

. Note that $w_{ij}$ only depends on $p_{\boldsymbol{C_i}\boldsymbol{C_j}}$, $s_1$ and $s_2$, therefore for some nodes $i,j,s,t$, if $\boldsymbol{C_i} = \boldsymbol{C_s}$ and $\boldsymbol{C_j} = \boldsymbol{C_t}$, then $\mathbf{W}(i,j) = \mathbf{W}(s,t)$. This fact tells that the weight is invariant within communities. Therefore, since $\mathbf{\Pi}$ keeps invariant of the community representation of all the nodes, it is easy to verify that $\mathbf{W}\circ\mathbf{\Pi}\mathbf{A} = \mathbf{\Pi}\mathbf{W}\circ\mathbf{A}$. Thus we have $\hat{\mathbf{A}} = \mathbf{W}\circ\mathbf{A}$ and similarly, $\hat{\mathbf{B}} = \mathbf{W}\circ\mathbf{B}$. Then the lemma holds naturally. □

### F. PROOF OF LEMMA 2

*Proof.* First we verify that $F_1(\mathbf{\Pi})$ is a convex function. One of the sufficient and necessary condition for a function whose variable is matrix is convex is that the Hessian matrix of this function is positive semi-definite. The Hessian matrix of $F(\mathbf{\Pi})$ can be obtained by taking the second derivative over $\mathbf{\Pi}$ on $F(\mathbf{\Pi})$, we denote it as $\nabla^2 F(\mathbf{\Pi})$. Therefore we can obtain the Hessian matrix of $F_1(\mathbf{\Pi})$ by

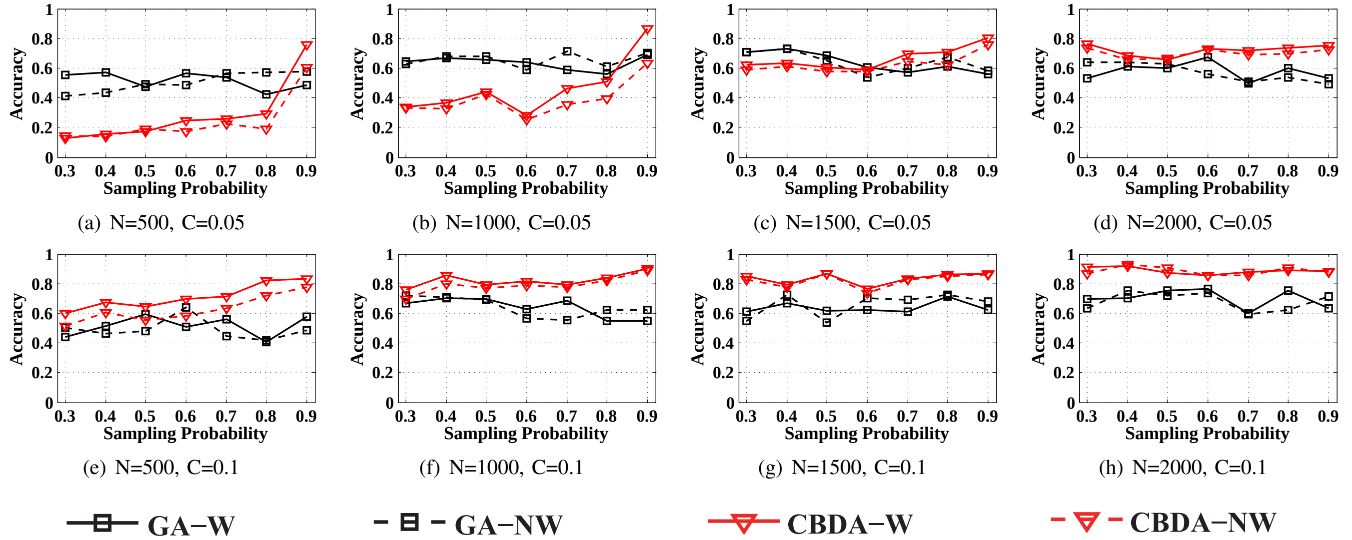
$$\nabla^2 F_1(\mathbf{\Pi}) = \nabla^2 F_0(\mathbf{\Pi}) - \lambda_{min}\mathbf{I}.$$

where $\mathbf{I}$ is the identity matrix[11]. Note that $\lambda_{min}$ is the minimum eigenvalue of $\nabla^2 F_0(\mathbf{\Pi})$, therefore all the eigenvalues of $\nabla^2 F_0(\mathbf{\Pi}) - \lambda_{min}\mathbf{I}$ are equal to or larger than 0. Hence $\nabla^2 F_1(\mathbf{\Pi})$ is a nonnegative definite matrix and $F_1(\mathbf{\Pi})$ is a convex function.

Meanwhile, one of the sufficient and necessary conditions for a function whose variable is matrix is concave is that the Hessian matrix of this function is negative semi-definite. Similar to the analysis of $F_1(\mathbf{\Pi})$, we can verify that $F_2(\mathbf{\Pi})$ is a concave function. Thus we complete the proof. □

[10]This is only a very special situation, which we use it to make an intuitive example to explain how $\chi$ causes the gap of approximation ratio. It is not necessarily the same as real situations

[11]The identity matrix $I$ means all the elements on the diagonal of $I$ are all 1s while others are all 0s. Note that here $I$ is an $n^2\times n^2$ matrix since the first order derivative of a function whose variable is a matrix is a $n\times n$ matrix, thus the second derivative of $F_0$ ($F_1$) is $n^2\times n^2$ matrix.




(a) N=500, C=0.05 (b) N=1000, C=0.05 (c) N=1500, C=0.05 (d) N=2000, C=0.05

(e) N=500, C=0.1 (f) N=1000, C=0.1 (g) N=1500, C=0.1 (h) N=2000, C=0.1

Fig. 11: Experiments on Weighted and Non-weighted Cost Function.

## G. Proof of Lemma 3

For the inner loop part,we focus on $F_\xi(\mathbf{\Pi_{k+1}})$ and $F_\xi(\mathbf{\Pi}^\xi)$. Since $\mathbf{\Pi_{k+1}} = \mathbf{\Pi_k} + \gamma_k(\mathbf{X}_\perp - \mathbf{\Pi_k})$, according to Taylor's Theorem,

$$\begin{aligned}F_\xi(\mathbf{\Pi_{k+1}}) &= F_\xi(\mathbf{\Pi_k} + \gamma_k(\mathbf{X}_\perp - \mathbf{\Pi_k}))\\ &= F_\xi(\mathbf{\Pi_k}) + \gamma_k \mathbf{tr}(\nabla F_\xi^T(\mathbf{\Pi_k})(\mathbf{X}^\perp - \mathbf{\Pi_k})) + \gamma_k \mathbf{R_k}\\ &\le F_\xi(\mathbf{\Pi_k}) + \gamma_k \mathbf{tr}(\nabla F_\xi^T(\mathbf{\Pi_k})(\mathbf{\Pi}^\xi - \mathbf{\Pi_k})) + \gamma_k \mathbf{R_k},\end{aligned} \tag{47}$$

where $\gamma_k \mathbf{R_k}$ is the remainder of this Taylor series, and this form makes sense since the remainder must contain a multiplicative factor of $\gamma_k$. The last inequality holds since $\mathbf{X}^\perp$ is the minimizer of $\mathbf{tr}(\nabla F_\xi^T(\mathbf{\Pi_k})(\mathbf{\Pi}^\xi - \mathbf{\Pi_k}))$.

In terms of $F_\xi(\mathbf{\Pi}^\xi)$, we have

$$\begin{aligned}&F_\xi(\mathbf{\Pi}^\xi) = F_\xi(\mathbf{\Pi_k} + \mathbf{\Pi}^\xi - \mathbf{\Pi_k})\\ &= F_\xi(\mathbf{\Pi_k}) + \mathbf{tr}(\nabla F_\xi^T(\mathbf{\Pi_k})(\mathbf{\Pi}^\xi - \mathbf{\Pi_k})) + \mathbf{R'_k},\end{aligned} \tag{48}$$

where $\mathbf{R'_k}$ is the remainder of this Taylor series.

Combining Eqn. (47) and (48), we can obtain

$$F_\xi(\mathbf{\Pi_{k+1}}) \le F_\xi(\mathbf{\Pi_k}) + \gamma_k(F_\xi(\mathbf{\Pi}^\xi) - F_\xi(\mathbf{\Pi_k})) + \gamma_k(\mathbf{R_k} - \mathbf{R'_k}). \tag{49}$$

Denote $\Delta\mathbf{R_k} = \mathbf{R_k} - \mathbf{R'_k}$ and by simple transformation of Inequality (49), we obtain

$$F_\xi(\mathbf{\Pi_{k+1}}) - F_\xi(\mathbf{\Pi}^\xi) \le (1-\gamma_k)(F_\xi(\mathbf{\Pi_k}) - F_\xi(\mathbf{\Pi}^\xi)) + \gamma_k \Delta\mathbf{R_k}. \tag{50}$$

Note that Inequality (50) builds up the relationship between $F_\xi(\mathbf{\Pi_{k+1}})$ and $F_\xi(\mathbf{\Pi_k})$, and we obtain

$$\begin{aligned}&F_\xi(\mathbf{\Pi_{k+1}}) - F_\xi(\mathbf{\Pi}^\xi)\\ &\le \prod_{i=1}^{k}(1-\gamma_i)(F_\xi(\mathbf{\Pi_1}) - F_\xi(\mathbf{\Pi}^\xi)) + \sum_{i=1}^{k}\gamma_i \prod_{j=1}^{k-i}(1-\gamma_j)\Delta\mathbf{R}_i.\end{aligned} \tag{51}$$

For $F_\xi(\mathbf{\Pi_1}) - F_\xi(\mathbf{\Pi}^\xi)$, note that $\mathbf{\Pi_1} = \mathbf{\Pi}^{\xi - \mathbf{\Delta}\xi}$, then

$$\begin{aligned}F_\xi(\mathbf{\Pi}^\xi) &= F_0(\mathbf{\Pi}^\xi) + \xi(n - ||\mathbf{\Pi}^\xi||_F^2)\\ &= F_0(\mathbf{\Pi}^\xi) + (\xi - \Delta\xi)(n - ||\mathbf{\Pi}^\xi||_F^2) - \Delta\xi(n - ||\mathbf{\Pi}^\xi||_F^2)\\ &\ge F_0(\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}) + (\xi - \Delta\xi)(n - ||\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}||_F^2)\\ &\quad - \Delta\xi(n - ||\mathbf{\Pi}^\xi||_F^2)\\ &= F_0(\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}) + \xi(n - ||\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}||_F^2)\\ &\quad + \Delta\xi(||\mathbf{\Pi}^\xi||_F^2 - ||\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}||_F^2)\\ &= F_\xi(\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}) + \Delta\xi(||\mathbf{\Pi}^\xi||_F^2 - ||\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}||_F^2).\end{aligned} \tag{52}$$

Hence

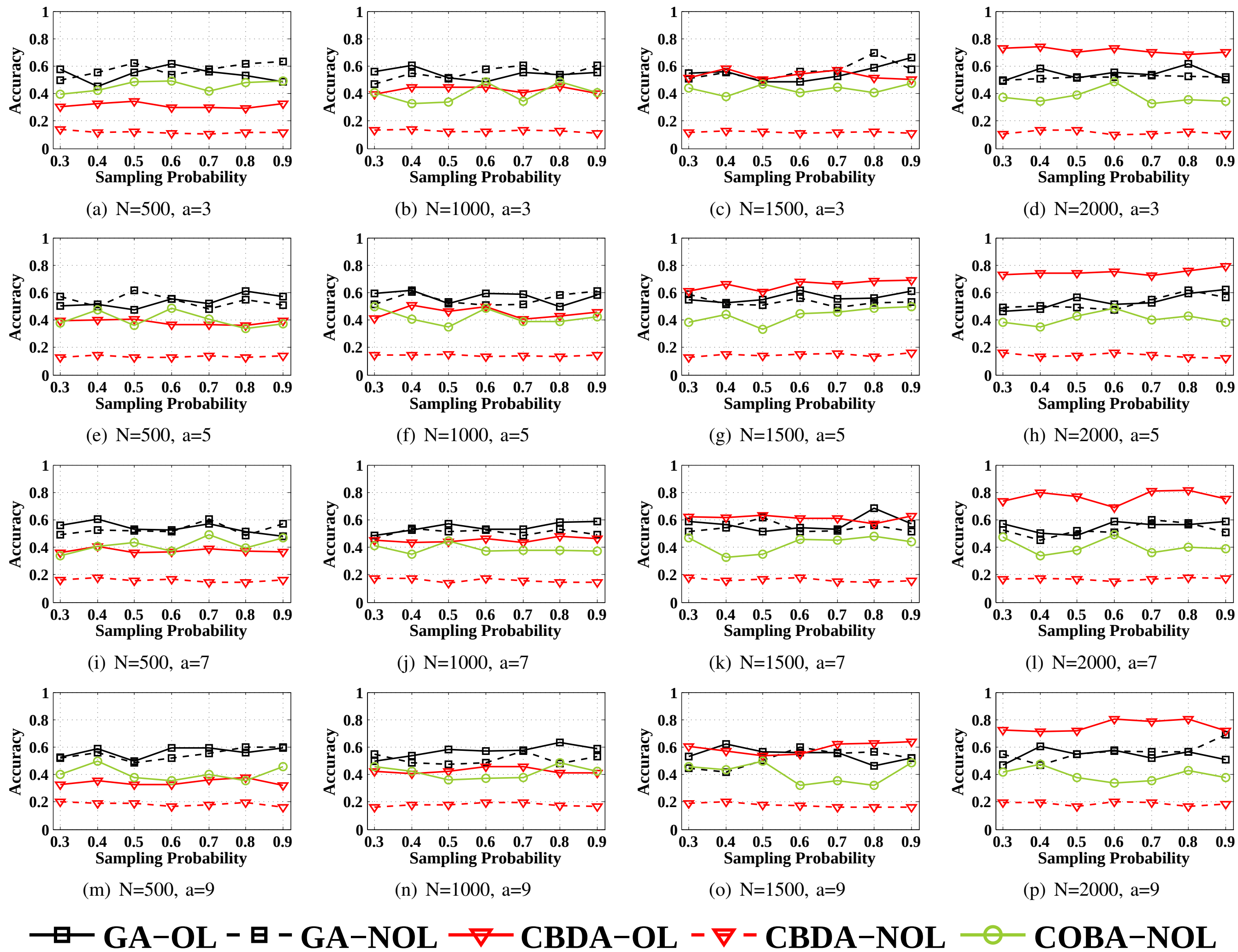
$$F_\xi(\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}) - F_\xi(\mathbf{\Pi}^\xi) \le \Delta\xi(||\mathbf{\Pi}^{\xi-\mathbf{\Delta}\xi}||_F^2 - ||\mathbf{\Pi}^\xi||_F^2). \tag{53}$$

Therefore by combining Inequalities (53) and (51), we can obtain if $\Delta\xi$ is small enough, or if $k \to \infty$, then the term $\prod_{i=1}^{k}(1-\gamma_i)(F_\xi(\mathbf{\Pi_1}) - F_\xi(\mathbf{\Pi}^\xi))$ in last expression of Inequality (51) goes to 0.

For the second term $\sum_{i=1}^{k}\gamma_i \prod_{j=1}^{k-i}(1-\gamma_j)\Delta\mathbf{R}_i$, we note that when $k \to \infty$, then $\forall\epsilon > 0, \exists K > 0, \delta_1 > 0$, when $i > K$, $\gamma_i \prod_{j=1}^{k-i}(1-\gamma_j) < \gamma_i < \frac{\epsilon}{2^{\delta_1 i}}$, and meanwhile when $i \le K$, $\gamma_k(1-\gamma_j) < \prod_{j=1}^{k-i}(1-\gamma_j) < \frac{\epsilon}{2^{\delta_2 i}}$. Setting $\delta^* = \min\{\delta_1, \delta_2\}$, then we can upper bound the sum $\sum_{i=1}^{k}\gamma_i \prod_{j=1}^{k-i}(1-\gamma_j)\Delta\mathbf{R}_i \le \sum_{i=1}^{\infty}\frac{\epsilon}{2^{\delta i}} = 0$. Therefore we prove that the inner loop converges.

## H. Supplementary Empirical Results

We provide Figs. 11, 12, 13 and 14 here to show more empirical evaluation of the effect of weight matrix $\mathbf{W}$ and results on both synthetic and sampled real social networks.



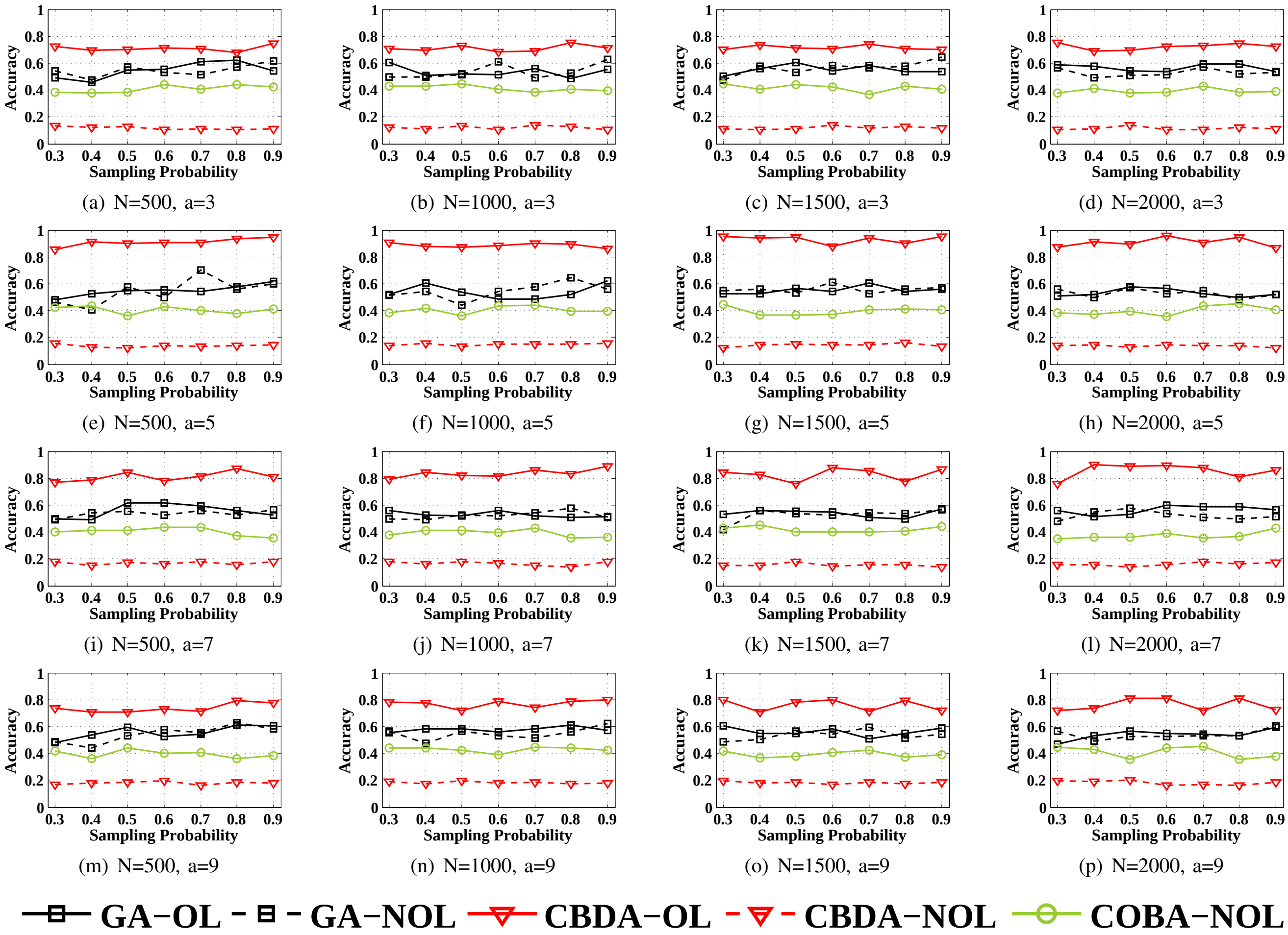

(a) N=500, a=3 (b) N=1000, a=3 (c) N=1500, a=3 (d) N=2000, a=3

(e) N=500, a=5 (f) N=1000, a=5 (g) N=1500, a=5 (h) N=2000, a=5

(i) N=500, a=7 (j) N=1000, a=7 (k) N=1500, a=7 (l) N=2000, a=7

(m) N=500, a=9 (n) N=1000, a=9 (o) N=1500, a=9 (p) N=2000, a=9

Fig. 12: Experiments on Synthetic Networks with $\eta = 0.05$.



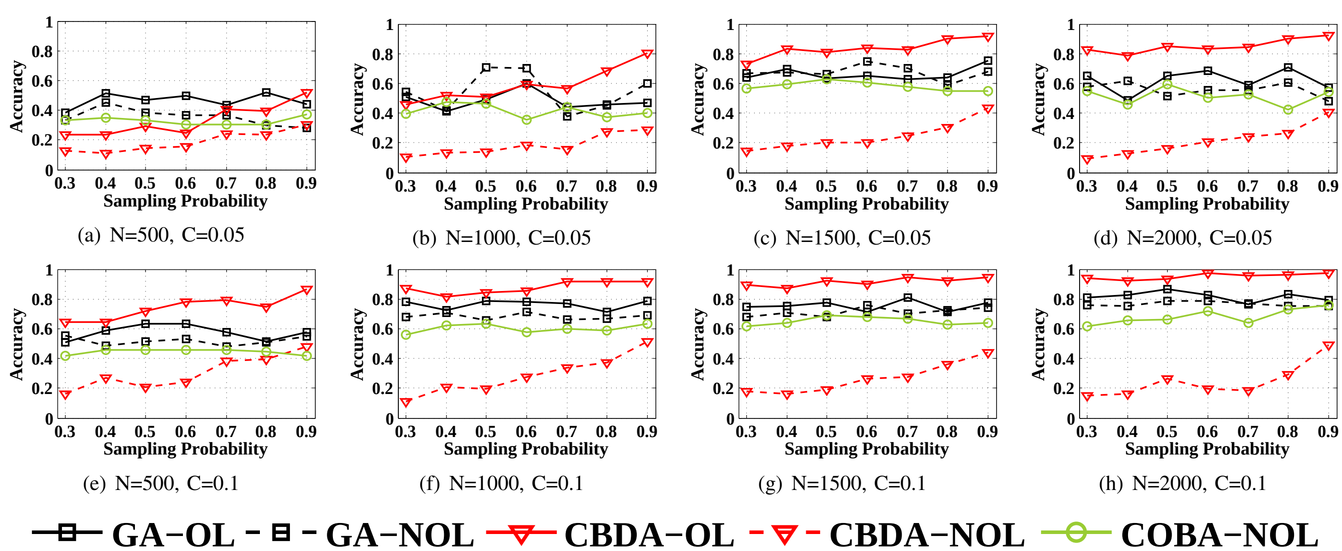

(a) N=500, a=3 (b) N=1000, a=3 (c) N=1500, a=3 (d) N=2000, a=3

(e) N=500, a=5 (f) N=1000, a=5 (g) N=1500, a=5 (h) N=2000, a=5

(i) N=500, a=7 (j) N=1000, a=7 (k) N=1500, a=7 (l) N=2000, a=7

(m) N=500, a=9 (n) N=1000, a=9 (o) N=1500, a=9 (p) N=2000, a=9

Fig. 13: Experiments on Synthetic Network with $\eta = 0.1$.




(a) N=500, C=0.05 (b) N=1000, C=0.05 (c) N=1500, C=0.05 (d) N=2000, C=0.05

(e) N=500, C=0.1 (f) N=1000, C=0.1 (g) N=1500, C=0.1 (h) N=2000, C=0.1

Fig. 14: Experiments on Sampled Real Social Networks.