

On the Feasible Region of Efficient Algorithms for Attributed Graph Alignment

Ziao Wang^{ID}, Graduate Student Member, IEEE, Ning Zhang, Weina Wang^{ID}, Member, IEEE,
and Lele Wang^{ID}, Member, IEEE

Abstract—Graph alignment aims at finding the vertex correspondence between two correlated graphs, a task that frequently occurs in graph mining applications such as social network analysis. Attributed graph alignment is a variant of graph alignment, in which publicly available side information or attributes are exploited to assist graph alignment. Existing studies on attributed graph alignment focus on either theoretical performance without computational constraints or empirical performance of efficient algorithms. This motivates us to investigate efficient algorithms with theoretical performance guarantee. In this paper, we propose two polynomial-time algorithms that exactly recover the vertex correspondence with high probability. The feasible region of the proposed algorithms is *near optimal* compared to the information-theoretic limits. When specialized to the seeded graph alignment problem under the seeded Erdős–Rényi graph pair model, the proposed algorithms *extends* the best known feasible region for exact alignment by polynomial-time algorithms.

Index Terms—Graph theory, statistics, inference algorithms.

I. INTRODUCTION

THE graph alignment problem, also referred to as the graph matching or noisy graph isomorphism problem, is the problem of finding the correspondence between the vertices of two correlated graphs. This problem has been given increasing attention for its applications in social network de-anonymization. For instance, datasets of social networks are typically anonymized for privacy protection. However, an attacker may be able to de-anonymize the dataset by aligning its user-user connection graph with that of publicly

available data. *Attributed graph alignment* is a variant of graph alignment in which side information, referred to as attributes of vertices, is also publicly available in addition to the user-user connection information. This variant is motivated by the fact that there might exist publicly available attribute information in social networks. For example, the anonymized network Netflix has users' movie-watching history and ratings publicly available. Moreover, the examination of the proposed model provides both achievability and converse results for various well-known variations of the Erdős–Rényi graph alignment problems. These variations include the traditional graph alignment problem without attributes, the seeded graph alignment problem, and the bipartite alignment problem. Consequently, the study of attributed graph alignment introduces a novel perspective for investigating these problems within an integrated framework.

In this paper, we focus on the attributed graph alignment problem under the attributed Erdős–Rényi pair model $\mathcal{G}(n, p, s_u; m, q, s_a)$, first proposed in [1]. In this model, a base graph G is generated on the vertex set $[n + m]$ where the vertices from the set $[n]$ represent *users* and the rest of the vertices represent *attributes*. Between each pair of users, an edge is generated independently and identically with probability p to represent their connection. For each user-attribute pair, an edge is generated independently and identically with probability q to represent their association. Note that there are no edges between attributes. The graph G is then independently subsampled to two graphs G_1 and G_2 , where each user-user edge is subsampled with probability s_u and each user-attribute edge is subsampled with probability s_a . To model the anonymization procedure, a random permutation Π^* chosen uniformly at random is applied to the *users* in G_2 to generate an anonymized version G'_2 . Our goal in this model is to achieve exact alignment, i.e., to exactly recover the permutation Π^* using G_1 and G'_2 .

For the attributed graph alignment problem, and the graph alignment problem in general, two often asked questions are the following. First, *for what region of graph statistics is exact alignment feasible with unlimited computational power?* This region is usually referred to as the information-theoretically feasible region or the information-theoretic limits. Second, *for what region of graph statistics is exact alignment feasible with polynomial-time algorithms?* This region is usually referred to as the feasible region of polynomial-time algorithms. Characterizing these two feasible regions and their relationship is of

Manuscript received 16 March 2023; revised 16 December 2023; accepted 22 December 2023. Date of publication 8 January 2024; date of current version 23 April 2024. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery under Grant RGPIN-2019-05448, in part by the NSERC Collaborative Research and Development under Grant CRDPJ 54367619, and in part by NSF under Grant CNS-2007733. An earlier version of this paper was presented in part at the 2022 IEEE International Symposium on Information Theory [DOI: 10.1109/ISIT50566.2022.9834398]. (Corresponding author: Ziao Wang.)

Ziao Wang and Lele Wang are with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada (e-mail: ziaow@ece.ubc.ca; lelewang@ece.ubc.ca).

Ning Zhang was with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada. She is now with the Department of Statistics, University of Oxford, OX1 3LB Oxford, U.K. (e-mail: ning.zhang@stats.ox.ac.uk).

Weina Wang is with the Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: weinaw@cs.cmu.edu).

Communicated by L. Lai, Associate Editor for Signal Processing and Source Coding.

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TIT.2024.3351107>.

Digital Object Identifier 10.1109/TIT.2024.3351107

0018-9448 © 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See <https://www.ieee.org/publications/rights/index.html> for more information.

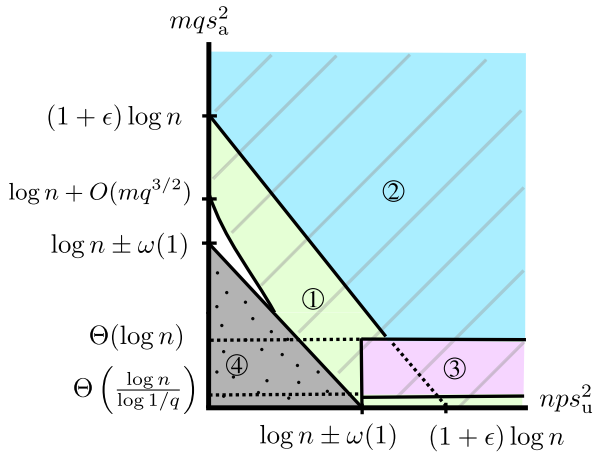


Fig. 1. Comparison between the feasible regions of the proposed algorithms and the information-theoretic limits: the shaded area (①+②+③) represents the information-theoretically feasible region given in [1]; area ② is the feasible region for Algorithm ATTRRICH and area ③ is the feasible region for Algorithm ATTRSPARSE; area ④ is the information-theoretically infeasible region given in [1].

utmost importance to developing a fundamental understanding of the graph alignment problem.

There has been extensive studies on these two questions under the Erdős–Rényi pair model without attributes. A line of research focuses on the information-theoretic limits of exact alignment [2], [3], [4], [5]. Roughly speaking, it is shown that exact alignment is information-theoretically feasible when the intersection graph is dense enough. A sharp threshold of exact alignment has been established, while there still exists some gap between the converse and the achievability results. Another line of research focuses on polynomial-time algorithms for exact alignment [6], [7], [8], [9], [10]. Compared to the information-theoretic limits, the existing polynomial-time algorithms further require higher edge correlation between the pair of graphs to achieve exact alignment. The question of whether there exists polynomial-time algorithms that achieve the known information-theoretic limits is still left open.

For the attributed graph alignment problem, the information-theoretic limit has been studied in [1], where the feasible region (achievability results) and infeasible region (converse results) are characterized with a gap in between in some regimes. However, the feasible region of polynomial-time algorithms for attributed graph alignment has not been studied before, and it is the focus of this paper. In this work, we propose two polynomial-time algorithms for attributed graph alignment and characterize their feasible regions. The two algorithms are designed for two different regimes of parameters based on the richness of attribute information: the algorithm ATTRRICH is designed for the regime where $mqs_a^2 = \Omega(\log n)$, referred to as the attribute-information rich regime; and the algorithm ATTRSPARSE is designed for the regime where $mqs_a^2 = o(\log n)$, referred to as the attribute-information sparse regime. In both algorithms, we first explore the user-attribute connections to align a set of anchor users, and then utilize the user-user connections to the anchors to align the rest of users. Due to the regime difference, ATTRRICH is

able to generate a much larger set of anchors in the first step than ATTRSPARSE. Therefore, ATTRRICH and ATTRSPARSE make use of the anchors differently in the second step: ATTRRICH explores one-hop user-user connections to align the rest of users, while ATTRSPARSE explores one-hop or multiple-hop user-user connections to align the rest of users depending on the sparsity of user-user connections. This idea of matching vertices based on common neighbor witnesses has been explored to construct efficient graph alignment algorithms under the context of seeded graph alignment [11], [12], [13], [14]. In this work, we employ this idea in a two-step procedure under the setting of attributed graphs, and analyze its performance through a careful treatment of the dependency between the two steps.

Our characterizations of the feasible regions of ATTRRICH and ATTRSPARSE are illustrated in Figure 1 as areas ② and ③, respectively. The information-theoretically feasible and infeasible regions given in [1] are also illustrated in the figure for comparison. We can see that there is a gap between the feasible region achieved by ATTRRICH and ATTRSPARSE and the known information-theoretically feasible region. It is left open whether this gap is a fundamental limit of polynomial-time algorithms¹.

In addition, we specialize the feasible region of the proposed algorithms to the context of the seeded Erdős–Rényi graph alignment problem, and demonstrate that the specialized feasible region includes certain range of parameters that is not known to be achievable in the literature.

Our results reveal that attributes possibly facilitate graph alignment in a much more significant way when computational efficiency is a priority. We demonstrate this possible impact of attributes under the sparse regime $np = \Theta(\log n)$ in Figure 2. In Figure 2a, we let $mqs_a^2 = 0$ so there is no information from the attributes, which is equivalent to the graph alignment problem without attributes; in Figure 2b, we let $mqs_a^2 = 0.1 \log n$. We keep $p = o(1)$, and assume the value of $nps_u / \log n$ can be an arbitrarily large constant but not tending to infinity in both settings for ease of comparison. Comparing these two figures, we can see that when mqs_a^2 increases from 0 to $0.1 \log n$, the information-theoretic limits in terms of user-user parameters (n, p and s_u) improve a bit as expected. However, a more fundamental improvement is in the feasible region achievable by polynomial-time algorithms. In Figure 2a, the green region above $s_u = \sqrt{\alpha}$ is achievable by a polynomial-time algorithm proposed in [10], where $\alpha \approx 0.338$ is known as the Otter's constant. The red region below $s_u = \sqrt{\alpha}$ is not known to be achievable by any polynomial-time algorithms. Moreover, in [18], this red region is conjectured to be infeasible by any polynomial-time algorithm.² In comparison, in Figure 2b, the

¹We comment that in the line of work [15], [16], [17], the authors have explored efficient algorithms for a closely-related problem known as attributed network alignment. In this problem, attributes are attached to both vertices and edges, in contrast to the exclusive association with vertices in attributed graph alignment. The focus in this line of work is the empirical performance rather than the theoretical feasible region.

²We note that the conjecture presented in [18] pertains to the sparse regime where $np = \Theta(\log n)$. In contrast, for the dense regime, it is conjectured in [19] that no polynomial-time algorithm can achieve exact alignment if $s_u \leq 1/\text{poly}(\log n)$.

green region above the information theoretic limit, which is presented as the dotted curve, is achievable by our proposed polynomial-time algorithm ATTRICH (assuming $q = o(1)$ and $s_a = \Theta(1)$). Therefore, if $\sqrt{\alpha}$ is indeed the right threshold between achievable and impossible under polynomial-time algorithms for graph alignment without attributes in the sparse regime, a small amount of attribute information lowers this threshold to an arbitrarily small constant.

II. MODEL

In this section, we describe a random process that generates a pair of correlated graphs, which we refer to as the attributed Erdős–Rényi pair model $\mathcal{G}(n, p, s_u; m, q, s_a)$. Under this model, we define the exact alignment problem.

A. Base Graph Generation

We first generate a base graph G , whose vertex set $\mathcal{V}(G)$ consists of two disjoint sets, the *user vertex set* $\mathcal{V}^u = \{1, 2, \dots, n\}$ and the *attribute vertex set* $\mathcal{V}^a = \{n+1, n+2, \dots, n+m\}$. There are two types of edges in the base graph G , the user-user edges (edges connecting a pair of users) and the user-attribute edges (edges connecting a user vertex and an attribute vertex). The user-user edges are generated independently and identically with probability p , and the user-attribute edges are generated independently and identically with probability q . Throughout this paper, we assume that $p = o(1)$ and $q = o(1)$. We write $i \stackrel{\mathcal{G}}{\sim} j$ if vertices i and j are connected in graph G .

B. Edge Subsampling

From the base graph G , we obtain two correlated graphs G_1 and G_2 by subsampling the edges in G independently. More specifically, we get G_1 and G_2 by independently including each user-user edge in G with probability s_u and independently including each user-attribute edge with probability s_a . Throughout this paper, we assume that $s_u = \Theta(1)$ and $s_a = \Theta(1)$.

C. Anonymization

From the G_2 generated as above, we get an anonymized graph G'_2 by applying an unknown permutation Π^* on the user vertices of G_2 , where Π^* is drawn uniformly at random from the set of all possible permutations on $\mathcal{V}^u$. We use $\mathcal{V}_2^u$ to denote the user vertex set of G'_2 and use $\mathcal{V}_1^u$ to denote the user vertex set of G_1 . Finally, we remark that this subsampling process is a special case of an earlier described attributed Erdős–Rényi pair model in [1].

D. Exact Alignment

Given an observable pair (G_1, G'_2) , our goal is to recover the unknown permutation Π^* , which allows us to recover the original labels of user vertices in the anonymized graph G'_2 . We say exact alignment is achieved *with high probability* (w.h.p.) if $\lim_{n \rightarrow \infty} \mathbb{P}(\hat{\Pi} \neq \Pi^*) = 0$. It is worth mentioning that $\mathbb{P}(\hat{\Pi} \neq \Pi^*) = \mathbb{P}(\hat{\Pi} \neq \Pi^* | \Pi^* = \pi_{\text{id}})$ due to the symmetry among user vertices. Thus, we later assume without loss of generality that the underlying true permutation is the identity permutation.

E. Related Random Graph Models

A closely related graph generation model is the correlated stochastic block model [20], [21], [22]. In this model, the base graph G is instead generated from the stochastic block model, where vertices are grouped into latent blocks, and edges are formed between blocks with specific probabilities, capturing the underlying community structure in the network. The base graph G is then subsampled into the two generated graphs G_1 and G_2 . In the proposed attributed Erdős–Rényi graph pair models, users and attributes can be viewed as two communities in the correlated stochastic block model. However, a key distinction lies in the fact that the attributed Erdős–Rényi graph pair model specifies the correspondence between attributes in the two graphs, whereas the correlated stochastic block model does not disclose such information.

Another well-studied graph alignment model with side information is the seeded Erdős–Rényi graph pair model [11], [12], [13], [14], where we have access to part of the true correspondence between user vertices. To make a comparison between the seeded Erdős–Rényi graph pair model and the proposed model, here we describe the seeded Erdős–Rényi pair model $\mathcal{G}(N, \alpha, p, s)$ in detail. We first sample a base graph G from the Erdős–Rényi graph on N vertices with edge probability p . Then two correlated copies G_1 and G_2 are obtained by independently subsampling the edges in the base graph where each edge is preserved with probability s . The anonymized graph G'_2 is obtained by applying an unknown permutation Π^* on G_2 , where Π^* is drawn uniformly at random. Let $\mathcal{V}(G_1)$ and $\mathcal{V}(G'_2)$ denote the vertex sets of G_1 and G'_2 respectively. Then, a subset $\mathcal{V}_s \subset \mathcal{V}(G_1)$ of size $\lfloor N\alpha \rfloor$ is chosen uniformly at random and we define the vertex pairs $\mathcal{I}_0 = \{(v_1, \Pi^*(v_1)) : v_1 \in \mathcal{V}_s\}$ as the *seed set*. The graph pair (G_1, G'_2) together with the seed set $\mathcal{I}_0$ are given and the goal of the exact alignment is to recover the underlying permutation for the remaining vertices w.h.p.

Comparing the seeded Erdős–Rényi pair model and the attributed Erdős–Rényi pair model, we can see that the seed set and the attribute set both provide side information to assist the alignment of the remaining vertices. Nevertheless, there are two main differences between the two models. First, in the attributed Erdős–Rényi pair model, we allow different edge probabilities and subsampling probabilities for user-user edges and user-attribute edges, whereas in the seeded Erdős–Rényi pair model, the edge probability is identical for all edges and so is the subsampling probability. Second, while there are edges between seeds in seeded Erdős–Rényi pair model, there are no attribute-attribute edges in the attributed Erdős–Rényi pair model. However, it can be shown that the existence of edges between seeds has no influence on the information-theoretic limits for exact alignment in the seeded Erdős–Rényi pair model (see Lemma 1 in [23]). This further suggests that regarding the task of exact alignment, the information-theoretic limits on attributed graph alignment recover the information-theoretic limits on seeded Erdős–Rényi graph alignment if we specialize $p = q$ and $s_u = s_a$ in the attributed Erdős–Rényi pair model $\mathcal{G}(n, p, s_u; m, q, s_a)$.

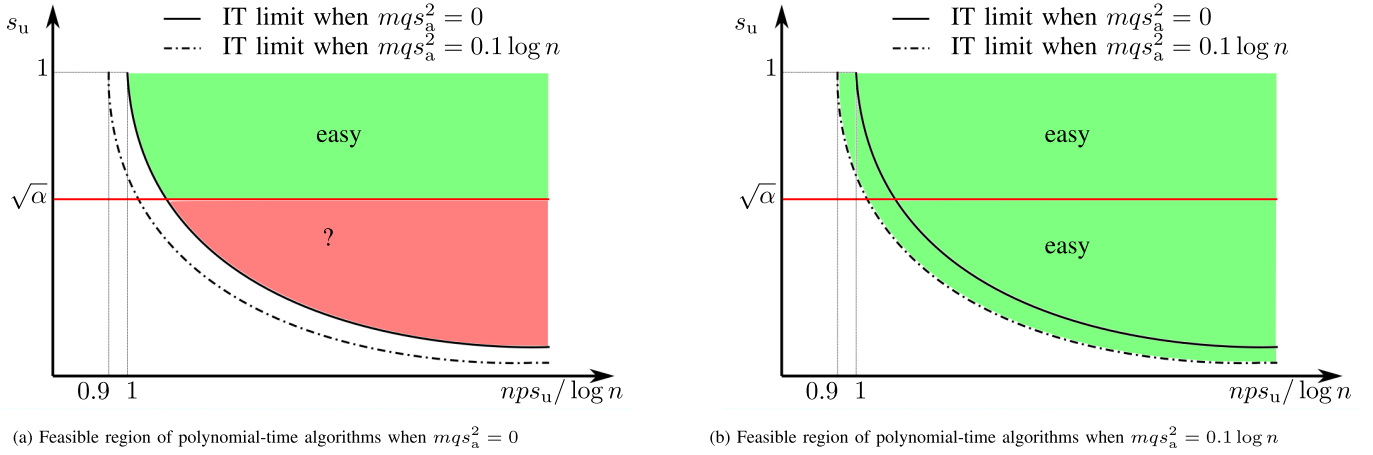


Fig. 2. Comparison between feasible regions of polynomial-time algorithms when $mqs_a^2 = 0$ and when $mqs_a^2 = 0.1 \log n$. Subgraph (a) captures the case when $mqs_a^2 = 0$. The green region is known to be feasible by a polynomial-time algorithm in [10], while no polynomial-time algorithms are known to be feasible in the red region. Subgraph (b) captures the case when $mqs_a^2 = 0.1 \log n$. The green region is feasible by the proposed algorithm ATTRICH.

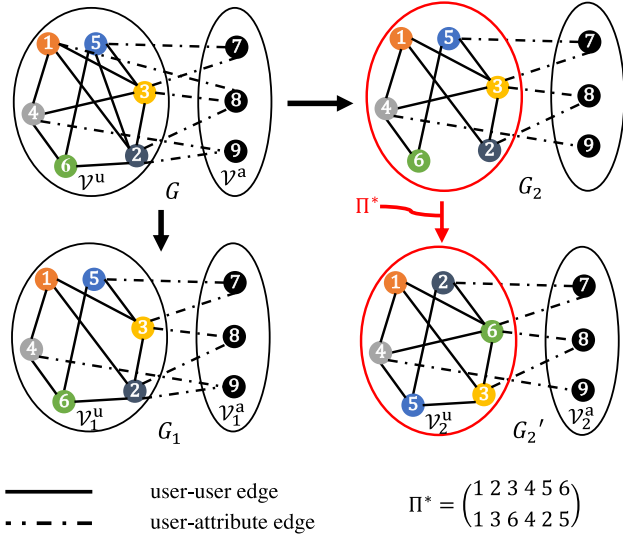


Fig. 3. An illustration of attributed Erdős-Rényi pair model. We first sample a base graph G . Then we get G_1 and G_2 through edge subsampling G . The anonymized graph G_2' is obtained through apply the permutation Π^* on the user vertex set of G_2 .

F. Other Notation

Our algorithms rely on exploring the neighborhood similarity of user vertices in G_1 and G_2' . Here we introduce our notation of local neighborhoods. We define $\mathcal{N}_1^a(i) \triangleq \{j \in \mathcal{V}_1^a : i \stackrel{G_1}{\sim} j\}$ as the set of attribute neighbors of a user vertex i in G_1 and $\mathcal{N}_2^a(i) \triangleq \{j \in \mathcal{V}_2^a : i \stackrel{G_2}{\sim} j\}$ as the set of attribute neighbors of a user vertex i in G_2 . For two user vertices i and j in the same graph, let $d(i, j)$ be the length of the shortest path connecting i and j via user-user edges. For a user vertex $i \in \mathcal{V}_1^u$, we define the set of l -hop user neighbors of vertex i as $\mathcal{N}_1^u(i, l) \triangleq \{j \in \mathcal{V}_1^u : d(i, j) \leq l\}$ for any positive integer l . By convention, when $l = 1$, we simply write $\mathcal{N}_1^u(i) \equiv \mathcal{N}_1^u(i, 1)$. The quantities $\mathcal{N}_2^u(i, l)$ and $\mathcal{N}_2^u(i)$ are defined similarly for user vertices in G_2' .

Reminder of the Landau notation.

Notation	Definition
$f(n) = \omega(g(n))$	$\lim_{n \rightarrow \infty} \frac{ f(n) }{ g(n) } = \infty$
$f(n) = o(g(n))$	$\lim_{n \rightarrow \infty} \frac{ f(n) }{ g(n) } = 0$
$f(n) = O(g(n))$	$\limsup_{n \rightarrow \infty} \frac{ f(n) }{ g(n) } < \infty$
$f(n) = \Omega(g(n))$	$\liminf_{n \rightarrow \infty} \frac{ f(n) }{ g(n) } > 0$
$f(n) = \Theta(g(n))$	$f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$

III. MAIN RESULTS

In this section, we propose two polynomial-time algorithms for the attributed graph alignment problem. Their feasible regions are characterized in the following two theorems.

Theorem 1: Consider the attributed Erdős-Rényi pair $\mathcal{G}(n, p, s_u; m, q, s_a)$ with $p = o(1)$, $q = o(1)$, $s_u = \Theta(1)$, and $s_a = \Theta(1)$. Assume that

$$mqs_a^2 = \Omega(\log n) \quad (1)$$

and that there exists some constant $\epsilon > 0$ such that

$$mqs_a^2 + nps_u^2 \geq (1 + \epsilon) \log n. \quad (2)$$

Then there exists a polynomial-time algorithm, namely, Algorithm ATTRICH with the parameters chosen in (8) and (9), that achieves exact alignment w.h.p.

Theorem 2: Consider the attributed Erdős-Rényi pair $\mathcal{G}(n, p, s_u; m, q, s_a)$ with $p = o(1)$, $q = o(1)$, $s_u = \Theta(1)$, and $s_a = \Theta(1)$. Assume that

$$mqs_a^2 = o(\log n), \quad (3)$$

$$nps_u^2 - \log n \rightarrow +\infty, \quad (4)$$

and that there exists some constant $\tau > 0$ such that

$$mqs_a^2 \geq \frac{2 \log n}{\tau \log \frac{1}{q}}. \quad (5)$$

Then there exists a polynomial-time algorithm, namely, Algorithm ATTRSPARSE with the parameters chosen in (11), (12) and (13), that achieves exact alignment w.h.p.

The proofs of Theorems 1 and 2 are deferred to Sections V and VI respectively.

A. Algorithm ATTRICH

In this subsection, we propose the first algorithm that leads to the achievable region in Theorem 1. This algorithm is designed for the attribute-information rich regime, hence named ATTRICH.

- **Input:** The graph pair (G_1, G'_2) and thresholds x and y .
- **Step 1: Align through attribute neighbors.** In this step, we only consider the edge connections between users and attributes, and use these information to find the matching for a set of vertices which will be later referred to as anchors. For each pair of users $i \in \mathcal{V}_1^u$ and $j \in \mathcal{V}_2^u$, compute the number of common attribute neighbors

$$C_{ij} \triangleq |\mathcal{N}_1^a(i) \cap \mathcal{N}_2^a(j)|. \quad (6)$$

If $C_{ij} > x$, add (i, j) into $\mathcal{S}_1$. We refer to vertex pairs in the set $\mathcal{S}_1$ as anchor. If there exists conflicting pairs in $\mathcal{S}_1$, i.e., two distinct pairs (i_1, j_1) and (i_2, j_2) with $i_1 = i_2$ or $j_1 = j_2$, set $\mathcal{S}_1 = \emptyset$ and declare failure. Otherwise, set $\hat{\pi}(i) = j$ for all pairs $(i, j) \in \mathcal{S}_1$.

- **Step 2: Align through user neighbors.** In the previous step, we have aligned the anchors using the edges between users and attributes. In this step, we will align the non-anchor vertices by their edge connections to the anchors. Let

$$\mathcal{U}_1 \triangleq \{i \in \mathcal{V}_1^u : (i, j) \notin \mathcal{S}_1, \forall j \in \mathcal{V}_2^u\}$$

denote the set of all unmatched vertices in G_1 and let

$$\mathcal{U}_2 \triangleq \{j \in \mathcal{V}_2^u : (i, j) \notin \mathcal{S}_1, \forall i \in \mathcal{V}_1^u\}$$

denote the set of all unmatched vertices in G_2 . For each unmatched pair $i \in \mathcal{U}_1$ and $j \in \mathcal{U}_2$, consider the user neighbors of i and the user neighbors of j that are matched as pairs in $\mathcal{S}_1$, and compute the number of such matched pairs for (i, j)

$$W_{ij} \triangleq \sum_{k \in \mathcal{N}_1^u(i), l \in \mathcal{N}_2^u(j)} \mathbb{1}_{\{(k, l) \in \mathcal{S}_1\}}. \quad (7)$$

For each $i \in \mathcal{U}_1$, if $W_{ij} > y|\mathcal{S}_1|$ for a unique $j \in \mathcal{U}_2$, set $\hat{\pi}(i) = j$. Otherwise, declare failure. If $\hat{\pi}$ is not a bijection from $\mathcal{V}_1^u$ to $\mathcal{V}_2^u$, declare failure.

- **Output:** The estimated permutation $\hat{\pi}$.

In this algorithm, there are two threshold parameters x and y . In the following analysis, we choose

$$x = (1 - \delta_x)mqs_a^2, \quad (8)$$

where $1 - \delta_x = \frac{\Delta_x}{\log \frac{1}{\delta_x}}$ with constant $\Delta_x \geq \max\{1, \frac{3 \log n}{mqs_a^2}\}$, and

$$y = (1 - \delta_y)ps_u^2, \quad (9)$$

where $1 - \delta_y = \frac{\Delta_y}{\log \frac{1}{\delta_y}}$ with constant $\Delta_y \geq 2$.

Remark 1 (Complexity of Algorithm ATTRICH): In Algorithm ATTRICH, the time complexity for computing C_{ij} for all pairs $(i, j) \in \mathcal{V}_1^u \times \mathcal{V}_2^u$ is $O(n^2m)$ since there are

n^2 pairs and for each pair, there are m attributes to consider. Similarly, the time complexity for computing W_{ij} for all pairs $(i, j) \in \mathcal{U}_1 \times \mathcal{U}_2$ is at most $O(n^3)$. Therefore, if $m = \omega(n)$, the time complexity of Algorithm ATTRICH is $O(n^2m)$ and if $m = O(n)$, the time complexity of Algorithm ATTRICH is $O(n^3)$.

B. Algorithm ATTRSPARSE

In this subsection, we propose the second algorithm that leads to the achievable region in Theorem 2. This algorithm is designed for the attribute-information sparse regime, hence named ATTRSPARSE. In Step 2 of this algorithm, we consider two different cases. In the case when the user-user connection is dense, we perform a similar process as in Step 2 of Algorithm ATTRICH. In the case when the user-user connections is sparse, we call a seeded alignment algorithm proposed in [13], which is restated in Subsection III-C.

- **Input:** The graph pair (G_1, G'_2) , three thresholds y, z and η , an integer l , and the model parameters n and p .
- **Step 1: Align through attribute neighbors.** Similar to Step 1 of Algorithm ATTRICH, for each pair of users $i \in \mathcal{V}_1^u$ and $j \in \mathcal{V}_2^u$, we compute the quantity

$$C_{ij} = |\mathcal{N}_1^a(i) \cap \mathcal{N}_2^a(j)|. \quad (10)$$

Unlike Step 1 of Algorithm ATTRICH, we create an anchor set using a different threshold z . If $C_{ij} > z$, add (i, j) into $\mathcal{S}_2$. We refer to vertex pairs in the set $\mathcal{S}_2$ as anchors. If there exists conflicting pairs in $\mathcal{S}_2$, i.e., two distinct pairs (i_1, j_1) and (i_2, j_2) with $i_1 = i_2$ or $j_1 = j_2$, set $\mathcal{S}_2 = \emptyset$ and declare failure.

- **Step 2: Align through user-user edges.**

– If $np > n^{1/7}$, we perform the similar process as in Step 2 of Algorithm ATTRICH to align the non-anchor vertices. Define

$$\mathcal{U}_3 \triangleq \{i \in \mathcal{V}_1^u : (i, j) \notin \mathcal{S}_2, \forall j \in \mathcal{V}_2^u\}$$

and

$$\mathcal{U}_4 \triangleq \{j \in \mathcal{V}_2^u : (i, j) \notin \mathcal{S}_2, \forall i \in \mathcal{V}_1^u\}.$$

For each unmatched pair $i \in \mathcal{U}_3$ and $j \in \mathcal{U}_4$, compute W_{ij} as defined in (7). For each $i \in \mathcal{U}_3$, if $W_{ij} > y|\mathcal{S}_2|$ for a unique $j \in \mathcal{U}_4$, set $\hat{\pi}(i) = j$. Otherwise, declare failure. If $\hat{\pi}$ is not a bijection from $\mathcal{V}_1^u$ to $\mathcal{V}_2^u$, declare failure.

– If $np \leq n^{1/7}$, run Algorithm III-C with the induced subgraphs on user vertices $\mathcal{V}_1^u$ and $\mathcal{V}_2^u$, the seed set $\mathcal{I}_0 = \mathcal{S}_2$, and parameters l and η .

- **Output:** The estimated permutation $\hat{\pi}$.

In this algorithm, there are four parameters y, z, l and η that we can choose. In the following analysis, we choose y to be the same value as in (9),

$$z = (1 + \tau)mqs_a^2 \quad (11)$$

(cf. the same τ as in Theorem 2),

$$l = \left\lceil \frac{(6/7) \log n}{\log(np)} \right\rceil, \quad (12)$$

and

$$\eta = 4^{2l+2}n^{-2/7}. \quad (13)$$

Remark 2 (Complexity of Algorithm ATTRSPARSE): In Algorithm ATTRSPARSE, the time complexity for computing C_{ij} for all pairs $(i, j) \in \mathcal{V}_1^u \times \mathcal{V}_2^u$ is $O(n^2m)$ since there are n^2 pairs and for each pair, there are m attributes to consider. The time complexity of computing W_{ij} for all pairs $(i, j) \in \mathcal{U}_3 \times \mathcal{U}_4$ is $O(n^3)$ and that of Algorithm III-C is $O(n^{37/7})$ as given in [13], which can be improved with better data structures. Therefore, if $np > n^{1/7}$, the complexity of Algorithm ATTRSPARSE is $O(n^2m + n^3)$; Otherwise its complexity is $O(n^2m + n^{37/7})$.

C. Seeded Alignment in the Sparse Regime [13, Algorithm 3]

Except for the two graphs G_1 and G'_2 , this algorithm takes a seed set $\mathcal{I}_0$ as input. The seed set $\mathcal{I}_0$ consists of vertex pairs (i, j) such that $\pi^*(i) = j$. The algorithm utilizes this seed set to align the rest of vertices.

- **Input:** The graph pair (G_1, G'_2) , the seed set $\mathcal{I}_0$, a threshold η , and an integer l .
- **Align high-degree vertices.** Let

$$\mathcal{J}_1 \triangleq \{i \in \mathcal{V}(G_1) : (i, j) \notin \mathcal{I}_0, \forall j \in \mathcal{V}(G'_2)\},$$

and

$$\mathcal{J}_2 \triangleq \{j \in \mathcal{V}(G'_2) : (i, j) \notin \mathcal{I}_0, \forall i \in \mathcal{V}(G_1)\}.$$

For each pair of unseeded vertices $u \in \mathcal{J}_1$ and $v \in \mathcal{J}_2$, and for each pair of their neighbors $i \in \mathcal{N}_1^u(u) \setminus \{u\}$ and $j \in \mathcal{N}_2^v(v) \setminus \{v\}$, compute

$$\lambda_{i,j}^{u,v} = \min_{x \in \mathcal{V}(G_1), y \in \mathcal{V}(G'_2)} \{|\{(k_1, k_2) \in \mathcal{I}_0 : k_1 \in \mathcal{N}_{G_1 \setminus \{u, x\}}^u(i, l), k_2 \in \mathcal{N}_{G'_2 \setminus \{v, y\}}^v(j, l)\}|\},$$

where $\mathcal{N}_{G \setminus S}^u(i_1, l)$ denotes the set of user vertices i_2 such that $d(i_1, i_2) \leq l$ in the induced subgraph G with the set of vertices S removed. Let

$$Z_{u,v} = \sum_{i \in \mathcal{N}_1^u(u) \setminus \{u\}} \sum_{j \in \mathcal{N}_2^v(v) \setminus \{v\}} \mathbb{1}_{\{\lambda_{i,j}^{u,v} \geq \eta |\mathcal{I}_0|\}}.$$

If $Z_{u,v} \geq \log n / \log \log n - 1$, add (u, v) into set $\mathcal{T}$. Add all the vertex pairs from $\mathcal{I}_0$ to $\mathcal{T}$. If there exists conflicting pairs in $\mathcal{T}$, i.e., two distinct pairs (i_1, j_1) and (i_2, j_2) with $i_1 = i_2$ or $j_1 = j_2$, set $\mathcal{T} = \emptyset$ and declare failure.

- **Align low-degree vertices.** Let

$$\mathcal{J}_3 \triangleq \{i \in \mathcal{V}(G_1) : (i, j) \notin \mathcal{T}, \forall j \in \mathcal{V}(G'_2)\},$$

and

$$\mathcal{J}_4 \triangleq \{j \in \mathcal{V}(G'_2) : (i, j) \notin \mathcal{T}, \forall i \in \mathcal{V}(G_1)\}.$$

For all pairs of unmatched vertices $i_1 \in \mathcal{J}_3$ and $i_2 \in \mathcal{J}_4$, if i_1 is adjacent to a user vertex j_1 in G_1 and i_2 is adjacent to a user vertex j_2 in G'_2 such that $(j_1, j_2) \in \mathcal{T}$, then set $\hat{\pi}(i_1) = i_2$.

- **Finalize and output:** For each vertex pair $(i, j) \in \mathcal{T}$, set $\hat{\pi}(i) = j$. If $\hat{\pi}$ is a bijection from $\mathcal{V}(G_1)$ to $\mathcal{V}(G'_2)$, output $\hat{\pi}$, otherwise declare failure.

IV. DISCUSSION

In this section, we compare the feasible region in Theorems 1 and 2 to existing works. In Section IV-A, we compare the feasible region with the information-theoretic limits in [1]. It is shown that there still exists a gap between the feasible region of the proposed algorithms and the information-theoretic limit. In Section IV-B, we specialize the feasible region of the proposed algorithm to the context of seeded graph alignment problem, and compare the specialized feasible region to the information-theoretic limits as well as the best-known feasible regions by polynomial-time algorithms for exact alignment in literature given in [13] and [14]. It is shown that while having a gap to the information-theoretic limit, the proposed Algorithms ATTRICH and ATTRSPARSE achieves exact recovery in certain range of parameters that is unknown to be feasible by any existing efficient algorithms. In Section IV-C, we consider the bipartite alignment problem which is another special case. We show that the proposed Algorithm ATTRICH provides an alternative polynomial-time algorithm to the theoretically optimal Hungarian Algorithm [24], with a slightly lower complexity.

A. Comparison to the Information Theoretic Limits

The information-theoretic limits of attributed graph alignment were established in [1]. The version for the subsampling model is stated as follows.

Theorem 3 (Theorem 1 in [1]): Consider the attributed graph pair $\mathcal{G}(n, p, s_u; m, q, s_a)$ with $1 - p = \Theta(1)$, $1 - q = \Theta(1)$, $s_u = \Theta(1)$, and $s_a = \Theta(1)$.

Achievability: In the regime where $q = O\left(\frac{1}{(\log n)^2}\right)$, if

$$nps_u^2 + mqs_a^2 \geq \log n + \omega(1),$$

then exact alignment is achievable w.h.p.

In the regime where $q = \omega\left(\frac{1}{(\log n)^2}\right)$, if

$$nps_u^2 + mqs_a^2 - a_n \geq \log n + \omega(1),$$

where $a_n \triangleq m\left(\sqrt{qs_a^2(1 - q + q(1 - s_a)^2)} - qs_a(1 - s_a)\right)^2 - mqs_a^2 = O(mq^{3/2})$,

then exact alignment is achievable w.h.p.

Converse: If

$$nps_u^2 + mqs_a^2 \leq \log n - \omega(1),$$

then no algorithm achieves exact alignment w.h.p.

From Theorem 3 we can see that when $q = \omega\left(\frac{1}{(\log n)^2}\right)$, the achievability and converse differ by at most some constant times $mq^{3/2}$; when $q = O\left(\frac{1}{(\log n)^2}\right)$, the achievability and converse are tight, because in this regime we have $mq^{3/2} = \omega(1)$. We visualize the information-theoretic limits and the computation feasible regions in Fig. 1.

B. Specialization to the Seeded Erdős–Rényi Graph Alignment

In this subsection, we specialize the feasible region of proposed algorithm to the seeded Erdős–Rényi graph alignment problem, and compare the feasible region with that of

the existing algorithms. Consider an attributed Erdős–Rényi graph pair $(G_1, G'_2) \sim \mathcal{G}(n, p, s_u; m, q, s_a)$ with $p = q$ and $s_a = s_u \triangleq s$. Then these m attributes can be viewed as m seeds and (G_1, G'_2) can be viewed as a graph pair generated from a seeded Erdős–Rényi pair model $\mathcal{G}(n + m, \frac{m}{m+n}, p, s)$, and the edges between the these m vertices are all removed.³ For simplicity, we write

$$N \triangleq m + n \quad \text{and} \quad \alpha \triangleq \frac{m}{m+n}$$

later on. In our comparisons, we always assume that $(1 - \alpha)N = \omega(1)$, $s = \Theta(1)$ and $p = o(1)$.

We first specialize the feasible regions of Algorithm ATTRICH in Theorem 1 and Algorithm ATTRSPARSE in Theorem 2 to the seeded graph alignment problem.

Corollary 1 (Feasible region of Algorithm ATTRICH):

Consider the seeded graph pair $\mathcal{G}(N, \alpha, p, s)$ with $p = o(1)$, $s = \Theta(1)$ and $(1 - \alpha)N = \omega(1)$. Assume that

$$\alpha N p s^2 = \Omega(\log((1 - \alpha)N)) \quad (14)$$

and that there exists some constant $\epsilon > 0$ such that

$$N p s^2 \geq (1 + \epsilon) \log((1 - \alpha)N). \quad (15)$$

Then there exists a polynomial time algorithm, namely, Algorithm ATTRICH with parameters

$$x = (1 - \delta_x) \alpha N p s^2 \quad \text{and} \quad y = (1 - \delta_y) p s^2$$

that achieves exact alignment w.h.p. Here, $1 - \delta_x = \frac{\Delta_x}{\log \frac{1}{p}}$ with constant $\Delta_x \geq \max\{1, \frac{3 \log((1 - \alpha)N)}{\alpha N p s^2}\}$ and $1 - \delta_y = \frac{\Delta_y}{\log \frac{1}{p}}$ with constant $\Delta_y \geq 2$.

Corollary 2 (Feasible region of Algorithm ATTRSPARSE):

Consider the seeded graph pair $\mathcal{G}(N, \alpha, p, s)$ with $p = o(1)$, $s = \Theta(1)$ and $(1 - \alpha)N = \omega(1)$. Assume that

$$\alpha N p s^2 = o(\log((1 - \alpha)N)), \quad (16)$$

$$(1 - \alpha) N p s^2 - \log((1 - \alpha)N) \geq \omega(1), \quad (17)$$

and that there exists some constant $\tau > 0$ such that

$$\alpha N p s^2 \geq \frac{2 \log((1 - \alpha)N)}{\tau \log \frac{1}{p}}. \quad (18)$$

Then there exists a polynomial time algorithm, namely, Algorithm ATTRSPARSE with parameters

$$z = (1 + \tau) \alpha N p s^2,$$

$$L = \left\lfloor \frac{(6/7) \log(1 - \alpha)N}{\log(1 - \alpha)N p} \right\rfloor,$$

and

$$\eta = 4^{2L+2} ((1 - \alpha)N)^{-2/7},$$

that achieves exact alignment w.h.p.

Remark 3: As expected, the feasible region of proposed algorithms ATTRICH and ATTRSPARSE is a strict subset of the information-theoretic feasible region established in [1]. We postpone the detailed comparison to Section VIII-A.

³The information theoretic limit of exact alignment in the seeded graph alignment problem remains the same after removing the edges between attributes (see Lemma 1 in [23]).

Now we compare the proposed algorithms to the best known polynomial-time algorithms in the literature [13] and [14]. The comparison of their feasible regions is summarized in the following theorem, the proof of which is postponed to Section VIII-B.

Theorem 4 (Comparison of polynomial-time algorithms):

Consider the seeded Erdős–Rényi pair model $\mathcal{G}(N, \alpha, p, s)$ with $p = o(1)$, $s = \Theta(1)$ and $(1 - \alpha)N = \omega(1)$. Assume that parameters N, α, p , and s satisfy any of the following four sets of conditions:

$$\textcircled{1} \begin{cases} p \log N \log \left(\frac{1}{p} \right) = \omega(1), \\ \alpha = \Omega \left(\frac{\log((1 - \alpha)N)}{N p s^2 \log \frac{1}{p}} \right), \\ \alpha < \frac{2 \log N}{N I(p, s)}; \end{cases}$$

$$\textcircled{2} \begin{cases} p \log N \log \left(\frac{1}{p} \right) = O(1), \\ p \log N \log^2 \left(\frac{1}{p} \right) = \omega(1), \\ \alpha = \Omega \left(\frac{\log((1 - \alpha)N)}{N p s^2 \log \frac{1}{p}} \right), \\ \alpha = O \left(\frac{1}{N I^2(p, s)} \right); \end{cases}$$

$$\textcircled{3} \begin{cases} p \log N \log^2 \left(\frac{1}{p} \right) = O(1), \\ N p > \frac{s N^{1/2}}{16(2 - s)^2}, \\ \alpha = \Omega \left(\frac{\log((1 - \alpha)N)}{N p s^2 \log \frac{1}{p}} \right), \\ \alpha < \frac{300 \log N}{N p s^2}; \end{cases}$$

$$\textcircled{4} \begin{cases} N = \Omega(((1 - \alpha)N)^{1+\epsilon}), \\ N p s^2 - \log N = O(1), \\ N p s^2 \geq (1 + \epsilon) \log((1 - \alpha)N) \end{cases}$$

for some positive constant ϵ . Then the proposed algorithms achieve exact alignment w.h.p., while none of existing algorithms in [13] and [14] is known to achieve exact alignment w.h.p. On the other hand, when

$$\frac{\log N + \omega(1)}{s^2} \leq N p \leq \frac{s N^{1/2}}{16(2 - s)^2},$$

the feasible region of the proposed algorithms is a strict subset of that in [13].

Remark 4: The polynomial-time algorithms for graph alignment under the unseeded Erdős–Rényi graph pair model proposed in [9] and [10] trivially imply polynomial-time algorithm for seeded graph alignment. However, the feasible regions of algorithms in [9] and [10] both require additional conditions on the subsampling probability s . Therefore, we do not include them to the comparison in this section.

Remark 5: As shown in Corollaries 1 and 2, the two proposed algorithms achieve exact alignment in two mutually exclusive regimes. Theorem 4 summarizes the comparison between the *union* of the feasible regions of the two proposed algorithms and the *union* of the feasible regions in [13] and [14]. Regions ①, ②, and ③ in Theorem 4 correspond

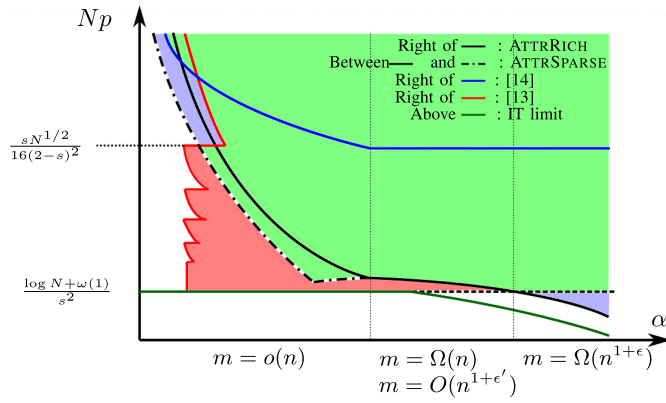


Fig. 4. Comparison of the feasible region of Corollaries 1 and 2 to the feasible region in [13] and [14]. On the top-left corner and bottom-right corner, the two blue regions are feasible for the proposed algorithms but not for any existing works. The red region is feasible for existing works, but not for the proposed algorithms. The green region is the overlap of our feasible region with the feasible region in the existing works.

to the top left blue area in Fig. 4. Region ④ in Theorem 4 corresponds to the bottom right blue area in Fig. 4.

As stated in Theorem 4, Corollaries 1 and 2 introduce some new feasible region in the regime $Np > \frac{s}{16(2-s)^2} N^{1/2}$ (the top-left blue corner in Fig. 4). To understand the improvement over the feasible region in [13], recall that the proposed algorithms in this work align vertices two steps: initially aligning a group of anchor vertices by exploring the seeds within their one-hop neighborhood, and subsequently aligning the remaining unmatched vertices with the assistance of the anchors in their respective one-hop neighborhood. In contrast, the algorithm proposed in [13] for this scenario aligns all vertices by exploring the one-hop seed neighbors of the non-seed vertices, which closely resembles the first step of the proposed algorithms. This is why the proposed algorithms expand the feasible region introduced in [13]. To grasp the improvements made over the feasible region in [14], note that the algorithm introduced in [14] performs a similar two-step process as the algorithms in this work. The primary distinction lies in the second step of matching. In [14], the matched vertices from the first step, along with the seeds, act as anchors in the second step. In contrast, our proposed algorithms utilize only the matched vertices from the first step as anchors. We comment that the improvement of the proposed algorithms over the algorithm in [14] mainly comes from the tightness of analysis.

For completeness of the comparison, we restate the feasible regions of algorithms in [13] and [14] as follows.

Theorem 5 (Theorem 4 in [13]): Consider the seeded Erdős-Rényi pair model $\mathcal{G}(N, \alpha, p, s)$. Suppose Np can be written as $Np = bN^a$ for some constants a and b such that

$$0 < a \leq 1 \quad \text{and} \quad 0 < b \leq \frac{s}{16(2-s)^2}.$$

Assume that

$$\alpha \geq \frac{300 \log N}{(Nps^2)^{1/a}}. \quad (19)$$

Then there exists a polynomial-time algorithm, namely, Algorithm 2 in [13], that achieves exact alignment w.h.p. Moreover, the algorithm runs in $O(N^3)$ time.

Theorem 6 (Theorem 3 in [13]): Consider the seeded Erdős-Rényi pair model $\mathcal{G}(N, \alpha, p, s)$ with $Np \leq N^\beta$ for a fixed constant $\beta < 1/6$, and $s = \Theta(1)$. Assume that

$$Nps^2 \geq \log N + \omega(1) \quad (20)$$

and

$$\alpha \geq N^{-1+3\beta}. \quad (21)$$

Then Algorithm III-C with the parameters

$$l = \left\lfloor \frac{(6/7) \log N}{\log(Np)} \right\rfloor \quad \text{and} \quad \eta = 4^{2l+2} N^{-2/7}$$

achieves exact alignment w.h.p. Moreover, the algorithm runs in $O(n^{5+2\beta})$ time.

Theorem 7 (Theorem 2 in [14]): Consider the seeded Erdős-Rényi pair model $\mathcal{G}(N, \alpha, p, s)$ with $p = o(1)$ and $s = \Theta(1)$. Assume that

$$\alpha \geq \omega\left(\frac{1}{NI(p, s)^2}\right) \quad (22)$$

and

$$\alpha \geq \frac{2 \log N}{NI(p, s)}, \quad (23)$$

where $I(p, s) \triangleq 2ps \log \frac{1}{ps} + (2-2ps) \log \frac{1}{1-ps} + ps^2 \log ps^2 + 2ps(1-s) \log(ps-ps^2) + (1+ps^2-2ps) \log(1+ps^2-2ps) = (1+o(1))s^2 p \log \frac{1}{p}$ denotes the mutual information between a pair of correlated edges in G_1 and G'_2 . Then there exists a polynomial-time algorithm, namely, the TMS algorithm in [14], that achieves exact alignment w.h.p.

C. Specialization to the Bipartite Alignment

Consider an attributed Erdős-Rényi graph pair $(G_1, G_2) \sim \mathcal{G}(n, p, s_u; m, q, s_a)$ with $p = 0$ or $s_u = 0$. Then G_1 and G_2 reduce to two bipartite graphs with edges connected only between users and attributes. In this special case, conditions (1) and (2) reduce to a single condition: If there exists some positive constant $\epsilon > 0$ such that

$$mq s_a^2 \geq (1 + \epsilon) \log n, \quad (24)$$

then Algorithm ATTRICH achieves exact alignment w.h.p. In contrast, Corollary 1 in [1] implies that the maximum likelihood estimator exactly recovers π^* w.h.p if

$$mq s_a^2 \geq \log n + \omega(1). \quad (25)$$

Moreover, the maximum likelihood estimator can be computed in polynomial time by first computing the similarity score for each pair of vertices $(u, v) \in \mathcal{V}_1^u \times \mathcal{V}_2^u$ as $Z_{u,v} = |\mathcal{N}_1^a(u) \cap \mathcal{N}_2^a(v)|$ and then solving the balanced assignment problem using the famous Hungarian Algorithm first proposed in [24]. From conditions (24) and (25), we can see that the feasible region of Algorithm ATTRICH is completely covered by the feasible region of the Hungarian Algorithm. By the above argument, the maximum likelihood estimator requires $O(n^2 m)$

time complexity to compute the similarity score for all pairs of vertices and the Hungarian Algorithm can be implemented with $O(n^3)$ time complexity [25]. Therefore, if $m = \omega(n)$, the time complexity of the maximum likelihood estimator is $O(n^2m)$ and if $m = O(n)$, the time complexity of the maximum likelihood estimator is $O(n^3)$. In comparison, the time complexity of Algorithm ATTRICH in this special case is always $O(n^2m)$.

V. PROOF OF THEOREM 1

Recall that our algorithm consists of the following two steps: Step 1 (Align through attribute neighbors) produces a set $\mathcal{S}_1$ that we refer to as the set of anchor pairs based on the C_{ij} defined in (6); Step 2 (Align through user neighbors) aligns the remaining vertices based on the W_{ij} defined in (7). Moreover, recall that we assume without generality that the true underlying permutation π^* is the identity permutation. Our proof of Theorem 1 analyzes the following corresponding error events.

• Step 1 Error Event. Define

$$\mathcal{E}_1 \triangleq \{\exists(i, j) \in \mathcal{V}_1^u \times \mathcal{V}_2^u \text{ s.t. } i \neq j \text{ and } C_{ij} > x\}.$$

The event $\mathcal{E}_1^c$ guarantees that the anchor set found in the first step only contains correctly matched pairs.

• Step 2 Error Events. Define

$$\mathcal{E}_2 \triangleq \{\exists(i, i) \in \mathcal{U}_1 \times \mathcal{U}_2 \text{ s.t. } W_{ii} \leq y|\mathcal{S}_1|\}$$

and

$$\mathcal{E}_3 \triangleq \{\exists(i, j) \in \mathcal{U}_1 \times \mathcal{U}_2 \text{ s.t. } i \neq j \text{ and } W_{ij} > y|\mathcal{S}_1|\}.$$

In the special case that $\mathcal{U}_1 \times \mathcal{U}_2 = \emptyset$, i.e., all the vertices are matched in Step 1, we set events $\mathcal{E}_2$ and $\mathcal{E}_3$ to be empty by convention and thus $P(\mathcal{E}_2) = P(\mathcal{E}_3) = 0$. In the case that $\mathcal{U}_1 \times \mathcal{U}_2 \neq \emptyset$, event $\mathcal{E}_2^c \cap \mathcal{E}_3^c$ corresponds to the event that all non-anchor vertices are correctly matched through their user neighbors.

We first show that $\mathcal{E}_1^c \cap \mathcal{E}_2^c \cap \mathcal{E}_3^c$ implies that Algorithm ATTRICH does not declare failure and it outputs $\hat{\pi} = \pi^*$. Under event $\mathcal{E}_1^c$, it follows that there exists no conflicting pairs in $\mathcal{S}_1$, so Algorithm ATTRICH does not declare failure at Step 1 and for each vertex $i \in \mathcal{V}_1^u \setminus \mathcal{U}_1$, we have $\hat{\pi}(i) = i$. Furthermore, $\mathcal{E}_1^c$ implies that $\mathcal{U}_1 = \mathcal{U}_2$. Now we consider two different cases to complete the proof.

- $\mathcal{U}_1 = \mathcal{U}_2 = \emptyset$: In this case, all the vertices are correctly aligned through attribute neighbors. Therefore, Algorithm ATTRICH terminates at Step 1 and outputs $\hat{\pi} = \pi^*$.
- $\mathcal{U}_1 = \mathcal{U}_2 \neq \emptyset$: In this case, not all the vertices are aligned through attribute neighbors. Then event $\mathcal{E}_2^c \cap \mathcal{E}_3^c$ guarantees that for each $i \in \mathcal{U}_1$, we have

$$W_{ii} > (1 - \delta_y)|\mathcal{S}_1|ps_u^2$$

and

$$W_{ij} \leq (1 - \delta_y)|\mathcal{S}_1|ps_u^2, \forall j \in \mathcal{U}_2, i \neq j.$$

Thus, the algorithm does not declare failure at Step 2 and we have for each $i \in \mathcal{U}_1$, $\hat{\pi}(i) = i$. Finally, it follows that $\hat{\pi}$ is a bijection and $\hat{\pi} = \pi^*$.

Next, to prove Theorem 1, it suffices to show that $P(\mathcal{E}_1^c \cap \mathcal{E}_2^c \cap \mathcal{E}_3^c) = 1 - o(1)$. However, both events $\mathcal{E}_2$ and $\mathcal{E}_3$ are based on two random sets $\mathcal{U}_1$ and $\mathcal{U}_2$. If we apply the union bound to upper bound the probability of $\mathcal{E}_2$ and $\mathcal{E}_3$ without any restriction on the size of $\mathcal{U}_1$ and $\mathcal{U}_2$, the bound will be very loose. Our key idea of the proof is to analyze the error events of Step 2 by conditioning on a carefully chosen event. Specifically, let the set of vertex pairs *correctly matched* through attribute neighbors be denoted as

$$\mathcal{S}'_1 = \{(i, i) \in \mathcal{V}_1^u \times \mathcal{V}_2^u \text{ s.t. } C_{ii} > x\}.$$

Note that $\mathcal{S}'_1 \subseteq \mathcal{S}_1$ since $\mathcal{S}'_1$ only counts the correctly matched pairs. Then we consider the following auxiliary event:

$$\mathcal{A} \triangleq \{n - |\mathcal{S}'_1| < n^c\},$$

where $c \triangleq \max\left\{1 - \frac{mq s_a^2}{(1+\epsilon/2)\log n}, 0\right\}$ (cf. the constant ϵ in Theorem 1). To prove that $P(\mathcal{E}_1^c \cap \mathcal{E}_2^c \cap \mathcal{E}_3^c) = 1 - o(1)$, it then suffices to prove that $P(\mathcal{A} \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c \cap \mathcal{E}_3^c) = 1 - o(1)$. Notice that event $\mathcal{E}_1^c$ guarantees that the anchor set found in the first step only contains correctly matched pairs. It follows that event $\mathcal{E}_1^c$ implies $\mathcal{S}_1 = \mathcal{S}'_1$. Therefore, event $\mathcal{A} \cap \mathcal{E}_1^c$ restricts the size of the set $\mathcal{U}_1 \times \mathcal{U}_2$, so we can apply a much tighter union bound to upper bound the probability of $\mathcal{E}_2$ and $\mathcal{E}_3$ if we condition on the event $\mathcal{A} \cap \mathcal{E}_1^c$.

We now analyze $P(\mathcal{A} \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c \cap \mathcal{E}_3^c)$. Note that

$$\begin{aligned} P(\mathcal{A} \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c \cap \mathcal{E}_3^c) &= P(\mathcal{A} \cap \mathcal{E}_1^c)P(\mathcal{E}_2^c \cap \mathcal{E}_3^c | \mathcal{A} \cap \mathcal{E}_1^c) \\ &= (1 - P(\mathcal{A}^c \cup \mathcal{E}_1))(1 - P(\mathcal{E}_2 \cup \mathcal{E}_3 | \mathcal{A} \cap \mathcal{E}_1^c)) \\ &\geq (1 - P(\mathcal{A}^c) - P(\mathcal{E}_1))(1 - P(\mathcal{E}_2 | \mathcal{A} \cap \mathcal{E}_1^c) - P(\mathcal{E}_3 | \mathcal{A} \cap \mathcal{E}_1^c)). \end{aligned} \quad (26)$$

where (26) follows by the union bound. With Lemmas 1–4 below, it is easy to see that $P(\mathcal{A} \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c \cap \mathcal{E}_3^c) = 1 - o(1)$.

Lemma 1: Consider $\mathcal{G}(n, p, s_u; m, q, s_a)$ with $q = o(1)$ and $s_a = \Theta(1)$. Assume that

$$mq s_a^2 = \Omega(\log n).$$

Then $P(\mathcal{E}_1) = o(n^{-1/2})$.

Lemma 2: Consider $\mathcal{G}(n, p, s_u; m, q, s_a)$ with $q = o(1)$ and $s_a = \Theta(1)$. Assume that

$$mq s_a^2 = \Omega(\log n).$$

Then $P(\mathcal{A}) = 1 - o(1)$.

Lemma 3: Consider $\mathcal{G}(n, p, s_u; m, q, s_a)$ with $p = o(1)$ and $s_u = \Theta(1)$. Assume there exists some constant $\epsilon > 0$ such that

$$mq s_a^2 + nps_u^2 \geq (1 + \epsilon) \log n, \quad (27)$$

$$mq s_a^2 = \Omega(\log n). \quad (28)$$

Then $P(\mathcal{E}_2 | \mathcal{A} \cap \mathcal{E}_1^c) = n^{-\Theta(1)}$.

Lemma 4: Consider $\mathcal{G}(n, p, s_u; m, q, s_a)$ with $p = o(1)$ and $s_u = \Theta(1)$. Assume there exists some constant $\epsilon > 0$ such that

$$mq s_a^2 + nps_u^2 \geq (1 + \epsilon) \log n, \quad (29)$$

$$mqs_a^2 = \Omega(\log n). \quad (30)$$

Then $P(\mathcal{E}_3 | \mathcal{A} \cap \mathcal{E}_1^c) = n^{-\Theta(1)}$.

VI. PROOF OF THEOREM 2

Recall that Algorithm ATTRSPARSE consists of the following two steps: In Step 1 (Align through attribute neighbors), the algorithm produces a set $\mathcal{S}_2$, which we refer to as the set of anchor pairs based on C_{ij} defined in (6); In Step 2 (Align through user neighbors), the algorithm performs two different processes depending on the sparsity of the user-user connections. In case when the user-user connection is sparse $np \leq n^{1/7}$, the algorithm treats the anchors found in Step 1 as seeds and runs Algorithm III-C. In case the user-user connection is dense $np > n^{1/7}$, the algorithm explores the anchor vertices in the one-hop neighborhood of each non-anchor user vertex to align them. To prove Theorem 2, we first consider two error events associated with Step 1, and then we separately consider the two different cases for Step 2.

Define

$$\mathcal{E}_4 \triangleq \{\exists(i, j) \in \mathcal{V}_1^u \times \mathcal{V}_2^u \text{ s.t. } i \neq j \text{ and } C_{ij} \geq z\}.$$

Event $\mathcal{E}_4^c$ guarantees the anchors found in Step 1 only contain correctly matched pairs. Let

$$\mathcal{S}'_2 = \{(i, i) \in \mathcal{V}_1^u \times \mathcal{V}_2^u \text{ s.t. } C_{ii} \geq z\}$$

and define

$$\mathcal{E}_5 \triangleq \{|\mathcal{S}'_2| < n^{7/8}\}.$$

Event $\mathcal{E}_4^c \cap \mathcal{E}_5^c$ guarantees the size of anchor set is large. Given $\mathcal{E}_4^c \cap \mathcal{E}_5^c$, we first show that algorithm ATTRSPARSE outputs the correct permutation w.h.p., i.e., $P(\hat{\Pi} \neq \Pi^* | \mathcal{E}_4^c \cap \mathcal{E}_5^c) = o(1)$. We consider two different cases for np .

- $np \leq n^{1/7}$. Recall that in this regime of np , Algorithm III-C is applied in Step 2. Notice that the choice of set $\mathcal{S}_2$ only depends on the user-attribute edges, so it is independent of all user-user edges. Therefore, by symmetry, when we run Algorithm III-C, we can view the seed set $\mathcal{I}_0$ as a set chosen uniformly at random over all subsets with size $|\mathcal{I}_0|$ and thus Theorem 6 can directly be applied. So it suffices to show that the conditions in Theorem 6 are satisfied under event $\mathcal{E}_4^c \cap \mathcal{E}_5^c$. Event $\mathcal{E}_4^c$ implies that $\mathcal{I}_0 = \mathcal{S}_2 = \mathcal{S}'_2$ and event $\mathcal{E}_5^c$ implies that $|\mathcal{I}_0| \geq n^{7/8}$. Moreover, because we assume that $np s_a^2 - \log n = \omega(1)$, we have $P(\hat{\Pi} \neq \Pi^* | \mathcal{E}_4^c \cap \mathcal{E}_5^c) = o(1)$ by Theorem 6 with $\beta = 1/7$.
- $np > n^{1/7}$. Recall that in this regime of np , we perform a similar process as Step 2 of Algorithm ATTRICH. We align the non-anchor vertices by exploring the anchors in their one-hop neighborhood. Similarly to the analysis for Theorem 1, we define two error events associated with this steps. Define

$$\mathcal{E}_6 \triangleq \{\exists(i, i) \in \mathcal{U}_3 \times \mathcal{U}_4 \text{ s.t. } W_{ii} \leq y|\mathcal{S}_2|\}$$

and

$$\mathcal{E}_7 \triangleq \{\exists(i, j) \in \mathcal{U}_3 \times \mathcal{U}_4 \text{ s.t. } i \neq j \text{ and } W_{ij} > y|\mathcal{S}_2|\}.$$

From the proof of Theorem 1, we know that $\{\hat{\Pi} \neq \Pi^* | \mathcal{E}_4^c \cap \mathcal{E}_5^c\} \subset \{\mathcal{E}_6^c \cap \mathcal{E}_7^c | \mathcal{E}_4^c \cap \mathcal{E}_5^c\}$. Therefore, we have

$$\begin{aligned} P(\hat{\Pi} \neq \Pi^* | \mathcal{E}_4^c \cap \mathcal{E}_5^c) &\leq P(\mathcal{E}_6 \cup \mathcal{E}_7 | \mathcal{E}_4^c \cap \mathcal{E}_5^c) \\ &\leq P(\mathcal{E}_6 | \mathcal{E}_4^c \cap \mathcal{E}_5^c) + P(\mathcal{E}_7 | \mathcal{E}_4^c \cap \mathcal{E}_5^c). \end{aligned}$$

The statement $P(\hat{\Pi} \neq \Pi^* | \mathcal{E}_4^c \cap \mathcal{E}_5^c) = o(1)$ follows from the Lemma below.

Lemma 5: Consider $\mathcal{G}(n, p, s_u; m, q, s_a)$ with $p = o(1)$ and $s_u = \Theta(1)$. Suppose that $np > n^{1/7}$. Then we have

$$P(\mathcal{E}_6 | \mathcal{E}_4^c \cap \mathcal{E}_5^c) = n^{-\Theta(1)}.$$

Lemma 6: Consider $\mathcal{G}(n, p, s_u; m, q, s_a)$ with $p = o(1)$ and $s_u = \Theta(1)$. Suppose that $np > n^{1/7}$. Then we have

$$P(\mathcal{E}_7 | \mathcal{E}_4^c \cap \mathcal{E}_5^c) = n^{-\Theta(1)}.$$

Finally, we have

$$\begin{aligned} P(\hat{\Pi} \neq \Pi^*) &= P(\hat{\Pi} \neq \Pi^*, \mathcal{E}_4^c \cap \mathcal{E}_5^c) + P(\hat{\Pi} \neq \Pi^*, \mathcal{E}_4 \cup \mathcal{E}_5) \\ &\leq P(\hat{\Pi} \neq \Pi^* | \mathcal{E}_4^c \cap \mathcal{E}_5^c) + P(\mathcal{E}_4 \cup \mathcal{E}_5) \\ &\leq o(1) + P(\mathcal{E}_4) + P(\mathcal{E}_5). \end{aligned}$$

With Lemmas 7 and 8 below, it is easy to see that $P(\hat{\Pi} \neq \Pi^*) = o(1)$.

Lemma 7: Consider $\mathcal{G}(n, p, s_u; m, q, s_a)$ with $q = o(1)$ and $s_a = \Theta(1)$. Assume that there exists some constant $\tau > 0$ such that

$$mqs_a^2 \geq \frac{2 \log n}{\tau \log \frac{1}{q}}. \quad (31)$$

Then $P(\mathcal{E}_4) = o(1)$.

Lemma 8: Consider $\mathcal{G}(n, p, s_u; m, q, s_a)$ with $q = o(1)$ and $s_a = \Theta(1)$. Assume that

$$mqs_a^2 = o(\log n).$$

Then $P(\mathcal{E}_5) = o(1)$.

VII. PROOF OF LEMMAS

In this section, we prove the lemmas stated in the proofs of Theorems 1 and 2. To this end, we first state two technical lemmas that bound the binomial tail probability.

Lemma 9 (Theorem 1 in [26]): Let $X \sim \text{Binom}(n, \theta)$. Then we have:

- $P(X \geq n\theta + \lambda) \leq \exp(-nD_{\text{KL}}(\theta + \frac{\lambda}{n} || \theta))$ for $0 < \lambda < n - n\theta$;
- $P(X \leq n\theta - \lambda) \leq \exp(-nD_{\text{KL}}(\theta - \frac{\lambda}{n} || \theta))$ for $0 < \lambda < n\theta$,

where $D_{\text{KL}}(x || y) \triangleq x \log \frac{x}{y} + (1 - x) \log \frac{1-x}{1-y}$ denotes the Kullback–Leibler divergence between $\text{Bern}(x)$ and $\text{Bern}(y)$.

Lemma 10 (Lemma 4.7.2 in [27]): Let $X \sim \text{Binom}(n, \theta)$. Then we have

$$P(X \geq \lambda) \geq \frac{1}{\sqrt{8\lambda(1 - \lambda/n)}} \exp\{-nD_{\text{KL}}(\frac{\lambda}{n} || \theta)\}$$

for any $n\theta < \lambda \leq n$.

The following lemma gives an approximation for the Kullback–Leibler divergence term in Lemmas 9 and 10.

Lemma 11 (Approximation of KL-Divergence): Assume that $\theta = o(1)$, $s = \Theta(1)$, and $\Delta = \Theta(1)$. Then we have

$$D_{\text{KL}}\left(\frac{\Delta}{\log 1/\theta} \theta s^2 \middle| \theta s^2\right) = \theta s^2 + o(\theta), \quad (32)$$

$$D_{\text{KL}}\left(\frac{\Delta}{\log 1/\theta} \theta s^2 \middle| \theta^2 s^2\right) = \Delta \theta s^2 + o(\theta). \quad (33)$$

With the three technical lemmas above, we are ready to prove Lemmas 1-8.

A. Proof of Lemma 1

Recall that we defined error event $\mathcal{E}_1 \triangleq \{\exists(i, j) \in \mathcal{V}_1^u \times \mathcal{V}_2^u \text{ s.t. } i \neq j \text{ and } C_{ij} > x\}$. Here we prove $P(\mathcal{E}_1) = o(n^{-1/2})$ under the assumption that $mq s_a^2 = \Omega(\log n)$. To bound the probability of $\mathcal{E}_1$, we first consider distribution of random variable C_{ij} . For two different vertices $i \in \mathcal{V}_1^u$ and $j \in \mathcal{V}_2^u$ such that $i \neq j$, it follows from the definition of the model $\mathcal{G}(n, p, s_u; m, q, s_a)$ that $C_{ij} \sim \text{Binom}(m, q^2 s_a^2)$. Moreover, notice that

$$x = \frac{\Delta_x}{\log \frac{1}{q}} mq s_a^2 = \omega(mq^2 s_a^2) = \omega(E[C_{ij}]),$$

because $\Delta_x = \Theta(1)$ and $q = o(1)$. Now, we can upper bound the probability of error event $\mathcal{E}_1$ as

$$P(\mathcal{E}_1) = P\{\exists(i, j) \in \mathcal{V}_1^u \times \mathcal{V}_2^u : i \neq j \text{ and } C_{ij} > x\} \leq (n^2 - n)P(C_{12} > x) \quad (34)$$

$$\leq n^2 \exp(-m D_{\text{KL}}\left(\frac{x}{m} \middle| q^2 s_a^2\right)) \quad (35)$$

$$= n^2 \exp\left(-m D_{\text{KL}}\left(\frac{\Delta_x}{\log \frac{1}{q}} q s_a^2 \middle| q^2 s_a^2\right)\right) = n^2 \exp(-m(\Delta_x q s_a^2 + o(q))) \quad (36)$$

$$= \exp(2 \log n - m(\Delta_x q s_a^2 + o(q))) = o(n^{-1/2}), \quad (37)$$

where (34) follows from the union bound, (35) follows from Lemma 9, (36) follows from Lemma 11 and (37) follows since constant $\Delta_x \geq \frac{3 \log n}{mq s_a^2}$.

B. Proof of Lemma 2

Recall that we defined auxiliary event $\mathcal{A} \triangleq \{n - |S'_1| < n^c\}$, where $S'_1 = \{(i, i) \in \mathcal{V}_1^u \times \mathcal{V}_2^u \text{ s.t. } C_{ii} > x\}$. Here we prove $P(\mathcal{A}) = 1 - o(1)$ under the assumption that $mq s_a^2 = \Omega(\log n)$. To show this, we first consider the distribution of random variable C_{ii} and upper bound the probability of event $\{C_{ii} \leq x\}$. For each vertex $i \in \mathcal{V}_1^u$, it follows from the definition of the model $\mathcal{G}(n, p, s_u; m, q, s_a)$ that $C_{ii} \sim \text{Binom}(m, q s_a^2)$. Notice that

$$x = \frac{\Delta_x}{\log \frac{1}{q}} mq s_a^2 = o(mq s_a^2) = o(E[C_{ii}]),$$

because $q = o(1)$ and $\Delta_x = \Theta(1)$. We can upper bound the probability of the tail event $\{C_{ii} \leq x\}$ using Lemma 9:

$$P(C_{ii} \leq x) \leq \exp\left(-m D_{\text{KL}}\left(\frac{x}{m} \middle| q s_a^2\right)\right) = \exp\left(-m D_{\text{KL}}\left(\frac{\Delta_x}{\log \frac{1}{q}} q s_a^2 \middle| q s_a^2\right)\right)$$

$$= \exp(-m(q s_a^2 + o(q))), \quad (38)$$

where (38) follows from Lemma 11. For simplicity, we denote $\exp(-m D_{\text{KL}}(\frac{\Delta_x}{\log \frac{1}{q}} q s_a^2 \middle| q s_a^2))$ by γ from this point. By (38) and the assumption that $mq s_a^2 = \Omega(\log n)$, we have $\gamma = o(1)$.

Furthermore, notice that for each different $i \in \mathcal{V}_1^u$, the random variable C_{ii} are independent and identically distributed. Therefore, the number of vertices $i \in \mathcal{V}_1^u$ with $C_{ii} \leq x$ is distributed according to the Binomial distribution

$$n - |S'_1| \sim \text{Binom}(n, P(C_{ii} \leq x)).$$

By the upper bound (38), for any positive number z , we have

$$P(n - |S'_1| \geq z) \leq P(\text{Binom}(n, \gamma) \geq z). \quad (39)$$

Recall that $c \triangleq \max\left\{1 - \frac{mq s_a^2}{(1+\epsilon/2) \log n}, 0\right\}$, where ϵ is a positive constant. We have

$$\frac{n^{c-1}}{\gamma} \geq n^{-\frac{mq s_a^2}{(1+\epsilon/2) \log n}} \exp\left(m D_{\text{KL}}\left(\frac{\Delta_x}{\log \frac{1}{q}} q s_a^2 \middle| q s_a^2\right)\right) = \exp\left(-\frac{mq s_a^2}{1+\epsilon/2} + mq s_a^2 + o(mq)\right) \quad (40)$$

$$= \omega(1), \quad (41)$$

where (40) follows from (38), and (41) follows since $mq s_a^2 = \Omega(\log n)$. Finally, we can upper bound the probability of event $\mathcal{A}^c$ as

$$P(\mathcal{A}^c) = P(n - |S'_1| \geq n^c) \leq P(\text{Binom}(n, \gamma) \geq n^c) \quad (42)$$

$$\leq \exp(-n D_{\text{KL}}(n^{c-1} \middle| \gamma)) \quad (43)$$

$$= \exp\left(-n\left(n^{c-1} \log \frac{n^{c-1}}{\gamma} + (1 - n^{c-1}) \log \frac{1 - n^{c-1}}{1 - \gamma}\right)\right)$$

$$= \exp\left(-n\left(n^{c-1} \log \frac{n^{c-1}}{\gamma} + (1 - n^{c-1}) \frac{\gamma - n^{c-1}}{1 - \gamma} (1 + o(1))\right)\right) \quad (44)$$

$$= \exp(-n(\omega(n^{c-1}) - n^{c-1}(1 + o(1)))) \quad (45)$$

$$= \exp(-n \cdot \omega(n^{c-1}))$$

$$= o(1), \quad (46)$$

where (42) follows from (39), (43) follows from Lemma 9, (44) follow from the Taylor expansion of the function $\log x$ and the fact that $n^{c-1} = o(1)$ and $\gamma = o(1)$, (45) follows since $n^{c-1} = \omega(\gamma)$ and finally, (46) follows from the fact that $c \geq 0$.

C. Proof of Lemma 3

Here we prove that the conditions (27) and (28) imply $P(\mathcal{E}_2 | \mathcal{A} \cap \mathcal{E}_1^c) = n^{-\Theta(1)}$. Recall that we defined events

$$\mathcal{E}_2 = \{\exists(i, i) \in \mathcal{U}_1 \times \mathcal{U}_2 : W_{ii} \leq y | S_1\},$$

$$\mathcal{A} = \{n - |S'_1| < n^c\},$$

$$\mathcal{E}_1 = \{\exists(i, j) \in \mathcal{V}_1^u \times \mathcal{V}_2^u \text{ s.t. } i \neq j \text{ and } C_{ij} > x\}.$$

To analyze the conditional event $\mathcal{E}_2 | \mathcal{A} \cap \mathcal{E}_1^c$, first notice that $\mathcal{E}_1^c$ implies that only identical pairs are in the anchor set, i.e., $\mathcal{U}_1 =$

$\mathcal{U}_2$. Thus, we obtain a simplified expression of the conditional event as

$$\mathcal{E}_2 | \mathcal{A} \cap \mathcal{E}_1^c = \{\exists i \in \mathcal{U}_1 : W_{ii} \leq y | \mathcal{S}_1\} | \mathcal{A} \cap \mathcal{E}_1^c.$$

The condition on the auxiliary event $\mathcal{A}$ further implies that the number of identical pairs that are not discovered in the anchor set is at most n^c , i.e., $|\mathcal{U}_1| < n^c$. Here, recall that $c = \max \left\{ 1 - \frac{mq s_a^2}{(1+\epsilon/2) \log n}, 0 \right\}$. We note that $c = 0$ implies that the unmatched users set $\mathcal{U}_1 = \emptyset$, which is taken care of by a separate analysis in the proof of Theorem 1. For the remaining analysis in this Lemma, we only consider the case where $\mathcal{U}_1 \neq \emptyset$, and consequently we have $c = 1 - \frac{mq s_a^2}{(1+\epsilon/2) \log n}$.

Applying the union bound on the conditional error event, we have

$$\begin{aligned} P(\mathcal{E}_2 | \mathcal{A} \cap \mathcal{E}_1^c) &= P(\exists i \in \mathcal{U}_1, W_{ii} \leq y | \mathcal{S}_1 | \mathcal{A} \cap \mathcal{E}_1^c) \\ &= \sum_{k=0}^{n^c} P(\exists i \in \mathcal{U}_1, W_{ii} \leq y | \mathcal{S}_1 | |\mathcal{U}_1| = k, \mathcal{E}_1^c) \\ &\quad \cdot P(|\mathcal{U}_1| = k | \mathcal{A} \cap \mathcal{E}_1^c) \end{aligned} \quad (47)$$

$$\leq \sum_{k=0}^{n^c} k P(W_{11} \leq y | \mathcal{S}_1 | |\mathcal{U}_1| = k, \mathcal{E}_1^c) P(|\mathcal{U}_1| = k | \mathcal{A} \cap \mathcal{E}_1^c) \quad (48)$$

$$\leq \max_{k \in [0, n^c]} \{k P(W_{11} \leq y | \mathcal{S}_1 | |\mathcal{U}_1| = k, \mathcal{E}_1^c)\}. \quad (49)$$

In (47), we have n^c as the upper limit in the summation because of conditioning on $\mathcal{A}$. Equation (48) follows from the union bound.

Next, we upper bound $P(W_{11} \leq y | \mathcal{S}_1 | |\mathcal{U}_1| = k, \mathcal{E}_1^c)$. Recall

$$W_{11} = \sum_{v \in \mathcal{N}_1^u(1), u \in \mathcal{N}_2^u(1)} \mathbb{1}_{\{(v,u) \in \mathcal{S}_1\}}$$

counts the number of aligned anchor neighbors of a user vertex. To see the conditional distribution of W_{11} , notice that conditioned on events $\mathcal{E}_1^c$ and $|\mathcal{U}_1| = k$, the whole anchor set $\mathcal{S}_1$ only contains identical pairs and is of size $n - k$. Thus, we get the following simpler expression

$$\begin{aligned} W_{11} | \{|\mathcal{U}_1| = k, \mathcal{E}_1^c\} &= \sum_{v \in \mathcal{N}_1^u(1), v \in \mathcal{N}_2^u(1)} \mathbb{1}_{\{(v,v) \in \mathcal{S}_1\}} \\ &= \sum_{(v,v) \in \mathcal{S}_1} \mathbb{1}_{\{v \in \mathcal{N}_1^u(1), v \in \mathcal{N}_2^u(1)\}}. \end{aligned}$$

Here this random variable $W_{11} | \{|\mathcal{U}_1| = k, \mathcal{E}_1^c\}$ is the summation of $|\mathcal{S}_1| = n - k$ independent and identically distributed Bernoulli random variables. Each Bernoulli random variable takes value 1 when a pair of anchor vertices connect to vertex 1 in both G_1 and G_2 , and this happens with probability ps_u^2 . We therefore have $W_{11} | \{|\mathcal{U}_1| = k, \mathcal{E}_1^c\} \sim \text{Binom}(n-k, ps_u^2)$. To upper bound the probability $P(W_{11} \leq y | \mathcal{S}_1 | |\mathcal{U}_1| = k, \mathcal{E}_1^c)$ in (49), we use the Chernoff bound from Lemma 9 and get

$$\begin{aligned} P(W_{11} \leq y | \mathcal{S}_1 | |\mathcal{U}_1| = k, \mathcal{E}_1^c) \\ \leq \exp \left\{ -(n-k) D_{\text{KL}} \left(\frac{\Delta_y}{\log 1/p} ps_u^2 || ps_u^2 \right) \right\}. \end{aligned} \quad (50)$$

Plugging (50) into the previous (49), we finally have

$$P(\mathcal{E}_2 | \mathcal{A} \cap \mathcal{E}_1^c)$$

$$\begin{aligned} &\leq \max_{k \in [0, n^c]} \left\{ k \exp \left\{ -(n-k) D_{\text{KL}} \left(\frac{\Delta_y}{\log 1/p} ps_u^2 || ps_u^2 \right) \right\} \right\} \\ &= n^c \exp \left\{ -(n-n^c) D_{\text{KL}} \left(\frac{\Delta_y}{\log 1/p} ps_u^2 || ps_u^2 \right) \right\} \\ &= \exp \{ c \log n - (n-n^c) (ps_u^2 + o(ps_u^2)) \} \end{aligned} \quad (51)$$

$$\leq \exp \left\{ \log n - \frac{mq s_a^2}{(1+\epsilon/2)} - nps_u^2 + o(np) \right\} \quad (52)$$

$$\leq \exp \left\{ \log n - \frac{mq s_a^2 + nps_u^2}{(1+\epsilon/2)} \right\} \quad (53)$$

$$\begin{aligned} &\leq \exp \left\{ -\frac{\epsilon/2}{1+\epsilon/2} \log n \right\} \\ &= n^{-\Theta(1)}. \end{aligned} \quad (54)$$

Here (51) follows from the KL-divergence approximation (32) in Lemma 11. We get (52) by plugging in $c = 1 - \frac{mq s_a^2}{(1+\epsilon/2) \log n}$ and applying assumption (28) $mq s_a^2 = \Omega(\log n)$. Equation (53) follows because $o(np) \leq \frac{\epsilon/2}{1+\epsilon/2} nps_u^2 = \Theta(np)$. Equation (54) follows from the condition (27), which requires that $mq s_a^2 + nps_u^2 \geq (1+\epsilon) \log n$ for a constant ϵ .

D. Proof of Lemma 4

Here we prove that conditions (29) and (30) imply $P(\mathcal{E}_3 | \mathcal{A} \cap \mathcal{E}_1^c) = n^{-\Theta(1)}$. The conditioned events here are exactly the same as those of Lemma 3, and we use a similar proof strategy. Recall that we defined the event

$$\mathcal{E}_3 \triangleq \{\exists (i, j) \in \mathcal{U}_1 \times \mathcal{U}_2 \text{ s.t. } i \neq j \text{ and } W_{ij} > y | \mathcal{S}_1\}.$$

To analyze the conditional event $\mathcal{E}_3 | \mathcal{A} \cap \mathcal{E}_1^c$, we reuse the same observation from the proof of Lemma 3, that only identical pairs are in the anchor set. Thus, we are able to simplify the expression as

$$\mathcal{E}_3 | \mathcal{A} \cap \mathcal{E}_1^c = \{\exists i, j \in \mathcal{U}_1, i \neq j : W_{ij} > y | \mathcal{S}_1\} | \mathcal{A} \cap \mathcal{E}_1^c,$$

where the number of unaligned identical pairs $|\mathcal{U}_1| < n^c$. Here, we also have $c = 1 - \frac{mq s_a^2}{(1+\epsilon/2) \log n}$, because $c = 0$ implies that the unmatched users set $\mathcal{U}_1 = \emptyset$, which is taken care of by a separate analysis in the proof of Theorem 1.

Applying the union bound on the conditional error event, we get

$$\begin{aligned} P(\mathcal{E}_3 | \mathcal{A} \cap \mathcal{E}_1^c) &= P(\exists i, j \in \mathcal{U}_1, i \neq j, W_{ij} > y | \mathcal{S}_1 | \mathcal{A} \cap \mathcal{E}_1^c) \\ &= \sum_{k=0}^{n^c} P(\exists i, j \in \mathcal{U}_1, i \neq j, W_{ij} > y | \mathcal{S}_1 | |\mathcal{U}_1| = k, \mathcal{E}_1^c) \\ &\quad \times P(|\mathcal{U}_1| = k | \mathcal{A} \cap \mathcal{E}_1^c) \\ &\leq \sum_{k=0}^{n^c} k^2 P(W_{12} > y | \mathcal{S}_1 | |\mathcal{U}_1| = k, \mathcal{E}_1^c) P(|\mathcal{U}_1| = k | \mathcal{A} \cap \mathcal{E}_1^c) \end{aligned} \quad (55)$$

$$\leq \max_{k \in [0, n^c]} \{k^2 P(W_{12} > y | \mathcal{S}_1 | |\mathcal{U}_1| = k, \mathcal{E}_1^c)\}. \quad (56)$$

Here we further upper bound $P(W_{12} > y | \mathcal{S}_1 | |\mathcal{U}_1| = k, \mathcal{E}_1^c)$ in (56). Recall that

$$W_{12} = \sum_{v \in \mathcal{N}_1^u(1), u \in \mathcal{N}_2^u(2)} \mathbb{1}_{\{(v,u) \in \mathcal{S}_1\}},$$

represents the number of aligned anchor neighbors of user vertex 1 in G_1 and user vertex 2 in G_2 . Notice that conditioned on events $\mathcal{E}_1^c$ and $|\mathcal{U}_1| = k$, the anchor set $\mathcal{S}_1$ only contains identical pairs and is of size $n - k$. Thus, we get the following simpler expression

$$\begin{aligned} W_{12} \mid \{|\mathcal{U}_1| = k, \mathcal{E}_1^c\} &= \sum_{v \in \mathcal{N}_1^u(1), v \in \mathcal{N}_2^u(2)} \mathbb{1}_{\{(v,v) \in \mathcal{S}_1\}} \\ &= \sum_{(v,v) \in \mathcal{S}_1} \mathbb{1}_{\{v \in \mathcal{N}_1^u(1), v \in \mathcal{N}_2^u(2)\}} \end{aligned}$$

Here the random variable $W_{12} \mid \{|\mathcal{U}_1| = k, \mathcal{E}_1^c\}$ is the summation of $n - k$ independent and identically distributed Bernoulli random variables. Each Bernoulli random variable takes value 1 when a pair of anchor vertices in $\mathcal{S}_1$ connect to vertex 1 in G_1 and vertex 2 in G_2 and this happens with probability $p^2 s_u^2$. Therefore, we have $W_{12} \mid \{|\mathcal{U}_1| = k, \mathcal{E}_1^c\} \sim \text{Binom}(n - k, p^2 s_u^2)$. We then apply Chernoff bound in Lemma 9 and get

$$\begin{aligned} &\mathbb{P}(W_{12} > y \mid \mathcal{S}_1 \mid |\mathcal{U}_1| = k, \mathcal{E}_1^c) \\ &\leq \exp \left\{ -(n - k) D_{\text{KL}} \left(\frac{\Delta_y}{\log 1/p} p s_u^2 \parallel p^2 s_u^2 \right) \right\} \end{aligned} \quad (57)$$

Plugging (57) into (56), we have

$$\begin{aligned} &\mathbb{P}(\mathcal{E}_3 \mid \mathcal{A} \cap \mathcal{E}_1^c) \\ &\leq \max_{k \in [0, n^c]} \left\{ k^2 \exp \left\{ -(n - k) D_{\text{KL}} \left(\frac{\Delta_y}{\log 1/p} p s_u^2 \parallel p^2 s_u^2 \right) \right\} \right\} \\ &= n^{2c} \exp \left\{ -(n - n^c) D_{\text{KL}} \left(\frac{\Delta_y}{\log 1/p} p s_u^2 \parallel p^2 s_u^2 \right) \right\} \\ &\leq \exp \{ 2c \log n - (n - n^c) (\Delta_y p s_u^2 + o(p)) \} \end{aligned} \quad (58)$$

$$\leq \exp \left\{ 2 \log n - \frac{2mq s_a^2}{1 + \epsilon/2} - \Delta_y n p s_u^2 + o(pn) \right\} \quad (59)$$

$$\leq \exp \left\{ 2 \log n - \frac{2mq s_a^2 + 2np s_u^2}{1 + \epsilon/2} \right\} \quad (60)$$

$$\leq \exp \left\{ -\frac{\epsilon}{1 + \epsilon/2} \log n \right\} \quad (61)$$

$$= n^{-\Theta(1)}. \quad (62)$$

Here (58) follows from the KL-divergence approximation formula (33) in Lemma 11. We get (59) by plugging in $c = 1 - \frac{mq s_a^2}{(1 + \epsilon/2) \log n}$ and applying assumption (30) $mq s_a^2 = \Omega(\log n)$. Equation (60) follows from the fact that $o(np) \leq \frac{\epsilon/2 \Delta_y}{1 + \epsilon/2} n p s_u^2 = \Theta(np)$, and condition that $\Delta_y \geq 2$. Equation (61) follows from the condition (29).

E. Proof of Lemma 5

Recall that we defined events

$$\mathcal{E}_4 = \{\exists(i, j) \in \mathcal{V}_1^u \times \mathcal{V}_2^u \text{ s.t. } i \neq j \text{ and } C_{ij} \geq z\},$$

$$\mathcal{E}_5 = \{|\mathcal{S}'_2| < n^{7/8}\},$$

$$\mathcal{E}_6 = \{\exists(i, i) \in \mathcal{U}_3 \times \mathcal{U}_4 \text{ s.t. } W_{ii} \leq y \mid \mathcal{S}_2\}.$$

Here we prove that $\mathbb{P}(\mathcal{E}_6 \mid \mathcal{E}_4^c \cap \mathcal{E}_5^c) = o(1)$ under the assumption that $np > n^{1/7}$. By the analogous argument as for equations (49) and (50) respectively, we have that

$$\mathbb{P}(\mathcal{E}_6 \mid \mathcal{E}_4^c \cap \mathcal{E}_5^c)$$

$$\leq \max_{k \in [0, n - n^{7/8}]} \{k \mathbb{P}(W_{11} \leq y \mid \mathcal{S}_2 \mid |\mathcal{U}_3| = k, \mathcal{E}_4^c)\}, \quad (63)$$

and

$$\begin{aligned} &\mathbb{P}(W_{11} \leq y \mid \mathcal{S}_2 \mid |\mathcal{U}_3| = k, \mathcal{E}_4^c) \\ &\leq \exp \left\{ -(n - k) D_{\text{KL}} \left(\frac{\Delta_y}{\log 1/p} p s_u^2 \parallel p s_u^2 \right) \right\}. \end{aligned} \quad (64)$$

Finally, plugging equation (64) into equation (63) gives

$$\begin{aligned} &\mathbb{P}(\mathcal{E}_6 \mid \mathcal{E}_4^c \cap \mathcal{E}_5^c) \\ &\leq \max_{k \in [0, n - n^{7/8}]} \left\{ k \exp \left\{ -(n - k) D_{\text{KL}} \left(\frac{\Delta_y}{\log 1/p} p s_u^2 \parallel p s_u^2 \right) \right\} \right\} \\ &\leq n \exp \left(-n^{7/8} D_{\text{KL}} \left(\frac{\Delta_y}{\log 1/p} p s_u^2 \parallel p s_u^2 \right) \right) \\ &\leq \exp(\log n - n^{7/8} (p s_u^2 + o(p s_u^2))) \end{aligned} \quad (65)$$

$$= n^{-\Theta(1)}, \quad (66)$$

where (65) follows from the KL-divergence approximation (32) in Lemma 11 and (66) follows because $np > n^{1/7}$.

F. Proof of Lemma 6

Recall that we defined event

$$\mathcal{E}_7 \triangleq \{\exists(i, j) \in \mathcal{U}_3 \times \mathcal{U}_4 \text{ s.t. } i \neq j \text{ and } W_{ij} > y \mid \mathcal{S}_2\}.$$

Here we prove that $\mathbb{P}(\mathcal{E}_7 \mid \mathcal{E}_4^c \cap \mathcal{E}_5^c) = o(1)$ under the assumption that $np > n^{1/7}$. By the analogous argument as for equations (56) and (57), we have that

$$\begin{aligned} &\mathbb{P}(\mathcal{E}_7 \mid \mathcal{E}_4^c \cap \mathcal{E}_5^c) \\ &\leq \max_{k \in [0, n - n^{7/8}]} \{k^2 \mathbb{P}(W_{12} > y \mid \mathcal{S}_2 \mid |\mathcal{U}_3| = k, \mathcal{E}_4^c)\}, \end{aligned} \quad (67)$$

and

$$\begin{aligned} &\mathbb{P}(W_{12} > y \mid \mathcal{S}_2 \mid |\mathcal{U}_3| = k, \mathcal{E}_4^c) \\ &\leq \exp \left\{ -(n - k) D_{\text{KL}} \left(\frac{\Delta_y}{\log 1/p} p s_u^2 \parallel p^2 s_u^2 \right) \right\}. \end{aligned} \quad (68)$$

Plugging equation (68) into equation (67) gives

$$\begin{aligned} &\mathbb{P}(\mathcal{E}_7 \mid \mathcal{E}_4^c \cap \mathcal{E}_5^c) \\ &\leq \max_{k \in [0, n - n^{7/8}]} \left\{ k^2 \exp \left\{ -(n - k) D_{\text{KL}} \left(\frac{\Delta_y}{\log 1/p} p s_u^2 \parallel p^2 s_u^2 \right) \right\} \right\} \\ &\leq n^2 \exp \left(-n^{7/8} D_{\text{KL}} \left(\frac{\Delta_y}{\log 1/p} p s_u^2 \parallel p^2 s_u^2 \right) \right) \\ &\leq \exp(2 \log n - n^{7/8} (\Delta_y p s_u^2 + o(p))) \end{aligned} \quad (69)$$

$$= o(1), \quad (70)$$

where (69) follows from the KL-divergence approximation formula (33) in Lemma 11 and (70) follows because $np > n^{1/7}$ and $\Delta_y = \Theta(1)$.

G. Proof of Lemma 7

Recall that we defined error event $\mathcal{E}_4 \triangleq \{\exists(i, j) \in \mathcal{V}_1^u \times \mathcal{V}_2^u \text{ s.t. } i \neq j \text{ and } C_{ij} \geq z\}$. Here we prove $\mathbb{P}(\mathcal{E}_4) = o(1)$ under the assumption that $mq s_a^2 \geq \frac{2 \log n}{\tau \log \frac{1}{q}}$. To bound the probability of $\mathcal{E}_4$, we first consider the distribution of random variable C_{ij} . For two different vertices $i \in \mathcal{V}_1^u$ and $j \in \mathcal{V}_2^u$ such that $i \neq j$, it follows from the definition of the model

$\mathcal{G}(n, p, s_u; m, q, s_a)$ that $C_{ij} \sim \text{Binom}(m, q^2 s_a^2)$. Moreover, notice that

$$z = (1 + \tau)mqs_a^2 = \omega(mq^2 s_a^2) = \omega(\mathbb{E}[C_{ij}]),$$

since $q = o(1)$. Therefore, we can upper bound the probability of error event $\mathcal{E}_4$ as

$$\begin{aligned} \mathbb{P}(\mathcal{E}_4) &= \mathbb{P}\{\exists(i, j) \in \mathcal{V}_1^u \times \mathcal{V}_2^u : i \neq j \text{ and } C_{ij} \geq z\} \\ &\leq (n^2 - n)\mathbb{P}(C_{12} \geq z) \end{aligned} \quad (71)$$

$$\begin{aligned} &\leq n^2 \exp(-mD_{\text{KL}}(\frac{z}{m} \| q^2 s_a^2)) \\ &= n^2 \exp(-mD_{\text{KL}}((1 + \tau)qs_a^2 \| q^2 s_a^2)) \\ &= n^2 \exp\left(-m(1 + \tau)qs_a^2 \log \frac{1 + \tau}{q}\right) \\ &\quad \cdot \exp\left(-m(1 - (1 + \tau)qs_a^2) \log \frac{1 - (1 + \tau)qs_a^2}{1 - q^2 s_a^2}\right) \\ &= n^2 \exp\left(-m(1 + \tau)qs_a^2 \log \frac{1 + \tau}{q}\right) \\ &\quad \cdot \exp\left((1 + o(1))m(1 - (1 + \tau)qs_a^2) \frac{(1 + \tau)qs_a^2 - q^2 s_a^2}{1 - q^2 s_a^2}\right) \\ &= n^2 \exp\left(-m(1 + \tau)qs_a^2 \log \frac{1 + \tau}{q}\right) \\ &\quad \cdot \exp((1 + o(1))m(1 + \tau)qs_a^2) \\ &= \exp\left(2 \log n - m(1 - o(1))(1 + \tau)qs_a^2 \log \frac{1}{q}\right) \\ &= \exp(-\Omega(\log n)) \\ &= n^{-\Omega(1)}, \end{aligned} \quad (72)$$

where (71) follows by the union bound, (72) follows from Lemma 9, and (73) follows since in (5) we assume that $m q \log \frac{1}{q} \tau s_a^2 \geq 2 \log n$ and $\tau = \Theta(1)$.

H. Proof of Lemma 8

Recall that we defined error event $\mathcal{E}_5 \triangleq \{|S'_2| < n^{7/8}\}$, where $S'_2 = \{(i, i) \in \mathcal{V}_1^u \times \mathcal{V}_2^u \text{ s.t. } C_{ii} \geq z\}$. Here we prove $\mathbb{P}(\mathcal{E}_5) = o(1)$ under the assumption that $mqs_a^2 = o(\log n)$. To show this, we first consider the distribution of random variable C_{ii} and lower bound the probability of event $\{C_{ii} \geq z\}$. For each vertex $i \in \mathcal{V}_1^u$, it follows from the definition of the model $\mathcal{G}(n, p, s_u; m, q, s_a)$ that $C_{ii} \sim \text{Binom}(m, qs_a^2)$. Notice that

$$z = (1 + \tau)mqs_a^2 > \mathbb{E}[C_{ii}].$$

We can lower bound the probability of the tail event $\{C_{ii} \geq z\}$ using Lemma 10 as

$$\begin{aligned} \mathbb{P}(C_{ii} \geq z) &\geq \frac{1}{\sqrt{8z(1 - z/m)}} \exp(-mD_{\text{KL}}(\frac{z}{m} \| qs_a^2)) \\ &\geq \frac{1}{\sqrt{8z}} \exp(-mD_{\text{KL}}((1 + \tau)qs_a^2 \| qs_a^2)) \\ &= \exp\left(-\frac{1}{2} \log(8(1 + \tau)mqs_a^2)\right) \\ &\quad \cdot \exp(-m(1 + \tau)qs_a^2 \log(1 + \tau)) \end{aligned}$$

$$\begin{aligned} &\cdot \exp\left(-m(1 - (1 + \tau)qs_a^2) \log \frac{1 - (1 + \tau)qs_a^2}{1 - qs_a^2}\right) \\ &= \exp\left(-\frac{1}{2} \log(8(1 + \tau)mqs_a^2)\right) \\ &\quad \cdot \exp(-m(1 + \tau)qs_a^2 \log(1 + \tau)) \\ &\quad \cdot \exp\left((1 + o(1))(1 - (1 + \tau)qs_a^2)m \frac{\tau qs_a^2}{1 - qs_a^2}\right) \\ &= \exp\left(-\frac{1}{2} \log(8(1 + \tau)mqs_a^2)\right) \\ &\quad \cdot \exp(-(1 + o(1))((1 + \tau) \log(1 + \tau) - \tau)mqs_a^2) \\ &= \exp(-o(\log n)) \\ &\geq n^{-1/9}, \end{aligned} \quad (73)$$

where (74) follows since $mqs_a^2 = o(\log n)$ and that $\tau = \Theta(1)$. Furthermore, notice that for each different $i \in \mathcal{V}_1^u$, the random variable C_{ii} are independent and identically distributed. Therefore, the number of vertices $i \in \mathcal{V}_1^u$ with $C_{ii} \geq z$ is distributed according to the Binomial distribution

$$|S'_2| \sim \text{Binom}(n, \mathbb{P}(C_{ii} \geq z)).$$

By the lower bound (75), for any positive number w , we have

$$\mathbb{P}(|S'_2| \leq w) \leq \mathbb{P}\left(\text{Binom}(n, n^{-1/9}) \leq w\right). \quad (76)$$

Finally we can upper bound the probability of event $\mathcal{E}_5$ as

$$\begin{aligned} \mathbb{P}(\mathcal{E}_5) &= \mathbb{P}(|S'_2| < n^{7/8}) \\ &\leq \mathbb{P}\left(\text{Binom}(n, n^{-1/9}) \leq n^{7/8}\right) \\ &\leq \exp\left(-nD_{\text{KL}}(n^{-1/8} \| n^{-1/9})\right) \\ &= \exp\left(-n \cdot n^{-1/8} \log(n^{-1/72})\right) \\ &\quad \cdot \exp\left(-n(1 - n^{-1/8}) \log \frac{1 - n^{-1/8}}{1 - n^{-1/9}}\right) \\ &= \exp\left(\frac{1}{72} n^{7/8} \log n - (1 + o(1))n^{8/9}\right) \\ &= o(1), \end{aligned} \quad (77)$$

where (77) follows from (76) and (78) follows from Lemma 9.

I. Proof of Lemma 11

(1) Proof for equation (32): From the definition of the KL-divergence between $\text{Bern}\left(\frac{\Delta}{\log 1/\theta} \theta s^2\right)$ and $\text{Bern}(\theta s^2)$, we have

$$\begin{aligned} D_{\text{KL}}\left(\frac{\Delta}{\log 1/\theta} \theta s^2 \middle| \theta s^2\right) &= \frac{\Delta}{\log 1/\theta} \theta s^2 \log\left(\frac{\Delta}{\log 1/\theta}\right) \\ &\quad + \left[1 - \left(\frac{\Delta}{\log 1/\theta}\right) \theta s^2\right] \log\left(\frac{1 - (\frac{\Delta}{\log 1/\theta}) \theta s^2}{1 - \theta s^2}\right) \\ &= \frac{\Delta}{\log 1/\theta} \theta s^2 \log\left(\frac{\Delta}{\log 1/\theta}\right) \\ &\quad + \left[1 - \left(\frac{\Delta}{\log 1/\theta}\right) \theta s^2\right] \left(\left(1 - \frac{\Delta}{\log 1/\theta}\right) \frac{\theta s^2}{1 - \theta s^2} + o(\theta)\right) \end{aligned} \quad (79)$$

Here equation (79) follows from the Taylor series for the logarithm $\log(1+x) = x + o(x)$ for $x = o(1)$. Within the terms from (79), we have that $\Delta = \Theta(1)$ and $\theta = o(1)$, we thus get $\frac{\Delta}{\log 1/\theta} = o(1)$ and $\left(\frac{\Delta}{\log 1/\theta}\right) \log\left(\frac{\Delta}{\log 1/\theta}\right) = o(1)$. We therefore have

$$\begin{aligned} D_{\text{KL}}\left(\frac{\Delta}{\log 1/\theta} \theta s^2 \middle| \theta s^2\right) \\ = o(\theta) + \left(\frac{\theta s^2}{1 - \theta s^2}\right) \\ = o(\theta) + \theta s^2. \end{aligned} \quad (80)$$

Here, the last equality (80) follows because $\theta = o(1)$.

(2) Proof for equation (33): From the definition of the KL-divergence between $\text{Bern}\left(\frac{\Delta}{\log 1/\theta} \theta s^2\right)$ and $\text{Bern}(\theta^2 s^2)$, we have

$$\begin{aligned} D_{\text{KL}}\left(\frac{\Delta}{\log 1/\theta} \theta s^2 \middle| \theta^2 s^2\right) \\ = \frac{\Delta}{\log 1/\theta} \theta s^2 \log\left(\frac{\frac{\Delta}{\log 1/\theta}}{\theta}\right) \\ + \left[1 - \left(\frac{\Delta}{\log 1/\theta}\right) \theta s^2\right] \log\left(\frac{1 - (\frac{\Delta}{\log 1/\theta}) \theta s^2}{1 - \theta^2 s^2}\right) \\ = \left(\frac{\Delta}{\log 1/\theta}\right) \theta s^2 \log\left(\frac{\frac{\Delta}{\log 1/\theta}}{\theta}\right) \\ + \left[1 - \left(\frac{\Delta}{\log 1/\theta}\right) \theta s^2\right] \left(-\frac{(\frac{\Delta}{\log 1/\theta}) \theta s^2 - \theta^2 s^2}{1 - \theta^2 s^2} + o(\theta)\right) \end{aligned} \quad (81)$$

$$= \left(\frac{\Delta}{\log 1/\theta}\right) \theta s^2 \log\left(\frac{\frac{\Delta}{\log 1/\theta}}{\theta}\right) + o(\theta) \quad (82)$$

$$\begin{aligned} &= \Delta \theta s^2 \frac{\log \Delta + \log 1/\theta - \log \log 1/\theta}{\log 1/\theta} + o(\theta) \\ &= \Delta \theta s^2 + o(\theta) \end{aligned} \quad (83)$$

Here (81) follows from the Taylor expansion of $\log(1+x)$. Equation (82) follows because $\frac{\Delta}{\log 1/\theta} = o(1)$. Equation (83) follows because $\Delta = \Theta(1)$ and $\log 1/\theta = \omega(1)$.

VIII. DETAILED DISCUSSION IN SEEDED ALIGNMENT

In this section, we provide further details on the discussion in the seeded graph alignment problem in Section IV-B. We prove Remark 3 and Theorem 4 by comparing the feasible regions in Corollaries 1 and 2 to the information theoretic feasible region and the best known feasible region by efficient algorithms in literature.

A. Comparison to the Information-Theoretic Limits

Let us first compare the feasible region of Algorithms ATTRICH and ATTRSPARSE to the information theoretic limit of seeded alignment. The next corollary follows readily from Theorem 3.

Corollary 3 (Information-Theoretic Limits on Seeded Graph Alignment): Consider the seeded Erdős-Rényi pair model $\mathcal{G}(N, \alpha, p, s)$, with $1 - p = \Theta(1)$, $s = \Theta(1)$, and $(1 - \alpha)N = \omega(1)$.

Achievability: In the regime where $p = O\left(\frac{1}{[\log(N(1-\alpha))]^2}\right)$, if

$$Nps^2 \geq \log(N(1-\alpha)) + \omega(1), \quad (84)$$

then exact alignment is achievable w.h.p.

In the regime where $p = \omega\left(\frac{1}{[\log(N(1-\alpha))]^2}\right)$, if

$$Nps^2 - a_N \geq \log(N(1-\alpha)) + \omega(1), \quad (85)$$

where $a_N \triangleq \alpha N \left(\sqrt{ps^2(1-p+p(1-s)^2)} - ps(1-s)\right)^2 - N\alpha ps^2 = O(N\alpha p^{3/2})$, then exact alignment is achievable w.h.p.

Converse: If

$$Nps^2 \leq \log(N(1-\alpha)) - \omega(1),$$

then no algorithm achieves exact alignment w.h.p.

To compare the feasible region of Corollary 3 to the feasible regions of Corollaries 1 and 2, we consider two different regimes for p .

1) $p = O\left(\frac{1}{[\log(N(1-\alpha))]^2}\right)$. In this regime, the feasible region in Corollary 3 is given by condition (84): $Nps^2 \geq \log(N(1-\alpha)) + \omega(1)$. In Corollary 1, condition (15) requires that $Nps^2 \geq (1+\epsilon)\log((1-\alpha)N)$. Since we assume that $n = (1-\alpha)N = \omega(1)$, we can see that condition (15) implies condition (84). Therefore, the feasible region in Corollary 1 is a subset of the feasible region in Corollary 3. Similarly, in Corollary 2, condition (17) can be written as $Nps^2 \geq \frac{\log((1-\alpha)N) + \omega(1)}{1-\alpha}$, which also implies condition (84) since $1-\alpha < 1$. Therefore, the feasible region in Corollary 2 is also a subset of the feasible region in Corollary 3.

2) $p = \omega\left(\frac{1}{[\log(N(1-\alpha))]^2}\right)$. In this regime, the feasible region in Corollary 3 is given by condition (85): $Nps^2 \geq \log(N(1-\alpha)) + a_N + \omega(1)$. Since $s = \Theta(1)$ and $p = \omega\left(\frac{1}{[\log(N(1-\alpha))]^2}\right)$, it follows that $Nps^2 = \omega(\log(N(1-\alpha)))$. Moreover, since $p = o(1)$ and $a_N = O(N\alpha p^{3/2})$, we have $Nps^2 = \omega(a_N)$. Therefore, condition (85) is satisfied and thus Corollary 3 applies in this regime while Corollaries 1 and 2 requires further conditions to apply. Therefore, the feasible regions in Corollaries 1 and 2 are both subsets of the feasible region in Corollary 3.

B. Comparison to Existing Efficient Algorithms

In this subsection, we prove Theorem 4 by thoroughly comparing the feasible region of the proposed algorithms in Corollaries 1 and 2 to that of the existing algorithms in Theorems 5, 6 and 7. We discuss three regimes for m .

Case 1: $m = o(n)$. Recall that $\alpha \triangleq \frac{m}{m+n}$. Therefore, we have $\alpha = o(1)$ in this regime. For the discussion in this regime, we first show that both Corollaries 1 and 2 apply in this regime.

We then show that the feasible region in Corollaries 1 and 2 extends the feasible region for exact alignment in Theorems 5, 6 and 7.

To see Corollary 1 applies in this regime, note that when $\alpha = o(1)$, condition (14) $\alpha Nps^2 = \Omega(\log((1-\alpha)N))$ implies condition (15) $Nps^2 \geq (1+\epsilon)\log((1-\alpha)N)$. Therefore, the feasible region in Corollary 1 simplifies to condition (14), which is non-empty when p , s and α are large enough. To see Corollary 2 applies in this regime, we notice that conditions (16) $\alpha Nps^2 = o(\log((1-\alpha)N))$ and (17) $(1-\alpha)Nps^2 - \log((1-\alpha)N) \geq \omega(1)$ together imply that $\alpha = o(1)$ which is exactly the case in this regime.

Now we move on to compare

the feasible region in Corollaries 1 and 2 to that in Theorems 5, 6 and 7. We consider three different regimes for Np .

Case 1.1: $Np \leq N^{1/7}$.

For this regime of Np , both Theorems 5 and 6 apply. We will show that the feasible regions in Corollaries 1 and 2 are covered by the feasible region in Theorem 6 to illustrate that the proposed algorithms do not introduce any new feasible region in this regime. In this and all later cases, we always set $\beta = 1/7$ when we apply Theorem 6.

We first argue that the feasible region in Corollary 1 is a subset of that in Theorem 6. This is because condition (14) $\alpha = \Omega(\frac{\log((1-\alpha)N)}{Nps^2})$ implies condition (21) $\alpha \geq N^{-4/7}$ under the assumption that $\alpha = o(1)$ and $Np \leq N^{1/7}$, and condition (15) $Nps^2 \geq (1+\epsilon)\log((1-\alpha)N)$ implies condition (20) $Nps^2 \geq \log N + \omega(1)$ under the assumption that $\alpha = o(1)$.

Now we move on to show that the feasible region in Corollary 2 is also a subset of that in Theorem 6. This is because conditions (17) $(1-\alpha)Nps^2 - \log((1-\alpha)N) \geq \omega(1)$ and (18) $\alpha Nps^2 \geq \frac{2\log((1-\alpha)N)}{\tau \log \frac{1}{p}}$ together imply condition (21) $\alpha \geq N^{-4/7}$ under the assumption that $\alpha = o(1)$ and $Np \leq N^{1/7}$, and condition (17) $(1-\alpha)Nps^2 - \log((1-\alpha)N) \geq \omega(1)$ implies condition (20) $Nps^2 \geq \log N + \omega(1)$ under the assumption that $\alpha = o(1)$.

Case 1.2: $N^{1/7} < Np \leq \frac{s}{16(2-s)^2} N^{1/2}$. In this case, both Theorems 5 and 6 apply. We will show that the feasible regions in Corollaries 1 and 2 are covered by the feasible region in Theorem 5 to illustrate that the proposed algorithms do not introduce any new feasible region in this regime.

Firstly, we argue that the feasible region in Corollary 1 is a subset of that in Theorem 5. Note that because $N^{1/7} < Np \leq \frac{s}{16(2-s)^2} N^{1/2}$, we can write $Np = bN^a$ for some a and b such that $a, b = \Theta(1)$ and

$0 < a \leq \frac{1}{2}$ and $0 < b \leq \frac{s}{16(2-s)^2}$. Because $a \leq \frac{1}{2}$, we have $\lfloor 1/a \rfloor \geq 2$. Then the statement follows because condition (14) $\alpha Nps^2 = \Omega(\log((1-\alpha)N))$ implies condition (19) $\alpha \geq \frac{300 \log N}{(Nps^2)^{\lfloor 1/a \rfloor}}$ under the assumption that $\alpha = o(1)$.

Now we move on to show that the feasible region in Corollary 2 is also covered by the feasible region in Theorem 5. This is because again we have $\lfloor 1/a \rfloor \geq 2$ and it follows that conditions (17) $(1-\alpha)Nps^2 - \log((1-\alpha)N) \geq \omega(1)$ and (18) $\alpha Nps^2 \geq \frac{2\log((1-\alpha)N)}{\tau \log \frac{1}{p}}$ together imply condition (19) $\alpha \geq \frac{300 \log N}{(Nps^2)^{\lfloor 1/a \rfloor}}$ under that assumption that $\alpha = o(1)$.

Case 1.3: $Np > \frac{s}{16(2-s)^2} N^{1/2}$. In this regime, Theorem 6 does not apply since $Np = \Omega(N^{1/2})$. We compare the feasible regions in Corollaries 1 and 2 to those in Theorems 5 and 7 to show that the proposed algorithms *strictly improve* the best known feasible region in this regime.

To compare with the existing results, we first consider the feasible region achieved by the proposed algorithms. Because we are in the dense regime $Np = \Omega(N^{1/2})$, conditions (15) $Nps^2 \geq (1+\epsilon)\log((1-\alpha)N)$ and (17) $(1-\alpha)Nps^2 - \log((1-\alpha)N) \geq \omega(1)$ are satisfied. Then conditions for Corollaries 1 and 2 reduce to constraints on α . The constraint in Corollary 1 is given by (14) $\alpha Nps^2 = \Omega(\log((1-\alpha)N))$, and the constraint in Corollary 2 is given by (16) $\alpha Nps^2 = o(\log((1-\alpha)N))$ and (18), which can be written as $\alpha = \Omega(\frac{\log((1-\alpha)N)}{Nps^2 \log \frac{1}{p}})$. Taking the union of the two constraints on α , the feasible region of the proposed algorithm simplifies to

$$\alpha = \Omega\left(\frac{\log((1-\alpha)N)}{Nps^2 \log \frac{1}{p}}\right). \quad (86)$$

Now, we move on to consider the feasible region in Theorem 5. Because $Np > \frac{s}{16(2-s)^2} N^{1/2}$, if we write $Np = bN^a$ for some $a, b = \Theta(1)$ and $0 < b \leq \frac{s}{16(2-s)^2}$, we must have $\frac{1}{2} < a < 1$. Therefore, we have $\lfloor 1/a \rfloor = 1$ in this regime and the feasible region in Theorem 5

reduces to

$$\alpha \geq \frac{300 \log N}{Nps^2}. \quad (87)$$

From the above argument, we see that if condition (86) is satisfied while condition (87) and at least one of conditions (22) $\alpha = \omega(\frac{1}{NI(p,s)^2})$ and (23) $\alpha \geq \frac{2\log N}{NI(p,s)}$ in Theorem 7 are not, then the proposed algorithms achieve exact w.h.p. while none of the existing works [13] and [14] do. In the following, we further consider three sub-cases for p to discuss which one of conditions (22), (23) and (87) is the bottleneck condition in the existing works [13] and [14] and to see the range of α that is feasible by the proposed algorithms while not by [13] and [14].

- Suppose $p \log N \log \frac{1}{p} = \omega(1)$. In this sub-case, condition (23) $\alpha \geq \frac{2\log N}{NI(p,s)}$ implies condition (22) $\alpha = \omega(\frac{1}{NI^2(p,s)})$ and condition (87) $\alpha \geq \frac{300 \log N}{Nps^2}$ implies condition (23). Then the bottleneck condition is condition (23). Therefore, if condition (86) is satisfied and $\alpha < \frac{2\log N}{NI(p,s)}$, the proposed algorithms achieve exact alignment w.h.p. while the algorithms in [13] and [14] do not. Such α exists because $I(p,s) = (1+o(1))s^2 p \log \frac{1}{p}$ and thus $\frac{2\log N}{NI(p,s)} = \Theta(\frac{\log((1-\alpha)N)}{Nps^2 \log \frac{1}{p}})$.
- Suppose $p \log N \log \frac{1}{p} = O(1)$ and $p \log N \log^2 \frac{1}{p} = \omega(1)$. In this sub-case, condition (22) $\alpha = \omega(\frac{1}{NI^2(p,s)})$ implies condition (23) $\alpha \geq \frac{2\log N}{NI(p,s)}$ and condition (87) $\alpha \geq \frac{300 \log N}{Nps^2}$ implies condition (22). Then the bottleneck condition is condition (22). Therefore, if condition (86) is satisfied and $\alpha = O(\frac{1}{NI^2(p,s)})$, the proposed algorithms achieve exact alignment w.h.p. while the algorithms in [13] and [14] do

not. Such α exists because $\frac{\log((1-\alpha)N)}{Nps^2 \log \frac{1}{p}} = O(\frac{1}{NT^2(p,s)})$ under the assumption that $p \log N \log \frac{1}{p} = O(1)$.

- Suppose $p \log N \log^2 \frac{1}{p} = O(1)$. In this sub-case, condition (22) $\alpha = \omega(\frac{1}{NT^2(p,s)})$ implies condition (23) $\alpha \geq \frac{2 \log N}{NT(p,s)}$ and condition (22) implies condition (87) $\alpha \geq \frac{300 \log N}{Nps^2}$. Then the bottleneck condition is condition (87). Therefore, if condition (86) is satisfied and $\alpha < \frac{300 \log N}{Nps^2}$, the proposed algorithms achieve exact alignment w.h.p while the algorithms in [13] and [14] do not. Such α exists because $\frac{\log((1-\alpha)N)}{Nps^2 \log \frac{1}{p}} = o(\frac{300 \log N}{Nps^2})$.

To summarize, in the regime of $m = o(n)$, the proposed Algorithms ATTRICH and ATTRSPARSE together *strictly improve* the best known feasible region for exact alignment when $Np > \frac{s}{16(2-s)^2} N^{1/2}$. In the other two regimes for Np , the feasible region of the proposed algorithms is a strict subset of that in [13].

Case 2: $m = \Omega(n)$ and $m = O(n^{1+\epsilon'})$ for some positive constant ϵ' that satisfies $\epsilon' < \epsilon$ and $\epsilon - \epsilon' = \Theta(1)$ (cf. the constant ϵ in Corollary 1). In this regime, we have $\alpha = \Theta(1)$. For the discussion in this regime, we first show that Corollary 1 applies in this regime while Corollary 2 does not.

We then show that the feasible region in Corollary 1 is a subset of the feasible regions in Theorems 5 and 6 from existing work [13].

Since we have $\alpha = \Theta(1)$, it follows that $Nps^2 \geq (1 + \epsilon) \log((1 - \alpha)N)$ implies $\alpha Nps^2 = \Omega(\log((1 - \alpha)N))$, i.e., condition (15) implies (14). Therefore, the feasible region in Corollary 1 simplifies to condition (15) which holds true when p and s are large enough.

As we argued in Case 1, Corollary 2 requires $\alpha = o(1)$, so Corollary 2 does not apply in this regime.

Now we move on to compare the feasible region in Corollary 1 to that in Theorems 5 and 6. We further consider two different regimes for Np .

Case 2.1: $Np = \exp(o(\log N))$. We can see that Theorem 6 applies in this regime since $Np = \exp(o(\log N)) \leq N^{1/7}$. Condition (21) $\alpha \geq N^{-1+3/7}$ is satisfied without any further assumptions since $\alpha = \Theta(1)$. Moreover, condition (15) $Nps^2 \geq (1 + \epsilon) \log((1 - \alpha)N)$ implies condition (20) $Nps^2 \geq \log N + \omega(1)$ under the assumption that $m = O(n^{1+\epsilon'})$ and thus the feasible region in Corollary 1 is a subset of the feasible region in Theorem 6.

Case 2.2: $Np = \exp(\Omega(\log N))$. Since $Np = \exp(\Omega(\log N))$ and we assume that $p = o(1)$, we can write $Np = bN^a$, where $a, b = \Theta(1)$, $0 < a \leq 1$, and $0 < b \leq \frac{s}{16(2-s)^2}$. Therefore, Theorem 5 applies in this regime. Furthermore, condition (19) in Theorem 5 is satisfied since $\alpha = \Theta(1) \geq \frac{300 \log N}{(Nps^2)^{1/a}}$. We can see that all the conditions in Theorem 5 are satisfied in this regime without any further assumptions.

For Corollary 1, we have $Nps^2 = \exp(\Omega(\log N)) \geq (1 + \epsilon) \log((1 - \alpha)N)$, i.e., condition (15) is satisfied. We can see that all conditions in Corollary 1 are also satisfied in this regime without any further assumptions.

To summarize, in the regime $m = \Omega(n)$ and $m = O(n^{1+\epsilon'})$, the feasible region of the proposed algorithms is a subset of the feasible region given by existing work [13].

Case 3: $m = \Omega(n^{1+\epsilon})$. In this regime, we have $\alpha = 1 - o(1)$. For the same reason as in Case 2, Corollary 1 applies while Corollary 2 does not.

We will show that the feasible region in Corollary 1 extends the best known region for exact alignment in Theorems 5, 6 and 7. We discuss two regimes for Np .

Case 3.1: $Np = \exp(o(\log N))$. We first point out that Theorem 7 does not apply in this regime. This is because $I(p, s) = (1 + o(1))s^2 p \log \frac{1}{p}$ and $p \log \frac{1}{p} = \frac{\exp(o(\log N))}{N} \log \frac{N}{\exp(o(\log N))} = o(\frac{1}{\sqrt{N}})$, so condition (22) $\alpha = \omega(\frac{1}{NT(p,s)^2})$ cannot be satisfied.

Therefore, we focus on showing that the feasible region in Corollary 1 strictly improves in the feasible regions in Theorem 6. This is because $m = \Omega(n^{1+\epsilon})$ is equivalent as $N = \Omega(((1 - \alpha)N)^{1+\epsilon})$, and it follows that condition (20) $Nps^2 \geq \log N + \omega(1)$ implies condition (15) $Nps^2 \geq (1 + \epsilon) \log((1 - \alpha)N)$. So in the case when $Nps^2 - \log N = O(1)$ and $Nps^2 \geq (1 + \epsilon) \log((1 - \alpha)N)$ the proposed Algorithm ATTRICH achieves exact alignment w.h.p., while the algorithms in [13] does not.

Case 3.2: $Np = \exp(\Omega(\log N))$.

This case is included in the feasible region of both Corollary 1 and Theorem 5 for the same reason as in Case 2. Theorem 7 further requires conditions (22) and (23) to apply in this regime.

To summarize, in the regime $m = \Omega(n^{1+\epsilon})$, the proposed Algorithm ATTRICH

strictly improves the best known feasible region for exact alignment in the case when $Np = \exp(o(\log N))$. In the case when $Np = \exp(\Omega(\log N))$, the feasible region by the proposed algorithms is a subset of that in existing literature [13].

From the above discussion, when specialized to the seeded Erdős-Rényi graph pair model, the feasible region in Corollaries 1 and 2 strictly improves the best known feasible region in [13] and [14], as shown in the blue area in Fig. 4. We note, however, that there is also some region that is feasible by existing results but not feasible by the proposed algorithms in this paper, as shown in the red area in Fig. 4.

REFERENCES

- [1] N. Zhang, W. Wang, and L. Wang, "Attributed graph alignment," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2021, pp. 1829–1834.
- [2] P. Pedarsani and M. Grossglauser, "On the privacy of anonymized networks," in *Proc. 17th ACM SIGKDD Conf. Knowl. Discovery Data Mining (KDD)*, Aug. 2011, pp. 1235–1243.
- [3] D. Cullina and N. Kiyavash, "Improved achievability and converse bounds for Erdős-Rényi graph matching," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 44, no. 1, pp. 63–72, Jun. 2016.
- [4] D. Cullina and N. Kiyavash, "Exact alignment recovery for correlated Erdős-Rényi graphs," 2017, *arXiv:1711.06783*.
- [5] Y. Wu, J. Xu, and S. H. Yu, "Settling the sharp reconstruction thresholds of random graph matching," *IEEE Trans. Inf. Theory*, vol. 68, no. 8, pp. 5391–5417, Aug. 2022.
- [6] O. E. Dai, D. Cullina, N. Kiyavash, and M. Grossglauser, "Analysis of a canonical labeling algorithm for the alignment of correlated Erdős-Rényi graphs," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 47, no. 1, pp. 96–97, Dec. 2019.
- [7] J. Ding, Z. Ma, Y. Wu, and J. Xu, "Efficient random graph matching via degree profiles," *Probab. Theory Rel. Fields*, vol. 179, nos. 1–2, pp. 29–115, Feb. 2021.
- [8] Z. Fan, C. Mao, Y. Wu, and J. Xu, "Spectral graph matching and regularized quadratic relaxations: Algorithm and theory," in *Proc. 37th Int. Conf. Mach. Learn.*, in Proceedings of Machine Learning Research, vol. 119, H. D. III and A. Singh, Eds., Jul. 2020, pp. 2985–2995.

- [9] C. Mao, M. Rudelson, and K. Tikhomirov, "Exact matching of random graphs with constant correlation," *Probab. Theory Rel. Fields*, vol. 186, nos. 1–2, pp. 327–389, Jun. 2023.
- [10] C. Mao, Y. Wu, J. Xu, and S. H. Yu, "Random graph matching at Otter's threshold via counting chandeliers," in *Proc. 55th Annu. ACM Symp. Theory Comput.*, 2023, pp. 1345–1356.
- [11] N. Korula and S. Lattanzi, "An efficient reconciliation algorithm for social networks," *Proc. VLDB Endowment*, vol. 7, no. 5, pp. 377–388, Jan. 2014.
- [12] L. Yartseva and M. Grossglauser, "On the performance of percolation graph matching," in *Proc. 1st ACM Conf. Online Social Netw. (COSN)*. New York, NY, USA: Association for Computing Machinery, Oct. 2013, p. 119, doi: [10.1145/2512938.2512952](https://doi.org/10.1145/2512938.2512952).
- [13] E. Mossel and J. Xu, "Seeded graph matching via large neighborhood statistics," *Random Struct. Algorithms*, vol. 57, no. 3, pp. 570–611, Oct. 2020.
- [14] F. Shirani, S. Garg, and E. Erkip, "Seeded graph matching: Efficient algorithms and theoretical guarantees," in *Proc. 51st Asilomar Conf. Signals, Syst., Comput.*, Oct. 2017, pp. 253–257.
- [15] S. Zhang and H. Tong, "FINAL: Fast attributed network alignment," in *Proc. 22nd ACM SIGKDD Conf. Knowl. Discovery Data Mining (KDD)*. New York, NY, USA: Association for Computing Machinery, Aug. 2016, p. 1345.
- [16] S. Zhang and H. Tong, "Attributed network alignment: Problem definitions and fast solutions," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 9, pp. 1680–1692, Sep. 2019.
- [17] Q. Zhou, L. Li, X. Wu, N. Cao, L. Ying, and H. Tong, *Attent: Active Attributed Network Alignment*. New York, NY, USA: Association for Computing Machinery, 2021, p. 3896–3906.
- [18] S. H. Yu, "Matching in networks: Fundamental limits and efficient algorithms," Ph.D. dissertation, Duke Univ., Durham, NC, USA, 2023.
- [19] C. Mao, Y. Wu, J. Xu, and S. H. Yu, "Testing network correlation efficiently via counting trees," 2021, *arXiv:2110.11816*.
- [20] E. Onaran, S. Garg, and E. Erkip, "Optimal de-anonymization in random graphs with community structure," in *Proc. 50th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2016, pp. 709–713.
- [21] D. Cullina, K. Singhal, N. Kiyavash, and P. Mittal, "On the simultaneous preservation of privacy and community structure in anonymized networks," 2016, *arXiv:1603.08028*.
- [22] V. Lyzinski, D. E. Fishkind, M. Fiori, J. T. Vogelstein, C. E. Priebe, and G. Sapiro, "Graph matching: Relax at your own risk," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 1, pp. 60–73, Jan. 2016.
- [23] N. Zhang, Z. Wang, W. Wang, and L. Wang, "Attributed graph alignment," 2021, *arXiv:2102.00665*.
- [24] H. W. Kuhn and B. Yaw, "The Hungarian method for the assignment problem," *Nav. Res. Logist. Quart.*, vol. 2, nos. 1–2, pp. 83–97, Mar. 1955.
- [25] J. Munkres, "Algorithms for the assignment and transportation problems," *J. Soc. Ind. Appl. Math.*, vol. 5, no. 1, pp. 32–38, Mar. 1957. [Online]. Available: <http://www.jstor.org/stable/2098689>
- [26] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *J. Amer. Stat. Assoc.*, vol. 58, no. 301, pp. 13–30, Mar. 1963.
- [27] R. B. Ash, *Information Theory* (Dover Books on Advanced Mathematics). New York, NY, USA: Dover, 1990.

Ziao Wang (Graduate Student Member, IEEE) received the B.E. degree in electrical and computer engineering from Nanyang Technological University, Singapore, in 2019, and the M.A.Sc. degree in electrical and computer engineering from The University of British Columbia, Vancouver, BC, Canada, in 2021, where he is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering. His research interests include information theory and statistical inference on graphs.

Ning Zhang received the bachelor's degree in physics from Nankai University, China, in 2019, and the M.A.Sc. degree in electrical and computer engineering from The University of British Columbia, Vancouver, BC, Canada, in 2022. She is currently pursuing the D.Phil. degree in computational discovery with the Department of Statistics, Oxford University. Her research interests include statistics and computation, with a current focus on the development and analysis of algorithms for large networks.

Weina Wang (Member, IEEE) received the bachelor's degree from the Department of Electronic Engineering, Tsinghua University, in 2009, and the Ph.D. degree in electrical engineering from Arizona State University in 2016. From 2016 to 2018, she was a joint Post-Doctoral Research Associate with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, and the School of Electrical, Computer and Energy Engineering, Arizona State University. She is currently an Assistant Professor with the Department of Computer Science, Carnegie Mellon University. Her research interests include applied probability and stochastic systems, with applications in cloud computing, data centers, and privacy-preserving data analytics. Her dissertation received the Dean's Dissertation Award from the Ira A. Fulton Schools of Engineering, Arizona State University, in 2016. She received the Kenneth C. Sevcik Outstanding Student Paper Award from the ACM SIGMETRICS 2016, the Best Paper Award from the ACM MobiHoc 2022, the NSF CAREER Award in 2022, and the ACM SIGMETRICS Rising Star Research Award in 2023.

Lele Wang (Member, IEEE) received the B.E. degree from Tsinghua University, China, in 2009, and the Ph.D. degree in communication theory and systems from the University of California, San Diego, in 2015. She also attended the Academic Talent Program. From 2015 to 2017, she was a Post-Doctoral Researcher with Stanford University and Tel Aviv University. She is currently an Assistant Professor with the Department of Electrical and Computer Engineering, The University of British Columbia (UBC), Vancouver, BC, Canada. Before joining UBC, she was an NSF Center for Science of Information Post-Doctoral Fellow. Her research interests include information theory, coding theory, communication theory, statistical inference on graphs, high dimensional statistics, and generative models in machine learning. She was a recipient of the 2013 UCSD Shannon Memorial Fellowship, the Qualcomm Innovation Fellowship from 2013 to 2014, and the 2017 NSF Center for Science of Information Postdoctoral Fellowship. Her Ph.D. thesis "Channel Coding Techniques for Network Communication" Won the 2017 IEEE Information Theory Society Thomas M. Cover Dissertation Award. She serves as the Vice President for the Canadian Society of Information Theory from 2022 to 2025, an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS from 2023 to 2026, and the Guest Associate Editor for the IEEE JOURNAL IN SELECTED AREAS ON INFORMATION THEORY in 2023.