Efficient Joint Object Matching via Linear Programming

Antonio De Rosa * Aida Khajavirad †

November 25, 2022

Abstract

Joint object matching, also known as multi-image matching, namely, the problem of finding consistent partial maps among all pairs of objects within a collection, is a crucial task in many areas of computer vision. This problem subsumes bipartite graph matching and graph partitioning as special cases and is NP-hard, in general. We develop scalable linear programming (LP) relaxations with theoretical performance guarantees for joint object matching. We start by proposing a new characterization of consistent partial maps; this in turn enables us to formulate joint object matching as an integer linear programming (ILP) problem. To construct strong LP relaxations, we study the facial structure of the convex hull of the feasible region of this ILP, which we refer to as the joint matching polytope. We present an exponential family of facet-defining inequalities that can be separated in strongly polynomial time, hence obtaining a partial characterization of the joint matching polytope that is both tight and cheap to compute. To analyze the theoretical performance of the proposed LP relaxations, we focus on permutation group synchronization, an important special case of joint object matching. We show that under the random corruption model for the input maps, a simple LP relaxation, that is, an LP containing only a very small fraction of the proposed facet-defining inequalities, recovers the ground truth with high probability if the corruption level is below 40%. Finally, via a preliminary computational study on synthetic data, we show that the proposed LP relaxations outperform a popular SDP relaxation both in terms of recovery and tightness.

Key words. *Joint object matching; Convex relaxations; Linear programming; Recovery guarantee.*

AMS subject classifications. *90C05, 90C10, 05C70, 65D19, 68Q87.*

1 Introduction

Object matching techniques are widely used in many areas of computer vision such as image analysis, object recognition, robotics, biomedical identification, and object tracking. While there is a rich literature on finding isomorphisms between a pair of objects, the task of *joint object matching*, also known as *multi-image matching*, i.e., finding consistent maps among *all* pairs of objects within a collection is underdeveloped. The most classic example for joint object matching is the problem of matching feature points among many images of the same object, a step used for instance in image recognition [13] and in structure from motion [3]. Almost all early approaches to tackle joint object matching relied on sequential matchings of pairs of objects often yielding erroneous results when the input data is noisy. Joint object matching is NP-hard in general; common techniques for tackling this problem are graph neural networks [27, 38], spectral methods [32, 28, 36], and semidefinite programming (SDP) relaxations [22, 7, 39, 21]. Indeed,

*Department of Mathematics, University of Maryland, 4176 Campus Dr, College Park, MD 20742, USA. E-mail: anderosa@umd.edu.

†Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA 18015, USA. E-mail: aida@lehigh.edu.

to date, the only existing convex relaxations for joint object matching are SDP relaxations. It is well-understood that in spite of their polynomial-time complexity, SDPs are too expensive to solve and are often impractical for large-scale problems. In this paper, we develop scalable LP relaxations with theoretical performance guarantees for joint object matching; we start by formally defining the problem.

1.1 Problem Statement

We consider the problem of joint object matching in its full generality; that is, when the objects are only partially similar and the input is possibly incomplete. To formally define this problem, we introduce some terminology which is mostly adapted from [7]. Suppose that we have a collection of n objects $\mathcal{S}_i$, $i \in [n] := \{1, \dots, n\}$ each of which consists of d_i elements for some $n \geq 3$ and $d_i \geq 1$ such that $\max_{i\in[n]} d_i \geq 2$. The case with $n = 2$ is equivalent to bipartite graph matching (see for example chapter 8 of [34]) and the case with $d_i = 1$ for all $i \in [n]$ is equivalent to graph partitioning (see for example [9]), both of which are well-studied problems in combinatorial optimization. Given two discrete sets $\mathcal{S}$ and $\mathcal{S}'$, a subset $\phi \subset \mathcal{S} \times \mathcal{S}'$ is called a *partial map* if each element of $\mathcal{S}$ (resp. $\mathcal{S}'$) is paired with at most one element of $\mathcal{S}'$ (resp. $\mathcal{S}$); in particular, not all elements need to be paired. We denote by $\phi_{ij} : \mathcal{S}_i \to \mathcal{S}_j$ the partial map between the pair of objects $\mathcal{S}_i$ and $\mathcal{S}_j$. In the special case where all elements in $\mathcal{S}_i$ and $\mathcal{S}_j$ are paired, we say that the two objects have *full similarity*; otherwise, we say that the two objects have *partial similarity*. Note that two objects with full similarity must contain the same number of elements.

In joint object matching the input consists of *noisy* pair-wise partial maps ϕ^{in}_{ij} between some of the objects $\mathcal{S}_i$ and $\mathcal{S}_j$, $i, j \in [n]$. These input maps are obtained using off-the-shelf pairwise graph matching algorithms and often contain some erroneous information. An undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is called a *map graph* for the objects $\mathcal{S}_i, i \in [n]$ with $\mathcal{V} := \{\mathcal{S}_1, \mathcal{S}_2, \cdots, \mathcal{S}_n\}$ and $(\mathcal{S}_i, \mathcal{S}_j) \in \mathcal{E}$ whenever an input partial map ϕ^{in}_{ij} is available. For notational simplicity, throughout this paper instead of $(\mathcal{S}_i, \mathcal{S}_j) \in \mathcal{E}$, we write $(i, j) \in \mathcal{E}$. If the map graph is not a complete graph, we say that the input is *incomplete*. The objective is to find a collection of *consistent* partial maps ϕ_{ij} for all $1 \leq i < j \leq n$ that are close to the input maps. By consistent partial maps, we imply that for any $i, j, k \in [n]$, whenever an element $s \in \mathcal{S}_i$ is paired with an element $s' \in \mathcal{S}_j$ which in turn is paired with an element $s'' \in \mathcal{S}_k$, then s and s'' are paired as well. This condition is often referred to as *cycle consistency* in the literature [22], as consistency can be achieved by requiring that the composition of maps between two objects is independent of the connecting path.

In the following, we formulate joint object matching as a mathematical optimization problem. To this end, we use a binary $d_i \times d_j$ matrix $X(i, j)$ to encode ϕ_{ij}; that is, the s, s' entry of $X(i, j)$, denoted by $X_{ss'}(i, j)$, equals one if and only if $(s, s') \in \phi_{ij}$. It then follows that a binary matrix $X(i, j)$ satisfying

$$X(i,j)\mathbf{1}_{d_j} \leq \mathbf{1}_{d_j}, \qquad X^T(i,j)\mathbf{1}_{d_i} \leq \mathbf{1}_{d_i}, \tag{1}$$

corresponds to a partial map, where $\mathbf{1}_{d_i} \in \mathbb{R}^{d_i}$ denotes a vector of all ones. Moreover, cycle consistency is achieved by requiring:

$$\begin{aligned} &X(i,j)X(j,k) \leq X(i,k), \\ &X(j,i)X(i,k) \leq X(j,k), \quad \forall 1 \leq i < j < k \leq n, \\ &X(i,k)X(k,j) \leq X(i,j), \end{aligned} \tag{2}$$

where $X(j, i) = X^T(i, j)$ for any $i < j$, $X(i, j)X(j, k)$ denotes the standard matrix product between $X(i, j)$ and $X(j, k)$, and inequalities are all component-wise. In case of full similarity among all objects, $X(i, j)$ are permutation matrices for all $1 \leq i < j \leq n$; this special case is

often referred to as *permutation group synchronization* in the literature (see for example [32]). It can be checked that cycle consistency in this case is obtained by imposing:

$$X(i,j)X(j,k) = X(i,k) \quad \forall 1 \le i < j < k \le n. \tag{3}$$

The following example gives an illustration of partial maps and cycle consistency.

Example 1. *Let $n = 3$, $d_1 = 1$, $d_2 = d_3 = 2$. Then the following are consistent partial maps:*

$$X(1,2) = \begin{pmatrix} 0 & 1 \end{pmatrix}, \qquad X(1,3) = \begin{pmatrix} 0 & 0 \end{pmatrix}, \qquad X(2,3) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{4}$$

indicating that the only element of object 1 corresponds to the second element of object 2, and the first element of object 2 corresponds to the first element of object 3. However, the following partial maps are not consistent:

$$X(1,2) = \begin{pmatrix} 1 & 0 \end{pmatrix}, \qquad X(1,3) = \begin{pmatrix} 0 & 1 \end{pmatrix}, \qquad X(2,3) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{5}$$

This is because we have $X_{11}(1,2) = X_{12}(1,3) = 1$, which by cycle consistency implies $X_{12}(2,3) = 1$. Indeed by letting $X_{12}(2,3) = 1$, we obtain consistent partial maps indicating that the only element of object 1 corresponds to the first element of object 2 and to the second element of object 3, and the second element of object 2 corresponds to the second element of object 3.

Let $X^{\text{in}}(i,j) \in \{0,1\}^{d_i \times d_j}$ denote the matrix representation of the input partial map ϕ_{ij}^{in}. We would like to find consistent partial maps $X(i,j)$, $1 \le i < j \le n$, so as to minimize

$$f = \sum_{(i,j)\in\mathcal{E}} \left\| X^{\text{in}}(i,j) - X(i,j) \right\|_F^2, \tag{6}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. Denote by $\mathcal{N}$ the number of matched pairs in the input, i.e., $\mathcal{N} = \sum_{(i,j)\in\mathcal{E}} \sum_{t\in[d_i]} \sum_{q\in[d_j]} X_{tq}^{\text{in}}(i,j)$. Since matrices $X^{\text{in}}(i,j)$ and $X(i,j)$ are binary-valued, the objective function (6) can be written as:

$$f = \mathcal{N} + \sum_{(i,j)\in\mathcal{E}} \left\langle \mathbf{1}_{d_i}\mathbf{1}_{d_j}^T - 2X^{\text{in}}(i,j),\ X(i,j) \right\rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the standard matrix inner product. It then follows that joint object matching can be formulated as follows:

$$\begin{aligned} \min \quad & \sum_{(i,j)\in\mathcal{E}} \left\langle \mathbf{1}_{d_i}\mathbf{1}_{d_j}^T - 2X^{\text{in}}(i,j),\ X(i,j) \right\rangle \qquad \text{(JOM)} \\ \text{s.t.} \quad & X(i,j)\mathbf{1}_{d_j} \le \mathbf{1}_{d_j}, \quad X^T(i,j)\mathbf{1}_{d_i} \le \mathbf{1}_{d_i}, \quad \forall 1 \le i < j \le n, \\ & X(i,j)X(j,k) \le X(i,k), \\ & X(j,i)X(i,k) \le X(j,k), \quad \forall 1 \le i < j < k \le n, \\ & X(i,k)X(k,j) \le X(i,j), \\ & X(i,j) \in \{0,1\}^{d_i \times d_j}, \quad \forall 1 \le i < j \le n. \end{aligned}$$

In this paper, we are interested in the quality of convex relaxations for Problem (JOM). To date, the only existing convex relaxations for joint object matching are SDP relaxations. In the following, we describe a widely-used SDP relaxation of Problem (JOM) first proposed in [22].

1.2 SDP relaxations

Assume that there exists a universe $\mathcal{U}$ consisting of m elements such that each object $\mathcal{S}_i$ is a (partial) image of $\mathcal{U}$ and each element in $\mathcal{U}$ is contained in at least one object $\mathcal{S}_i$. We should remark that the set $\mathcal{U}$ or even its size is not known a priori. We only rely on its existence to construct the SDP relaxation. Let the matrix $Y(i) \in \{0,1\}^{d_i \times m}$ encode the correspondences between $\mathcal{S}_i$ and $\mathcal{U}$, i.e., for any $s \in \mathcal{S}_i$ and $s' \in \mathcal{U}$, the s, s' entry of $Y(i)$ equals one if and only if s corresponds to s'. It then follows that $X(i,j) = Y(i)Y^T(j)$ for all $1 \leq i < j \leq n$. Define $X(i,i) = I_{d_i}$ for all $i \in [n]$, where I_{d_i} denotes a $d_i \times d_i$ identity matrix. Let $\bar{d} = \sum_{i\in[n]} d_i$; denote by $X \in \{0,1\}^{\bar{d}\times\bar{d}}$ a matrix whose (i,j)-th block is given by $X(i,j)$, where we let $X(j,i) = X^T(i,j)$. Let $Y = (Y^T(1), \cdots, Y^T(n))^T$. It can be checked that $X = YY^T$ which implies X is a rank-m positive semidefinite matrix. Similarly denote by X^{in} a $\bar{d} \times \bar{d}$ matrix whose (i,j)-th block is given by $X^{\text{in}}(i,j)$, where we let $X^{\text{in}}(i,i) = I_{d_i}$. Then an SDP relaxation of Problem (JOM) is given by:

$$\begin{aligned} \min \quad & \sum_{(i,j)\in\mathcal{E}} \left\langle \mathbf{1}_{d_i}\mathbf{1}_{d_j}^T - 2X^{\text{in}}(i,j),\ X(i,j) \right\rangle \\ \text{s.t.} \quad & X \succeq 0, \quad X \geq 0, \\ & X(i,i) = I_{d_i}, \quad \forall i \in [n], \\ & X(i,j)\mathbf{1}_{d_j} \leq \mathbf{1}_{d_j}, \quad X^T(i,j)\mathbf{1}_{d_i} \leq \mathbf{1}_{d_i}, \quad \forall 1 \leq i < j \leq n, \end{aligned} \tag{7}$$

where $X \succeq 0$ and $X \geq 0$ mean that X is positive semidefinite and component-wise nonnegative, respectively. In case where $X(i,j)$, $1 \leq i < j \leq n$, are all permutation matrices, the authors of [22] proved that the positive semidefiniteness condition $X \succeq 0$ is equivalent to cycle consistency. This in turn implies that in case of fully similar objects, replacing $X \geq 0$ by $X \in \{0,1\}^{\bar{d}\times\bar{d}}$ in Problem (7), one obtains an exact reformulation of Problem (JOM) as a binary SDP.

It is important to emphasize that the size of the universe m is *not* known a priori; the authors of [7] proposed a spectral technique to *estimate* m using the input data X^{in}. They first *trim* X^{in} to remove the bias from over represented rows or columns. Let us denote by $\tilde{X}^{\text{in}}$ the trimmed matrix. Denote by λ_k the k-th largest eigenvalue of $\tilde{X}^{\text{in}}$. Then they let $\hat{m} = \arg\max_{d_{\max}\leq k<\bar{d}} |\lambda_k - \lambda_{k+1}|$, where $d_{\max} = \max_{i\in[n]} d_i$, and where $\hat{m}$ denotes the estimate of m. Subsequently, they used this estimate to further strengthen the SDP relaxation (7); that is, they replaced $X \succeq 0$ by the following constraint:

$$\begin{pmatrix} \hat{m} & \mathbf{1}_{\bar{d}}^T \\ \mathbf{1}_{\bar{d}} & X \end{pmatrix} \succeq 0. \tag{8}$$

In [39], the authors employed Burer-Monteiro factorization [6] by letting $X = YY^T$ to tackle the SDP. Recall that Y is a $\bar{d} \times m$ matrix; hence, to benefit from the low-rank approach, the availability of a good upper bound on the size of the universe is essential. The authors of [39] observed that the estimation technique of [7] is inaccurate when the input is noisy and incomplete. They then chose to use the upper bound $2\hat{m}$, where $\hat{m}$ is obtained by the spectral method of [7] outlined above. To conclude, constraint (8) should only be used if a reliable upper bound on the size of the universe is available.

1.3 Recovery guarantees under stochastic models

To perform a theoretical analysis of various existing and new algorithms for data science applications, a recent stream of research in mathematical data science is focused on obtaining sufficient conditions for recovery of the ground truth under various stochastic models for the input (see for example [7, 2, 19, 12, 11, **?**]). We say that an optimization algorithm *recovers*

the ground truth, whenever its unique optimal solution coincides with the ground truth. In the context of joint object matching, the question can be formally stated as follows: given a probabilistic model for the noise in input partial maps, what is the maximum level of corruption under which the (optimization) algorithm *recovers* the ground truth partial maps with high probability? Throughout this paper, by high probability, we imply the probability tending to 1 as the number of objects $n \to \infty$.

Huang and Guibas [22] considered the SDP relaxation (7) for the permutation group synchronization problem and obtained a deterministic sufficient condition for recovery, implying under a random model recovery is possible with high probability when the corruption level remains below 50%. In [7], the authors considered the general joint object matching with partially similar objects, together with the SDP relaxation (7) enhanced by constraint (8). They first significantly improved their earlier deterministic recovery guarantee in [22]. Subsequently, they focused on the *random corruption model*, roughly defined as follows: each observed $X^{\text{in}}(i,j)$ coincides with ground truth independently with probability p_{true}, and each observed but incorrect $X^{\text{in}}(i,j)$ is independently drawn from a set of partial maps satisfying $\mathbb{E}[X^{\text{in}}(i,j)] = \frac{\mathbf{1}\mathbf{1}^{\mathbf{T}}}{m}$, where $\mathbb{E}[\cdot]$ denotes the expectation of a random variable. Note that, in this model, the authors assume that the size of the universe m is known a priori, an assumption which often does not hold. They proved that the SDP relaxation recovers the ground truth with high probability if $p_{\text{true}} > C\frac{\log^2(nm)}{\sqrt{n}}$, for some universal large constant C. In [4, 28], the authors considered spectral methods for solving permutation group synchronization under the random corruption model. They proved that the spectral algorithm recovers the ground truth with high probability, if $p_{\text{true}} > C\sqrt{\frac{\log(nd)}{n}}$, where d denotes the number of elements in each object. This recovery guarantee is nearly optimal in terms of information theoretical limits [8]. In [31], the authors studied the recovery properties of a spectral algorithm for permutation group synchronization under the additive Gaussian noise.

To summarize, for joint object matching under the random corruption model, both SDP relaxations and spectral methods exhibit near optimal recovery thresholds. It is well understood that solving SDPs is computationally prohibitive for large-scale problems. Spectral methods on the other hand, only require estimating the leading eigenvector of a matrix, and are quite efficient in practice. However, works like [30, 33] suggest that spectral methods are extremely sensitive to slight modification in the generative model. Indeed, convex relaxations enjoy robustness to adversarial corruptions of the inputs for statistical problems that spectral methods do not. Therefore, it is of great interest to understand the most efficient convex relaxation algorithms for solving joint object matching.

1.4 Our contribution

In this paper, we propose scalable LP relaxations with theoretical performance guarantees for joint object matching. To this end, we first present an alternative characterization of consistent partial maps. This in turn enables us to formulate joint object matching as an ILP. Subsequently, with the objective of constructing strong LP relaxations, we study the facial structure of the convex hull of the feasible region of the ILP. As part of this polyhedral study, to effectively approximate cycle consistency, we introduce consistency inequalities, an exponential family of facet-defining inequalities that can be separated in strongly polynomial time. We next study the theoretical properties of the proposed LP relaxation for the permutation group synchronization problem; we show that under the random corruption model, a simple LP relaxation containing only a very small fraction of consistency inequalities, recovers the ground truth with high probability if $p_{\text{true}} > 0.585$. While in the asymptotic regime our recovery guarantee is suboptimal, in many cases of practical interest, the corruption level is below %40. It is for such applications that our proposed LP provides a robust and efficient matching algorithm. In fact, our numerical experiments suggest that for moderate values of n and d, the proposed LP

relaxation outperforms the SDP relaxation in recovering the ground truth.

The remainder of the paper is organized as follows. In Section 2 we present a novel LP relaxation for joint object matching. Subsequently, in Section 3 we focus on the permutation group synchronization problem under the random corruption model and obtain a recovery guarantee for the proposed LP relaxation. We present our numerical experiments in Section 4. Section 5 contains further results regarding the facial structure of the joint matching polytope that were omitted from Section 2.

2 Linear programming relaxation

In this section, we propose a scalable LP relaxation with performance guarantees for joint object matching. We start by presenting an ILP formulation for Problem (JOM). Let us first obtain a linear characterization of cycle consistency constraints (2). A collection of partial maps $X(i,j)$, $1 \leq i < j \leq n$, is consistent if for any $1 \leq i < j < k \leq n$ and for any $l \in [d_i], t \in [d_j], q \in [d_k]$, whenever two out of the three elements $X_{lt}(i,j)$, $X_{tq}(j,k)$, $X_{lq}(i,k)$ equal one, the third one equals one as well. It is simple to check that this condition is enforced by the following system of inequalities:

$$\begin{cases} -X_{lt}(i,j) + X_{tq}(j,k) + X_{lq}(i,k) \leq 1, \\ X_{lt}(i,j) - X_{tq}(j,k) + X_{lq}(i,k) \leq 1, \\ X_{lt}(i,j) + X_{tq}(j,k) - X_{lq}(i,k) \leq 1, \end{cases} \quad \forall l \in [d_i], t \in [d_j], q \in [d_k],\ 1 \leq i < j < k \leq n. \tag{9}$$

Hence an ILP formulation for Problem (JOM) is given by:

$$\begin{aligned} \min \quad & \sum_{(i,j)\in\mathcal{E}} \left\langle \mathbf{1}_{d_i}\mathbf{1}_{d_j}^T - 2X^{\text{in}}(i,j),\ X(i,j) \right\rangle && \text{(IP)} \\ \text{s.t.} \quad & X(i,j)\mathbf{1}_{d_j} \leq \mathbf{1}_{d_j}, \quad X^T(i,j)\mathbf{1}_{d_i} \leq \mathbf{1}_{d_i}, \quad \forall 1 \leq i < j \leq n, \\ & \begin{cases} -X_{lt}(i,j) + X_{tq}(j,k) + X_{lq}(i,k) \leq 1, \\ X_{lt}(i,j) - X_{tq}(j,k) + X_{lq}(i,k) \leq 1, \\ X_{lt}(i,j) + X_{tq}(j,k) - X_{lq}(i,k) \leq 1, \end{cases} \quad \forall l \in [d_i], t \in [d_j], q \in [d_k],\ \forall 1 \leq i < j < k \leq n, \\ & X(i,j) \in \{0,1\}^{d_i \times d_j}, \quad \forall 1 \leq i < j \leq n. \end{aligned}$$

Remark 1. *Inequalities* (9) *can be considered as a generalization of triangle inequalities used for instance in graph partitioning problems [17, 9]. Given a graph with n nodes, the goal in graph partitioning is to partition the nodes of the graph into at most K subsets such that some similarity measure across different partitions is minimized. In this context, for each pair of nodes $i, j \in [n]$ a variable y_{ij} is defined as follows: $y_{ij} = 1$ if nodes i and j belong to the same partition and $y_{ij} = 0$, otherwise. The triangle inequalities are then defined as*

$$y_{ij} + y_{ik} - y_{jk} \leq 1, \quad \forall i \neq j \neq k \in [n].$$

The above inequality states that if i and j are in the same cluster, and i and k are in the same cluster, then also j and k must be in the same cluster. It then follows that inequalities (9) *are a generalization of triangle inequalities for the case node i of the graph represents object $\mathcal{S}_i$ consisting of d_i elements where $d_i \geq 2$ for some $i \in [n]$.*

Now consider the simple LP relaxation of Problem (IP), that is, the LP obtained by replacing $X(i,j) \in \{0,1\}^{d_i \times d_j}$ by the constraint $X(i,j) \geq 0$. Let us refer to this LP as the *basic LP*. In the next example, we show that inequalities (9) are not implied by the SDP relaxation (7).

Example 2. *Let $n = 3$ and $d_1 = d_2 = d_3 = 2$; it can be checked that the following is feasible for Problem* (7)*:*

$$X(1,2) = X(1,3) = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}, \qquad X(2,3) = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}. \tag{10}$$

Now consider the inequality obtained by letting $l = t = q = 1$ *in the second inequality of system* (9)*:*

$$X_{11}(1,2) - X_{11}(2,3) + X_{11}(1,3) \leq 1.$$

Substituting (10) *in the above inequality yields* $\frac{3}{4} - \frac{1}{4} + \frac{3}{4} \not\leq 1$.

In the following, we improve the strength of the basic LP relaxation by obtaining a partial linear characterization of cycle consistency. Subsequently, we establish the strength of the proposed inequalities by showing that they define facets of the convex hull of the feasible region of Problem (IP).

2.1 Consistency inequalities

We now present a generalization of inequalities (9) for joint object matching.

Proposition 1. *Let* $1 \leq i < j < k \leq n$. *Then following inequalities are valid for the feasible region of Problem* (IP)*:*

$$\begin{aligned}
&-\sum_{t\in D_1}\sum_{q\in D_2} X_{tq}(i,j) + \sum_{q\in D_2} X_{ql}(j,k) + \sum_{t\in D_1} X_{tl}(i,k) \leq 1, \\
&\qquad\qquad \forall D_1 \subseteq [d_i],\ D_2 \subseteq [d_j],\ l \in [d_k],\ D_1, D_2 \neq \emptyset \\
&\sum_{t\in D_1} X_{lt}(i,j) - \sum_{t\in D_1}\sum_{q\in D_2} X_{tq}(j,k) + \sum_{q\in D_2} X_{lq}(i,k) \leq 1, \\
&\qquad\qquad \forall D_1 \subseteq [d_j],\ D_2 \subseteq [d_k],\ l \in [d_i],\ D_1, D_2 \neq \emptyset \qquad (11) \\
&\sum_{t\in D_1} X_{tl}(i,j) + \sum_{q\in D_2} X_{lq}(j,k) - \sum_{t\in D_1}\sum_{q\in D_2} X_{tq}(i,k) \leq 1, \\
&\qquad\qquad \forall D_1 \subseteq [d_i],\ D_2 \subseteq [d_k],\ l \in [d_j],\ D_1, D_2 \neq \emptyset.
\end{aligned}$$

Proof. Without loss of generality, consider the third inequality in (11), for some $l \in [d_j]$. To see the validity of this inequality, notice that by condition (1), at most one term in $\sum_{t\in D_1} X_{tl}(i,j)$ and at most one term in $\sum_{q\in D_2} X_{lq}(j,k)$ equal one. Now suppose that we have $X_{\hat{t}l}(i,j) = 1$ for some $\hat{t} \in D_1$ and $X_{l\hat{q}}(j,k) = 1$ for some $\hat{q} \in D_2$. Then by inequalities (9), we have $X_{\hat{t}\hat{q}}(i,k) = 1$ and this completes the proof of validity. □

Notice that by letting $|D_1| = |D_2| = 1$ in inequalities (11), we obtain inequalities (9). Henceforth, we refer to inequalities (11) as *consistency inequalities*. The following example demonstrates that consistency inequalities (11) strengthen the basic LP relaxation of Problem (IP).

Example 3. *Let* $n = 3$ *and* $d_1 = d_2 = d_3 = 2$*; it can be checked that the following satisfies inequalities* (9)*:*

$$X(1,2) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{pmatrix}, \qquad X(2,3) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad X(1,3) = \begin{pmatrix} 0 & \frac{1}{2} \\ 0 & 0 \end{pmatrix}. \qquad (12)$$

Now consider the inequality obtained by letting $l = 1$, $D_1 = \{1,2\}$, *and* $D_2 = \{2\}$ *in the second inequality of system* (11)*:*

$$X_{11}(1,2) + X_{12}(1,2) - X_{12}(2,3) - X_{22}(2,3) + X_{12}(1,3) \leq 1.$$

Substituting (12) *in the above inequality yields* $\frac{1}{2} + \frac{1}{2} - 0 - 0 + \frac{1}{2} \not\leq 1$. *Next consider*

$$X(1,2) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad X(2,3) = \begin{pmatrix} 0 & \frac{1}{3} \\ 0 & \frac{1}{3} \end{pmatrix}, \qquad X(1,3) = \begin{pmatrix} 0 & \frac{1}{3} \\ 0 & \frac{1}{3} \end{pmatrix}. \qquad (13)$$

It can be checked that (13) *satisfies all consistency inequalities with* $|D_1| = 1$ *or* $|D_2| = 1$. *Now consider the consistency inequality with* $l = 2$, $|D_1| = |D_2| = 2$ *given by:*

$$-X_{11}(1,2) - X_{12}(1,2) - X_{21}(1,2) - X_{22}(1,2) + X_{12}(2,3) + X_{22}(2,3) + X_{12}(1,3) + X_{22}(1,3) \leq 1.$$

Substituting (13) *in the above inequality yields* $-0 - 0 - 0 - 0 + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} \not\leq 1$.

The number of consistency inequalities (11) is exponential in the number of elements d_i, $i \in [n]$. That is, for each $1 \leq i < j < k \leq n$, we have $d_i(2^{d_j} + 2^{d_k} - 2) + d_j(2^{d_i} + 2^{d_k} - 2) + d_k(2^{d_i} + 2^{d_j} - 2)$ consistency inequalities. However, as we detail next, separating over these inequalities can be done in a number of operations that is polynomial in n and d_i, $i \in [n]$.

Separation of consistency inequalities. Let us start by defining the separation problem:

The separation problem. Let $\tilde{X}(i,j) \in [0,1]^{d_i \times d_j}$ for all $1 \leq i < j \leq n$ satisfy inequalities (1). Decide whether $\tilde{X}$ satisfies all consistency inequalities (11) or not, and in the latter case find a consistency inequality that is violated by $\tilde{X}$.

Proposition 2. *There exists a strongly polynomial time algorithm that solves the separation problem over all consistency inequalities.*

Proof. Let $1 \leq i < j < k \leq n$ and let $l \in [d_j]$; consider the third inequality in (11) for all nonempty $D_1 \subseteq [d_i]$, $D_2 \subseteq [d_k]$. We claim that the inequality that is most violated by $\tilde{X}(i,j) \in [0,1]^{d_i \times d_j}$ for all $1 \leq i < j \leq n$ (if one exists) among all such $2^{d_i} + 2^{d_k} - 2$ inequalities can be found by an algorithm that is strongly polynomial in d_i, d_k. This in turn implies that the separation problem over all consistency inequalities can be solved in strongly polynomial time with respect to n, d_i, $i \in [n]$. Define binary variables y_t, z_q for all $t \in [d_i], q \in [d_k]$ as follows: for each $t \in [d_i]$, we let $y_t = 1$ if $t \in D_1$ and $y_t = 0$, otherwise; similarly, for each $q \in [d_k]$, we let $z_q = 1$ if $q \in D_2$ and $z_q = 0$, otherwise. Then to find the most violated consistency inequality, it suffices to solve the following optimization problem:

$$\begin{aligned} \max \quad & \sum_{t \in [d_i]} \tilde{X}_{tl}(i,j) y_t + \sum_{q \in [d_k]} \tilde{X}_{lq}(j,k) z_q - \sum_{t \in [d_i]} \sum_{q \in [d_k]} \tilde{X}_{tq}(i,k) y_t z_q \\ \text{s.t.} \quad & y \in \{0,1\}^{d_i},\ z \in \{0,1\}^{d_k}. \end{aligned} \tag{14}$$

If the optimal value of the above problem is greater than one, the maximizer provides us with D_1, D_2 corresponding to a most violated inequality; otherwise, no violated consistency inequality exists. Now let us examine the complexity of solving Problem (14). Denote by $f(y,z)$ the objective function of Problem (14). Defining $\bar{z}_q = 1 - z_q$ for all $q \in [d_k]$, we get

$$\begin{aligned} f(y,\bar{z}) &= \sum_{t \in [d_i]} \tilde{X}_{tl}(i,j) y_t + \sum_{q \in [d_k]} \tilde{X}_{lq}(j,k)(1 - \bar{z}_q) - \sum_{t \in [d_i]} \sum_{q \in [d_k]} \tilde{X}_{tq}(i,k) y_t (1 - \bar{z}_q) \\ &= \sum_{t \in [d_i]} \big(\tilde{X}_{tl}(i,j) - \sum_{q \in [d_k]} \tilde{X}_{tq}(i,k)\big) y_t + \sum_{q \in [d_k]} \tilde{X}_{lq}(j,k)(1 - \bar{z}_q) + \sum_{t \in [d_i]} \sum_{q \in [d_k]} \tilde{X}_{tq}(i,k) y_t \bar{z}_q \end{aligned}$$

Since by assumption $\tilde{X}_{tq}(i,k) \geq 0$ for all $t \in [d_i], q \in [d_k]$ and for all $1 \leq i < k \leq n$, it can be checked that $f(y,\bar{z})$, $y \in \{0,1\}^{d_i}, \bar{z} \in \{0,1\}^{d_k}$ is a super-modular function (see for example [20]). It then follows that Problem (14) can be transformed into the problem of maximizing a super-modular function and hence can be solved in strongly polynomial time in $d_i + d_k$ [35]. Hence, solving the separation problem over all consistency inequalities amounts to (in the worst case) solving $3\tilde{d}\binom{n}{3}$ optimization problems of the form (14), where $\tilde{d} = \max_{i \in [n]} d_i$. □

Remark 2. *It can be shown that Problem* (14)*, i.e., the problem of maximizing a super-modular quadratic function, can be equivalently solved by solving the following LP [5]:*

$$\begin{aligned} \max \quad & \sum_{t\in[d_i]} \tilde{X}_{tl}(i,j)y_t + \sum_{q\in[d_k]} \tilde{X}_{lq}(j,k)z_q - \sum_{t\in[d_i]}\sum_{q\in[d_k]} \tilde{X}_{tq}(i,k)w_{tq} \\ \text{s.t.} \quad & w_{tq} \geq y_t + z_q - 1,\ w_{tq} \geq 0, \quad \forall t \in [d_i], q \in [d_k] \\ & y \in [0,1]^{d_i},\ z \in [0,1]^{d_k}, \end{aligned} \tag{15}$$

which can be readily solved using a generic LP solver. Note that the above LP has $d_i d_k + d_i + d_k$ *variables; i.e.,* y_t, z_q, w_{tq} *for all* $t \in [d_i], q \in [d_k]$*. As we detail in Section 4, to solve the separation problem over consistency inequalities, we solve Problem* (15)*.*

Notice that in order to construct Problem (14) we do not require D_1 and D_2 to be nonempty, since if at least one of the two subsets is empty, the resulting inequality is implied by inequalities (1) and hence is trivially satisfied.

Block consistency inequalities. Consistency inequalities (11) can be significantly generalized as follows:

Proposition 3. *Let* $1 \leq i < j < k \leq n$*. Then following inequalities are valid for the feasible region of Problem* (IP)*:*

$$\begin{aligned} -\sum_{t\in D_1}\sum_{q\in D_2} X_{tq}(i,j) + \sum_{q\in D_2}\sum_{l\in D_3} X_{ql}(j,k) + \sum_{t\in D_1}\sum_{l\in D_3} X_{tl}(i,k) \leq |D_3|, \\ \forall D_1 \subseteq [d_i],\ D_2 \subseteq [d_j],\ D_3 \subseteq [d_k], D_1, D_2, D_3 \neq \emptyset. \\ \sum_{l\in D_3}\sum_{t\in D_1} X_{lt}(i,j) - \sum_{t\in D_1}\sum_{q\in D_2} X_{tq}(j,k) + \sum_{l\in D_3}\sum_{q\in D_2} X_{lq}(i,k) \leq |D_3|, \\ \forall D_1 \subseteq [d_j],\ D_2 \subseteq [d_k],\ D_3 \subseteq [d_i], D_1, D_2, D_3 \neq \emptyset \\ \sum_{t\in D_1}\sum_{l\in D_3} X_{tl}(i,j) + \sum_{l\in D_3}\sum_{q\in D_2} X_{lq}(j,k) - \sum_{t\in D_1}\sum_{q\in D_2} X_{tq}(i,k) \leq |D_3|, \\ \forall D_1 \subseteq [d_i],\ D_2 \subseteq [d_k],\ D_3 \subseteq [d_j], D_1, D_2, D_3 \neq \emptyset. \end{aligned} \tag{16}$$

Proof. Without loss of generality, let us consider the third inequality and let us rewrite this inequality as

$$\sum_{l\in D_3}\Big(\sum_{t\in D_1} X_{tl}(i,j) + \sum_{q\in D_2} X_{lq}(j,k)\Big) - \sum_{t\in D_1}\sum_{q\in D_2} X_{tq}(i,k) \leq |D_3|.$$

By inequalities (1) for each $l \in D_3$ we have $\sum_{t\in D_1} X_{tl}(i,j) \leq 1$ and $\sum_{q\in D_2} X_{lq}(j,k) \leq 1$. By proof of Proposition 1, for each $l \in D_3$, if $X_{\hat{t}l}(i,j) = X_{l\hat{q}}(j,k) = 1$, for some $\hat{t} \in D_1$ and $\hat{q} \in D_2$, then $X_{\hat{t}\hat{q}}(i,k) = 1$. Now suppose that there exists $l_1 \neq l_2 \in D_3$ for which this condition is satisfied; i.e., $X_{\hat{t}l_1}(i,j) = X_{l_1\hat{q}}(j,k) = 1$ and $X_{\tilde{t}l_2}(i,j) = X_{l_2\tilde{q}}(j,k) = 1$. Then by inequalities (1) it follows that $\hat{t} \neq \tilde{t}$ and $\hat{q} \neq \tilde{q}$ implying that by inequalities (9), two distinct terms $X_{\hat{t}\hat{q}}(i,k)$ and $X_{\tilde{t}\tilde{q}}(i,k)$ in the third summation equal one as well. A recursive application of this argument completes the proof of validity. □

Notice that by letting $|D_3| = 1$ in inequalities (16), we obtain consistency inequalities (11). Henceforth, we refer to inequalities (16) as *block consistency inequalities*. If $|D_1| + |D_2| \leq |D_3|$, then by (1), block consistency inequalities are redundant as the value of the left-hand side of these inequalities does not exceed $|D_1| + |D_2|$. However, in general, block consistency inequalities (16) with $|D_3| > 1$ are not implied by the consistency inequalities (11). The following example illustrates this fact:

Example 4. *Let $n = 3$ and $d_1 = d_2 = d_3 = 2$; it can be checked that the following satisfies inequalities* (11)*:*

$$X(1,2) = \begin{pmatrix} 0 & \frac{1}{2} \\ 0 & 0 \end{pmatrix}, \qquad X(2,3) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \qquad X(1,3) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{pmatrix}. \tag{17}$$

Now consider the inequality obtained by letting $D_1 = \{1\}$, $D_2 = D_3 = \{1,2\}$ in the first inequality of system (16)*:*

$$-X_{11}(1,2) - X_{12}(1,2) + X_{11}(2,3) + X_{12}(2,3) + X_{21}(2,3) + X_{22}(2,3) + X_{11}(1,3) + X_{12}(1,3) \leq 2.$$

Substituting (17) *in the above inequality yields $0 - \frac{1}{2} + \frac{4}{2} + \frac{2}{2} \not\leq 2$.*

We will further detail on the theoretical strength of block consistency inequalities in the next section. From a computational perspective however, unlike consistency inequalities (11), the separation problem for this more general class cannot be reduced to the problem of maximizing a binary supermodular quadratic function. Indeed, there does not seem to be a reduction to any known class of polynomially-solvable binary quadratic program [**?**, **?**, 5, **?**]. While we leave the complexity of this separation problem as an open question, we suspect that it is a difficult problem. Moreover, inequalities (16) are very dense; that is they contain many variables with nonzero coefficients; a feature that is not computationally desirable. Hence, in this paper, we do not consider block consistency inequalities for our computational experiments.

2.2 The Joint Matching Polytope

In this section, we establish the strength of the proposed inequalities by performing a brief polyhedral study of the convex hull of the feasible region of Problem (IP). For simplicity, throughout this section we assume that all objects have the same number of elements d. We assume that $n \geq 3$ and $d \geq 2$. Now consider the feasible region of Problem (IP), that is, the union of all consistent partial maps corresponding to n objects each of which consists of d elements. Any consistent partial map $X(i,j)$, $1 \leq i < j \leq n$, can be equivalently represented by a binary vector $\bar{X}$ in $\mathbb{R}^{\binom{n}{2}d^2}$ defined as

$$\bar{X} = (\text{vec}(X(1,1))^T, \text{vec}(X(1,2))^T, \cdots, \text{vec}(X(1,n))^T, \text{vec}(X(2,1))^T \cdots, \text{vec}(X(n,n))^T)^T, \tag{18}$$

where $\text{vec}(X(i,j))$ denotes the vectorization of matrix $X(i,j)$. We refer to the convex hull of such binary vectors corresponding to all consistent partial maps as the *joint matching polytope* and denote it by $\mathcal{C}_{n,d}$. In the following, we refer to a vertex of $\mathcal{C}_{n,d}$ by specifying the corresponding partial maps, with the understanding that transformation (18) is applied to construct a vertex from the partial maps. For notational simplicity, we sometimes use a set theoretic notation to specify the partial maps. Namely, for any $q \in [d]$ and $i \in [n]$, we denote by i_q the q-th element of object $\mathcal{S}_i$. The set $\mathcal{M}_X$ then represents the consistent partial maps $X(i,j)$, $1 \leq i < j \leq n$, as follows: $(\hat{i}_t, \hat{j}_q, \hat{k}_l) \in \mathcal{M}_X$ if and only if $\hat{i}_t$, $\hat{j}_q$ and $\hat{k}_l$ correspond to an element u in the universe $\mathcal{U}$, i.e., $X_{tq}(\hat{i},\hat{j}) = X_{ql}(\hat{j},\hat{k}) = X_{tl}(\hat{i},\hat{k}) = 1$. Moreover, no other element of $\cup_{i\in[n]}\mathcal{S}_i$, corresponds to u. We also assume that $\hat{i}_t \neq \hat{j}_q \neq \hat{k}_l$, where we define $\hat{i}_t = \hat{j}_q$ when $\hat{i} = \hat{j}$ and $t = q$.

With the objective of constructing strong LP relaxations for joint object matching, in the following, we conduct a brief polyhedral study of the joint matching polytope. Namely, we obtain several classes of facet-defining inequalities for this polytope. To this end, we first establish a fundamental property of the joint matching polytope.

Proposition 4. *The joint matching polytope $\mathcal{C}_{n,d}$, $n \geq 3$, $d \geq 2$ is full dimensional, i.e.,* $\dim(\mathcal{C}_{n,d}) = \binom{n}{2}d^2$.

Proof. To prove the statement, it suffices to show that $\mathcal{C}_{n,d}$ contains $\binom{n}{2}d^2+1$ affinely independent points. Let the first point correspond to partial maps with no paired elements, i.e., $X(i,j)=0$ for all $1\leq i<j\leq n$. Next for each $\hat{i},\hat{j}$ with $1\leq\hat{i}<\hat{j}\leq n$, and each $\hat{t},\hat{q}\in[d]$ consider partial maps with $\mathcal{M}_X=\{(\hat{i}_{\hat{t}},\hat{j}_{\hat{q}})\}$, i.e., $X_{\hat{t}\hat{q}}(\hat{i},\hat{j})=1$ and $X_{tq}(i,j)=0$, otherwise. Clearly, we have a total number of $\binom{n}{2}d^2$ such partial maps. It then follows that together with the zero vertex, the corresponding vertices constitute a collection of $\binom{n}{2}d^2+1$ affinely independent points in $\mathcal{C}_{n,d}$. □

Proposition 5. *The nonnegativity constraints $X_{tq}(i,j)\geq 0$ for all $t,q\in[d]$ and for all $1\leq i<j\leq n$, are facet-defining for $\mathcal{C}_{n,d}$.*

Proof. Without loss of generality, we show that inequality $X_{11}(1,2)\geq 0$ defines a facet of $\mathcal{C}_{n,d}$. Consider a nontrivial valid inequality $\sum_{1\leq i<j\leq n}\sum_{t,q\in[d]}\alpha_{tq}(i,j)X_{tq}(i,j)\leq\beta$ for $\mathcal{C}_{n,d}$ that is satisfied tightly by all consistent partial maps $X(i,j)$, $1\leq i<j\leq n$ in which $X_{11}(1,2)=0$. In the following, we show that the two inequalities $X_{11}(1,2)\geq 0$ and $\alpha X\leq\beta$ coincide up to a positive scaling which by full dimensionality of $\mathcal{C}_{n,d}$ (see Proposition 4) implies $X_{11}(1,2)\geq 0$ is facet-defining.

First consider the partial maps corresponding to no matched pairs, i.e., $X(i,j)=0$ for all $1\leq i<j\leq n$. Substituting this point in $\alpha X=\beta$ yields $\beta=0$. Next consider some $\hat{t},\hat{q}\in[d]$ and $1\leq\hat{i}<\hat{j}\leq n$ such that $(\hat{i},\hat{j},\hat{t},\hat{q})\neq(1,2,1,1)$ and consider the partial maps with $\mathcal{M}_X=\{(\hat{i}_{\hat{t}},\hat{j}_{\hat{q}})\}$; substituting this point in $\alpha X=\beta$ gives $\alpha_{\hat{t}\hat{q}}(\hat{i},\hat{j})=0$. It then follows that $\alpha_{tq}(i,j)=0$ for all $t,q\in[d]$ and, for all $1\leq i<j\leq n$ such that $(i,j,t,q)\neq(1,2,1,1)$. Hence, inequality $\alpha X\leq\beta$ can be written as $\alpha_{11}(1,2)X_{11}(1,2)\leq 0$. Since this inequality is valid for $\mathcal{C}_{n,d}$, we have $\alpha_{11}(1,2)<0$ and this completes the proof. □

Proposition 6. *Inequalities* (1) *define facets of the joint matching polytope.*

Proof. Without loss of generality, we show that inequality $\sum_{q\in[d]}X_{1q}(1,2)\leq 1$ defines a facet of $\mathcal{C}_{n,d}$. As in the proof of Proposition 5, denote by $\alpha X\leq\beta$ a nontrivial valid inequality for $\mathcal{C}_{n,d}$ that is binding at all consistent partial maps satisfying $\sum_{q\in[d]}X_{1q}(1,2)=1$. Let $\hat{q}\in[d]$ and consider the partial maps with $\mathcal{M}_X=\{(1_1,2_{\hat{q}})\}$. Substituting this point in $\alpha X=\beta$ yields:

$$\alpha_{1q}(1,2)=\beta,\quad\forall q\in[d]. \tag{19}$$

Next, for each $\hat{q}\in[d]$, consider any partial maps of the form $\mathcal{M}_X=\{(1_1,2_{\hat{q}}),(i_t,j_l)\}$. Notice that by definition of $\mathcal{M}_X$, we have $i_t\neq 1_1$, $i_t\neq 2_{\hat{q}}$, $j_l\neq 1_1$, and $j_l\neq 2_{\hat{q}}$. Substituting such points in $\alpha X=\beta$ and using (19) give:

$$\alpha_{tq}(i,j)=0,\quad\forall 1\leq i<j\leq n,\ \forall t,q\in[d]\quad\text{such that}\quad(i,t)\neq(1,1). \tag{20}$$

Finally for any $j\in[n]\setminus\{1,2\}$ and any $\hat{t},\hat{q}\in[d]$, consider the partial maps with $\mathcal{M}_X=\{(1_1,2_{\hat{t}},j_{\hat{q}})\}$. Substituting such points in $\alpha X=\beta$ and using (19) and (20) give:

$$\alpha_{1q}(1,j)=0,\quad\forall 2<j\leq n,\ \forall q\in[d]. \tag{21}$$

By (19)-(21), inequality $\alpha X\leq\beta$ can be written as $\beta\sum_{q\in[d]}X_{1q}(1,2)\leq\beta$. Moreover, since this inequality is valid for $\mathcal{C}_{n,d}$, it follows that $\beta>0$ and this completes the proof. □

The following proposition establishes the strength of consistency inequalities; that is, consistency inequalities (11) are facet-defining for $\mathcal{C}_{n,d}$.

Proposition 7. *Consistency inequalities* (11) *define facets of the joint matching polytope for all nonempty $D_1,D_2\subseteq[d]$.*

Proof. Without loss of generality, we show that for any nonempty $D_1, D_2 \subseteq [d]$, the inequality

$$\sum_{t\in D_1} X_{1t}(1,2) - \sum_{t\in D_1}\sum_{q\in D_2} X_{tq}(2,3) + \sum_{q\in D_2} X_{1q}(1,3) \leq 1, \tag{22}$$

defines a facet of $\mathcal{C}_{n,d}$. Denote by $\alpha X \leq \beta$ a nontrivial valid inequality for $\mathcal{C}_{n,d}$ that is binding at all consistent partial maps satisfying inequality (22) at equality. Notice that inequality (22) is binding at a point if in the corresponding partial maps there exists $e \in \mathcal{M}_X$ with $1_1 \in e$ such that (i) $2_t \in e$ for some $t \in D_1$ and $3_q \notin e$ for all $q \in D_2$ or (ii) $2_t \notin e$ for all $t \in D_1$ and $3_q \in e$ for some $q \in D_2$ or (iii) $2_t \in e$ for some $t \in D_1$ and $3_q \in e$ for some $q \in D_2$. First consider the partial maps with $\mathcal{M}_X = \{(1_1, 2_t)\}$ for some $t \in D_1$ (resp. $\mathcal{M}_X = \{(1_1, 3_q)\}$ for some $q \in D_2$). Substituting in $\alpha X = \beta$ yields:

$$\alpha_{1t}(1,2) = \alpha_{1q}(1,3) = \beta, \quad \forall t \in D_1,\ \forall q \in D_2. \tag{23}$$

Next, consider the partial maps with $\mathcal{M}_X = \{(1_1, 2_t, 3_q)\}$ for some $t \in D_1$ and some $q \in D_2$; substituting this point in $\alpha X = \beta$ and using (23) gives:

$$\alpha_{tq}(2,3) = -\beta, \quad \forall t \in D_1, q \in D_2. \tag{24}$$

Now let $t \in D_1$, consider any partial maps of the form $\mathcal{M}_X = \{(1_1, 2_t), (i_r, j_s)\}$ (resp. let $q \in D_2$, and consider any partial map of the form $\mathcal{M}_X = \{(1_1, 3_q), (i_r, j_s)\}$). Substituting these points in $\alpha X = \beta$ and using (23) yield:

$$\begin{aligned} \alpha_{tq}(i,j) = 0, \quad & \forall 1 \leq i < j \leq n,\ \forall t, q \in [d], \text{ such that } (t,i) \neq (1,1) \text{ and} \\ & (t,q,i,j) \neq (\tilde{t}, \tilde{q}, 2, 3), \text{ for some } \tilde{t} \in D_1,\ \tilde{q} \in D_2. \end{aligned} \tag{25}$$

Let $t \in D_1$, for any $3 < j \leq n$ and any $q \in [d]$, consider the partial maps of the form $\mathcal{M}_X = \{(1_1, 2_t, j_q)\}$. Substituting in $\alpha X = \beta$ and using (23) and (25) yield:

$$\alpha_{1q}(1,j) = 0, \quad \forall 3 < j \leq n,\ \forall q \in [d]. \tag{26}$$

Finally, let $t \in D_1$; for any $q \in [d] \setminus D_2$, consider the partial maps $\mathcal{M}_X = \{(1_1, 2_t, 3_q)\}$. Substituting in $\alpha X = \beta$ and using (23) and (25) yield

$$\alpha_{1q}(1,3) = 0, \quad \forall q \in [d] \setminus D_2. \tag{27}$$

Similarly, let $t \in D_2$; for any $q \in [d] \setminus D_1$, consider the partial maps $\mathcal{M}_X = \{(1_1, 2_q, 3_t)\}$ to obtain

$$\alpha_{1q}(1,2) = 0, \quad \forall q \in [d] \setminus D_1. \tag{28}$$

From (23)-(28), it follows that $\alpha X \leq \beta$ can be written as

$$\beta\Big(\sum_{t\in D_1} X_{1t}(1,2) - \sum_{t\in D_1}\sum_{q\in D_2} X_{tq}(2,3) + \sum_{q\in D_2} X_{1q}(1,3)\Big) \leq \beta.$$

By the validity of the latter we have $\beta > 0$ and this completes the proof. □

Finally, we consider block consistency inequalities (16). The next proposition indicates that all non-redundant block consistency inequalities (16); namely, those with $|D_1| + |D_2| > |D_3|$, are facet-defining. The proof is given in Section 5.1.

Proposition 8. *Inequalities* (16) *define facets of the joint matching polytope for all nonempty* $D_1, D_2, D_3 \subseteq [d]$ *such that* $|D_1| + |D_2| > |D_3|$.

It is interesting to note that the polytope defined by facet-defining inequalities of Propositions 5-8 does *not* coincide with the joint matching polytope even for $n = 3$ and $d = 2$. The following example demonstrates this fact:

Example 5. *Let $n = 3$ and $d = 2$; it can be checked that the following satisfies all facet-defining inequalities of Propositions 5-8:*

$$X(1,2) = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}, \qquad X(2,3) = \begin{pmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \qquad X(1,3) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}. \tag{29}$$

Now consider the following inequality:

$$\begin{aligned} &-X_{11}(1,2) - X_{12}(1,2) - X_{22}(1,2) + X_{11}(2,3) + X_{21}(2,3) + X_{22}(2,3) + X_{11}(1,3) \\ &+X_{12}(1,3) + X_{22}(1,3) \leq 2. \end{aligned}$$

To see the validity of the above inequality, note that by inequalities (1) *we have $X_{11}(2,3) + X_{21}(2,3) + X_{22}(2,3) \leq 2$ and $X_{11}(1,3) + X_{12}(1,3) + X_{22}(1,3) \leq 2$. Moreover, if $X_{11}(2,3) + X_{21}(2,3) + X_{22}(2,3) = 2$, then we must have $X_{11}(2,3) = X_{22}(2,3) = 1$. This in turn implies that if $X_{11}(1,3) = 1$ (resp. $X_{12}(1,3) = 1$ and $X_{22}(1,3) = 1$), then $X_{11}(1,2) = 1$ (resp. $X_{12}(1,2) = 1$ and $X_{22}(1,2) = 1$). By symmetry, the validity of the inequality for the case with $X_{11}(1,3) + X_{12}(1,3) + X_{22}(1,3) = 2$ follows. Substituting* (29) *in the above inequality yields $-\frac{1}{2} + \frac{3}{2} + \frac{3}{2} \not\leq 2$, implying the proposed facet defining inequalities do not characterize $\mathcal{C}_{3,2}$.*

Obtaining a linear characterization of cycle consistency for joint object matching remains an open question and is a subject of future research.

2.3 A new LP relaxation for joint object matching

We propose an LP relaxation of Problem (IP) whose feasible set is defined by facet-defining inequalities of Propositions 5, 6 and 7:

$$\begin{aligned} \min \ & \sum_{(i,j)\in\mathcal{E}} \left\langle \mathbf{1}_{d_i}\mathbf{1}_{d_j}^T - 2X^{\text{in}}(i,j),\ X(i,j) \right\rangle && \text{(LPF)} \\ \text{s.t. } & X(i,j)\mathbf{1}_{d_j} \leq \mathbf{1}_{d_j},\ X^T(i,j)\mathbf{1}_{d_i} \leq \mathbf{1}_{d_i}, \quad \forall 1 \leq i < j \leq n \\ & -\sum_{t\in D_1}\sum_{q\in D_2} X_{tq}(i,j) + \sum_{q\in D_2} X_{ql}(j,k) + \sum_{t\in D_1} X_{tl}(i,k) \leq 1, \\ & \qquad \forall D_1 \subseteq [d_i],\ D_2 \subseteq [d_j],\ l \in [d_k],\ D_1, D_2 \neq \emptyset,\ 1 \leq i < j < k \leq n \\ & \sum_{t\in D_1} X_{lt}(i,j) - \sum_{t\in D_1}\sum_{q\in D_2} X_{tq}(j,k) + \sum_{q\in D_2} X_{lq}(i,k) \leq 1, \\ & \qquad \forall D_1 \subseteq [d_j],\ D_2 \subseteq [d_k],\ l \in [d_i],\ D_1, D_2 \neq \emptyset,\ 1 \leq i < j < k \leq n \\ & \sum_{t\in D_1} X_{tl}(i,j) + \sum_{q\in D_2} X_{lq}(j,k) - \sum_{t\in D_1}\sum_{q\in D_2} X_{tq}(i,k) \leq 1, \\ & \qquad \forall D_1 \subseteq [d_i],\ D_2 \subseteq [d_k],\ l \in [d_j],\ D_1, D_2 \neq \emptyset,\ 1 \leq i < j < k \leq n \\ & X(i,j) \geq 0, \quad \forall 1 \leq i < j \leq n. \end{aligned}$$

Recall that the number of consistency inequalities (11) in Problem (LPF) is exponential in the number of elements d_i in object i. However, by Proposition 2, separating over these inequalities can be done in time that is polynomial in n, d_i, $i \in [n]$. By polynomial equivalence of separation and optimization [16], it then follows that Problem (LPF) is solvable in polynomial time.

As we detailed in Section 1, in joint object matching the size of the universe is not known a priori. However, some studies [7, 39] have proposed to employ *an upper bound estimate* on the size of the universe to tighten their SDP relaxation. Such an upper bound estimate can also be used to tighten our proposed LP relaxation. Since a reliable estimation of $\hat{m}$ is based on a good understanding of the specific application and the input data, and hence is beyond the scope of this paper, we do not include it in our numerical experiments. However, for completeness, in Section 5.2 we present a class of valid inequalities that can significantly tighten the LP relaxation when an estimate on the size of the universe is available.

3 Recovery for permutation group synchronization

In this section, we study the recovery properties of the proposed LP relaxation under a popular stochastic model for the input maps. We consider the case with $d_i = d = m$ for all $i \in [n]$; this special case is often referred to as permutation group synchronization in the literature [32, 31, 28]. For simplicity, we assume that the input consists of all possible pair-wise matchings, that is, we assume that the map graph is complete. Define $a_{tq}(i,j) := 1 - 2X^{\text{in}}_{tq}(i,j)$ for all $t, q \in [d]$ and for all $1 \le i < j \le n$. As the first step, in this paper, we focus on the basic LP relaxation, that is, the LP consisting of all consistency inequalities with $|D_1| = |D_2| = 1$:

$$\min \quad \sum_{1\le i<j\le n} \sum_{t,q\in[d]} a_{tq}(i,j) X_{tq}(i,j) \tag{P}$$

$$\text{s.t.} \quad \sum_{q\in[d]} X_{tq}(i,j) = 1, \quad \forall t \in [d],\ \forall 1 \le i < j \le n, \tag{30}$$

$$\sum_{t\in[d]} X_{tq}(i,j) = 1, \quad \forall q \in [d],\ \forall 1 \le i < j \le n, \tag{31}$$

$$\begin{cases} X_{lt}(i,j) + X_{tq}(j,k) - X_{lq}(i,k) \le 1, \\ X_{lt}(i,j) - X_{tq}(j,k) + X_{lq}(i,k) \le 1, \\ -X_{lt}(i,j) + X_{tq}(j,k) + X_{lq}(i,k) \le 1, \end{cases} \quad \forall l,t,q \in [d],\ \forall 1 \le i < j < k \le n, \tag{32}$$

$$X_{tq}(i,j) \ge 0, \quad \forall t,q \in [d],\ \forall 1 \le i < j \le n. \tag{33}$$

The random corruption model. To analyze the quality of the proposed LP relaxation, we consider the *random corruption model* [7, 28], defined as follows:

$$X^{\text{in}}(i,j) = \begin{cases} Y(i)Y^T(j) & \text{with probability } p_{\text{true}}, \\ \mathbf{P}_{ij} & \text{with probability } 1 - p_{\text{true}}, \end{cases}$$

where $Y(i) \in \{0,1\}^{d\times d}$, $i \in [n]$ encodes the correspondences between the i-th object and the universe, $\mathbf{P}_{ij}$ is an independent random permutation matrix uniformly sampled from $d!$ permutation matrices, and $p_{\text{true}} \in [0,1]$ denotes the probability that each $X^{\text{in}}(i,j)$ coincides with the ground truth. Without loss of generality, we assume that the elements of each object are ordered so that the q-th element of all objects are matched with each other for all $q \in [d]$; that is, $X^{\text{in}}(i,j) = I_d$, whenever the input matches the ground truth. For each $1 \le i < j \le n$, denote by z_{ij} a Bernoulli random variable with parameter p_{true}, independently drawn from $\mathbf{P}_{ij}$. It then follows that:

$$X^{\text{in}}(i,j) = z_{ij} I_d + (1 - z_{ij})\mathbf{P}_{ij}, \quad \forall 1 \le i < j \le n. \tag{34}$$

We are interested in addressing the following question:

What is the maximum level of corruption under which the basic LP relaxation recovers the ground truth with high probability?

Recall that we say an optimization problem (e.g. the LP relaxation) recovers the ground truth if its unique optimal solution coincides with the ground truth. Moreover, by high probability we imply that the probability tends to 1 as the number of objects $n \to \infty$. The following theorem states that the basic LP relaxation recovers the ground truth with high probability provided that the corruption level is below 40%:

Theorem 1. *Suppose that input maps $X^{\text{in}}(i,j)$ for all $1 \le i < j \le n$ are generated according to the random corruption model* (34) *and assume that $d \in o((n/\log(n))^{\frac{1}{4}})$. Then Problem* (P) *recovers the ground truth maps with high probability if $p_{\text{true}} > 0.585$.*

In the remainder of this section, we prove Theorem 1 via a number of intermediate steps stated in six lemmata. Recall that for our ground truth, we assume that the elements of each object are ordered in a way that for each $q \in [d]$, q-th elements of all objects are matched with each other. That is, for each $1 \le i < j \le n$, we have $\bar{X}_{qq}(i,j) = 1$ for all $q \in [d]$ and $\bar{X}_{tq}(i,j) = 0$ for all $t \neq q \in [d]$. We first obtain a deterministic sufficient condition under which $\bar{X}$ is the unique optimal solution of Problem (P). Subsequently, we focus on the random corruption model and prove that our deterministic condition implies the recovery guarantee of Theorem 1.

3.1 A deterministic condition for recovery

In the following we obtain a deterministic sufficient condition under which the ground truth is the unique optimal solution of the basic LP. First, in Lemma 1, we obtain a sufficient condition under which $\bar{X}$ is an optimal solution of Problem (P), where we assume that each input map $X^{\text{in}}(i,j)$, $1 \le i < j \le n$ is a permutation matrix. Subsequently, in Lemma 2, we examine the uniqueness of $\bar{X}$. In the following, we define $a(j,i) := a^T(i,j)$ for all $1 \le i < j \le n$.

Lemma 1. *For each $1 \le i < j \le n$ and $t \in [d]$, define*

$$\begin{aligned}\kappa_t(i,j) := \frac{1}{n} \sum_{l \in [d]\setminus\{t\}} \sum_{k \in [n]\setminus\{i,j\}} \min\left\{ \frac{a_{lt}(k,i)}{\alpha} + \frac{a_{tt}(k,j)}{\beta d}, \frac{a_{lt}(k,j)}{\alpha} + \frac{a_{tt}(k,i)}{\beta d} \right\} \\ + \frac{1}{n}\Big(\frac{1}{2\alpha} - \frac{d-1}{2\beta d}\Big)\Big(\sum_{k \in [n]\setminus\{i\}} a_{tt}(k,i) + \sum_{k \in [n]\setminus\{j\}} a_{tt}(k,j)\Big) - \frac{(d-2)}{\alpha} + 1 \\ + \frac{1}{2n} \sum_{l \in [d]\setminus\{t\}} \Big(\frac{a_{tl}(i,j) + a_{tl}(j,i)}{\alpha} + 2\frac{a_{tt}(i,j)}{\beta d}\Big),\end{aligned} \tag{35}$$

and for each $1 \le i < j \le n$ and $t, q \in [d]$, define

$$\delta_{t\to q}^{i\to j} := \frac{1}{n} \sum_{\substack{k \in [n]\setminus\{i,j\}: \\ a_{tt}(k,i) = -1}} \Big(\frac{a_{tq}(k,j) - a_{tq}(i,j)}{\alpha} + \frac{a_{tt}(i,j) - a_{tt}(k,j)}{\beta d}\Big). \tag{36}$$

Let $\alpha > 0$ and $\beta > 0$ be parameters independent of n, d. Then $\bar{X}$ is an optimal solution of Problem (P), if for each $t, q \in [d]$ and for each $1 \le i < j \le n$, the following conditions are satisfied:

(i) if $a_{tt}(i,j) = a_{qq}(i,j) = a_{tq}(i,j) = 1$, then

$$\delta_{t\to q}^{i\to j} \ge 0, \quad \delta_{q\to t}^{j\to i} \ge 0, \tag{C1}$$

(ii) if $a_{tt}(i,j) = a_{qq}(i,j) = 1$, $a_{tq}(i,j) = -1$, then

$$\delta_{t\to q}^{i\to j} \ge 1, \quad \delta_{q\to t}^{j\to i} \ge 1, \tag{C2}$$

(iii) if $a_{tt}(i,j) = a_{qq}(i,j) = -1$, $a_{tq}(i,j) = 1$, then

$$\kappa_t(i,j) + \frac{1}{2}\Big(\delta_{t\to q}^{i\to j} + \delta_{q\to t}^{j\to i}\Big) \ge 0, \quad \kappa_q(i,j) + \frac{1}{2}\Big(\delta_{t\to q}^{i\to j} + \delta_{q\to t}^{j\to i}\Big) \ge 0, \tag{C3}$$

(iv) if $a_{tt}(i,j) = 1$, $a_{qq}(i,j) = -1$, $a_{tq}(i,j) = 1$, then

$$\kappa_q(i,j) + \delta_{t\to q}^{i\to j} + \delta_{q\to t}^{j\to i} \ge 0, \tag{C4}$$

(v) if $a_{tt}(i,j) = -1$, $a_{qq}(i,j) = 1$, $a_{tq}(i,j) = 1$, *then*

$$\kappa_t(i,j) + \delta^{i\to j}_{t\to q} + \delta^{j\to i}_{q\to t} \geq 0. \tag{C5}$$

Proof. We start by constructing the dual of Problem (P): define dual variables $r_t(i,j)$ for all $t \in [d]$ (resp. $c_q(i,j)$ for all $q \in [d]$) and for all $1 \leq i < j \leq n$ associated with equalities (30) (resp. equalities (31)). Define dual variables $\lambda^1_{ltq}(i,j,k)$, $\lambda^2_{ltq}(i,j,k)$, $\lambda^3_{ltq}(i,j,k)$ for all $l,t,q \in [d]$ and for all $1 \leq i < j < k \leq n$ for the first, the second and the third inequalities of system (32), respectively. Finally, for inequalities (33), define dual variables $\mu_{tq}(i,j)$ for all $t,q \in [d]$ and for all $1 \leq i < j \leq n$. It then follows that the dual of Problem (P) is given by:

$$\begin{aligned}
\max \quad & -\sum_{1\leq i<j<k\leq n}\ \sum_{l,t,q\in[d]} \Big(\lambda^1_{ltq}(i,j,k) + \lambda^2_{ltq}(i,j,k) + \lambda^3_{ltq}(i,j,k)\Big) && \text{(D)}\\
& -\sum_{1\leq i<j\leq n}\ \sum_{t\in[d]} \Big(r_t(i,j) + c_t(i,j)\Big)\\
\text{s.t.} \quad & a_{tq}(i,j) + r_t(i,j) + c_q(i,j) - \mu_{tq}(i,j) + \sum_{l\in[d]} \Big(\sum_{\substack{k\in[n]:\\ k<i}} \big(\lambda^1_{ltq}(k,i,j) - \lambda^2_{ltq}(k,i,j) + \lambda^3_{ltq}(k,i,j)\big)\\
& + \sum_{\substack{k\in[n]:\\ i<k<j}} \big(-\lambda^1_{tlq}(i,k,j) + \lambda^2_{tlq}(i,k,j) + \lambda^3_{tlq}(i,k,j)\big)\\
& + \sum_{\substack{k\in[n]:\\ k>j}} \big(\lambda^1_{tql}(i,j,k) + \lambda^2_{tql}(i,j,k) - \lambda^3_{tql}(i,j,k)\big)\Big) = 0, \quad \forall t,q\in[d], \forall 1\leq i<j\leq n, && \text{(37)}\\
& \lambda^1_{ltq}(i,j,k) \geq 0,\ \lambda^2_{ltq}(i,j,k) \geq 0,\ \lambda^3_{ltq}(i,j,k) \geq 0, \quad \forall l,t,q\in[d],\ \forall 1\leq i<j<k\leq n,\\
& \mu_{tq}(i,j) \geq 0, \quad \forall t,q\in[d],\ \forall 1\leq i<j\leq n. && \text{(38)}
\end{aligned}$$

To prove the statement, it suffices to construct a dual feasible point $(\bar r, \bar c, \bar\lambda^1, \bar\lambda^2, \bar\lambda^3, \bar\mu)$ that satisfies complementary slackness, i.e., for all $1 \leq i < j < k \leq n$ and all $l,t,q \in [d]$, we have

- $\bar\lambda^1_{ltq}(i,j,k) = 0$ if $l \neq t$ and $t \neq q$,
- $\bar\lambda^2_{ltq}(i,j,k) = 0$ if $l \neq t$ and $l \neq q$,
- $\bar\lambda^3_{ltq}(i,j,k) = 0$ if $l \neq q$ and $t \neq q$,

and $\bar\mu_{tt}(i,j) = 0$ for all $1 \leq i < j \leq n$ and for all $t \in [d]$.

Moreover, to construct the dual certificate, we make the following simplifications:

- $\bar\lambda^1_{ttt}(i,j,k) = \bar\lambda^2_{ttt}(i,j,k) = \bar\lambda^3_{ttt}(i,j,k) = 0$, for all $t \in [d]$ an for all $1 \leq i < j < k \leq n$,
- $\bar r_t(i,j) = \bar c_t(i,j)$ for all $t \in [d]$ and for all $1 \leq i < j \leq n$.

Hence by complementary slackness and our simplifications stated above, for each $1 \leq i < j \leq n$, if $t = q$, constraint (37) simplifies to

$$a_{tt}(i,j) + 2\bar r_t(i,j) + \Sigma_t(i,j) = 0, \tag{39}$$

where we define

$$\begin{aligned}
\Sigma_t(i,j) := \sum_{l\in[d]\setminus\{t\}} \Big(& \sum_{k<i} (\bar\lambda^1_{ltt}(k,i,j) + \bar\lambda^3_{ltt}(k,i,j)) + \sum_{i<k<j} (\bar\lambda^2_{tlt}(i,k,j) + \bar\lambda^3_{tlt}(i,k,j))\\
& + \sum_{k>j} (\bar\lambda^1_{ttl}(i,j,k) + \bar\lambda^2_{ttl}(i,j,k))\Big),
\end{aligned} \tag{40}$$

and for each $1 \leq i < j \leq n$, if $t \neq q$, using inequalities (38) to project out $\mu_{tq}(i,j)$, constraint (37) simplifies to:

$$a_{tq}(i,j) + \bar{r}_t(i,j) + \bar{r}_q(i,j) + \Delta_{tq}(i,j) \geq 0, \tag{41}$$

where we define

$$\begin{aligned}\Delta_{tq}(i,j) := &\sum_{k<i} \big(\bar{\lambda}^1_{ttq}(k,i,j) - \bar{\lambda}^2_{ttq}(k,i,j) + \bar{\lambda}^3_{qtq}(k,i,j) - \bar{\lambda}^2_{qtq}(k,i,j)\big) \\ &+ \sum_{i<k<j} \big(\bar{\lambda}^2_{ttq}(i,k,j) - \bar{\lambda}^1_{ttq}(i,k,j) + \bar{\lambda}^3_{tqq}(i,k,j) - \bar{\lambda}^1_{tqq}(i,k,j)\big) \\ &+ \sum_{k>j} \big(\bar{\lambda}^1_{tqq}(i,j,k) - \bar{\lambda}^3_{tqq}(i,j,k) + \bar{\lambda}^2_{tqt}(i,j,k) - \bar{\lambda}^3_{tqt}(i,j,k)\big).\end{aligned} \tag{42}$$

Hence, we need to determine nonnegative $\bar{\lambda}^1$, $\bar{\lambda}^2$, $\bar{\lambda}^3$ satisfying(39) and (41). To this end, for each $1 \leq i < j \leq n$ and each $t \in [d]$ such that $a_{tt}(i,j) = -1$, do the following:

- for each $1 \leq k < i$ and for each $q \in [d] \setminus \{t\}$, let

$$\begin{cases} \bar{\lambda}^1_{qtt}(k,i,j) = & \frac{1}{n} \max\left\{ \frac{a_{tq}(j,k) - a_{tq}(i,k)}{\alpha} + \frac{a_{tt}(i,k) - a_{tt}(j,k)}{\beta d},\ 0 \right\}, \\ \bar{\lambda}^3_{qtt}(k,i,j) = & \frac{1}{n} \max\left\{ \frac{a_{tq}(i,k) - a_{tq}(j,k)}{\alpha} + \frac{a_{tt}(j,k) - a_{tt}(i,k)}{\beta d},\ 0 \right\}, \end{cases} \tag{43}$$

- for each $i < k < j$ and for each $l \in [d] \setminus \{t\}$, let

$$\begin{cases} \bar{\lambda}^2_{tqt}(i,k,j) = & \frac{1}{n} \max\left\{ \frac{a_{tq}(j,k) - a_{tq}(i,k)}{\alpha} + \frac{a_{tt}(i,k) - a_{tt}(j,k)}{\beta d},\ 0 \right\}, \\ \bar{\lambda}^3_{tqt}(i,k,j) = & \frac{1}{n} \max\left\{ \frac{a_{tq}(i,k) - a_{tq}(j,k)}{\alpha} + \frac{a_{tt}(j,k) - a_{tt}(i,k)}{\beta d},\ 0 \right\}, \end{cases} \tag{44}$$

- for each $j < k \leq n$ and for each $l \in [d] \setminus \{t\}$, let

$$\begin{cases} \bar{\lambda}^1_{ttq}(i,j,k) = \frac{1}{n} \max\left\{ \frac{a_{tq}(i,k) - a_{tq}(j,k)}{\alpha} + \frac{a_{tt}(j,k) - a_{tt}(i,k)}{\beta d},\ 0 \right\}, \\ \bar{\lambda}^2_{ttq}(i,j,k) = \frac{1}{n} \max\left\{ \frac{a_{tq}(j,k) - a_{tq}(i,k)}{\alpha} + \frac{a_{tt}(i,k) - a_{tt}(j,k)}{\beta d},\ 0 \right\}. \end{cases} \tag{45}$$

Moreover, for each $1 \leq i < j \leq n$ and for each $t \in [d]$ such that $a_{tt}(i,j) = 1$, do the following:

- for each $1 \leq k < i$ and for each $q \in [d] \setminus \{t\}$, let

$$\bar{\lambda}^1_{qtt}(k,i,j) = \bar{\lambda}^3_{qtt}(k,i,j) = 0, \tag{46}$$

- for each $i < k < j$ and for each $q \in [d] \setminus \{t\}$, let

$$\bar{\lambda}^2_{tqt}(i,k,j) = \bar{\lambda}^3_{tqt}(i,k,j) = 0, \tag{47}$$

- for each $j < k \leq n$ and for each $q \in [d] \setminus \{t\}$, let

$$\bar{\lambda}^1_{ttq}(i,j,k) = \bar{\lambda}^2_{ttq}(i,j,k) = 0. \tag{48}$$

Substituting (43)-(45) in (40), gives:

$$\Sigma_t(i,j) = 2 - 2\kappa_t(i,j), \qquad \text{if } a_{tt}(i,j) = -1 \tag{49}$$

where $\kappa_t(i,j)$ is defined by (35), and substituting (46)-(48) in (40), yields

$$\Sigma_t(i,j) = 0, \qquad \text{if } a_{tt}(i,j) = 1. \tag{50}$$

Moreover, substituting (43)-(48) in (42) gives

$$\Delta_{tq}(i,j) = \delta^{i\to j}_{t\to q} + \delta^{j\to i}_{q\to t}, \tag{51}$$

where $\delta^{i\to j}_{t\to q}$ is defined by (36). Finally, substituting (49) and (51) in (39) and (41), we obtain the following five cases:

(i) $a_{tt}(i,j) = a_{qq}(i,j) = a_{tq}(i,j) = 1$: in this case, by (50) we have $\Sigma_t(i,j) = \Sigma_q(i,j) = 0$ which by (39) gives $\bar{r}_t(i,j) = \bar{r}_q(i,j) = -1/2$. Hence, inequality (41) simplifies to $\Delta_{tq}(i,j) \geq 0$. By (51), the validity of this inequality follows from condition (C1).

(ii) $a_{tt}(i,j) = a_{qq}(i,j) = 1$, $a_{tq}(i,j) = -1$: in this case, by (50) we have $\Sigma_t(i,j) = \Sigma_q(i,j) = 0$ and hence by (39) we get $\bar{r}_t(i,j) = \bar{r}_q(i,j) = -1/2$. Hence, inequality (41) simplifies to $\Delta_{tq}(i,j) \geq 2$. By (51), the validity of this inequality follows from condition (C2).

(iii) $a_{tt}(i,j) = a_{qq}(i,j) = -1$, $a_{tq}(i,j) = 1$: in this case by (49) we have $\Sigma_t(i,j) = 2 - 2\kappa_t(i,j)$ and $\Sigma_q(i,j) = 2 - 2\kappa_q(i,j)$. Hence, from (39) it follows that $\bar{r}_t(i,j) = \frac{2\kappa_t(i,j)-1}{2}$ and $\bar{r}_q(i,j) = \frac{2\kappa_q(i,j)-1}{2}$. Then, inequality (41) simplifies to $\kappa_t(i,j) + \kappa_q(i,j) + \Delta_{tq}(i,j) \geq 0$. By (51) this inequality is implied by condition (C3).

(iv) $a_{tt}(i,j) = 1$, $a_{qq}(i,j) = -1$, $a_{tq}(i,j) = 1$: in this case, by (50) we have $\Sigma_t(i,j) = 0$, and by (49) we have $\Sigma_q(i,j) = 2 - 2\kappa_q(i,j)$. Hence by (39) we get $\bar{r}_t(i,j) = -1/2$, and $\bar{r}_q(i,j) = \frac{2\kappa_q(i,j)-1}{2}$. Then, inequality (41) simplifies to $\kappa_q(i,j) + \Delta_{tq}(i,j) \geq 0$. Finally, by (51), this inequality is equivalent to condition (C4).

(v) $a_{tt}(i,j) = -1$, $a_{qq}(i,j) = 1$, $a_{tq}(i,j) = 1$: by symmetry, this case follows from case (iv).

□

We now provide a sufficient condition under which the ground truth $\bar{X}$ is the unique optimal solution of Problem (P). To this end, we use Mangasarian's characterization of uniqueness of solution in LP [29]:

Proposition 9 (Part (iv) of Theorem 2 in [29])**.** *Consider an LP whose feasible region is defined by $Ax = b$ and $Cx \leq d$, where $x \in \mathbb{R}^n$ denotes the vector of optimization variables and, b, d, A, C are vectors and matrices of appropriate dimensions. Let $\bar{x}$ be an optimal solution of this LP and denote by $\bar{u}$ the dual optimal solution corresponding to the inequality constraints. Let C_i denote the i-th row of C. Define $K = \{i : C_i\bar{x} = d_i,\ \bar{u}_i > 0\}$, $L = \{i : C_i\bar{x} = d_i,\ \bar{u}_i = 0\}$. Let C_K and C_L be the matrices whose rows are C_i, $i \in K$ and C_i, $i \in L$, respectively. Then $\bar{x}$ is the unique optimal solution of the LP, if there exists no x different from the zero vector satisfying*

$$Ax = 0, \quad C_K x = 0, \quad C_L x \leq 0. \tag{52}$$

We are now ready to establish our uniqueness result:

Lemma 2. *Suppose that all inequalities in conditions* (C1)-(C5) *are strictly satisfied. Then $\bar{X}$ is the unique optimal solution of Problem* (P)*.*

Proof. Consider the dual certificate $(\bar{r}, \bar{c}, \bar{\lambda}^1, \bar{\lambda}^2, \bar{\lambda}^3, \bar{\mu})$ constructed in Lemma 1; to achieve uniqueness, we perturb this point by requiring $\bar{\mu}_{tq}(i,j) > 0$, for all $t \neq q \in [d]$ and for all $1 \leq i < j \leq n$; this can be enforced by replacing inequalities (41) by strict inequalities, i.e.,

$$a_{tq}(i,j) + \bar{r}_t(i,j) + \bar{r}_q(i,j) + \Delta_{tq}(i,j) > 0, \quad \forall t \neq q \in [d], \quad \forall 1 \leq i < j \leq n.$$

It can be checked that the above condition is satisfied if inequalities (C1)-(C5) are strictly satisfied. Hence it remains to show that if $\bar{\mu}_{tq}(i,j) > 0$, for all $t \neq q$ and for all $1 \leq i < j \leq n$, then $\bar{X}$ is the unique optimal solution. To this end, we make use of Proposition 9; suppose that $\bar{X}$ is not the unique optimal solution of Problem (P). Then there exists $\tilde{X}$ not identically zero, satisfying condition (52). Since $\bar{\mu}_{tq}(i,j) > 0$, from condition (52), it follows that $\tilde{X}_{tq}(i,j) = 0$ for all $t \neq q \in [d]$ and for all $1 \leq i < j \leq n$. Moreover, by condition (52), we must have $\sum_{q \in [d]} \tilde{X}_{tq}(i,j) = 0$ for all $t \in [d]$. Therefore, $\tilde{X}_{tt}(i,j) = 0$ for all $t \in [d]$ and for all $1 \leq i < j \leq n$, which contradicts with the assumption that $\tilde{X}$ must have a nonzero element. Hence $\bar{X}$ is the unique optimal solution of Problem (P). □

3.2 Recovery under the random corruption model

We now consider the random corruption model (34); we prove that our deterministic recovery condition given by conditions (C1)-(C5) implies that under the random corruption model, Problem (P) recovers the ground truth with high probability, if $p_{\text{true}} > 0.585$. In this section, for an event A, we denote by $\mathbb{P}(A)$ the probability of A with respect to the random corruption model. For a random variable Y, we denote by $\mathbb{E}[Y]$ the expected value of Y with respect to the random corruption model. We denote by $A^1 - A^5$ the events that conditions (C1)-(C5) are strictly satisfied, respectively. We denote by A_{recovery} the event that Problem (P) recovers the ground truth under the random corruption model; then by Lemma 2 we have

$$\mathbb{P}(A_{\text{recovery}}) \geq \mathbb{P}\left(A^1 \cap A^2 \cap A^3 \cap A^4 \cap A^5\right).$$

Since A_{recovery} contains the intersection of a constant number of events A^i, to prove recovery with high probability, it suffices to prove that each A^i, $i \in \{1, \dots 5\}$ occurs with high probability. We establish this in the next four lemmata.

Lemma 3. *Suppose that $X^{\text{in}}(i,j)$ for all $1 \leq i < j \leq n$ are generated according to the random corruption model (34). If*

$$p_{\text{true}} \geq \frac{\beta}{\alpha + \beta}, \tag{A1}$$

and $d \in o(\sqrt{n/\log(n)})$, then event A^1 occurs with high probability.

Proof. For each $1 \leq i < j \leq n$, for each $t, q \in [d]$, and for each $k \in [n] \setminus \{i, j\}$ define:

$$Y^1_{t,q,k}(i,j) := \frac{n-2}{n} \begin{cases} \frac{a_{tq}(k,j)-1}{\alpha} + \frac{1-a_{tt}(k,j)}{\beta d} & \text{if } a_{tt}(k,i) = -1 \\ 0 & \text{otherwise,} \end{cases}$$

and

$$Y^2_{t,q,k}(i,j) := \frac{n-2}{n} \begin{cases} \frac{a_{tq}(i,k)-1}{\alpha} + \frac{1-a_{qq}(i,k)}{\beta d} & \text{if } a_{qq}(k,j) = -1 \\ 0 & \text{otherwise.} \end{cases}$$

It then follows that event A^1 occurs, i.e., condition (C1) is strictly satisfied, if for all $t, q \in [d]$ and all $1 \leq i < j \leq n$ with $a_{tt}(i,j) = a_{qq}(i,j) = a_{tq}(i,j) = 1$, we have

$$\frac{1}{n-2} \sum_{k \in [n] \setminus \{i,j\}} Y^1_{t,q,k}(i,j) > 0, \tag{53}$$

$$\frac{1}{n-2} \sum_{k \in [n] \setminus \{i,j\}} Y^2_{t,q,k}(i,j) > 0. \tag{54}$$

We start by showing that under assumption (A1), the following inequalities hold for every $t, q \in [d]$ and for every $1 \leq i < j \leq n$:

$$\epsilon_1(d) := \mathbb{E}\left[\frac{1}{n-2} \sum_{k \in [n] \setminus \{i,j\}} Y^1_{t,q,k}(i,j)\right] > 0, \quad \epsilon_2(d) := \mathbb{E}\left[\frac{1}{n-2} \sum_{k \in [n] \setminus \{i,j\}} Y^2_{t,q,k}(i,j)\right] > 0. \tag{55}$$

By direct computation we get that for every $t, q \in [d]$ and for every $1 \leq i < j \leq n$:

$$\begin{aligned} \epsilon_1(d) =& \mathbb{E}\left[\frac{1}{n-2} \sum_{k \in [n] \setminus \{i,j\}} Y^1_{t,q,k}(i,j)\right] = \mathbb{E}\left[\frac{1}{n-2} \sum_{k \in [n] \setminus \{i,j\}} Y^2_{t,q,k}(i,j)\right] = \epsilon_2(d) \\ =& \frac{2}{d}\Big(p_{\text{true}} + \frac{1-p_{\text{true}}}{d}\Big)\Big(\frac{p_{\text{true}}}{\beta} - (1-p_{\text{true}})(\frac{1}{\alpha} - \frac{1}{\beta d})\Big) = g_1 g_2, \end{aligned}$$

where $g_1 = \frac{2}{d}(p_{\text{true}} + \frac{1-p_{\text{true}}}{d})$ and $g_2 = \frac{p_{\text{true}}}{\beta} - (1-p_{\text{true}})(\frac{1}{\alpha} - \frac{1}{\beta d})$. Clearly, $g_1 > 0$ for all d and p_{true}. Moreover, g_2 is strictly decreasing in d since its derivative with respect to d is $-\frac{(1-p_{\text{true}})}{\beta d^2}$. Therefore, to show that $g_2 > 0$ for every d, it is enough to show that

$$\frac{p_{\text{true}}}{\beta} - \frac{1-p_{\text{true}}}{\alpha} = \lim_{d\to\infty} \frac{p_{\text{true}}}{\beta} - (1-p_{\text{true}})(\frac{1}{\alpha} - \frac{1}{\beta d}) \geq 0,$$

which follows from assumption (A1). Hence, $\epsilon_1(d)$ and $\epsilon_2(d)$ can be written as a product of two positive functions for every d and this completes the proof of inequalities (55).

We now show that event A^1 occurs with high probability. Let $A^1 = A_1^1 \cap A_2^1$, where A_1^1 (resp. A_2^1) denotes the event that inequalities (53) (resp. inequalities (54)) are satisfied. To show A^1 occurs with high probability, it suffices to show that A_1^1 and A_2^1 occur with high probability. In the following, we prove that A_1^1 occurs with high probability. The proof for A_2^1 follows from a similar line of arguments.

$$\begin{aligned}
&\mathbb{P}(A_1^1)\\
\geq&\mathbb{P}\left(\bigcap_{t,q,i,j}\Big\{\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^1_{t,q,k}(i,j) > 0\Big\}\right)\\
=&\mathbb{P}\left(\bigcap_{t,q,i,j}\Big\{\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^1_{t,q,k}(i,j) - \mathbb{E}\Big[\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^1_{t,q,k}(i,j)\Big] > -\epsilon_1(d)\Big\}\right)\\
\geq&\mathbb{P}\left(\bigcap_{t,q,i,j}\Big\{\Big|\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^1_{t,q,k}(i,j) - \mathbb{E}\Big[\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^1_{t,q,k}(i,j)\Big]\Big| < \epsilon_1(d)\Big\}\right)\\
=&1-\mathbb{P}\left(\bigcup_{t,q,i,j}\Big\{\Big|\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^1_{t,q,k}(i,j) - \mathbb{E}\Big[\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^1_{t,q,k}(i,j)\Big]\Big| \geq \epsilon_1(d)\Big\}\right)\\
\geq&1-\sum_{t,q,i,j}\mathbb{P}\left(\Big|\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^1_{t,q,k}(i,j) - \mathbb{E}\Big[\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^1_{t,q,k}(i,j)\Big]\Big| \geq \epsilon_1(d)\right)\\
\geq&1-2d^2(n-2)^2\exp\Big(-\frac{\epsilon_1^2(d)(n-2)}{8}\Big),
\end{aligned}$$

where the second and fourth lines follow by set inclusion and the sixth line follows from the union bound. The last line follows from the application of Hoeffding's inequality [37] by noting that for each (t,q,i,j) the random variables $Y^1_{t,q,k}(i,j)$ are independent for all $k \in [n]\setminus\{i,j\}$ and can be bounded as $|Y^1_{t,q,k}(i,j)| \leq 2$. Since $\epsilon_1(d) \in \Theta(\frac{1}{d})$ and by assumption $d \in o(\sqrt{n/\log(n)})$, we get

$$2d^2(n-2)^2\exp\Big(-\frac{\epsilon_1^2(d)(n-2)}{8}\Big) \to 0$$

as $n \to \infty$. □

Lemma 4. *Suppose that $X^{\text{in}}(i,j)$ for all $1 \leq i < j \leq n$ are generated according to the random corruption model (34). If*

$$p_{\text{true}} \geq \frac{\alpha}{2}, \tag{A2}$$

then event A^2 occurs with high probability.

Proof. For each $1 \leq i < j \leq n$, for each $t,q \in [d]$ and for each $k \in [n]\setminus\{i,j\}$ define:

$$Y^3_{t,q,k}(i,j) := \frac{n-2}{n}\begin{cases}\frac{a_{tq}(k,j)+1}{\alpha} + \frac{1-a_{tt}(k,j)}{\beta d} - 1 & \text{if } a_{tt}(k,i) = -1\\ -1 & \text{otherwise,}\end{cases}$$

and

$$Y^4_{t,q,k}(i,j) := \frac{n-2}{n}\begin{cases}\frac{a_{tq}(i,k)+1}{\alpha}+\frac{1-a_{qq}(i,k)}{\beta d}-1 & \text{if } a_{qq}(k,j)=-1\\ -1 & \text{otherwise.}\end{cases}$$

It can be checked that event A^2 occurs, i.e., condition (C2) is strictly satisfied, if for all $t,q\in[d]$ and all $1\le i<j\le n$ with $a_{tt}(i,j)=a_{qq}(i,j)=1, a_{tq}(i,j)=-1$, we have

$$\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^3_{t,q,k}(i,j) > 0, \tag{56}$$

$$\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^4_{t,q,k}(i,j) > 0. \tag{57}$$

We start by proving that under assumption (A2), the following inequalities hold for every $t,q\in[d]$ and for every $1\le i<j\le n$:

$$\epsilon_3(d) := \mathbb{E}\Big[\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^3_{t,q,k}(i,j)\Big] > 0, \quad \epsilon_4(d) := \mathbb{E}\Big[\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^4_{t,q,k}(i,j)\Big] > 0. \tag{58}$$

By direct computation we get that for every $t,q\in[d]$ and for every $1\le i<j\le n$:

$$\begin{aligned}\epsilon_3(d) =& \mathbb{E}\Big[\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^3_{t,q,k}(i,j)\Big] = \mathbb{E}\Big[\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^4_{t,q,k}(i,j)\Big] = \epsilon_4(d)\\ =& 2\Big(p_{\text{true}}+\frac{1-p_{\text{true}}}{d}\Big)\Big(\frac{p_{\text{true}}}{\beta d}-\frac{1-p_{\text{true}}}{d}(\frac{1}{\alpha}-\frac{1}{\beta d})+\frac{1}{\alpha}\Big)-1.\end{aligned}$$

It can be checked that the derivative of the above expression with respect to d is given by $-\frac{2(1-p_{\text{true}})^2}{\alpha d^2}(1-\frac{2}{d})-\frac{2p^2_{\text{true}}}{\beta d^2}-8\frac{p_{\text{true}}(1-p_{\text{true}})}{\beta d^3}-6\frac{(1-p_{\text{true}})^2}{\beta d^4}$, which is negative for all $0\le p_{\text{true}}\le 1$ and $d\ge 2$. Hence it is enough to show that

$$\frac{2p_{\text{true}}}{\alpha}-1 = \lim_{d\to\infty} 2\Big(p_{\text{true}}+\frac{1-p_{\text{true}}}{d}\Big)\Big(\frac{p_{\text{true}}}{\beta d}-\frac{1-p_{\text{true}}}{d}(\frac{1}{\alpha}-\frac{1}{\beta d})+\frac{1}{\alpha}\Big)-1 \ge 0,$$

which follows from assumption (A2).

We now show that event A^2 occurs with high probability. Let $A^2 = A^2_1\cap A^2_2$, where A^2_1 (resp. A^2_2) denotes the event for which inequalities (56) (resp. inequalities (57)) are satisfied. To show A^2 occurs with high probability, it suffices to show that A^2_1 and A^2_2 occur with high probability. Notice that for each $t,q\in[d]$ and each $1\le i<j\le n$, the random variables $Y^3_{t,q,k}, Y^4_{t,q,k}$ for all $k\in[n]\setminus\{i,j\}$ are independent and can be bounded as $-1\le Y^3_{t,q,k}(i,j)\le 3$ and $-1\le Y^4_{t,q,k}(i,j)\le 3$. Hence, the proof that A^2_1 (or A^2_2) occurs with high probability is very similar to that of A^1_1 stated in the proof of Lemma 3. That is, it can be checked that

$$\mathbb{P}(A^2_1) \ge 1-2d^2(n-2)^2\exp\Big(-\frac{\epsilon^2_3(d)(n-2)}{8}\Big).$$

Since $\epsilon_3(d)\in\Theta(1)$, it follows that $\mathbb{P}(A^2_1)\to 1$, as $n\to\infty$. □

Lemma 5. *Suppose that $X^{\text{in}}(i,j)$ for all $1\le i<j\le n$ are generated according to the random corruption model* (34)*. If*

$$\alpha\le\beta\le 2, \tag{A3}$$

$$\frac{p^2_{\text{true}}}{\beta}+(\frac{1}{\alpha}-\frac{1}{\beta})p_{\text{true}}+(\frac{1}{2}-\frac{1}{\alpha})>0, \tag{A4}$$

$$p_{\text{true}}\ge\frac{1}{4}, \tag{A5}$$

$$n^2>12d, \tag{A6}$$

and $d\in o((n/\log(n))^{\frac{1}{4}})$, then event A^3 occurs with high probability.

Proof. We start by defining a number of random variables associated with inequalities of condition (C3). For each $1 \leq i < j \leq n$, for each $t \in [d]$ and for each $k \in [n] \setminus \{i, j\}$ define:

$$Z_{t,k}(i,j) := 1 - \frac{(d-2)}{\alpha} + \frac{n-2}{n} \sum_{l \in [d] \setminus \{t\}} \min \left\{ \frac{a_{lt}(k,i)}{\alpha} + \frac{a_{tt}(k,j)}{\beta d}, \frac{a_{lt}(k,j)}{\alpha} + \frac{a_{tt}(k,i)}{\beta d} \right\}. \quad (59)$$

For each $1 \leq i < j \leq n$, for each $t \in [d]$ and for each $k \in [n]$ define:

$$Q_{t,k}(i,j) := \big(\frac{1}{2\alpha} - \frac{d-1}{2\beta d}\big)\big(a_{tt}(k,i) + a_{tt}(k,j)\big), \quad (60)$$

where we define $a_{tt}(i,i) = a_{tt}(j,j) = 0$. For each $1 \leq i < j \leq n$, for each $t \in [d]$ define:

$$W_t(i,j) := \frac{1}{2n} \sum_{l \in [d] \setminus \{t\}} \Big(\frac{a_{tl}(i,j) + a_{tl}(j,i)}{\alpha} + 2\frac{a_{tt}(i,j)}{\beta d} \Big). \quad (61)$$

Finally, for each $1 \leq i < j \leq n$, for each $t, q \in [d]$ and for each $k \in [n] \setminus \{i, j\}$ define:

$$Y^5_{t,q,k}(i,j) := \begin{cases} Z_{t,k}(i,j) + \frac{1}{2}\big(\frac{a_{tq}(k,j)-1}{\alpha} + \frac{-1-a_{tt}(k,j)}{\beta d} + \frac{a_{tq}(i,k)-1}{\alpha} + \frac{-1-a_{qq}(i,k)}{\beta d}\big) \\ \qquad\qquad \text{if } a_{tt}(k,i) = -1,\ a_{qq}(k,j) = -1 \\ Z_{t,k}(i,j) + \frac{1}{2}\big(\frac{a_{tq}(i,k)-1}{\alpha} + \frac{-1-a_{qq}(i,k)}{\beta d}\big) \\ \qquad\qquad \text{if } a_{tt}(k,i) = 1,\ a_{qq}(k,j) = -1 \\ Z_{t,k}(i,j) + \frac{1}{2}\big(\frac{a_{tq}(k,j)-1}{\alpha} + \frac{-1-a_{tt}(k,j)}{\beta d}\big) \\ \qquad\qquad \text{if } a_{tt}(k,i) = -1,\ a_{qq}(k,j) = 1 \\ Z_{t,k}(i,j) \qquad\qquad \text{if } a_{tt}(k,i) = 1,\ a_{qq}(k,j) = 1 \end{cases}$$

and

$$Y^6_{t,q,k}(i,j) := \begin{cases} Z_{q,k}(i,j) + \frac{1}{2}\big(\frac{a_{tq}(k,j)-1}{\alpha} + \frac{-1-a_{tt}(k,j)}{\beta d} + \frac{a_{tq}(i,k)-1}{\alpha} + \frac{-1-a_{qq}(i,k)}{\beta d}\big) \\ \qquad\qquad \text{if } a_{tt}(k,i) = -1,\ a_{qq}(k,j) = -1 \\ Z_{q,k}(i,j) + \frac{1}{2}\big(\frac{a_{tq}(i,k)-1}{\alpha} + \frac{-1-a_{qq}(i,k)}{\beta d}\big) \\ \qquad\qquad \text{if } a_{tt}(k,i) = 1,\ a_{qq}(k,j) = -1 \\ Z_{q,k}(i,j) + \frac{1}{2}\big(\frac{a_{tq}(k,j)-1}{\alpha} + \frac{-1-a_{tt}(k,j)}{\beta d}\big) \\ \qquad\qquad \text{if } a_{tt}(k,i) = -1,\ a_{qq}(k,j) = 1 \\ Z_{q,k}(i,j) \qquad\qquad \text{if } a_{tt}(k,i) = 1,\ a_{qq}(k,j) = 1 \end{cases}$$

It then follows that event A^3 occurs, i.e., condition (C3) is strictly satisfied, if for all $t, q \in [d]$ and all $1 \leq i < j \leq n$ with $a_{tt}(i,j) = a_{qq}(i,j) = -1, a_{tq}(i,j) = 1$, we have

$$W_t(i,j) + \frac{1}{n} \sum_{k \in [n]} Q_{t,k}(i,j) + \frac{1}{n-2} \sum_{k \in [n] \setminus \{i,j\}} Y^5_{t,q,k}(i,j) > 0, \quad (62)$$

$$W_q(i,j) + \frac{1}{n} \sum_{k \in [n]} Q_{q,k}(i,j) + \frac{1}{n-2} \sum_{k \in [n] \setminus \{i,j\}} Y^6_{t,q,k}(i,j) > 0. \quad (63)$$

We first prove the following inequalities hold for all $t, q \in [d]$ and for all $1 \leq i < j \leq n$:

$$\begin{aligned} \epsilon_5(d) &:= \mathbb{E}\Big[\frac{1}{n} \sum_{k \in [n]} Q_{t,k}(i,j) + \frac{1}{n-2} \sum_{k \in [n] \setminus \{i,j\}} Y^5_{t,q,k}(i,j)\Big] > 0, \\ \epsilon_6(d) &:= \mathbb{E}\Big[\frac{1}{n} \sum_{k \in [n]} Q_{q,k}(i,j) + \frac{1}{n-2} \sum_{k \in [n] \setminus \{i,j\}} Y^6_{t,q,k}(i,j)\Big] > 0. \end{aligned} \quad (64)$$

Notice that by symmetry, $\epsilon_5(d) = \epsilon_6(d)$; hence, it suffices to prove the validity of the first inequality in (64). By direct calculation we get:

$$\begin{aligned}
\epsilon_5(d) =&(d-1)\Big(p_{\text{true}}^2(\frac{1}{\alpha}-\frac{1}{\beta d})+2p_{\text{true}}(1-p_{\text{true}})(-\frac{1}{d}(\frac{1}{\alpha}+\frac{1}{\beta d})+(1-\frac{1}{d})(\frac{1}{\alpha}-\frac{1}{\beta d}))\\
&+(1-p_{\text{true}})^2(\frac{1}{\alpha}+\frac{1}{\beta d})(1-\frac{2}{d}-\frac{2}{d-1}(\frac{1}{d}-1)^2)\Big)\\
&+2\Big(p_{\text{true}}+\frac{1-p_{\text{true}}}{d}\Big)\Big(\frac{p_{\text{true}}}{\beta d}+\frac{1-p_{\text{true}}}{d}(\frac{1}{\beta d}-\frac{1}{\alpha})-\frac{1}{\beta d}\Big)\\
&-\Big(\frac{d-1}{\beta d}-\frac{1}{\alpha}\Big)\Big(1-2(p_{\text{true}}+\frac{1-p_{\text{true}}}{d})\Big)-\frac{(d-2)}{\alpha}+1\\
=&2\Big[\frac{p_{\text{true}}^2}{\beta}+(\frac{1}{\alpha}-\frac{1}{\beta})p_{\text{true}}+(\frac{1}{2}-\frac{1}{\alpha})\Big]+(\frac{4}{\alpha}-\frac{2}{\beta})\frac{(p_{\text{true}}-1)^2(d-1)}{d^2}+\frac{2p_{true}}{\beta d}\\
\geq&\frac{2p_{true}}{\beta d}\geq\frac{1}{4d},
\end{aligned} \tag{65}$$

where the inequalities in the last line follow from assumptions (A3), (A4), and (A5).

Next we show that event A^3 occurs with high probability. Denote by A_1^3 (resp. A_2^3) the event that inequalities (62) (resp. inequalities (63)) are satisfied. Then $A^3 = A_1^3 \cap A_2^3$. To show that A^3 occurs with high probability, it suffices to show that A_1^3 and A_2^3 occur with probability. In the following we show that A_1^3 occurs with high probability. The proof for A_2^3 follows from a similar line of arguments.

$$\begin{aligned}
&\mathbb{P}(A_1^3)\\
\geq&\mathbb{P}\Bigg(\bigcap_{t,q,i,j}\Big\{W_t(i,j)+\frac{1}{n}\sum_{k\in[n]}Q_{t,k}(i,j)+\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}}Y_{t,q,k}^5(i,j)>0\Big\}\Bigg)\\
=&\mathbb{P}\Bigg(\bigcap_{t,q,i,j}\Big\{W_t(i,j)+\frac{1}{n}\sum_{k\in[n]}Q_{t,k}(i,j)-\mathbb{E}\Big[\frac{1}{n}\sum_{k\in[n]}Q_{t,k}(i,j)\Big]+\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}}Y_{t,q,k}^5(i,j)\\
&-\mathbb{E}\Big[\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}}Y_{t,q,k}^5(i,j)\Big]>-\epsilon_5(d)\Big\}\Bigg)\\
\geq&\mathbb{P}\Bigg(\bigcap_{t,q,i,j}\Big\{\Big|W_t(i,j)+\frac{1}{n}\sum_{k\in[n]}Q_{t,k}(i,j)-\mathbb{E}\Big[\frac{1}{n}\sum_{k\in[n]}Q_{t,k}(i,j)\Big]+\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}}Y_{t,q,k}^5(i,j)\\
&-\mathbb{E}\Big[\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}}Y_{t,q,k}^5(i,j)\Big]\Big|<\epsilon_5(d)\Big\}\Bigg)\\
\geq&\mathbb{P}\Bigg(\bigcap_{t,i,j}\Big\{\Big|W_t(i,j)\Big|<\frac{\epsilon_5(d)}{3}\Big\}\cap\bigcap_{t,i,j}\Big\{\Big|\frac{1}{n}\sum_{k\in[n]}Q_{t,k}(i,j)-\mathbb{E}\Big[\frac{1}{n}\sum_{k\in[n]}Q_{t,k}(i,j)\Big]<\frac{\epsilon_5(d)}{3}\Big\}\\
\cap&\bigcap_{t,q,i,j}\Big\{\Big|\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}}Y_{t,q,k}^5(i,j)-\mathbb{E}\Big[\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}}Y_{t,q,k}^5(i,j)\Big]\Big|<\frac{\epsilon_5(d)}{3}\Big\}\Big|\Bigg),
\end{aligned} \tag{66}$$

where all inequalities follow by set inclusion. It can be checked that $|W_t(i,j)| \leq \frac{d}{n}$; hence by (A6) we get $|W_t(i,j)| \leq \frac{1}{12d}$. Moreover by (65) we have $\frac{\epsilon_5(d)}{3} \geq \frac{1}{12d}$. So we deduce that $|W_t(i,j)| \leq \frac{\epsilon_5(d)}{3}$, which in turn implies

$$\Big\{\Big|W_t(i,j)\Big|\geq\frac{\epsilon_5}{3}\Big\}=\emptyset. \tag{67}$$

Combining (66) and (67) we obtain:

$$
\begin{aligned}
&\mathbb{P}(A_1^3)\\
\geq& 1-\mathbb{P}\Bigg(\bigcup_{t,i,j}\Big\{\Big|\frac{1}{n}\sum_{k\in[n]}Q_{t,k}(i,j)-\mathbb{E}\Big[\frac{1}{n}\sum_{k\in[n]}Q_{t,k}(i,j)\Big]\Big|\geq\frac{\epsilon_5(d)}{3}\Big\}\\
&\cup\bigcup_{t,q,i,j}\Big\{\Big|\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}}Y^5_{t,q,k}(i,j)-\mathbb{E}\Big[\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}}Y^5_{t,q,k}(i,j)\Big]\Big|\geq\frac{\epsilon_5(d)}{3}\Big\}\Bigg)\\
\geq& 1-\sum_{t,i,j}\mathbb{P}\left(\Big|\frac{1}{n}\sum_{k\in[n]}Q_{t,k}(i,j)-\mathbb{E}\Big[\frac{1}{n}\sum_{k\in[n]}Q_{t,k}(i,j)\Big]\Big|\geq\frac{\epsilon_5(d)}{3}\right)\\
&-\sum_{t,q,i,j}\mathbb{P}\left(\Big|\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}}Y^5_{t,q,k}(i,j)-\mathbb{E}\Big[\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}}Y^5_{t,q,k}(i,j)\Big]\Big|\geq\frac{\epsilon_5(d)}{3}\right)\\
\geq& 1-2dn^2\exp\Big(-\frac{\epsilon_5^2(d)n}{9}\Big)-2d^2(n-2)^2\exp\Big(-\frac{\epsilon_5^2(d)(n-2)}{162d^2}\Big),
\end{aligned}
$$

where the second inequality follows from taking the union bound. The last inequality follows from the application of Hoeffding's inequality [37] by noting that for each (t,i,j) the random variables $Q_{t,k}(i,j)$ are independent for all $k\in[n]$ and can be bounded as $|Q_{t,k}(i,j)|\leq 1$. Moreover, for each (t,q,i,j) the random variables $Y^5_{t,q,k}(i,j)$ are independent for all $k\in[n]\setminus\{i,j\}$ and can be bounded as $|Y^5_{t,q,k}(i,j)|\leq 6d$. By (65), the result follows by letting $n\to\infty$, since $d\in o((n/\log(n))^{\frac{1}{4}})$. □

Lemma 6. *Suppose that $X^{\text{in}}(i,j)$ for all $1\leq i<j\leq n$ are generated according to the random corruption model* (34)*. If assumptions* (A1), (A3), (A4), (A5), *and* (A6) *are satisfied, and $d\in o((n/\log(n))^{\frac{1}{4}})$, then events A^4 and A^5 occur with high probability.*

Proof. We start by defining some random variables associated with inequalities of conditions (C4) and (C5). For each $1\leq i<j\leq n$, for each $t,q\in[d]$ and for each $k\in[n]\setminus\{i,j\}$ define:

$$
Y^7_{t,q,k}(i,j):=\begin{cases}
Z_{q,k}(i,j)+\frac{a_{tq}(k,j)-1}{\alpha}+\frac{1-a_{tt}(k,j)}{\beta d}+\frac{a_{tq}(i,k)-1}{\alpha}+\frac{-1-a_{qq}(i,k)}{\beta d} & \\
\qquad\qquad\text{if } a_{tt}(k,i)=-1,\ a_{qq}(k,j)=-1 & \\
Z_{q,k}(i,j)+\frac{a_{tq}(i,k)-1}{\alpha}+\frac{-1-a_{qq}(i,k)}{\beta d} & \\
\qquad\qquad\text{if } a_{tt}(k,i)=1,\ a_{qq}(k,j)=-1 & \\
Z_{q,k}(i,j)+\frac{a_{tq}(k,j)-1}{\alpha}+\frac{1-a_{tt}(k,j)}{\beta d} & \\
\qquad\qquad\text{if } a_{tt}(k,i)=-1,\ a_{qq}(k,j)=1 & \\
Z_{q,k}(i,j) & \text{if } a_{tt}(k,i)=1,\ a_{qq}(k,j)=1,
\end{cases}
$$

where $Z_{q,k}(i,j)$ is defined by (59). It can be checked that condition (C4) is strictly satisfied, if for all $t,q\in[d]$ and all $1\leq i<j\leq n$ with $a_{tt}(i,j)=1, a_{qq}(i,j)=-1, a_{tq}(i,j)=1$, we have

$$
W_q(i,j)+\frac{1}{n}\sum_{k\in[n]}Q_{q,k}(i,j)+\frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}}Y^7_{t,q,k}(i,j)>0,
$$

where $Q_{q,k}(i,j)$ $W_q(i,j)$ and given by (60) and (61), respectively. For each $1\leq i<j\leq n$, for

each $t, q \in [d]$ and for each $k \in [n] \setminus \{i, j\}$ define:

$$Y^8_{t,q,k}(i,j) := \begin{cases} Z_{t,k}(i,j) + \frac{a_{tq}(k,j)-1}{\alpha} + \frac{-1-a_{tt}(k,j)}{\beta d} + \frac{a_{tq}(i,k)-1}{\alpha} + \frac{1-a_{qq}(i,k)}{\beta d} \\ \qquad\qquad \text{if } a_{tt}(k,i) = -1,\, a_{qq}(k,j) = -1 \\ Z_{t,k}(i,j) + \frac{a_{tq}(i,k)-1}{\alpha} + \frac{1-a_{qq}(i,k)}{\beta d} \\ \qquad\qquad \text{if } a_{tt}(k,i) = 1,\, a_{qq}(k,j) = -1 \\ Z_{t,k}(i,j) + \frac{a_{tq}(k,j)-1}{\alpha} + \frac{-1-a_{tt}(k,j)}{\beta d} \\ \qquad\qquad \text{if } a_{tt}(k,i) = -1,\, a_{qq}(k,j) = 1 \\ Z_{t,k}(i,j) \qquad\qquad \text{if } a_{tt}(k,i) = 1,\, a_{qq}(k,j) = 1 \end{cases}$$

It then follows that condition (C5) is strictly satisfied, if for all $t, q \in [d]$ and all $1 \le i < j \le n$ with $a_{tt}(i,j) = -1, a_{qq}(i,j) = a_{tq}(i,j) = 1$, we have

$$W_t(i,j) + \frac{1}{n}\sum_{k\in[n]} Q_{t,k}(i,j) + \frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^8_{t,q,k}(i,j) > 0.$$

We first prove that the following inequalities hold for all $t, q \in [d]$ and for all $1 \le i < j \le n$:

$$\begin{aligned} \epsilon_7(d) &:= \mathbb{E}\Big[\frac{1}{n}\sum_{k\in[n]} Q_{q,k}(i,j) + \frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^7_{t,q,k}(i,j)\Big] > 0, \\ \epsilon_8(d) &:= \mathbb{E}\Big[\frac{1}{n}\sum_{k\in[n]} Q_{t,k}(i,j) + \frac{1}{n-2}\sum_{k\in[n]\setminus\{i,j\}} Y^8_{t,q,k}(i,j)\Big] > 0. \end{aligned} \tag{68}$$

Notice that by symmetry, $\epsilon_7(d) = \epsilon_8(d)$; hence, it suffices to prove the validity of the first inequality in (68). By direct computation we get

$$\begin{aligned} \epsilon_7(d) =& (d-1)\Big(p^2_{\text{true}}(\frac{1}{\alpha} - \frac{1}{\beta d}) + 2p_{\text{true}}(1-p_{\text{true}})(-\frac{1}{d}(\frac{1}{\alpha} + \frac{1}{\beta d}) + (1-\frac{1}{d})(\frac{1}{\alpha} - \frac{1}{\beta d})) \\ &+(1-p_{\text{true}})^2(\frac{1}{\alpha} + \frac{1}{\beta d})(1 - \frac{2}{d} - \frac{2}{d-1}(\frac{1}{d} - 1)^2)\Big) \\ &+4\Big(p_{\text{true}} + \frac{1-p_{\text{true}}}{d}\Big)\Big(\frac{p_{\text{true}}}{\beta d} + \frac{1-p_{\text{true}}}{d}(\frac{1}{\beta d} - \frac{1}{\alpha}) - \frac{1}{2\beta d}\Big) \\ &-\Big(\frac{d-1}{\beta d} - \frac{1}{\alpha}\Big)\Big(1 - 2(p_{\text{true}} + \frac{1-p_{\text{true}}}{d})\Big) - \frac{d-2}{\alpha} + 1 \\ =& \epsilon_5(d) + \frac{p_{\text{true}}}{\beta d} + \frac{1-p_{\text{true}}}{d}(\frac{1}{\beta d} - \frac{1}{\alpha}) \\ \ge& \epsilon_5(d), \end{aligned}$$

where $\epsilon_5(d)$ is defined by (64). The last inequality follows from the proof of Lemma 3 which in turn follows from assumption (A1). Hence, $\epsilon_7(d) > 0$ is implied by $\epsilon_5(d) > 0$ which by proof of Lemma 5 is valid if assumptions (A3), (A4), and (A5) hold. It remains to show that events A^4, A^5 occur with high probability. Note that for each (t, q, i, j), random variables $Y^7_{t,q,k}(i,j)$ (resp. $Y^8_{t,q,k}(i,j)$) are independent for all $k \in [n] \setminus \{i, j\}$ and can be bounded as $|Y^7_{t,q,k}(i,j)| \le 6d$ (resp. $|Y^8_{t,q,k}(i,j)| \le 6d$). Hence, by assumption (A6), the associated proof for event A^3_1 of Lemma 5 can be repeated verbatim to show A^4 and A^5 occur with high probability. □

We are now ready to prove the main result of this section.

Proof of Theorem 1. Let $\alpha = 1.172$ and $\beta = 1.657$; then it is simple to verify that for $p_{\text{true}} > 0.585$, assumptions (A1)-(A5) of Lemmata 3-6 are satisfied. Moreover, assumption $d \in o((n/\log(n))^{\frac{1}{4}})$ implies that (A6) is satisfied for n large enough and this completes the proof. □

In the proofs of Lemmata 3-6, to prove the positivity of expectations of certain random variables, we show that these expectations are non-increasing functions of d and hence we show their positivity for the limiting case $d \to \infty$. Indeed assumptions (A1)-(A5) can be slightly relaxed if such asymptotic analysis is avoided. That is, for each $d \geq 2$, we consider the following optimization problem:

$$\begin{aligned}
\min_{\alpha,\beta,p_{\text{true}}} \quad & p_{\text{true}} \\
\text{s.t.} \quad & \frac{p_{\text{true}}}{\beta} + (1-p_{\text{true}})(\frac{1}{\beta d} - \frac{1}{\alpha}) \geq 0, \\
& \Big(p_{\text{true}} + \frac{1-p_{\text{true}}}{d}\Big)\Big(\frac{p_{\text{true}}}{\beta d} + (\frac{1}{\beta d} - \frac{1}{\alpha})\frac{1-p_{\text{true}}}{d} + \frac{1}{\alpha}\Big) \geq \frac{1}{2}, \\
& 2\Big(\frac{p_{\text{true}}^2}{\beta} + (\frac{1}{\alpha} - \frac{1}{\beta})p_{\text{true}} + (\frac{1}{2} - \frac{1}{\alpha})\Big) + (\frac{4}{\alpha} - \frac{2}{\beta})\frac{(p_{\text{true}}-1)^2(d-1)}{d^2} + \frac{2p_{\text{true}}}{\beta d} \geq 0, \\
& \frac{1}{4} \leq p_{\text{true}} \leq 1, \\
& 0 \leq \alpha \leq \beta \leq 2,
\end{aligned} \tag{69}$$

where the first constraint is stated in the proof of Lemma 3, the second constraint is stated in the proof of Lemma 4, and the third constraint is stated in the proof of Lemma 5. Then for each $d \in [2, +\infty)$, the optimal value p^*_{true} of the above problem serves as a recovery threshold of Problem (P). While we are unable to obtain an analytical solution for Problem (69), we solve this problem numerically with high accuracy using the global solver `BARON` [25] for $d \in [2, 50]$. Results are depicted in Figure 1; for example, this figure indicate that for $d = 2$, the LP recovers the ground truth with high probability if $p_{\text{true}} > \frac{1}{3}$. While by performing a worst-case analysis, we have obtained a more conservative bound of $p_{\text{true}} > 0.585$ for the recovery threshold of the LP, Figure 1 suggests that this bound is a good approximation for $d \gtrsim 40$.

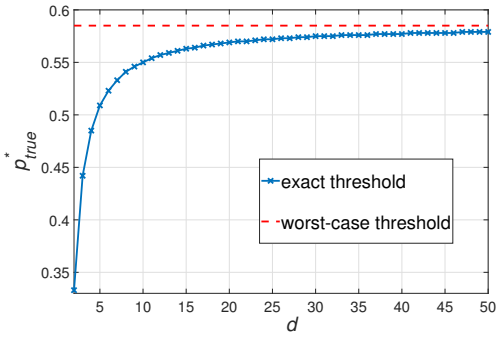


Figure 1: Recovery thresholds for Problem (P): the worst-case recovery threshold $p^*_{\text{true}} = 0.585$ is given by Theorem 1 and the exact recovery threshold is obtained by solving Problem (69) numerically for each $d \in [2, 50]$.

We conclude this section by acknowledging that this theoretical study serves as a starting point for understanding the recovery properties of LP relaxations for joint object matching. Obtaining recovery guarantees for the general problem with partially similar objects and incomplete map graphs together with investigating the impact of consistency inequalities (11) on the power of LP relaxations are topics of future research.

4 Numerical Experiments

In this section, we conduct a preliminary numerical study to demonstrate the desirable numerical properties of proposed LP relaxations for joint object matching. A comprehensive

Table 1: Exact solution of permutation group synchronization with $n = 20$ using the MIP solver `Gurobi`. All CPU times are reported in seconds and are averaged over 5 random instances. g_{rel} denotes the MIP relative gap upon termination and is averaged over 5 random instances.

(d, p_{true})	MIP time	g_{rel} (%)	LP time
(3, 0.3)	> 25000	11.5%	3.8
(3, 0.4)	> 25000	7.5%	2.2
(3, 0.5)	> 25000	10.5%	2.8
(4, 0.3)	> 25000	15.5%	5.3
(4, 0.4)	> 25000	24.9%	10.2
(4, 0.5)	> 25000	21.1%	11.4

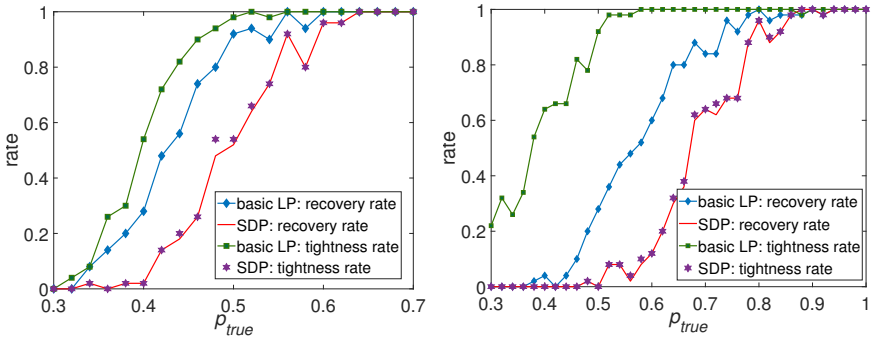
computational study that includes various real data sets from the literature is a topic of future research. Throughout this section, we focus on the special case of permutation group synchronization problem and generate problem instances according to the random corruption model (34). Moreover, our numerical experiments are all performed on the `NEOS` server [10].

4.1 Exact solution of the ILP

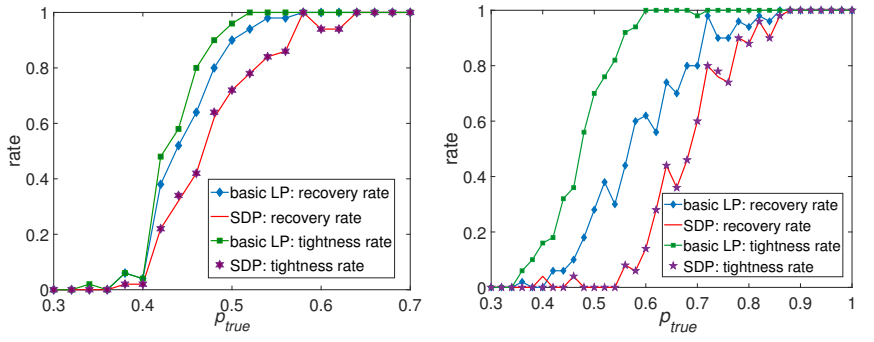
We start by demonstrating that even for small instances, the exact solution of joint object matching using the state-of-the-art MIP solvers is beyond reach. To this end, we solve Problem (IP) using `GAMS/Gurobi` [18]. We set a time limit of 25,000 seconds; all other parameters are set to their default values. We set $n = 20$, $d \in \{3, 4\}$, and $p_{\text{true}} \in \{0.30, 0.40, 0.50\}$. For each combination of (n, d, p_{true}), we run 5 random instances. We find that the MIP solver is unable to solve any of these instances to optimality within the time limit. More detailed results are shown in Table 1; for each (n, d, p_{true}), we provide the average relative gap of the MIP upon termination. The relative gap is defined as $g_{\text{rel}} = \frac{|BP-BF|}{|BF|}$, where BF is the objective function value of the current best integer solution, while BP is the best possible integer solution. For each case, we also report the average solution time of the corresponding basic LP given by Problem (P). As practical instances of joint object matching have larger values for (n, d), this experiment conveys the need for developing efficient convex relaxations for this problem.

4.2 The basic LP versus the SDP

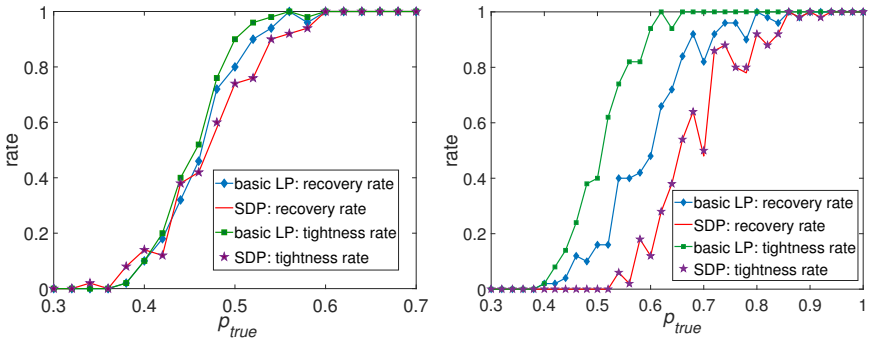
Next we compare recovery properties of the basic LP relaxation (P) with those of the popular SDP relaxation (7) strengthened with constraint (8). In the following, we consider a more general random model than the random corruption model given by (34); namely, we do not assume that the map graph is complete. Let $p_{\text{obs}} \in (0, 1]$; for each $1 \leq i < j \leq n$, with probability p_{obs} the input map $X^{\text{in}}(i, j)$ is generated according to (34); otherwise, the input map between objects i and j is not observed and we set $a_{tq}(i, j) = 0$ for all $t, q \in [d]$. For our numerical experiments, we set $n = 20$, $d \in \{3, 4, 5\}$, and $p_{\text{obs}} \in \{\frac{1}{2}, 1\}$. We set $p_{\text{true}} \in [\underline{p} : 0.02 : \bar{p}]$, where $[\underline{p} : 0.02 : \bar{p}]$ denotes a regularly-spaced vector between $\underline{p}$ and $\bar{p}$ using 0.02 as the increment between elements, and where $\underline{p}$ (resp. $\bar{p}$) is chosen small enough (resp. large enough) so that the recovery rate is zero (resp. one). For each combination of $(n, d, p_{\text{true}}, p_{\text{obs}})$ we conduct 50 random trials. We count the number of times the optimization algorithm returns the ground truth $\bar{X}$ as the optimal solution; dividing this number by total number of trials, we obtain the empirical recovery rate. In addition to the empirical recovery rate, we compute the empirical *tightness rate*. That is, we compute the fraction of times the optimization algorithm returns a binary solution. Recall that if the solution of the LP/SDP relaxation is binary, then it is feasible for Problem (JOM) and hence is optimal for the original nonconvex problem.



(a) $p_{\text{obs}} = 1.0,\ d = 3$ (b) $p_{\text{obs}} = 0.5,\ d = 3$



(c) $p_{\text{obs}} = 1.0,\ d = 4$ (d) $p_{\text{obs}} = 0.5,\ d = 4$



(e) $p_{\text{obs}} = 1.0,\ d = 5$ (f) $p_{\text{obs}} = 0.5,\ d = 5$

Figure 2: Empirical rates of recovery and tightness of the basic LP relaxation versus the SDP relaxation under the random corruption model.

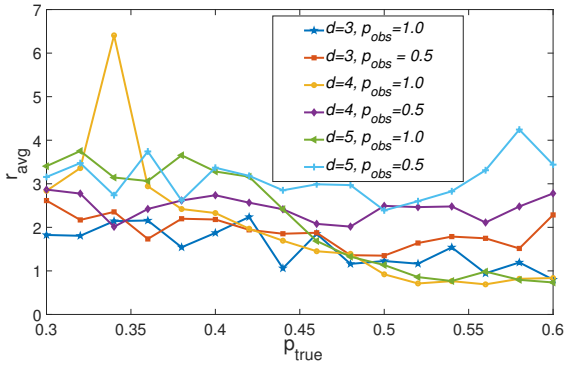


Figure 3: Comparing the CPU times of solving the basic LP relaxation (T_{LP}) vs solving the SDP relaxation (T_{SDP}), where we define $r_{\text{avg}} = \mathbb{E}[\frac{T_{\text{SDP}}}{T_{\text{LP}}}]$.

All LPs and SDPs are solved with GAMS/MOSEK [1]. Results are shown in Figure 2: in all cases, the basic LP outperforms the SDP in terms of recovery and tightness. Moreover, the CPU times are compared in Figure 3. As can be seen from this figure, in most of the experiments the LP solver is 2-4 times faster than the SDP solver. We should remark that we ran these experiments setting all options to their default values. That is, we did not use any technique to expedite either the LP or the SDP solver. In particular, for solving the LP relaxation employing a cutting plane type algorithm together with dual simplex often leads to significant speedups. From this experiment we make three important observations:

(i) Unlike the SDP relaxation, the quality of the basic LP relaxation degrades by increasing d; this is due to the fact that the basic LP includes only $\theta(d)$ of consistency inequalities which are $\theta(d2^{2d})$ in total; as we detail next, we address this shortcoming by considering a stronger LP relaxation, i.e., Problem (LPF).

(ii) Unlike the SDP relaxation, the quality of the basic LP relaxation does *not* quickly degrade by decreasing p_{obs}. For the random model described above and for the SDP relaxation, the authors of [7] obtain a recovery guarantee of the form $p_{\text{true}} > \theta(\frac{1}{\sqrt{p_{\text{obs}}}})$. Recall that our theoretical analysis in Section 3 is under the assumption that the map graph is complete, i.e., $p_{\text{obs}} = 1.0$. Our numerical experiments suggest that in case of the basic LP, as a function of p_{obs}, the recovery threshold grows slower than $\frac{1}{\sqrt{p_{\text{obs}}}}$; we plan to explore the exact form of this relationship as a next step.

(iii) Unlike the SDP relaxation, the basic LP relaxation returns a binary solution in many cases for which it fails in recovering the ground truth. Recall that if the LP solution is binary, then it is optimal for Problem (JOM). Indeed, in most practical settings, asking for the exact recovery of the ground truth is not realistic. Hence obtaining sufficient conditions for the tightness of the LP relaxation under proper stochastic models is of key importance and is a topic of future research.

4.3 The basic LP versus the double LP

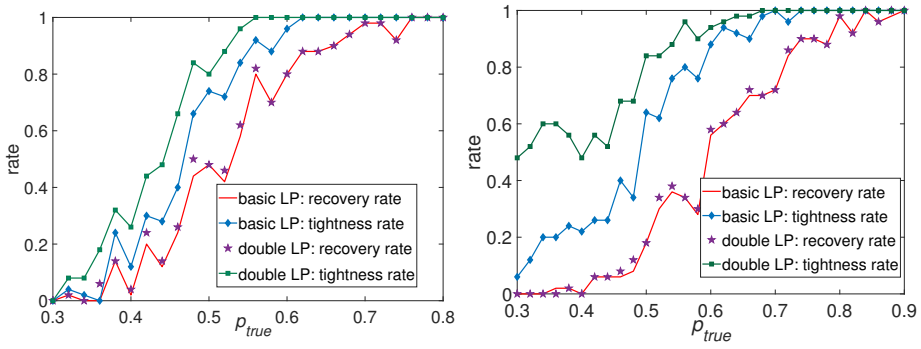
We now illustrate the impact of consistency inequalities (11) in strengthening the basic LP relaxation (P). To this end, we consider the following simple two-step algorithm:

1. The basic LP relaxation, i.e., Problem (P) is solved; if the optimal solution $\tilde{X}$ is binary-valued, then the algorithm terminates and returns $\tilde{X}$ as the optimal solution.

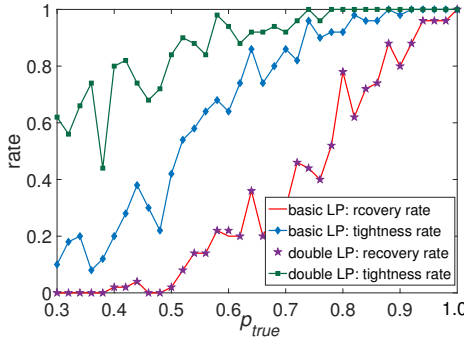
2. If $\tilde{X}$ is not binary-valued, then at most 1000 consistency inequalities that violate $\tilde{X}$ are generated by solving the separation problem (15). The cutting planes are added to the basic LP and the augmented LP is solved using the dual simplex algorithm with $\tilde{X}$ as the starting point. The solution of this augmented LP is reported as the optimal solution.

We refer to this two-step algorithm as the *double LP*. To understand the impact of consistency inequalities (11), we compare the performance of basic LP and double LP with respect to recovery and tightness rates.

We set $n = 10$, $d = 5$, $p_{\text{obs}} \in \{0.5, 0.75, 1\}$, and $p_{\text{true}} \in [0.3 : 0.02 : 1.0]$. As before, for each combination of $(n, d, p_{\text{true}}, p_{\text{obs}})$, we conduct 50 random trials. As the cutting plane algorithm described above cannot be efficiently implemented in the GAMS modeling language, the two-step algorithm is implemented in JuMP [14] and all corresponding LPs are solved with Gurobi [18]. Results are depicted in Figure 4; as can be seen from this figure, while there is no visible difference between the recovery rates of the basic LP and the double LP, the tightness rate of the double LP is significantly better than that of the basic LP. This improvement is more significant for sparser problems; that is, problems with smaller p_{obs}. We believe that



(a) $p_{\text{obs}} = 1.0$ (b) $p_{\text{obs}} = 0.75$



(c) $p_{\text{obs}} = 0.5$

Figure 4: Empirical rates of recovery and tightness of the basic LP relaxation versus the double LP relaxation under the random corruption model.

this phenomenon is due to the cutting plane strategy employed in Step 2 of double LP and is not an inherent property of consistency inequalities. Namely, in order to keep the overall computational cost low, in Step 2 of double LP, we add the first 1000 violated consistency inequalities obtained by solving the separation problem (15). For denser problems, that is, for problems with larger p_{obs}, in almost all cases, many more consistency inequalities are violated by $\tilde{X}$. Hence, an upper bound of 1000 for the number of violated inequalities is often too small. However, increasing this number will increase the cost of solving the augmented LP. Hence, for dense problems, a more elaborate strategy such as adding top 1000 most violated cuts or conducting multiple rounds of cut generation seems more appropriate.

We conclude by emphasizing that our goal in this experiment was to convey the effectiveness of consistency inequalities (11). Indeed, solving double LP is not equivalent to solving Problem (LPF) as double LP contains only a small subset of consistency inequalities (11). An efficient solution of Problem (LPF) via devising effective cut generation and cut management algorithms for consistency inequalities is a subject of future research. Finally, an effective incorporation of consistency inequalities in a branch-and-cut framework to solve Problem (JOM) to global optimality is an interesting future direction as well.

5 Appendix

5.1 Facetness for block consistency inequalities

Proof of Proposition 8. Without loss of generality, we prove for any nonempty $D_1, D_2, D_3 \subseteq [d]$ with $|D_1| + |D_2| > |D_3|$, the following defines a facet of the joint matching polytope $\mathcal{C}_{n,d}$:

$$\sum_{l \in D_3} \Big(\sum_{t \in D_1} X_{lt}(1,2) + \sum_{q \in D_2} X_{lq}(1,3) \Big) - \sum_{t \in D_1} \sum_{q \in D_2} X_{tq}(2,3) \leq |D_3|. \tag{70}$$

We start by identifying the set of consistent partial maps in $\mathcal{C}_{n,d}$ that satisfy inequality (70) tightly. Subsequently, we show that any nontrivial valid inequality $\alpha X \leq \beta$ for $\mathcal{C}_{n,d}$ that is satisfied tightly at all such maps coincides with (70) up to a positive scaling. Since $\mathcal{C}_{n,d}$ is full dimensional, this in turn implies that inequality (70) defines a facet of $\mathcal{C}_{n,d}$.

A consistent partial map is binding for inequality (70), if for every $l \in D_3$, there exists $e_l \in \mathcal{M}_X$ with $1_l \in e_l$ satisfying one of the following conditions:

(i) $2_t \in e_l$ for some $t \in D_1$ and $3_q \notin e_l$ for all $q \in D_2$,

(ii) $2_t \notin e_l$ for all $t \in D_1$ and $3_q \in e_l$ for some $q \in D_2$,

(iii) $2_t \in e_l$ for some $t \in D_1$ and $3_q \in e_l$ for some $q \in D_2$.

Now consider a consistent partial map $\mathcal{M}_X^1$ satisfying conditions (i) or (ii) above for all $l \in D_3$ such that for some $t' \in D_1$ we have $2_{t'} \notin e_l$ for all $l \in D_3$. Notice that such a consistent partial map exists since by assumption $|D_1| + |D_2| > |D_3|$. Moreover, suppose that $\mathcal{M}_X^1$ contains no matched pairs other than those required by conditions (i)-(ii), i.e., $\mathcal{M}_X^1 = \{e_l,\ l \in D_3\}$ with $|e_l| = 2$ for all $l \in D_3$. Next consider another consistent partial map obtained from $\mathcal{M}_X^1$ by replacing a matched pair of the form $(1_{\hat{l}}, 2_{\hat{t}})$ for some $\hat{t} \in D_1$ and $\hat{l} \in D_3$ by $(1_{\hat{l}}, 2_{t'})$, where $t' \in D_1$ is the index defined above. Notice that this *flipping* operation results in a consistent partial map that is also binding for (70). Substituting these two partial maps in $\alpha X = \beta$ yields $\alpha_{\hat{l}\hat{t}}(1,2) = \alpha_{\hat{l}t'}(1,2)$. Using a similar line of arguments for all possible consistent partial maps satisfying conditions (i) or (ii) for all $l \in D_3$ together with $2_{t'} \notin e_l$ for some $t' \in D_1$ ($3_{t'} \notin e_l$ for some $t' \in D_2$) for all $l \in D_3$, we conclude that for each $l \in D_3$, we have

$$\alpha_{lt}(1,2) = \alpha_{lq}(1,3), \quad \forall t \in D_1,\ \forall q \in D_2. \tag{71}$$

Next consider a consistent partial map $\mathcal{M}_X^2$ satisfying conditions (i) or (ii) above for all $l \in \mathcal{D}_3 \setminus \{\hat{l}\}$ and suppose that for $\hat{l}$ condition (iii) is satisfied, i.e., $e_{\hat{l}} = (1_{\hat{l}}, 2_{\hat{t}}, 3_{\hat{q}})$ for some $\hat{t} \in D_1$ and $\hat{q} \in D_2$. Consider a second consistent partial map satisfying conditions (i) or (ii) above for all $l \in \mathcal{D}_3 \setminus \{\tilde{l}\}$ where $\tilde{l} \neq \hat{l}$ and suppose that for $\tilde{l}$ condition (iii) is satisfied with $e_{\tilde{l}} = (1_{\tilde{l}}, 2_{\hat{t}}, 3_{\hat{q}})$. In addition, suppose that in both these maps, no additional matched pairs other than those required by conditions (i)-(iii) exist. Substituting these two maps in $\alpha X = \beta$ and using (71) yield $\alpha_{\hat{l}\hat{t}}(1,2) = \alpha_{\tilde{l}\hat{t}}(1,2)$ and $\alpha_{\hat{l}\hat{q}}(1,3) = \alpha_{\tilde{l}\hat{q}}(1,3)$. Using a similar line of arguments for all possible pairs of partial maps satisfying assumptions above, we obtain:

$$\alpha_{lt}(1,2) = \alpha_{lq}(1,3) = \frac{\beta}{|D_3|}, \quad \forall t \in D_1,\ \forall q \in D_2,\ \forall l \in D_3. \tag{72}$$

Moreover, substituting the partial map corresponding to $\mathcal{M}_X^2$ in $\alpha X = \beta$ and using (72) we obtain $\alpha_{\hat{t}\hat{q}}(2,3) = -\beta/|D_3|$. It then follows that

$$\alpha_{tq}(2,3) = -\frac{\beta}{|D_3|}, \quad \forall t \in D_1,\ \forall q \in D_2. \tag{73}$$

Next consider a consistent partial map $\mathcal{M}_X^3$ satisfying conditions (i) or (ii) above for all $l \in \mathcal{D}_3$ of the form $\mathcal{M}_X^3 = \{e_l : l \in D_3\}$ with $|e_l| = 2$ for all $l \in D_3$. Construct another consistent partial map of the form $\mathcal{M}_X^3 \cup \{(i_t, j_q)\}$ for some $(i,j,t,q) \in \mathcal{Q}$, where

$$\begin{aligned}
Q :=& \Big\{(i,j,t,q) : 4 \leq i < j \leq n, t \in [d], q \in [d]\Big\} \cup \Big\{(1,j,t,q) : j \geq 4, t \in [d] \setminus D_3, q \in [d]\Big\} \\
&\cup \Big\{(2,j,t,q) : 4 \leq j \leq n, t \in [d] \setminus D_1, q \in [d]\Big\} \cup \Big\{(3,j,t,q) : 4 \leq j \geq n, t \in [d] \setminus D_2, q \in [d]\Big\} \\
&\cup \Big\{(1,2,t,q) : t \in [d] \setminus D_3, q \in [d] \setminus D_1\Big\} \cup \Big\{(1,3,t,q) : t \in [d] \setminus D_3, q \in [d] \setminus D_2\Big\} \\
&\cup \Big\{(2,3,t,q) : t \in [d] \setminus D_1, q \in [d] \setminus D_2\Big\}.
\end{aligned}$$

Substituting these two partial maps in $\alpha X = \beta$ yields

$$\alpha_{tq}(i,j) = 0, \quad \forall (t,q,i,j) \in \mathcal{Q}. \tag{74}$$

Consider the consistent partial map $\mathcal{M}_X^3$ defined above with the additional assumption that for some $\hat{t} \in D_1$ (resp. $\hat{q} \in D_2$), we have $(1_l, 2_{\hat{t}}) \notin \mathcal{M}_X^3$ (resp. $(1_l, 3_{\hat{q}}) \notin \mathcal{M}_X^3$) for all $l \in D_3$. Notice that such a partial map always exists since by assumption $|D_1| + |D_2| > |D_3|$. Construct another consistent partial map of the form:

- $\bar{\mathcal{M}}_X^3 = \mathcal{M}_X^3 \cup \{(1_{\hat{l}}, 2_{\hat{t}})\}$ (resp. $\bar{\mathcal{M}}_X^3 = \mathcal{M}_X^3 \cup \{(1_{\hat{l}}, 3_{\hat{q}})\}$) for some $\hat{l} \notin D_3$. Substituting $\mathcal{M}_X^3, \bar{\mathcal{M}}_X^3$ in $\alpha X = \beta$ gives $\alpha_{\hat{l}\hat{t}}(1,2) = 0$ (resp. $\alpha_{\hat{l}\hat{t}}(1,3) = 0$). Using a similar line of arguments for all possible partial maps satisfying the assumptions above, we obtain

$$\alpha_{lt}(1,2) = \alpha_{lq}(1,3) = 0, \quad \forall l \in [d] \setminus D_3, \forall t \in D_1, \forall q \in D_2. \tag{75}$$

- $\bar{\mathcal{M}}_X^3 = \mathcal{M}_X^3 \cup \{(2_{\hat{t}}, j_s)\}$ (resp. $\bar{\mathcal{M}}_X^3 = \mathcal{M}_X^3 \cup \{(3_{\hat{q}}, j_s)\}$) for some $4 \leq j \leq n$ and some $s \in [d]$. Substituting $\mathcal{M}_X^3, \bar{\mathcal{M}}_X^3$ in $\alpha X = \beta$ gives $\alpha_{\hat{t}s}(2,j) = 0$ (resp. $\alpha_{\hat{q}s}(3,j) = 0$). Using a similar line of arguments for all possible partial maps satisfying the assumptions above, we obtain

$$\alpha_{ts}(2,j) = \alpha_{qs}(3,j) = 0, \quad \forall 4 \leq j \leq n, \forall t \in D_1, \forall q \in D_2, \forall s \in [d]. \tag{76}$$

- $\bar{\mathcal{M}}_X^3 = \mathcal{M}_X^3 \cup \{(2_{\hat{t}}, 3_q)\}$ for some $q \in [d] \setminus D_2$ (resp. $\bar{\mathcal{M}}_X^3 = \mathcal{M}_X^3 \cup \{(2_t, 3_{\hat{q}})\}$ for some $t \in [d] \setminus D_1$). Substituting $\mathcal{M}_X^3, \bar{\mathcal{M}}_X^3$ in $\alpha X = \beta$ gives $\alpha_{\hat{t}s}(2,j) = 0$ (resp. $\alpha_{\hat{q}s}(3,j) = 0$). Using a similar line of arguments for all possible partial maps satisfying the assumptions above, we obtain

$$\alpha_{tq}(2,3) = 0, \quad \forall t \in [d] \setminus D_1, q \in D_2 \text{ or } \forall t \in D_1, q \in [d] \setminus D_2. \tag{77}$$

Consider a consistent partial map $\mathcal{M}_X^4$ satisfying conditions (i) or (ii) above for all $l \in \mathcal{D}_3$ of the form $\mathcal{M}_X^4 = \{e_l : l \in D_3\}$ with $|e_l| = 2$ for all $l \in D_3$. Consider some $\hat{l} \in D_3$ for which $e_{\hat{l}} = (1_{\hat{l}}, 2_{\hat{t}})$ for some $\hat{t} \in D_1$ (resp. $e_{\hat{l}} = (1_{\hat{l}}, 3_{\hat{q}})$ for some $\hat{q} \in D_2$). Construct another consistent partial map of the form

- $\bar{\mathcal{M}}_X^4 = \mathcal{M}_X^4 \cup \{(1_{\hat{l}}, 3_{\hat{q}}), (2_{\hat{t}}, 3_{\hat{q}})\}$ for some $\hat{q} \in [d] \setminus D_2$ ($\bar{\mathcal{M}}_X^4 = \mathcal{M}_X^4 \cup \{(1_{\hat{l}}, 2_{\hat{t}}), (2_{\hat{t}}, 3_{\hat{q}})\}$ for some $\hat{t} \in [d] \setminus D_1$). Substituting these two partial maps in $\alpha X = \beta$ and using (77) we obtain $\alpha_{\hat{l}\hat{t}}(1,2) = 0$ (resp. $\alpha_{\hat{l}\hat{q}}(1,3) = 0$). More generally it can be checked that

$$\alpha_{lt}(1,2) = \alpha_{lq}(1,3) = 0, \quad \forall l \in D_3, \forall t \in [d] \setminus D_1, \forall q \in [d] \setminus D_2. \tag{78}$$

- $\bar{\mathcal{M}}_X^4 = \mathcal{M}_X^4 \cup \{(1_{\hat{l}}, j_s), (2_{\hat{t}}, j_s)\}$ (resp. $\bar{\mathcal{M}}_X^4 = \mathcal{M}_X^4 \cup \{(1_{\hat{l}}, j_s), (3_{\hat{q}}, j_s)\}$) for some $4 \leq j \leq n$ and $s \in [d]$. Substituting these two partial maps in $\alpha X = \beta$ and using (76) we obtain $\alpha_{\hat{l}s}(1,j) = 0$. More generally, it can be checked that

$$\alpha_{ls}(1,j) = 0, \quad \forall 4 \leq j \leq n, \forall l \in D_3, \forall s \in [d]. \tag{79}$$

From (70)-(79) it follows that the inequality $\alpha X \leq \beta$ can be equivalently written as

$$\frac{\beta}{|D_3|}\Big(\sum_{l \in D_3} \Big(\sum_{t \in D_1} X_{lt}(1,2) + \sum_{q \in D_2} X_{lq}(1,3) \Big) - \sum_{t \in D_1} \sum_{q \in D_2} X_{tq}(2,3) \Big) \leq \beta.$$

Since $\alpha X \leq \beta$ is nontrivial and valid, we have $\beta > 0$ and this completes the proof.

5.2 Size inequalities

In this section, we assume that an upper bound $\hat{m}$ on the size of the universe is available and we utilize $\hat{m}$ to improve our proposed LP relaxation. As in Section 2, consider a collection of n objects each consisting of d_i, $i \in [n]$ elements; that is, we have a total number of $\bar{d} = \sum_{i\in[n]} d_i$ elements. Let $\mathcal{N}$ denote the set consisting of all these elements, i.e., $\mathcal{N} = \cup_{i\in[n]} \cup_{t\in[d_i]} i_t$. Define $d_{\min} = \min_{i\in[n]} d_i$ and $d_{\max} = \max_{i\in[n]} d_i$, where we assume $d_{\max} \geq 2$. Let $\hat{m} \in \{d_{\max}, \ldots, \bar{d}-1\}$. Notice that $\hat{m} = \bar{d}$ is a trivial upper bound and cannot be exploited to improve the relaxation.

Proposition 10. *Consider a subset $\mathcal{N}' \subset \mathcal{N}$ of cardinality $\hat{m}+1$. Then the following inequality is valid for the feasible region of Problem* (IP)*:*

$$\sum_{1\leq i<j\leq n} \sum_{\substack{t\in[d_i] q\in[d_j]:\\ i_t, j_q \in \mathcal{N}'}} X_{tq}(i,j) \geq 1, \tag{80}$$

where i_t and j_q denote the t-th element of object $\mathcal{S}_i$ and the q-th element of object $\mathcal{S}_j$, respectively.

Proof. To see the validity of inequality (80), suppose that $X_{tq}(i,j) = 0$ for all $i_t, j_q \in \mathcal{N}'$. Since by assumption $|\mathcal{N}'| = \hat{m}+1$, it then follows that the size of the universe is at least $\hat{m}+1$, which contradicts with the assumption that $\hat{m}$ is an upper bound on the size of the universe. □

Henceforth, we refer to inequalities of the form (80) for all $\mathcal{N}' \subset \mathcal{N}$ as *size inequalities*. The following example demonstrates that size inequalities can tighten the proposed LP relaxation.

Example 6. *Let $n = 3$ and $d_1 = d_2 = d_3 = 2$; moreover, suppose that $\hat{m} = 3$. Then it can be checked that the following is feasible for Problem* (LPF)*:*

$$X(1,2) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad X(1,3) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad X(2,3) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{81}$$

Now consider a size inequality obtained by letting $\mathcal{N}' = \{1_1, 1_2, 2_1, 2_2\}$ in inequality (80)*:*

$$X_{11}(1,2) + X_{12}(1,2) + X_{21}(1,2) + X_{22}(1,2) \geq 1.$$

Substituting (81) *in the above inequality yields $0+0+0+0 \ngeq 1$.*

Next, suppose that $\hat{m} = 4$ and consider the point:

$$X(1,2) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad X(1,3) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad X(2,3) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{82}$$

Again, it can be checked that (82) *is feasible for Problem* (LPF)*. Consider now the size inequality obtained by letting $\mathcal{N}' = \{1_1, 2_1, 2_2, 3_1, 3_2\}$:*

$$X_{11}(1,2) + X_{12}(1,2) + X_{11}(1,3) + X_{12}(1,3) + X_{11}(2,3) + X_{12}(2,3) + X_{21}(2,3) + X_{22}(2,3) \geq 1.$$

Substituting (82) *in the above inequality yields $0+0+0+0+0+0+0+0 \ngeq 1$.*

Remark 3. *Let us revisit the graph partitioning problem described in Remark 1; namely, the problem of partitioning the nodes of a graph into at most K subsets [9]. Let $\tilde{N}$ denote a subset of $[n]$ with $|\tilde{N}| = K+1$. Then the* clique inequality *associated with $\tilde{N}$ is defined as*

$$\sum_{i,j\in\tilde{N}:i<j} y_{ij} \geq 1. \tag{83}$$

Now, let $t \in [d_{\min}]$ and consider a size inequality (80) *with $\mathcal{N}' = \{i_t : i \in I \subseteq [n]\}$, where $|I| = \hat{m}+1$; that is, the inequality:*

$$\sum_{i,j\in I:i<j} X_{tt}(i,j) \geq 1. \tag{84}$$

Comparing (83) *and* (84)*, it follows that for a particular type of* $\mathcal{N}'$*, the corresponding size inequalities have the same form as clique inequalities. However, we would like to remark that while a clique inequality* (83) *associated with a clique of size* r *contains* $\binom{r+1}{2}$ *variables, a size inequality* (84) *associated with* $\mathcal{N}'$ *of cardinality* r *may contain different number of variables.*

For instance, let $n = 5$*,* $d_i = 3$ *for all* $i \in \{1, \dots, 5\}$*, and suppose that* $\hat{m} = 4$*. Then letting* $\mathcal{N}' = \{1_1, 1_2, 1_3, 2_1, 2_2\}$ *in* (80)*, we obtain the size inequality*

$$X_{11}(1,2) + X_{12}(1,2) + X_{21}(1,2) + X_{22}(1,2) + X_{31}(1,2) + X_{32}(1,2) \geq 1,$$

consisting of six variables, while letting $\mathcal{N}' = \{1_1, 2_1, 3_1, 4_1, 4_2\}$*, we obtain a size inequality*

$$\begin{aligned} &X_{11}(1,2) + X_{11}(1,3) + X_{11}(1,4) + X_{12}(1,4) + X_{11}(2,3) + X_{11}(2,4) + X_{12}(2,4) \\ &+X_{11}(3,4) + X_{12}(3,4) \geq 1, \end{aligned}$$

consisting of nine variables.

Hence size inequalities (80) *can be considered a generalization of clique inequalities for joint object matching. It is well-known that the separation problem over clique inequalities* (83) *is NP-hard [15]. Various heuristics for separating clique inequalities have been proposed in the literature [15] and similar ideas could be developed to efficiently separate over size inequalities. However, such a computational study is beyond the scope of this paper.*

Remark 4. *Consider the following variant of joint object matching: find a collection of consistent partial maps* $X(i,j) \in \{0,1\}^{d_i \times d_j}$ *for all* $1 \leq i < j \leq n$ *corresponding to a universe of at most* $\hat{m}$ *elements so as to minimize* (6)*. It can be checked that this problem can be equivalently solved by solving the ILP obtained by adding all size inequalities* (80) *to Problem* (IP)*.*

Acknowledgements. Antonio De Rosa has been partially supported by the NSF DMS Grant No. 1906451, the NSF DMS Grant No. 2112311, and the NSF DMS CAREER Award No. 2143124.

References

[1] *MOSEK 9.2*, 2019. `http://docs.mosek.com/9.0/faq.pdf`.

[2] E. Abbe, A. S. Bandeira, and G. Hall. Exact Recovery in the Stochastic Block Model. *IEEE Transactions on information theory*, 62:471 – 487, 2016.

[3] S. Agarwal, Y. Furukawa, N. Snavely, I. Simon, B. Curless, S.M. Seitz, and R. Szeliski. Building rome in a day. *Communications of the ACM*, 54 (10), 2011.

[4] C. Bajaj, T. Gao, Z. He, Q. Huang, and Z. Liang. Smac: simultaneous mapping and clustering using spectral decompositions. *Proceedings of the 35th International Conference on Machine Learning (PMLR)*, 80:324–333, 2018.

[5] E. Boros and P.L. Hammer. Pseudo-boolean optimization. *Discrete Applied Mathematics*, 123(1):155–225, 2002.

[6] S. Burer and R. Monteiro. Local minima and convergence in low-rank semidefinite programming. *Mathematical Programming*, 103:427–444, 2005.

[7] Y. Chen, L. Guibas, and Q. Huang. Near-optimal joint object matching via convex relaxation. *Proceedings of the 31st International Conference on Machine Learning*, 32(2):100–108, 2014.

[8] Y. Chen, C. Suh, and A. J. Goldsmith. Information recovery from pairwise measurements. *IEEE Transactions on Information Theory*, 62:5881–5905, 2016.

[9] S. Chopra and M.R. Rao. The partition problem. *Mathematical Programming*, 59:87–115, 1993.

[10] J. Czyzyk, M. P. Mesnier, and J. J. More. The neos server. *IEEE Journal on Computational Science and Engineering*, 5(3):68 –75, 1998.

[11] A. De Rosa and A. Khajavirad. The ratio-cut polytope and K-means clustering. *SIAM Journal on Optimization*, 32: 173–203, 2022.

[12] A. Del Pia, A. Khajavirad, and D. Kunisky. Linear programming and community detection. *Mathematics of Operations Research*, DOI: 10.1287/moor.2022.1282, 2022.

[13] M.F. Demirci, A. Shokoufandeh, L. Keselman, Y. abd Bretzner, and S. Dickinson. Object recognition as many-to-many feature matching. *International Journal of Computer Vision*, 69:203–222, 2006.

[14] Iain Dunning, Joey Huchette, and Miles Lubin. JuMP: A modeling language for mathematical optimization. *SIAM Review*, 59(2):295–320, 2017.

[15] A Eisenblatter. Frequency assignment in gsm networks: Models, heuristics, and lower bounds. *PhD Thesis, Technical Universityof Berlin.*, 2001.

[16] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Princeton Mathematical Series. Springer-Verlag, Berlin, New York, 1988.

[17] M. Grötschel and Y. Wakabayashi. Facets of the clique partitioning polytope. *Mathematical Programming*, 47:367–387, 1990.

[18] Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2021.

[19] B. Hajek, Y. Wu, and J. Xu. Achieving exact cluster recovery threshold via semidefinite programming. *IEEE Transactions on Information Theory*, 62(5):2788 – 2797, 2016.

[20] P.L. Hammer, P. Hansen, and B. Simeone. Roof duality, complementation and persistency in quadratic $0-1$ optimization. *Mathematical Programming*, 28(2):121–155, 1984.

[21] N. Hu, Q. Huang, B. Thibert, and L. J. Guibas. Distributable consistent multi-object matching. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2463–2471, 2018.

[22] Q.-X. Huang and L. Guibas. Consistent shape maps via semidefinite programming. *Computer Graphics Forum*, 32(5):177–186, 2013.

[23] IBM. *CPLEX Optimizer*, 2016. `http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/`.

[24] T. Iguchi, D. G. Mixon, J. Peterson, and S. Villar. Probably certifiably correct K-means clustering. *Mathematical Programming*, 165:605–642, 2017.

[25] A. Khajavirad and N. V. Sahinidis. A hybrid LP/NLP paradigm for global optimization relaxations. *Mathematical Programming Computation*, 10:383–421, 2018.

[26] X. Li, Y. Li, S. Ling, T. Strohmer, and K. Wei. When do birds of a feather flock together? K-means, proximity, and conic programming. *Mathematical Programming*, 179:295–341, 2020.

[27] Y. Li, C. Gu, T. Dullien, O. Vinyals, and P. Kohli. Graph matching networks for learning the similarity of graph structured objects. *Proceedings of the 36th International Conference on Machine Learning (PMLR)*, 97:3835–3845, 2019.

[28] S. Ling. Near-optimal performance bounds for orthogonal and permutation group synchronization via spectral methods. *arXiv preprint arXiv:2008.05341*, 2020.

[29] O.L. Mangasarian. Uniqueness of solution in linear programming. *Linear Algebra and its Applications*, 25:151–162, 1979.

[30] A. Moitra, W. Perry, and A. S. Wein. How robust are reconstruction thresholds for community detection? *In STOC '16: Proceedings of the forty-eighth annual ACM Symposium on Theory of Computing*, pages 828–841, 2016.

[31] D. Pachauri, R. Kondor, G. Sargur, and V. Singh. Permutation diffusion maps with application to the image association problem in computer vision. *In Advances in Neural Information Processing Systems*, pages 541–549, 2014.

[32] D. Pachauri, R. Kondor, and V. Singh. Solving the multi-way matching problem by permutation synchronization. *In Advances in Neural Information Processing Systems*, pages 1860–1868, 2013.

[33] F. Ricci-Tersenghi, A. Javanmard, and A. Montanari. Performance of a community detection algorithm based on semidefinite programming. *Journal of Physics: Conference Series*, 699:12015–12025, 2016.

[34] A. Schrijver. *Theory of Linear and Integer Programming*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Chichester, Great Britain, 1986.

[35] A. Schrijver. A combinatorial algorithm minimizing submodular functions in strongly polynomial time. *Journal of Combinatorial Theory, Series B*, 80:346–355, 2000.

[36] Y. Shen, Q. Huang, N. Srebro, and S. Sanghavi. Normalized spectral map synchronization. *In Advances in Neural Information Processing Systems*, pages 4925–4933, 2016.

[37] R. Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge University Press, 2018.

[38] J. Yan, S. Yang, and E. R. Hancock. Learning for graph matching and related combinatorial optimization problems. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20)*, pages 4988–4996, 2020.

[39] X. Zhou, M. Zhu, and K. Daniilidis. Multi-image matching via fast alternating minimization. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 4032–4040, 2015.