Consensus-based Distributed Discrete Optimal Transport for Decentralized Resource Matching

Rui Zhang, Quanyan Zhu

Abstract—Optimal transport has been used extensively in resource matching to promote the efficiency of resources usages by matching sources to targets. However, it requires a significant amount of computations and storage spaces for large-scale problems. In this paper, we take a consensus-based approach to decentralize discrete optimal transport problems and develop fully distributed algorithms with alternating direction method of multipliers. We show that our algorithms guarantee certain levels of efficiency and privacy besides the distributed nature. We further derive primal and dual algorithms by exploring the primal and dual problems of discrete optimal transport with linear utility functions and prove the equivalence between them. We verify the convergence, online adaptability, and the equivalence between the primal algorithm and the dual algorithm with numerical experiments. Our algorithms reflect the bargaining between sources and targets on the amounts and prices of transferred resources and reveal an averaging principle which can be used to regulate resource markets and improve resource efficiency.

Index Terms—Optimal Transport, Consensus-Based Decentralization, Resource Allocation, Distributed Resource Matching, Alternating Direction Method of Multipliers, Resource Markets.

I. INTRODUCTION

Resource matching plays an important role in many applications, such as emergency response [1]–[3], smart grids [4], [5], wireless networks [6]–[8], and data centers [9], [10]. The objective of matching is to assign resources with targets in a way that promotes resource efficiency and increases social benefits. For example, proper assignments of workers with heterogeneous skills to jobs with heterogeneous characteristics can improve the total economic output [11]; efficient allocations of incoming customers to different parking slots can reduce the average time and costs of finding parking spaces and improve the overall parking capacity in smart cities [12].

Optimal transport is one of the centralized planning approaches to resource matching [11], [13]. A central planner finds the optimal scheme to transport or move resources from their sources to the targets that maximizes the social welfares. However, the computations required to solve such planning problem grow exponentially with the increase in the numbers of sources and targets in large-scale problems. Moreover, the communication overhead required for the planner to collect information and coordinate between sources and targets is also significant with a large number of participants (See Fig. 1(a)). In certain cases, targets and sources may not be willing to share any information to the central planner (See Fig. 1(b)). Therefore, it is computationally and structurally costly for the central planner to make a global planning for large-scale systems despite the recent efforts on developing methods to speed up the computations.

Decentralized methods and algorithms offer practical solutions to reduce the complexity of the computations and make large-scale matching problems feasible [14]–[16]. In this paper, we aim to decentralize discrete optimal transport problems and derive a fully distributed algorithm which dispenses with a central planner to develop the matching mechanism between sources and targets. We consider a consensus-based approach in which consensus constraints are used to capture the agreements between the participants and the central planner on the matching. This approach allows the derivations of a fully distributed algorithm for discrete optimal transport with alternating direction method of multipliers (ADMM) [17].

Our algorithm does not require central planners to compute the matching between sources and targets or store their information. Each participant computes its own problem and communicates with only the participants on the other side, which enables parallel computations and yields high efficiency compared to centralized algorithms. The consensus between participants and central planners is simplified into direct consensus between sources and targets. Each pair of source and target updates their matching until they reach an agreement. When all sources and targets reach consensus, the resulting
matching achieves the optimal social utility and coincides with the centralized solution.

Besides the property of distributed computation and the improved computational efficiency, our algorithm offers additional useful features on distributed resource matching. Our algorithm guarantees a certain level of privacy as there are no central planners to collect information and each participant only shares information with the participants on the other side. Our algorithm can be easily implemented online and does not require re-running the algorithm when participants leave or enter the matching platform. Instead, the algorithm can update the matching and adapt to changes in the preferences, resources, and targets.

We further explore the dual problem of discrete optimal transport with linear utilities which naturally provides a way to study pricing schemes of resource matching. We leverage the consensus-based approach to capture the bargaining and agreement process of the participants on the prices of the resources. We use ADMM to derive a dual algorithm which shares similar characteristics with the primal algorithm on fully distributed nature, efficiency, privacy, and online adaptability. We further prove the equivalence between the primal algorithm and the dual algorithm, and show that each sub-problem in the dual algorithm is the dual problem of each sub-problem in the primal algorithm.

Our algorithms provide useful insights of resource matching in market environments. The optimal matching is achieved through the bargaining between sources and targets on the amounts and prices of transferred resources in the primal algorithm and the dual algorithm, respectively. Each participant proposes its offers with the objective of maximizing its utility while minimizing its differences from the participants on the other side. When all pairs of sources and targets reach consensus on their motions, the achieved matching maximizes the overall surplus.

Furthermore, our algorithms reveal an averaging principle that sources and targets follow during their bargaining process. The core idea is that a source and a target should propose their current offers to be close to the average of their previous offers. The principle indicates a fair and unbiased negotiation between sources and targets. The convergence of our algorithms to the optimal matching shows that the sources and targets can maximize their total surplus by following the averaging principle. Thus, the principle can be used to regulate sources and targets in resource markets and improve the resource efficiency.

Our contributions can be summarized as follows:

- We propose a consensus-based approach to decentralize discrete optimal transport problems and further derive a fully distributed resource matching algorithm using ADMM. Each participant computes its own matching decisions and only communicates with participants on the other side.
- We demonstrate that our algorithm guarantees certain levels of efficiency and privacy and show that our algorithm is fully adaptive when changes occur in the preferences, resources, and targets.
- We derive a fully distributed dual algorithm that provides a pricing scheme of resource matching in distributed markets and reveal an averaging principle that a source and a target always compromise to the average of either amounts or prices of transferred resources between them.
- We use numerical experiments to demonstrate the convergence and online capability of our algorithm, and we also corroborate the equivalence of primal algorithm and dual algorithm.

A. Related Works

Our work is related to the literature on resource allocation, optimal transport, and distributed optimization. Resource allocation deals with the assignment or distribution of available resources in efficient ways [18], [19]. Recent applications include natural resource management such as lands and forests [20], [21], sensor networks [14], [22], smart grids [4], [23], human resources [24], and emergency response [2], [3].

Resource matching is one class of resource allocation with multiple sources and multiple targets, and it finds the optimal matching between multiple sources and multiple targets to transfer resources and maximize the total surplus of them [1]–[10]. Various approaches have been developed to study resource matching, which can be summarized into two distinct categories. Planning-based approaches focus on the designation of optimal resource matching to maximize social welfare [11], while market-based approaches study interactions between sources and targets in markets through the analysis of supply and demand of resources [25].

It is common for most planning-based frameworks to have a central planner identify resources, sources, and targets, and then design the matching [11]. However, this approach requires a significant amount of computation for large-scale problems. A lot of techniques such as approximation [26], belief propagation [27], barrier methods [28], and network simplex [29], [30], have been developed to speed up the computations.

Although market-based approaches do not require central planners to design the matching between sources and targets, they have their own disadvantages [25]. On one hand, it is difficult to study every source and target and further analyze their complex interactions for large-scale problems; on the other hand, it is impractical to derive the optimal resource matching as any matching pair is affected by all the other matching pairs through the interconnections between sources and targets.

In this paper, we build on the discrete optimal transport framework which is a centralized planning-based approach to resource matching, and aim to derive a fully distributed algorithm to find the optimal matching between sources and targets. Optimal transport studies the optimal allocation of resources and provides economic interpretations and analysis of assignment problems and their properties [11], [12]. Besides resource matching, it has also been applied to areas such as machine learning [31], image processing [32], and reflector design [33].

A few distributed methods have been introduced to decentralize resource matching. For example, Gao et al. have
developed a distributed algorithm to match resources to events in sensor networks by exploring the tree metric from the underlying network metric [14]. Buyya has proposed distributed computational economy as an effective metaphor for the management of resources and application scheduling in grid computing [34]. Hasan et al. have developed a distributed approach using stable matching to allocate radio resources for device-to-device communication [15].

We consider a consensus-based approach using alternating direction method of multipliers (ADMM) to decentralize discrete optimal transport problems [17], [35]. ADMM provides a natural methodology to develop distributed algorithms for decentralized applications [39]–[41]. For example, Zhang et al. have developed a consensus-based distributed algorithm with ADMM to study the interactions between attackers and machine learning learners in networked environments [39]; Zennaro et al. have addressed the discrete optimal transport problems [17], [35]. ADMM provides the method of multipliers (ADMM) to decentralize computational economy as an effective metaphor for the underlying network metric [14]; Buyya has proposed distributed learning learners in networked environments [39]; Zennaro et al. have developed a consensus-based distributed algorithm with ADMM [40]; Shen et al. have addressed the discrete optimal transport problems [17], [35]. ADMM provides the method of multipliers (ADMM) to decentralize computational economy as an effective metaphor for the underlying network metric [14]; Buyya has proposed distributed learning learners in networked environments [39]; Zennaro et al. have developed a consensus-based distributed algorithm with ADMM [40]; Shen et al. have addressed the discrete optimal transport problems [17], [35].

Appendices B, C, D, and E provide the proofs for Proposition 3, Theorem 2, Proposition 5, and Proposition 6, respectively.

In this section, we present a discrete optimal transport algorithm with ADMM [40] to solve power-constrained gas transportation [42]; Benamou, et al. have presented augmented Lagrangian schemes for solving discretized dynamical optimal transport [41]; Shen et al. have addressed the discrete optimal transport problems [17], [35]. ADMM provides the method of multipliers (ADMM) to decentralize computational economy as an effective metaphor for the underlying network metric [14]; Buyya has proposed distributed learning learners in networked environments [39]; Zennaro et al. have developed a consensus-based distributed algorithm with ADMM [40]; Shen et al. have addressed the discrete optimal transport problems [17], [35].

Let \(x, y\) and \(X, Y\) represent the utility functions for participants in the network and \(G\) be the underlying graph. In this paper, we focus on situations with feasible \(\Pi_x\), i.e., \(\{x, y\} \in G, x \in X\) and in that case, we cannot find the optimal matching by solving \(\Pi_x\). “Bipartite” indicates that the only connections in the graph are directed” indicates that if a target \(x\) gets and sources by an undirected bipartite graph \(G\), then the source \(y\) must know the target \(x\), i.e., \(x, y \in G\). “Bipartite” indicates that the only connections in the graph are between targets \(X\) and sources \(Y\) to transfer resources.

We can represent the possible matching between targets and sources by an undirected bipartite graph \(G := \{\{x, y\} | y \in Y, x \in X\} = \{\{x, y\} | x \in X, y \in Y\}\). “Undirected” indicates that if a target \(x\) knows a source \(y\), i.e., \(y \in Y\), then the source \(y\) must know the target \(x\), i.e., \(x, y \in G\). “Bipartite” indicates that the only connections in the graph are between targets \(X\) and sources \(Y\) to transfer resources.

The objective of discrete optimal transport is to find the matching \(\Pi_y\) that maximizes the total surplus of targets and sources, which can be achieved by solving the following optimization problem:

\[
\begin{align*}
\max_{\Pi_{xy}} & \sum_{x \in X} \sum_{y \in Y} f_{xy}(\pi_{xy}) + \sum_{y \in Y} \sum_{x \in X} g_{yx}(\pi_{yx}) \\
\text{s.t.} \quad & p_{x,y} \leq \pi_{xy} \leq p_{x,y}, \quad \forall x \in X; \\
& q_{y,x} \leq \pi_{yx} \leq q_{y,x}, \quad \forall y \in Y; \\
& \pi_{xy} \geq 0, \quad \forall \{x, y\} \in G.
\end{align*}
\]

The objective function captures the total surplus given the matching rules \(\Pi_{xy}\). Functions \(f_{xy}: \mathbb{R}_{\geq0} \to \mathbb{R}\) and \(g_{yx}: \mathbb{R}_{\geq0} \to \mathbb{R}\) represent the utility functions for \(x, y\) respectively. The constraints capture the participant’s boundaries on the amount of transferred resources. \(p_{x,y}\) and \(q_{x,y}\) represent the lower and upper bounds for \(x, y\) respectively, and they satisfy \(0 \leq p_{x,y} \leq p_{x,y}\) and \(0 \leq q_{y,x} \leq q_{y,x}\) respectively.

However, there may exist no \(\Pi_{xy}\) that satisfies the constraints, and in that case, we cannot find the optimal matching by solving problem (I). For example, when all sources \(Y\) together cannot provide sufficient resources to meet the minimum requirement of target \(x\), i.e., \(p_{x,y} > \sum_{y \in Y} q_{y,x}\), no \(\Pi_{xy}\) is feasible for \(x\). In this paper, we focus on situations with feasible \(\Pi_{xy}\).

**Condition 1.** (Necessity) The following inequalities must hold to guarantee the feasibility of \(\Pi_{xy}\):

\[(a). \quad p_{x,y} \leq \sum_{y \in Y} q_{y,x}, \quad \forall x \in X;\]
(b). \( \sum_{x \in X_y} p_{x,t} \leq \sum_{y \in Y} q_{y,h} \).

Condition 1(a) indicates that any target can acquire the minimum amount of resources. Condition 1(b) indicates that the demands are lower than or equal to the supplies.

**Condition 2.** (Sufficiency) The feasibility of \( \Pi_g \) is guaranteed if the following inequalities hold:

\[
q_{y,h} \geq \sum_{x \in X_y} p_{x,t}, \quad \forall y \in \mathcal{Y}.
\]

Condition 2 indicates that any supplier can provide sufficient resources to satisfy the requirements of all the targets it knows. Condition 2 is stronger than Condition 1. Note that problems that do not meet Condition 2 but satisfy Condition 1 can be still feasible.

We have the following assumption regarding the convexity of the utility functions \( f_{xy} \) and \( g_{xy} \).

**Assumption 1.** Both the utility functions \( f_{xy} \) and \( g_{xy} \) are concave on \( \pi_{xy} \), \( \forall \{x, y\} \in \mathcal{G} \).

We use the following remarks to provide examples and discuss the interpretations behind \( f_{xy} \) and \( g_{xy} \).

**Remark 1.** The target \( x \)'s utility function \( f_{xy} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R} \) can be expressed as follows:

\[
f_{xy}(\pi_{xy}) = U_{xy}^{(t)}(\pi_{xy}) - C_{xy}^{(t)}(\pi_{xy}),
\]

where \( U_{xy}^{(t)} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0} \) is a concave and increasing function which indicates the unit revenue gained from consuming resources, and \( C_{xy}^{(t)} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0} \) is a convex and increasing function which captures the costs of acquiring resources from \( y \). The superscript \( (t) \) indicates that the function belongs to the party of targets, and it does not indicate exponentiation. We provide several examples of \( U_{xy}^{(t)} \) and \( C_{xy}^{(t)} \) and their interpretations for reference.

- \( U_{xy}^{(t)} := \gamma_{xy} \pi_{xy} \) is a linear revenue function with \( \gamma_{xy} \) denoting the unit revenue gained from consuming resources;
- \( U_{xy}^{(t)} := \gamma_{xy} \min(\pi_{xy}, p_x) \) is a threshold revenue function. It captures the situation when \( x \) has a threshold of consuming resources and \( x \) cannot gain more revenues with more resources after the threshold.
- \( U_{xy}^{(t)} := \gamma_{xy} \log(\pi_{xy} + 1) \) is a concave log revenue function which captures situations when the revenues increase slower with more resources;
- \( C_{xy}^{(t)} := c_{xy}^{(t)} \pi_{xy} \) indicates a linear cost function with \( c_{xy}^{(t)} \) denoting the unit cost of acquiring resources from \( y \).
- \( C_{xy}^{(t)} := c_{xy}^{(t)} \pi_{xy}^2 \) indicates a quadratic cost function which captures situations when the costs increase faster with more resources.

**Remark 2.** The source \( y \)'s utility function \( g_{xy} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R} \) can be expressed as follows:

\[
g_{xy} = U_{xy}^{(s)}(\pi_{xy}) - C_{xy}^{(s)}(\pi_{xy}),
\]

where \( U_{xy}^{(s)} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0} \) is a concave and increasing function which indicates the incomes from providing resources to targets, and \( C_{xy}^{(s)} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0} \) is a convex and increasing function which captures the losses caused by resource reduction. The superscript \( (s) \) indicates that the function belongs to the party of sources, and it does not indicate exponentiation. We provide several examples of \( U_{xy}^{(s)} \) and \( C_{xy}^{(s)} \) and their interpretations for reference.

- \( U_{xy}^{(s)} := \delta_{xy} \pi_{xy} \) indicates a linear income function with \( \delta_{xy} \) denoting the unit income of providing resources to \( x \);
- \( C_{xy}^{(s)} := c_{xy}^{(s)} \pi_{xy} \) indicates a linear loss function and \( c_{xy}^{(s)} \) can be interpreted as the unit loss of resource reduction.
- \( C_{xy}^{(s)} := \pi_{xy}^2 \) indicates a quadratic loss function which captures situations when the losses increase faster with more resources.

We have the following theorem regarding the existence of solution to problem (1).

**Theorem 1.** There exists a solution (not necessarily unique) for problem (1) if it satisfies Condition 2 and has a feasible set.

**Proof.** From Section 4.2 in [45], since both \( f_{xy} \) and \( g_{xy} \) are concave, there exists a solution for problem (1). However, since they may not be strictly concave, the solution may not be unique.

Problem (1) captures the central planner’s objective of finding the optimal matching to transfer the resources between sources and targets. However, solving it requires a significant amount of computations when the number of participants is large. We aim to develop a fully distributed algorithm to find the optimal matching which does not require a central planner. We take a consensus-based approach and decentralize problem (1) with alternating direction method of multipliers (ADMM).

**III. DISTRIBUTED DISCRETE OPTIMAL TRANSPORT USING ADMM**

In the centralized discrete optimal transport problem (1), the central planner designs \( \Pi_g \) for sources and targets, we could also interpret it as that sources and targets reach consensus with the central planner on \( \Pi_g \). Recall that \( \pi_{xy} \) denotes the amount of transferred resources from \( y \) to \( x \). Let \( \pi_{xy}^{(t)} \) and \( \pi_{xy}^{(s)} \) represent the amounts of resources that target \( x \) requests from \( y \) and source \( y \) offers to \( x \), respectively, the consensus between the participants and the central planner can be captured with \( \pi_{xy}^{(t)} = \pi_{xy} \) and \( \pi_{xy}^{(s)} = \pi_{xy} \). We further define \( \Pi_g^{(t)} := \{\pi_{xy}^{(t)} | \{x, y\} \in \mathcal{G}\} \), \( \Pi_g^{(s)} := \{\pi_{xy}^{(s)} | \{x, y\} \in \mathcal{G}\} \), \( \Pi_T^{(t)} := \{\pi_{xy}^{(t)} | y \in \mathcal{Y}, x = x, \forall \{x, y\} \in \mathcal{G}\} \), and \( \Pi_T^{(s)} := \{\pi_{xy}^{(s)} | x \in \mathcal{X}, y = y, \forall \{x, y\} \in \mathcal{G}\} \). The superscripts \((t)\) and \((s)\) indicate that the variables or parameters belong to targets and sources, respectively, and they do not indicate exponentiation.

As a result, we can rewrite problem (1) into the form of ADMM as the following problem:

\[
\begin{array}{ll}
\min_{\{\Pi_g^{(t)} \in \mathcal{G}_T^{(t)}, \Pi_g^{(s)} \in \mathcal{G}_T^{(s)}\}} & - \sum_{t \in \mathcal{T}} \sum_{x \in \mathcal{X}_t} f_{xy}(\pi_{xy}^{(t)}) - \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}_t} g_{xy}(\pi_{xy}^{(s)}) \\
\text{s.t.} & \pi_{xy}^{(t)} = \pi_{xy}, \quad \forall \{x, y\} \in \mathcal{G}, \\
& \pi_{xy}^{(s)} = \pi_{xy}, \quad \forall \{x, y\} \in \mathcal{G},
\end{array}
\]

(4)
where
\[ \mathcal{U}_1 := \left\{ \Pi_{xy}^{(l)} \mid \Pi_{xy}^{(l)} \geq 0, p_{x,l} \leq \sum_{y \in \mathcal{Y}_x} \Pi_{xy}^{(l)} \leq p_{x,h}, \{x,y\} \in \mathcal{G} \right\}, \]
\[ \mathcal{U}_2 := \left\{ \Pi_{xy}^{(l)} \mid \Pi_{xy}^{(l)} \geq 0, q_{y,l} \leq \sum_{x \in \mathcal{X}_y} \Pi_{xy}^{(l)} \leq q_{y,h}, \{x,y\} \in \mathcal{G} \right\}. \]

Problem (4) is a minimization problem with the objective function from problem (1). Since both \( f_{xy} \) and \( g_{xy} \) are concave, the objective function of problem (4) is convex. The constraints in problem (1) have been included in the feasible sets \( \mathcal{U}_1 \) and \( \mathcal{U}_2 \). The consensus constraints in problem (4) capture the consensus between the participants and the central planner. Furthermore, we can achieve \( \Pi_{xy}^{(l)} = \pi_{xy} \), which indicates the direct consensus between sources and targets.

Let \( \alpha_{xy1} \) and \( \alpha_{xy2} \) denote the Lagrange multipliers corresponding to the constraints \( \Pi_{xy}^{(l)} = \pi_{xy} \) and \( \Pi_{xy}^{(l)} = \pi_{xy} \), respectively, and then the augmented Lagrangian related to problem (4) is
\[
\mathcal{L}(\Pi_{xy}^{(l)}, \Pi_{xy}, \{\alpha_{xy1}\}_{xy}, \{\alpha_{xy2}\}_{xy}) = -\sum_{x \in \mathcal{X}_y} \sum_{y \in \mathcal{Y}_x} f_{xy}(\Pi_{xy}^{(l)}) - \sum_{y \in \mathcal{Y}_x} \sum_{x \in \mathcal{X}_y} g_{xy}(\Pi_{xy}^{(l)})
+ \sum_{x \in \mathcal{X}_y} \sum_{y \in \mathcal{Y}_x} \alpha_{xy1}(\Pi_{xy}^{(l)} - \pi_{xy}) + \sum_{y \in \mathcal{Y}_x} \sum_{x \in \mathcal{X}_y} \alpha_{xy2}(\Pi_{xy}^{(l)} - \pi_{xy})
+ \frac{\eta}{2} \sum_{x \in \mathcal{X}_y} \sum_{y \in \mathcal{Y}_x} (\Pi_{xy}^{(l)} - \pi_{xy})^2 + \frac{\eta}{2} \sum_{y \in \mathcal{Y}_x} \sum_{x \in \mathcal{X}_y} (\Pi_{xy}^{(l)} - \pi_{xy})^2.
\]

Note that \( \mathcal{L} \) is strictly convex with the quadratic terms \[ \frac{\eta}{2} \sum_{x \in \mathcal{X}_y} \sum_{y \in \mathcal{Y}_x} (\Pi_{xy}^{(l)} - \pi_{xy})^2 \] and \[ \frac{\eta}{2} \sum_{y \in \mathcal{Y}_x} \sum_{x \in \mathcal{X}_y} (\Pi_{xy}^{(l)} - \pi_{xy})^2, \]
and thus guarantees the uniqueness of the optimal solution even when \( f_{xy} \) and \( g_{xy} \) are not strictly concave. The scalar \( \eta \) controls the speed of the convergence. We can achieve the following proposition after applying the iterations of ADMM to problem (4).

**Proposition 1.** We can obtain the following iterations after applying the iterations of ADMM to Problem (4):

\[
\Pi_{xy}^{(l+1)}(k+1) \in \arg \min_{\Pi_{xy}^{(l)} \in \mathcal{U}_1} \sum_{y \in \mathcal{Y}_x} f_{xy}(\Pi_{xy}^{(l)}) + \sum_{y \in \mathcal{Y}_x} \alpha_{xy1}(\Pi_{xy}^{(l)} - \pi_{xy}) + \frac{\eta}{2} \sum_{y \in \mathcal{Y}_x} (\Pi_{xy}^{(l)} - \pi_{xy})^2,
\]
\[
\Pi_{xy}^{(l+1)}(k+1) \in \arg \min_{\Pi_{xy}^{(l)} \in \mathcal{U}_2} \sum_{x \in \mathcal{X}_y} g_{xy}(\Pi_{xy}^{(l)}) - \sum_{x \in \mathcal{X}_y} \alpha_{xy2}(\Pi_{xy}^{(l)} - \pi_{xy}) + \frac{\eta}{2} \sum_{x \in \mathcal{X}_y} (\Pi_{xy}^{(l)} - \pi_{xy})^2.
\]

Algorithm 1 summarizes (11)-(14). Iterations (11) and (12) are updates of target \( x \) and source \( y \), respectively. Iterations (13) and (14) are updates of pair \( xy \). Iterations (11)-(14) are fully distributed iterations to solve problem (4). A detailed analysis of Proposition 2 and Algorithm 1 is provided in the following subsections.

**A. Convergence and Complexity**

Since \( f_{xy} \) and \( g_{xy} \) are concave functions, we have that \( -f_{xy}(\Pi_{xy}^{(l)}) - g_{xy}(\Pi_{xy}^{(l)}) \) is a convex function, and thus iterations (6)-(10) converge to the optimum of problem (4) from Lemma 1 in Appendix A.

\[ \text{Algorithm 1} \]
\[ \begin{array}{ll}
1: & \text{for } k = 0, 1, 2, \ldots \text{ do} \\
2: & \text{Each } x \in \mathcal{X} \text{ computes } \Pi_{xy}^{(l)}(k+1) \text{ via (11)} \\
3: & \text{Each } y \in \mathcal{Y} \text{ computes } \Pi_{xy}^{(l)}(k+1) \text{ via (12)} \\
4: & \text{Each pair } (x,y) \in \mathcal{G} \text{ computes } \pi_{xy}(k+1) \text{ via (13)} \\
5: & \text{End pairs } (x,y) \in \mathcal{G} \text{ computes } \alpha_{xy}(k+1) \text{ via (14)} \\
6: & \text{End for}
\end{array} \]

Iterations (6)-(10) can be simplified further as shown in the following Proposition 2.

**Proposition 2.** Iterations (6)-(10) can be simplified further into the following iterations:

\[
\Pi_{xy}^{(l)}(k+1) \in \arg \min_{\Pi_{xy}^{(l)} \in \mathcal{U}_1} \sum_{y \in \mathcal{Y}_x} f_{xy}(\Pi_{xy}^{(l)})
+ \sum_{y \in \mathcal{Y}_x} \alpha_{xy}(k)(\Pi_{xy}^{(l)} - \pi_{xy}) + \frac{\eta}{2} \sum_{y \in \mathcal{Y}_x} (\Pi_{xy}^{(l)} - \pi_{xy})^2,
\]
\[
\Pi_{xy}^{(l)}(k+1) \in \arg \min_{\Pi_{xy}^{(l)} \in \mathcal{U}_2} \sum_{x \in \mathcal{X}_y} g_{xy}(\Pi_{xy}^{(l)})
- \sum_{x \in \mathcal{X}_y} \alpha_{xy}(k)(\Pi_{xy}^{(l)} - \pi_{xy}) + \frac{\eta}{2} \sum_{x \in \mathcal{X}_y} (\Pi_{xy}^{(l)} - \pi_{xy})^2,
\]
\[
\pi_{xy}(k+1) = \left( \frac{1}{2} (\Pi_{xy}^{(l)}(k) + \alpha_{xy}(k)) \right) + \frac{\eta}{2} (\pi_{xy}(k) - \pi_{xy}^{(l)}(k+1))^2.
\]

where \( \mathcal{U}_1 := \{\Pi_{xy}^{(l)} \mid \Pi_{xy}^{(l)} \geq 0, y \in \mathcal{Y}_x \} \) and \( \mathcal{U}_2 := \{\Pi_{xy}^{(l)} \mid \Pi_{xy}^{(l)} \geq 0, x \in \mathcal{X}_y \} \).

Proof. See Appendix B.
Algorithm 1 enables parallel computations and is more efficient with (11) or (12), and each pair of target and source updates are obtained through matrix computations, which have the complexity of \( O(n^3) \), where \( n \) is the size of \( \Pi_y \). \( \alpha_{xy} \) and \( \pi_{xy} \) are obtained through matrix computations, which have the complexity of \( O(1) \). The complexity of Algorithm 1 is further affected by the number of participants and connections between targets and sources.

**B. Efficiency and Privacy**

In Algorithm 1, each participant solves its own problem with (11) or (12), and each pair of target and source updates itself with (13) and (14). Thus, it is a fully distributed algorithm which does not require central planners to collect the information of participants and compute the optimal matching. Algorithm 1 enables parallel computations and is more efficient for large-scale problems. Moreover, the only information that transmitted between participants are \( \pi_{xy} \), \( \pi_{yx} \), \( \pi_{xy} \), and \( \alpha_{xy} \), which reduces communication overhead and keeps privacy at the same time. Each participant’s personal information \( \{p_{x,t}, p_{x,h}, f_{xy}, \mathcal{F}_x\} \) or \( \{q_{y,t}, q_{y,h}, g_{xy}, \mathcal{F}_y\} \) is also kept private to itself. As a result, our algorithm guarantees certain levels of efficiency and privacy compared to centralized algorithms.

**C. Resource Market: Negotiations and Bargaining**

In problem (4), the consensus constraints capture the consensus between participants and central planners. Furthermore, they can also be viewed as the direct consensus between sources and targets through central planners. One interpretation of the consensus constraints is that sources and targets must agree with each other to form a matching in resource markets. Algorithm 1 further reflects the negotiations and bargaining between sources and targets to reach agreements on the amount of transferred resources.

\( \pi_{xy} \) and \( \pi_{yx} \) denote the amounts of resources that target \( x \) requests from source \( y \) and source \( y \) offers to target \( x \), respectively. They are proposed by \( x \) and \( y \) to meet their own objectives (11) and (12), respectively. \( \pi_{xy} \) captures the consensus between \( x \) and \( y \) that they must reach an agreement on the amount of transferred resources, i.e., \( \pi_{xy} = \pi_{yx} = \pi_{xy} \) from problem (4). We can see from (13) that the agreement is approached by continuous negotiations following an averaging principle, i.e., \( \pi_{xy} \) is taken as the average of \( \pi_{xy} \) and \( \pi_{yx} \). Moreover, we can see from (11) and (12) that \( x \) and \( y \) modify their proposals \( \pi_{xy} \) and \( \pi_{yx} \) to not only maximize their own utilities but also minimize the differences from their previous negotiations \( \pi_{xy} \), respectively.

\( \alpha_{xy} \) is the Lagrange multipliers on the consensus constraints. We can see from (14) that \( \alpha_{xy} \) captures cumulative differences between \( x \) and \( y \). A positive \( \alpha_{xy} \) indicates that \( x \) requests more than \( y \) offers, thus, \( x \) will decrease \( \pi_{xy} \) with term \( \alpha_{xy}(k)\pi_{yx} \) in (11), while \( y \) will increase \( \pi_{yx} \) with term \( -\alpha_{xy}(k)\pi_{xy} \) in (12), and vice versa. \( \alpha_{xy} \) identifies the directions for \( x \) and \( y \) to modify their proposals \( \pi_{xy} \) and \( \pi_{yx} \). It reveals the tendencies of sources and targets to reach agreements on the amounts of transferred resources between them.

\( \eta \) controls each participant’s trade-off between its high utility and a small total difference from others. We can see from the last terms of the objective functions in (11) and (12) that a larger \( \eta \) indicates that the participant cares more about reaching agreements with the participants on the other side, while a smaller \( \eta \) indicates that the participant pays more attention on maximizing its utility. Moreover, we can see from (14) that a large \( \eta \) amplifies the degree of disagreement \( \alpha_{xy} \), which will increase the costs of disagreement for the participant through the second terms of the objective functions in (11) and (12).

As a result, our algorithm reflects the negotiation and bargaining processes between targets and sources in market environments. Each participant aims to maximize its own utility and minimize the differences from others. The matching is reached by the bargaining between sources and targets on the amounts of transferred resources following an averaging principle. The averaging principle contains three aspects: 1) each pair of source and target always settles at their average amount of transferred resources after each negotiation; 2) the source and the target then modify their proposals by reducing their differences from the average amount of resources from their last round of negotiation; 3) the negotiations repeat until agreements are reached for all sources and targets. The averaging principle provides a way to regulate resource markets and improve resource efficiency as the convergence of our algorithm to the optimal matching is guaranteed.

**D. Online Distributed Resource Matching**

In many applications, targets and sources may change over time. For example, a source may increase its upper bound \( q_{y,h} \) when it has more resources to distribute; a target may achieve more utilities with the same amount of resources after improving the efficiency of resource consumption; new participants may join the matching and old participants may quit the matching. Our algorithm can be extended to its online form to address the real-time requirements for these applications.
Remark 3. Let $\mathcal{X}^{(k)}$, $\mathcal{Y}^{(k)}$, $f_{xy}^{(k)}$, $g_{xy}^{(k)}$, $\mathcal{P}_{xy}^{(k)}$, and $\mathcal{Q}_{xy}^{(k)}$ denote the discrete optimal transport problem at time $k$, the iterations of finding the optimal matching are:

$$\Pi_y^{(k)}(k + 1) \in \arg \min_{\Pi_y^{(k)}(k + 1)} \mathcal{P}_{xy}^{(k)}(\pi_{xy}^{(k)}) - \sum_{y \in \mathcal{Y}^{(k+1)}} \sum_{x \in \mathcal{X}^{(k+1)}} \mathcal{Q}_{xy}^{(k)}(\pi_{xy}^{(k)}) + \sum_{x \in \mathcal{X}^{(k+1)}} \alpha_{xy}(k) \pi_{xy}^{(k)} + \frac{\mu}{2} \sum_{y \in \mathcal{Y}^{(k+1)}} (\pi_{xy}^{(k)} - \pi_{xy}(k))^2,$$

where $\mathcal{P}_{xy}^{(k)}(\pi_{xy}^{(k)}) := \left\{ \pi_{xy}^{(k)} \mid \pi_{xy}^{(k)} \geq 0, \sum_{y \in \mathcal{Y}^{(k+1)}} \pi_{xy}^{(k)} \leq p_x(k), \sum_{x \in \mathcal{X}^{(k+1)}} \pi_{xy}^{(k)} \leq q_y(k) \right\}$.

Condition 3. There exists a solution of problem (27), which is not necessarily unique, if and only if the following equality holds:

$$\sum_{x \in \mathcal{X}} p_x = \sum_{y \in \mathcal{Y}} q_y. \quad (21)$$

We can easily verify that Condition 3 satisfies Condition 1.

The optimal value, i.e., maximum total surplus, is unique, which is also referred as the optimal transport distance:

$$D_1 := \sup_{\{\pi_{xy} \geq 0\} \in \mathcal{X}, \mathcal{Y}} \left\{ \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} \pi_{xy} (g_{xy} + \delta_x y) - \pi_{xy} = p_x, \pi_{xy} = q_y \right\}. \quad (22)$$

After introducing $\pi_{xy}^{(k)}$ and $\pi_{xy}^{(s)}$ with a similar approach in the previous section, we can rewrite problem (20) as follows:

$$\min_{\{\pi_{xy}^{(k)} \in \mathcal{X}, \mathcal{Y}, \pi_{xy}^{(s)} \in \mathcal{X}, \mathcal{Y}\}} - \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} \pi_{xy}^{(k)} g_{xy} \leq \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} \pi_{xy}^{(s)} \delta_y,$$

s.t.

$$\pi_{xy}^{(k)} = \pi_{xy}^{(s)}, \forall x \in \mathcal{X}, y \in \mathcal{Y}.$$

where $\mathcal{U}_1 := \left\{ \Pi_y^{(k)} | \pi_{xy}^{(k)} \geq 0, \sum_{y \in \mathcal{Y}} \pi_{xy}^{(k)} = p_x, \{x, y\} \in \mathcal{G} \right\}$ and $\mathcal{U}_2 := \left\{ \Pi_y^{(s)} | \pi_{xy}^{(s)} \geq 0, \sum_{x \in \mathcal{X}} \pi_{xy}^{(s)} = q_y, \{x, y\} \in \mathcal{G} \right\}$. By Proposition 3 we have:

Proposition 3. We can obtain the following iterations after applying the iterations of ADMM to Problem (27):

$$\Pi_y^{(k)}(k + 1) \in \arg \min_{\Pi_y^{(k)} \in \mathcal{Y}} \mathcal{P}_{xy}^{(k)}(\pi_{xy}^{(k)}) + \frac{\mu}{2} \sum_{y \in \mathcal{Y}} (\pi_{xy}^{(k)} - \pi_{xy}(k))^2,$$

where $\mathcal{U}_1 := \left\{ \Pi_y^{(k)} | \pi_{xy}^{(k)} \geq 0, \sum_{y \in \mathcal{Y}} \pi_{xy}^{(k)} = p_x, \{x, y\} \in \mathcal{G} \right\}$ and $\mathcal{U}_2 := \left\{ \Pi_y^{(s)} | \pi_{xy}^{(s)} \geq 0, \sum_{x \in \mathcal{X}} \pi_{xy}^{(s)} = q_y, \{x, y\} \in \mathcal{G} \right\}$. By Proposition 3 we have:

IV. Duality and Distributed Dual Discrete Optimal Transport with Linear Utilities

In this section, we consider a special case of discrete optimal transport problems with linear objective functions and equality constraints, i.e., $f_{xy}(\pi_{xy}) = g_{xy}(\pi_{xy}) = \delta_y \pi_{xy}$, $p_x, p_y \in \mathcal{Y}$, $q_y, q_x \in \mathcal{X}$. Parameters $\gamma_{xy} \in \mathbb{R}^{\geq 0}$ and $\delta_{xy} \in \mathbb{R}^{\geq 0}$ are the unit utilities for $x$ and $y$, respectively. Inequalities $p_x \leq \sum_{y \in \mathcal{Y}} \pi_{xy} \leq p_x + \mu$, $q_y \leq \sum_{x \in \mathcal{X}} \pi_{xy} \leq q_y + \mu$ become equalities $\sum_{y \in \mathcal{Y}} \pi_{xy} = p_x$ and $\sum_{x \in \mathcal{X}} \pi_{xy} = q_y$, respectively. Moreover, we assume that all sources match with all targets, i.e., $\mathcal{Y} = \mathcal{Y}$ and $\mathcal{X} = \mathcal{X}$.

The discrete optimal transport in this case aims to find the optimal matching by solving the following problem:

$$\max_{\{\pi_{xy} \mid x \in \mathcal{X}, y \in \mathcal{Y}\}} \mathcal{P}_{xy}^{(k)}(\pi_{xy}) + \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \gamma_{xy} \pi_{xy} + \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \delta_{xy} \pi_{xy}, \quad \text{s.t.} \quad \sum_{x \in \mathcal{X}} \pi_{xy} = p_x, \sum_{y \in \mathcal{Y}} \pi_{xy} = q_y, \quad \forall x \in \mathcal{X}, y \in \mathcal{Y}. \quad (20)$$

Problem (20) is a well-studied problem in optimal transport [11, 13]. We have the following condition regarding the existence of its solution which captures a situation when supply meets demand.
Proposition 3 captures the price that target $x$ pays to source $y$ as the compensation of transferring resources.

A. Dual Problem of Discrete Optimal Transport

We derive the dual problem of problem (23), which can be used to develop the distributed dual algorithm of discrete optimal transport and study the pricing scheme of resource matching.

Theorem 2. Let $u_x \in \mathbb{R}$, $v_y \in \mathbb{R}$, $w_{xy}^{(i)} \in \mathbb{R}$, and $w_{xy}^{(s)} \in \mathbb{R}$ denote the Lagrange multipliers with respect to $\sum_{y \in \mathcal{Y}} \pi_{xy} = p_x$, $\sum_{x \in \mathcal{X}} \pi_{xy} = q_y$, $\pi_{xy} = \pi_{xy}^{(i)}$, and $\pi_{xy} = \pi_{xy}^{(s)}$, respectively. The dual problem of problem (23) is

$$\min_{\{u_x, v_y, w_{xy}^{(i)}, w_{xy}^{(s)}\}} \sum_{x \in \mathcal{X}} u_x p_x + \sum_{y \in \mathcal{Y}} v_y q_y$$

s.t. $\gamma_{xy} - u_x - w_{xy}^{(i)} \leq 0$, $\delta_{xy} - v_y + w_{xy}^{(s)} \leq 0$, $w_{xy}^{(i)} = w_{xy}^{(s)}$, and $\forall \{x, y\} \in \mathcal{G}$. (28)

Proof. See Appendix C.

We summarize several interesting results of problems (23) and (28) with the following proposition.

Proposition 4. The following facts hold for the primal problem (23) and the dual problem (28):

(i) Problem (23) is the dual problem of problem (28).

(ii) (Monge-Kantorovich duality) The optimal value of problem (23) is equal to the optimal value of problem (28). The optimal transport distance (22) can also be written as

$$D_2 := \inf_{\{u_x, v_y, w_{xy}^{(i)}, w_{xy}^{(s)}\}} \left\{ \sum_{x \in \mathcal{X}} u_x p_x + \sum_{y \in \mathcal{Y}} v_y q_y \right\}$$

s.t. $\gamma_{xy} - u_x - w_{xy}^{(i)} \leq 0$, $\delta_{xy} - v_y + w_{xy}^{(s)} \leq 0$, $w_{xy}^{(i)} = w_{xy}^{(s)}$, and $\forall \{x, y\} \in \mathcal{G}$. (29)

(iii) The solution $\{\Pi_{xy}^{(i)}, \Pi_{xy}^{(s)}\}$ of problem (23) may not be unique but $\pi_{xy} = \pi_{xy}^{(i)} = \pi_{xy}^{(s)}$ always holds; the solution $\{u_x^{*}, v_y^{*}, w_{xy}^{(i)*}, w_{xy}^{(s)*}\}$ of problem (28) is not unique but $w_{xy}^{(i)*} = w_{xy}^{(s)*}$ always holds.

Let $C \in \mathbb{R}$ denote a constant, $u_x^{*} + C, v_y^{*} - C, w_{xy}^{(i)*} - C, w_{xy}^{(s)*} - C$ is also a solution of problem (28).

(iv) if $\pi_{xy}^{(i)>0}$ or $\pi_{xy}^{(s)>0}$, we have $u_x^{*} = \gamma_{xy} - w_{xy}^{(i)*}$, $v_y^{*} = \delta_{xy} + w_{xy}^{(s)*}$, and $u_x^{*} + v_y^{*} = \gamma_{xy} + \delta_{xy}$.

(v) $u_x^{*}, v_y^{*}, w_{xy}^{(i)*}$, and $w_{xy}^{(s)*}$ satisfy:

$$u_x^{*} = \max_{y \in \mathcal{Y}} \left\{ \gamma_{xy} - w_{xy}^{(i)*} \right\}, \quad v_y^{*} = \max_{x \in \mathcal{X}} \left\{ \delta_{xy} + w_{xy}^{(s)*} \right\}.$$ (30)

Proof. We can verify (i) by deriving the dual problem of problem (28). (ii) can be achieved by the minimax theorem [46]. To prove (iii), we notice that $\gamma_{xy} - (u_x + C) - (w_{xy} - C) = \gamma_{xy} - u_x - w_{xy}^{(i)*}$, $\delta_{xy} - (v_y - C) - (w_{xy}^{(s)} - C) = \delta_{xy} - v_y + w_{xy}^{(s)*}$, $w_{xy}^{(i)} - C = w_{xy}^{(s)} - C$, and $\sum_{x \in \mathcal{X}} (u_x + C)p_x + \sum_{y \in \mathcal{Y}} (v_y - C)q_y = \sum_{x \in \mathcal{X}} u_x p_x + \sum_{y \in \mathcal{Y}} v_y q_y + \sum_{x \in \mathcal{X}} C p_x + \sum_{y \in \mathcal{Y}} C q_y = \sum_{x \in \mathcal{X}} u_x p_x + \sum_{y \in \mathcal{Y}} v_y q_y$ as $\forall \{x, y\} \in \mathcal{G}$.

Thus, (iii) holds. We can obtain (iv) by complementary slackness [47]. To prove (v), we first notice that the constraints in problem (28) imply that $u_x^{*} \geq \max_{y} \{ \gamma_{xy} - w_{xy}^{(i)*} \}$ and $v_y^{*} \geq \max_{x} \{ \delta_{xy} + w_{xy}^{(s)*} \}$. However, if the inequalities hold strict, $u_x^{*}$ and $v_y^{*}$ will not be optimal as one can achieve a lower value of $\sum_{x \in \mathcal{X}} u_x p_x + \sum_{y \in \mathcal{Y}} v_y q_y$ by selecting $u_x^{*} = \max_{y} \{ \gamma_{xy} - w_{xy}^{(i)*} \}$ and $v_y^{*} = \max_{x} \{ \delta_{xy} + w_{xy}^{(s)*} \}$, and thus the inequalities are actually equalities, which proves (v).

The dual problem and Monge-Kantorovich duality has largely enriched the interpretation of optimal transport [11]. The dual variables $u_x$ and $v_y$ can be interpreted as the individual surpluses that $x$ and $y$ achieve in their matching with others, respectively; the dual consensus variables $w_{xy}^{(i)}$ and $w_{xy}^{(s)}$ can be interpreted as the proposed prices that $x$ pays and $y$ charges for transferring resources, respectively. Proposition 4(iii) indicates that a matching between $x$ and $y$ exists only if their individual surpluses $u_x + v_y$ are equal to their pair surpluses $\gamma_{xy} + \delta_{xy}$, which is also reflected in (28) on the minimization over the overall individual surpluses under constraints $u_x \geq \gamma_{xy} - w_{xy}^{(i)}$, $v_y \geq \delta_{xy} + w_{xy}^{(s)}$, and $w_{xy}^{(i)} = w_{xy}^{(s)}$, i.e., $u_x + v_y \geq \gamma_{xy} + \delta_{xy}$. The prices $w_{xy}^{(i)}$ and $w_{xy}^{(s)}$ only serve as the compensation for transferring resources and they do not change the total surplus $u_x^{*} + v_y^{*} = \gamma_{xy} + \delta_{xy}$ of $x$ and $y$. We can see from Proposition 4(v) that the objective of each participant is to maximize its individual surplus by taking into account both its pair surpluses and the prices it pays or charges.

The objective functions of problems (23) and (28) capture the total surplus by summing over pair surpluses and individual surpluses, respectively. On the one hand, we can see from problem (23) that the primal problem of discrete optimal transport aims to find the optimal matching rules that maximize the total surplus of all pairs between $\mathcal{X}$ and $\mathcal{Y}$; on the other hand, we can see from Proposition 4(v) that the dual problem of discrete optimal transport aims to find an equilibrium where all participants reach their highest individual surpluses. The Monge-Kantorovich duality in Proposition 4(ii) indicates that both interpretations return the same optimal values and reach the maximum total surplus of all sources and targets.

B. Distributed Dual Algorithm

We have obtained a fully distributed primal algorithm to solve problem (23) with ADMM as shown in Proposition 3 and we can also obtain a fully distributed dual algorithm to solve the dual problem (28) using similar methods. We first introduce consensus variables $w_{xy}$ to decompose $w_{xy}^{(i)} = w_{xy}^{(s)}$ as $w_{xy} = w_{xy}^{(i)}$ and $w_{xy} = w_{xy}^{(s)}$. Thus, problem (28) can be rewritten as

$$\min_{\{u_x, v_y, \{w_{xy}^{(i)}, w_{xy}^{(s)}\}\}} \sum_{x \in \mathcal{X}} u_x p_x + \sum_{y \in \mathcal{Y}} v_y q_y$$

s.t. $w_{xy}^{(i)} = w_{xy}^{(s)}$, $\forall \{x, y\} \in \mathcal{G}$. (31)

where $\mathcal{G}' := \{u_x, w_{xy}^{(i)} | u_x - w_{xy}^{(i)} \leq 0, \{x, y\} \in \mathcal{G}\}$, and

$\mathcal{G}_2 := \{v_y, w_{xy}^{(s)} | \delta_{xy} - v_y + w_{xy}^{(s)} \leq 0, \{x, y\} \in \mathcal{G}\}$.
and (33) are individual operations which are quadratic with (34) and (35), respectively. The iterations continue until convergence. We can obtain the following iterations after from Lemma 1 in Appendix A as the objective function section. The convergence of the dual algorithm is guaranteed applications following similar methods used in the previous application. We can extend the dual algorithm to an online version to address real-time and the dual algorithm, respectively. We can also extend on the amounts and prices of resources in the primal algorithm and the dual algorithm, respectively. The iterations in Proposition 3 and the iterations Remark 4. The primal algorithm and the dual algorithm share a lot similarities in terms of their structures and both of them indicates an averaging principle on the bargaining between targets and sources. We further prove the primal-dual relations between their iterations. We shall see that \( \alpha_{xy} = w_{xy}, \pi_y = \beta_{xy}, \) and \( \eta = 1/\bar{\eta} \) where \( \alpha_{xy}, \pi_y, \) and \( \eta \) come from the primal algorithm and \( w_{xy}, \beta_{xy}, \) and \( \bar{\eta} \) come from the dual algorithm.

**C. Relations Between The Primal Algorithm And The Dual Algorithm**

In this subsection, we prove the primal-dual relations between the iterations of the primal algorithm in Proposition 5 and the iterations of the dual algorithm in Proposition 5.

**Proposition 6.** Let \( \lambda_{xy}^{(i)} \in \mathbb{R} \) and \( \lambda_{xy}^{(s)} \in \mathbb{R} \) denote the Lagrange multipliers for (32) and (33), respectively. Iterations (32)-(35) can be rewritten as

\[
\Lambda_{x}^{(i)(k+1)} = \arg \min_{\Lambda_{x}^{(i)} \in \mathcal{Y}} - \sum_{y \in \mathcal{Y}} w_{xy}^{(i)}(k)^2 - \sum_{y \in \mathcal{Y}} (\lambda_{xy}^{(i)} - \beta_{xy}(k))^2, \\
\Lambda_{y}^{(s)(k+1)} = \arg \min_{\Lambda_{y}^{(s)} \in \mathcal{Y}} - \sum_{x \in \mathcal{X}} w_{xy}^{(s)}(k)^2 - \sum_{x \in \mathcal{X}} (\lambda_{xy}^{(s)} - \beta_{xy}(k))^2, \\
\Lambda_{xy}^{(i)(k+1)} = \arg \min_{\Lambda_{xy}^{(i)} \in \mathcal{Y}} - \sum_{x \in \mathcal{X}} w_{xy}^{(i)}(k)^2 - \sum_{x \in \mathcal{X}} (\lambda_{xy}^{(i)} - \beta_{xy}(k))^2, \\
\Lambda_{xy}^{(s)(k+1)} = \arg \min_{\Lambda_{xy}^{(s)} \in \mathcal{Y}} - \sum_{x \in \mathcal{X}} w_{xy}^{(s)}(k)^2 - \sum_{x \in \mathcal{X}} (\lambda_{xy}^{(s)} - \beta_{xy}(k))^2.
\]

where \( \mathcal{Y} := \{ p_{x}, \beta_{xy}^{(i)} | \gamma_{x} - u_{x} - w_{xy}^{(i)} \leq 0, y \in \mathcal{Y} \}, \mathcal{Y} := \{ p_{x}, \beta_{xy}^{(s)} | \gamma_{x} - v_{y} - w_{xy}^{(s)} \leq 0, x \in \mathcal{X} \}. \]

**Proof.** See Appendix E. \( \square \)

**Remark 4.** The iterations in Proposition 3 and the iterations in Proposition 4 are identical with \( \eta = 1/\bar{\eta}, \pi_y = \lambda_{xy}^{(i)}, \lambda_{xy}^{(s)} = \pi_y, \pi_y = \lambda_{xy}^{(i)}, \alpha_{xy} = w_{xy} \).

Thus, \( \alpha_{xy} \) in the primal algorithm captures the prices that target \( x \) pays to source \( y \) as the compensation of transferring resources, i.e., \( w_{xy} \) in the dual problem (31), \( \beta_{xy} \) in the dual algorithm captures the amount of transferred resources from.
y to x, i.e., \( \pi_{xy} \) in the primal problem (23). Both the primal algorithm and the dual algorithm solve problems (23) and (31) at the same time with updates on allocations \( \pi_{xy} \) and prices \( w_{xy} \), and the convergence to the optimum is guaranteed.

V. NUMERICAL EXPERIMENTS

In this section, we present numerical experiments of distributed discrete optimal transport. We first verify the convergence of Algorithm 1 to the solution of problem (1), and we also demonstrate the online adaptability of the online algorithm presented in Remark 3. The experiment settings are described in Fig. 4. Blue rectangles represent targets while green rectangles represent sources. We let all \( p_{x,y} = 100 \) which indicates that targets have high upper bounds on acquiring resources, and all \( d_{x,y} = 0 \) which indicates that sources can choose not to allocate any resources. We consider both linear and quadratic utility functions with \( \gamma_{xy} \) and \( \delta_{xy} \) denoting the parameters, i.e, 
\[
\begin{aligned}
f_{xy}(\pi_{xy}) &= C_{xy}^{(l)} + \gamma_{xy} \pi_{xy} + \delta_{xy} \pi_{xy}^2, \\
g_{xy}(\pi_{xy}) &= C_{xy}^{(q)} + \delta_{xy} \pi_{xy}.
\end{aligned}
\]
and the centralized optimal values \( \Phi_{xy} \) and the corresponding value \( \Phi' \) are achieved by directly solving problem (1). The residuals measure the differences between the current updates and the centralized solution.

We then show that the primal algorithm in Proposition 3 and the dual algorithm in Proposition 5 can be used to solve discrete optimal transport problems (20) with linear utility functions, and the convergence to the solutions is guaranteed. Let \( p_{x,y} \) and \( q_{x} \) be generated by i.i.d. uniform distributions on \( (0,1) \). We then normalize \( \{p_{x,y}\} \) and \( \{q_{x}\} \) to 1 so that \( \sum_{y \in Y} p_{x,y} = 1 = \sum_{x \in X} q_{x} \) as required in Condition 5 which guarantees the existence of a solution. \( \gamma_{xy} \) and \( \delta_{xy} \) are also generated by i.i.d. uniform distributions on \( (0,1) \). We consider three discrete optimal transport problems ‘OT1’, ‘OT2’, and ‘OT3’ with different numbers of sources and targets under this setting. ‘OT1’ has 20 targets and 20 sources; ‘OT2’ has 20 targets and 100 sources; ‘OT3’ has 1000 targets and 2 sources. The

![Fig. 4: Online Discrete Optimal Transport Problems.](image)

![Fig. 5: Evolution of Iterations (16)-(19) for the online resource matching problem shown in Fig. 4.]
Optimal Transport Distance Residual

Fig. 6. Evolution of the primal algorithm and the dual algorithm. The central solution is achieved by directly solving problem \( 20 \). The residuals measure the differences between the current updates and the centralized solution.

results are shown in Figure 6. We can see that both the primal algorithm and the dual algorithm converge to the centralized solution \( \Pi^* \) and the centralized optimal value \( D^* \) of problem \( 20 \).

VI. CONCLUSION

In this paper, we have studied a decentralized discrete optimal transport problem using a consensus-based approach and have derived distributed algorithms with ADMM to address resource matching between multiple sources and multiple targets. We have shown that our algorithms are fully distributed and do not require central planners to compute the optimal matching or store personal information of sources and targets, which provides guaranteed levels of efficiency and privacy at the same time. We have developed distributed primal and dual algorithms by leveraging the primal and dual problems of discrete optimal transport with linear utility functions and further proved the equivalence between them. We have demonstrated with numerical experiments the convergence of our algorithms to centralized solutions, the online adaptability of our algorithms for real-time applications, and the equivalence between the primal algorithm and the dual algorithm for cases with linear utility functions. We have shown that our algorithms offer useful insights of resource matching in market environments as the interactions between sources and targets reflect the bargaining between them on the amount and price of transferred resources. Furthermore, our algorithms have revealed an averaging principle that sources and targets follow during their bargaining process, which can be used to regulate resource markets and improve resource efficiency. One future direction is to deploy our algorithms to various applications and explore other decentralization methods. Another future direction is to investigate private resource matching through differential privacy.

APPENDIX A: ALTERNATING DIRECTION METHOD OF MULTIPLIERS

Alternating Direction Method of Multipliers (ADMM) is an algorithm for solving problems in the following form [17]:

\[
\begin{align*}
\min_{x \in \mathcal{X}, y \in \mathcal{Y}} & \quad f(x) + g(y) \\
\text{s.t.} & \quad Ax = y.
\end{align*}
\]

where \( f : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\} \) and \( g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} \) are convex functions, \( A \in \mathbb{R}^{m \times n} \) is a linear operator. The optimal value is denoted by

\[
P^* = \inf_{x \in \mathcal{X}, y \in \mathcal{Y}} \{ f(x) + g(y) \mid Ax = y \}.
\]

With \( \alpha \in \mathbb{R}^n \) donating the Lagrange multiplier corresponding to the constraint \( Ax = y \), the augmented Lagrangian is

\[
\mathcal{L}(x, y, \alpha) = f(x) + g(y) + \alpha^T (Ax - y) + \frac{\eta}{2} \|Ax - y\|^2_2,
\]

where the parameter \( \eta > 0 \) controls the impact of the constraint violation in 40. The ADMM iterations are given by

\[
\begin{align*}
x(k+1) & \in \arg \min_{x \in \mathcal{X}} \mathcal{L}(x, y(k), \alpha(k)); \\
y(k+1) & \in \arg \min_{y \in \mathcal{Y}} \mathcal{L}(x(k+1), y, \alpha(k)); \\
\alpha(k+1) & = \alpha(k) + \eta (Ax(k+1) - y(k+1)).
\end{align*}
\]

Furthermore, we have the following lemma regarding the convergence of ADMM from Section 3.2 in [17].

**Lemma 1.** Under the following assumptions:

(i) \( f : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\} \) and \( g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} \) are closed, proper, and convex;

(ii) The unaugmented Lagrangian \( \mathcal{L}_0(x, y, \alpha) = f(x) + g(y) + \alpha^T (Ax - y) \) has a saddle point, i.e., there exist \((x^*, y^*, \alpha^*)\), not necessarily unique, for which \( \mathcal{L}_0(x^*, y^*, \alpha^*) \leq \mathcal{L}_0(x, y, \alpha) \leq \mathcal{L}_0(x^*, y^*, \alpha^*) \) holds for all \( x, y, \alpha \), iterations 42-44 converge as follows:
• Residual convergence. \( r(k) = \|Ax(k) - y(k)\|_2 \to 0 \) as \( k \to \infty \), i.e., the iterations approach feasibility.

• Objective convergence. \( f(x(k)) + g(y(k)) \to p^* \) as \( k \to \infty \), i.e., the objective function of the iterations approaches the optimal value.

• Dual variable convergence. \( \alpha(k) \to \alpha^* \) as \( k \to \infty \), where \( \alpha^* \) is a dual optimal point.

**APPENDIX B: PROOF OF PROPOSITION 1**

We first need to write problem (4) into the form of (10). Let us define \( x = [\text{vec}(\Pi_s^{(t)})^T, \text{vec}(\Pi_x^{(t)})]^T \), \( y = [\text{vec}(\Pi_y)^T, \text{vec}(\Pi_y^{(s)})]^T \), and \( \alpha = [\text{vec}(\{\alpha_{xy}\})]^T, \text{vec}(\{\alpha_{xy}\})]^T \), where \( \text{vec} \) indicates vectorization. Note that \( x, y, \) and \( \alpha \) are all \( 2G \times 1 \) column vectors, where \( G \) is the number of connections between targets and sources, i.e., the size of \( \mathcal{G} \). Thus, the consensus constraints can be rewritten into its matrix form as \( Ax = y \), where \( A = [I_{2G, G}, 0_{2G, G}, 0_{2G, G}, I_{2G, G}] \). Note that \( 0_{2G, G} \) and \( I_{2G, G} \) denote a zero square matrix and an identity matrix of dimension \( G \), respectively. Moreover, we have \( x \in \mathcal{A}_x \) and \( y \in \mathcal{Y} \), where \( \mathcal{A}_x = \{x|\pi_{xy} \geq 0, \sum_{y \in \mathcal{Y}} \pi_{xy} \leq p_{x,h}, \forall \{x, y\} \in \mathcal{G}\} \) and \( \mathcal{Y} = \{y|\pi_{xy} \geq 0, \sum_{x \in \mathcal{A}_x} \pi_{xy} \leq q_{x,y}, \forall \{x, y\} \in \mathcal{G}\} \). As a result, we can use iterations (42)–(44) to solve problem (4).

Since there is no coupling among \( \Pi_s^{(t)}, \Pi_x^{(t)}, \alpha_{xy}, \) and \( \alpha_{xy2} \), we can decompose (42)–(44) as (6)–(10).

**APPENDIX C: PROOF OF THEOREM 2**

Note that problem (22) is a minimization problem with the objective function \( -\sum_{x \in \mathcal{A}_x} \sum_{y \in \mathcal{Y}} \pi_{xy} y - \sum_{x \in \mathcal{A}_x} \sum_{y \in \mathcal{Y}} \pi_{xy} \delta_{xy} \). We consider problem (23) as maximizing over \( \sum_{x \in \mathcal{A}_x} \sum_{y \in \mathcal{Y}} \pi_{xy} y + \sum_{x \in \mathcal{A}_x} \sum_{y \in \mathcal{Y}} \pi_{xy} \delta_{xy} \) in this proof. Let \( u_x, v_y, w_{xy}^{(t)}, w_{xy}^{(s)}, e^{(t)}_{xy}, \) and \( e^{(s)}_{xy} \) denote the Lagrange multipliers with respect to \( \sum_{y \in \mathcal{Y}} \pi_{xy} = p_x, \sum_{x \in \mathcal{A}_x} \pi_{xy} = q_x, \pi_{xy} = \pi_{xy}^{(t)}, \pi_{xy}^{(s)} \geq 0, \) and \( \pi_{xy}^{(t)} \geq 0, \) respectively, we can achieve the dual problem of problem (23) as [47]

\[
\min_{\{u_x, v_y, w_{xy}^{(t)}, w_{xy}^{(s)}, e^{(t)}_{xy}, e^{(s)}_{xy}\} \in \mathcal{Y}} \quad \sum_{y \in \mathcal{Y}} u_x p_x + \sum_{x \in \mathcal{A}_x} v_y q_y \\
\quad \text{s.t.} \\
\quad \gamma_{xy} - u_x - w_{xy}^{(t)} + e^{(t)}_{xy} = 0, \quad \forall \{x, y\} \in \mathcal{G}; \\
\quad \delta_{xy} - v_y + w_{xy}^{(s)} + e^{(s)}_{xy} = 0, \quad \forall \{x, y\} \in \mathcal{G}; \\
\quad e^{(t)}_{xy} \geq 0, e^{(s)}_{xy} \geq 0, \quad \forall \{x, y\} \in \mathcal{G};
\]

(45)

By introducing \( e^{(t)}_{xy} \geq 0 \) and \( e^{(s)}_{xy} \geq 0 \) into \( \gamma_{xy} - u_x - w_{xy}^{(t)} + e^{(t)}_{xy} = 0 \) and \( \delta_{xy} - v_y + w_{xy}^{(s)} + e^{(s)}_{xy} = 0 \), respectively, we can achieve problem (28).

**APPENDIX D: PROOF OF PROPOSITION 5**

We follow similar steps in the proofs of Proposition 1 and Proposition 2. Let \( x = [u_y^T, v_y^T, \text{vec}(w_y^{(t)})^T, \text{vec}(w_y^{(s)})]^T \) and \( y = [\text{vec}(w_y)^T, \text{vec}(w_y^{(s)})]^T \). Thus, the consensus constraints can be captured as \( Ax = y \), where \( A = [0_{NM, N+M}, I_{NM, NM}, 0_{NM, N+M}, 0_{NM, NM}, I_{NM, NM}] \). Note that \( N \) and \( M \) are the numbers of targets and sources, respectively. As a result, we can achieve the following iterations to solve (51).

\[
\left\{ u_x(k+1), w_{xy}^{(t)}(k+1) \right\} \in \arg \min_{u_x, w_{xy}^{(t)}} u_x p_x + \sum_{y \in \mathcal{Y}} \beta_{xy2}(k) w_{xy}^{(s)}(k) + \frac{\eta}{2} (w_{xy}^{(t)} - w_{xy}(k))^2, \\
\left\{ v_y(k+1), w_{xy}^{(s)}(k+1) \right\} \in \arg \min_{v_y, w_{xy}^{(s)}} v_y q_y - \sum_{x \in \mathcal{X}} \beta_{xy2}(k) w_{xy}^{(s)}(k) + \frac{\eta}{2} (w_{xy}^{(s)} - w_{xy}(k))^2, \\
\left\{ w_{xy}(k+1) \right\} \in \arg \min_{w_{xy}} w_{xy} + \beta_{xy2}(k) w_{xy} + \frac{\eta}{2} (w_{xy} - w_{xy}(k))^2.
\]

(46) (47) (48)

**APPENDIX E: PROOF OF PROPOSITION 6**

We first derive the dual problem of (32). Note that the Lagrange function of (32) is given by:

\[
L_x(u_x, w_{xy}^{(t)}, \lambda_{xy}^{(t)}) = u_x p_x + \sum_{y \in \mathcal{Y}} \beta_{xy2}(k) w_{xy}^{(s)} + \frac{\eta}{2} (w_{xy}^{(t)} - w_{xy}(k))^2 + \sum_{x \in \mathcal{A}_x} \lambda_{xy}^{(t)} (\gamma_{xy} - u_x - w_{xy}^{(t)}),
\]

(51)

where \( \lambda_{xy}^{(t)} \geq 0 \) are the Lagrange multipliers. By KKT conditions from Section 5.5.3 in [45], we have \( \frac{\partial L}{\partial w_{xy}} = p_x - \sum_{y \in \mathcal{Y}} \lambda_{xy}^{(t)} = 0; \frac{\partial L}{\partial \lambda_{xy}^{(t)}} = \beta_{xy2}(k) + \frac{\eta}{2} (w_{xy}^{(t)} - w_{xy}(k) - \lambda_{xy}^{(t)}) = 0. \) Thus, we can achieve \( \sum_{y \in \mathcal{Y}} \lambda_{xy}^{(t)} = p_x, w_{xy}^{(t)} = w_{xy}(k) + \frac{\eta}{2} (\lambda_{xy}^{(t)} - \beta_{xy2}(k)), \) and \( \lambda_{xy}^{(t)} \geq 0. \) By plugging these equalities into (51), we have \( \sum_{y \in \mathcal{Y}} \lambda_{xy}^{(t)} \gamma_{xy} - \sum_{y \in \mathcal{Y}} w_{xy}(k) (\lambda_{xy}^{(t)} - \beta_{xy2}(k)) = \frac{\eta}{2} \sum_{y \in \mathcal{Y}} (\lambda_{xy}^{(t)} - \beta_{xy2}(k))^2. \) As a result, the dual problem of (32) can be written as (50). Note that the terms \( w_{xy}(k) \beta_{xy2}(k) \) have been ignored as they are constant for \( \lambda_{xy}^{(t)}. \) After solving (50), we can find \( w_{xy}(k+1) \) with

\[
w_{xy}(k+1) = w_{xy}(k) + \frac{1}{\eta} (\lambda_{xy}^{(t)}(k+1) - \beta_{xy2}(k)).
\]

(52)
Following a similar method, we can prove that the dual problem of (33) is (37). After solving (37), we can find $w^{(s)}_{xy}(k+1)$ with:

$$w^{(s)}_{xy}(k+1) = w_{xy}(k) + \frac{1}{\eta} (\beta_{xy}(k) - \lambda_{xy}^{(s)}(k+1)). \quad (53)$$

By plugging (52) and (53) into (34) and (35), we can achieve (39) and (38), respectively.

**REFERENCES**

[1] H. Purohit, C. Castillo, F. Diaz, A. Sheth, and P. Meier, “Emergency-relief coordination on social media: Automatically matching resource requests and offers,” First Monday, vol. 19, no. 1, 2013.

[2] F. Fiedrich, F. Gehbauer, and U. Rickers, “Optimized resource allocation for emergency response after earthquake disasters,” Safety science, vol. 35, no. 1-3, pp. 41–57, 2000.

[3] H. H. Willis, “Guiding resource allocations based on terrorism risk,” Risk analysis, vol. 27, no. 3, pp. 597–606, 2007.

[4] T. K. Wijaya, A. Gantman, K. S. Taylor, and L. Chen, “Dynamic resource allocation and reservations,” in Sensor Networks, 2009. IPSN 2009. International Conference on, pp. 1–6, IEEE, 2013.

[5] P. Samadi, A.-H. Mohsenian-Rad, R. Schober, V. W. Wong, and J. Jatskevich, “Optimal real-time pricing algorithm based on utility maximization for smart grid,” in Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on, pp. 415–420, IEEE, 2010.

[6] L. Giagiadis, M. J. Neely, L. Tassiulas, et al., “Resource allocation and cross-layer control in wireless networks,” Foundations and Trends® in Networking, vol. 1, no. 1, pp. 1–144, 2006.

[7] J. Steenstra, A. Gantman, K. S. Taylor, and L. Chen, “Dynamic resource matching system,” Apr. 1 2014. US Patent 8,688,148.

[8] Y. Gu, W. Saad, M. Bennis, M. Debbah, and Z. Han, “Matching theory for future wireless networks: fundamentals and applications,” IEEE Communications Magazine, vol. 53, no. 5, pp. 52–59, 2015.

[9] J. M. Ferris and G. E. Riveros, “Methods and systems for matching resource requests with cloud computing environments,” Mar. 19 2013. US Patent 8,402,139.

[10] S. Ding, C. Xia, Q. Cai, K. Zhou, and S. Yang, “Qos-aware resource matching and recommendation for cloud computing systems,” Applied Mathematics and Computation, vol. 247, pp. 941–950, 2014.

[11] A. Galichon, “Optimal transport methods in economics,” Available at SSRN, 2015.

[12] Y. Geng and C. G. Cassandras, “New “smart parking” system based on resource allocation and reservations,” IEEE Transactions on Intelligent Transportation Systems, vol. 14, no. 3, pp. 1129–1139, 2013.

[13] C. Villani, Optimal transport: old and new, vol. 338. Springer Science & Business Media, 2008.

[14] J. Gao, L. Guibas, N. Milosavljevic, and D. Zhou, “Distributed resource management and matching in sensor networks,” in Information Processing in Sensor Networks, 2009. IPSN 2009. International Conference on, pp. 97–108, IEEE, 2009.

[15] M. Hasan and E. Hossain, “Distributed resource allocation for relay-aided device-to-device communication under channel uncertainties: A stable matching approach,” IEEE Transactions on Communications, vol. 63, no. 10, pp. 3882–3897, 2015.

[16] H. Ramani, M. Livny, and M. Solomon, “Matchmaking: Distributed resource management for high throughput computing,” in High Performance Distributed Computing, 1998. Proceedings. The Seventh International Symposium on, pp. 140–146, IEEE, 1998.

[17] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” Foundations and Trends® in Machine Learning, vol. 3, no. 1, pp. 1–122, 2011.

[18] R. Jain, D.-M. Chiu, and W. R. Hawe, A quantitative measure of fairness and discrimination for resource allocation in shared computer system, vol. 38. Eastern Research Laboratory, Digital Equipment Corporation Hudson, MA, 1984.

[19] K. N. Ross and R. Levaciu, Needs-Based Resource Allocation in Educational Virtual Learning Environments. ERIC, 1999.

[20] W. G. Abrahamsson, “Patterns of resource allocation in wildflower populations of fields and woods,” American Journal of Botany, vol. 71–79, 1979.

[21] H. E. Bouis and L. J. Haddad, Effects of agricultural commercialization on land tenure, household resource allocation, and nutrition in the Philippines, vol. 79. Intl Food Policy Res Inst, 1990.

[22] G. Mainland, D. C. Parkes, and M. Welsh, “Decentralized, adaptive resource allocation for sensor networks,” in Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation-Volume 2, pp. 315–328, USENIX Association, 2005.

[23] A. Farbod and T. D. Todd, “Resource allocation and outage control for solar-powered wlan mesh networks,” IEEE Transactions on Mobile Computing, vol. 6, no. 8, pp. 960–970, 2007.

[24] D. P. Lepak and S. A. Snell, “The human resource architecture: Toward a theory of human capital allocation and development,” Academy of Management Review, vol. 24, no. 1, pp. 31–48, 1999.

[25] S. H. Clearwater, Market-based control: A paradigm for distributed resource allocation. World Scientific, 1996.

[26] N. R. Devanur, K. Jain, B. Sivan, and C. A. Wilkens, “Near optimal online algorithms and fast approximation algorithms for resource allocation problems,” in Proceedings of the 12th ACM conference on Electronic commerce, pp. 29–38, ACM, 2011.

[27] B. Huang and T. Jehara, “Fast b-matching via sufficient selection belief propagation,” in Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, pp. 361–369, 2011.

[28] R. Madan, S. P. Boyd, and S. Lall, “Fast algorithms for resource allocation in wireless cellular networks,” IEEE/ACM Transactions on Networking (TON), vol. 18, no. 3, pp. 973–984, 2010.

[29] N. Bonneel, M. Van De Panne, S. Paris, and W. Heidrich, “Displacement interpolation using lagrangian mass transport,” in ACM Transactions on Graphics (TOG), vol. 30, p. 158, ACM, 2011.

[30] G. Peyré, M. Cuturi, et al., “Computational optimal transport,” tech. rep., 2017.

[31] M. Cuturi, “Sinkhorn distances: lightspeed computation of optimal transport,” in Advances in neural information processing systems, pp. 2292–2300, 2013.

[32] S. Haker, L. Zhu, A. Tannenbaum, and S. Angenent, “Optimal mass transport for registration and warping,” International Journal of computer vision, vol. 60, no. 3, pp. 225–240, 2004.

[33] T. Glimm, “A rigorous analysis using optimal transport theory for a two-reflector design problem with a point source,” Inverse Problems, vol. 26, no. 4, p. 045001, 2010.

[34] R. Buyya, Economic-based Distributed Resource Management and Scheduling for Grid Computing, PhD thesis, Monash University, Melbourne, Australia, 2002.

[35] E. Wei and A. Ozdaglar, “Distributed alternating direction method of multipliers,” in Decision and Control (CDC), 2012 IEEE 51st Annual Conference on, pp. 5445–5450, IEEE, 2012.

[36] P. A. Forero, A. Cano, and G. B. Giannakis, “Consensus-based distributed support vector machines,” Journal of Machine Learning Research, vol. 11, no. May, pp. 1663–1707, 2010.

[37] T.-H. Chang, M. Hong, and X. Wang, “Multi-agent distributed optimization via inexact consensus admm,” IEEE Transactions on Signal Processing, vol. 63, no. 2, pp. 482–497, 2015.

[38] K. Huang and N. D. Sidiroopoulos, “Consensus-admm for general quadratically constrained quadratic programming,” IEEE Transactions on Signal Processing, vol. 64, no. 20, pp. 5297–5310, 2016.

[39] R. Zhang and Q. Zhu, “A game-theoretic approach to design secure and resilient distributed support vector machines,” IEEE transactions on neural networks and learning systems, no. 99, pp. 1–16, 2018.

[40] D. Zennaro, E. Dall’Anese, T. Erseghe, and L. Vangelista, “Fast clock synchronization in wireless sensor networks via admm-based consensus,” in Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (モデリングと最適化移動、アドホックと無線ネットワーク), 2016.

[41] B. Geissler, A. Morsi, L. Schewe, and M. Schmidt, “Solving power-constrained gas transportation problems using an mip-based alternating direction method,” Computers & Chemical Engineering, vol. 82, pp. 303–317, 2015.
[45] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
[46] M. Sion et al., “On general minimax theorems,” *Pacific J. Math.*, vol. 8, no. 1, pp. 171–176, 1958.
[47] G. B. Dantzig, *Linear programming and extensions*. Princeton university press, 1998.