Differentially Private ADMM-Based Distributed Discrete Optimal Transport for Resource Allocation

Jason Hughes and Juntao Chen

***Abstract*— Optimal transport (OT) is a framework that can guide the design of efficient resource allocation strategies in a network of multiple sources and targets. To ease the computational complexity of large-scale transport design, we first develop a distributed algorithm based on the alternating direction method of multipliers (ADMM). However, such a distributed algorithm is vulnerable to sensitive information leakage when an attacker intercepts the transport decisions communicated between nodes during the distributed ADMM updates. To this end, we propose a privacy-preserving distributed mechanism based on output variable perturbation by adding appropriate randomness to each node's decision before it is shared with other corresponding nodes at each update instance. We show that the developed scheme is differentially private, which prevents the adversary from inferring the node's confidential information even knowing the transport decisions. Finally, we corroborate the effectiveness of the devised algorithm through case studies.**

***Index Terms*— Discrete Optimal Transport, Distributed Algorithm, Differential Privacy, Resource Allocation**

I. Introduction

The optimal transport (OT) paradigm can be leveraged to guide the most efficient allocation of a limited amount of resources from a set of sources to a set of targets by considering their heterogeneous preferences [1]. The standard OT framework computes the transport strategy in a centralized manner, which requires the source and target nodes to send their information to a centralized transport planner. This centralized computation mechanism is not scalable when the transport network includes a large number of participants. Thus, it is imperative to design a computationally efficient scheme that applies to large-scale transport design.

To this end, distributed algorithm based on the alternating direction method of multipliers (ADMM) can be used to achieve this goal [1], [2]. In the distributed computation scheme, each node communicates with the connected nodes directly regarding the transport decisions and reaches a consensus through iterative negotiations. Under this paradigm, the central planner is not necessary to coordinate the resource matching. On the one hand, the distributed OT design eliminates the necessity of a centralized communication network based on which each node reports their preference information to the central planner. Instead, the communication occurs between each pair of connected source and target nodes enabled by a peer-to-peer network. Thus, the ADMM-based distributed algorithm does not require sharing all the nodes' information over the network.

However, the distributed OT algorithm still faces adversarial threats [3]. Specifically, the nodes need to communicate their computed resource transport preferences with the connected nodes at each update step in the algorithm. This information could be intercepted by an adversary during its transmission over the communication network (e.g., through eavesdropping attack). The attacker can then use it to infer the private information at each participating node (e.g., node's utility parameters used for the design of transport plan).

The privacy concerns of the distributed OT motivate us to develop an efficient privacy-preserving mechanism that can protect the nodes' sensitive utility information. To do this, we resort to the powerful differential privacy technique [4], [5]. Specifically, we develop an output variable perturbation-based differentially private distributed OT scheme. In this algorithm, instead of sharing the authentic transport strategies directly between connected source and target nodes, each node perturbs their transport decisions by adding a random noise drawn from an appropriate distribution with specified parameters at each step. The proposed algorithm prevents leakage of sensitive information of participants in the network even if the transport strategies shared between nodes during updates are captured by the adversary.

The contributions of this paper are presented as follows.

1) We develop a distributed OT design framework based on the alternating direction method of multipliers to compute the OT strategies efficiently.
2) We incorporate privacy consideration into the distributed OT and propose a differentially private distributed OT algorithm based on an output variable perturbation mechanism.
3) We demonstrate the effectiveness of the developed algorithm through case studies and characterize the trade-off between a node's privacy and transport utility.

Related Works. Differential privacy has been applied to many fields, especially the ones in artificial intelligence and machine learning. For example, perturbation-based ADMM algorithms were developed to improve privacy in classification learning problems [6], [7]. Differential privacy has also been leveraged to investigate privacy issues in empirical risk minimization [8], [9], support vector machines [10] and deep learning [11]. Additionally, differential privacy has been applied to improve the privacy of fog computing [12], and safety of vehicle network [13]. In this work, we address

The authors are with the Department of Computer and Information Sciences, Fordham University, New York, NY, 10023 USA. E-mail: {jhughes50,jchen504}@fordham.edu

the privacy concerns in the ADMM-based distributed OT algorithm based on differential privacy and develop a scheme that has a theoretical guarantee to maintain the privacy of the information at each transport node.

The rest of this paper is organized as follows. Section II presents the basics of discrete optimal transport over a network and develops a distributed algorithm to compute the solution. Section III concerns the privacy of the OT framework and proposes a differentially private distributed OT algorithm. Section IV presents case studies to illustrate the results, and Section V concludes the paper.

II. Discrete Optimal Transport over Networks and Distributed Algorithm

This section presents the framework of discrete optimal transport over a network and then develops a distributed algorithm to compute the optimal transport plan.

A. Discrete Optimal Transport

We denote by $\mathscr{X} := \{1,...,|\mathscr{X}|\}$ a set of destination/target nodes that receive the resources, and $\mathscr{Y} := \{|\mathscr{X}|+1,...,|\mathscr{X}|+|\mathscr{Y}|\}$ a set of origin/source nodes that distribute resources to the targets over a transport network. Additionally, we define $\mathscr{P} = \mathscr{X} \cup \mathscr{Y}$ as the set of all nodes. Each source node $y \in \mathscr{Y}$ is connected to a number of target nodes denoted by $\mathscr{X}_y$, representing that y can choose to allocate its resources to a specific group of destinations $\mathscr{X}_y$. Similarly, each target node $x \in \mathscr{X}$ can receive resources from multiple source nodes, and this set of resource suppliers to target x is denoted by $\mathscr{Y}_x$. It can be seen that the resources are transported over a bipartite network, where one side of the network consists of all source nodes and the other includes all destination nodes. This bipartite network is not necessarily complete because of constrained matching policies between participants. We further denote by $\mathscr{E}$ the set of all feasible transport paths in the network, i.e., $\mathscr{E} := \{\{x,y\} | x \in \mathscr{X}_y, y \in \mathscr{Y}\}$. Here, $\mathscr{E}$ also refers to the set of all edges in the established bipartite graph for resource transportation.

We next denote by $\pi_{xy} \in \mathbb{R}_+$ the amount of resources transported from the origin node $y \in \mathscr{Y}$ to the destination node $x \in \mathscr{X}$, where $\mathbb{R}_+$ is the set of nonnegative real numbers. Let $\Pi := \{\pi_{xy}\}_{x\in\mathscr{X}, y\in\mathscr{Y}}$ be the designed transport plan. Then, the centralized optimal transport problem can be formulated as follows:

$$\begin{aligned} \max_{\Pi} \ & \sum_{x\in\mathscr{X}} \sum_{y\in\mathscr{Y}_x} t_{xy}(\pi_{xy}) + \sum_{y\in\mathscr{Y}} \sum_{x\in\mathscr{X}_y} s_{xy}(\pi_{xy}) \\ \text{s.t.} \ & \underline{p}_x \le \sum_{y\in\mathscr{Y}_x} \pi_{xy} \le \bar{p}_x, \ \forall x \in \mathscr{X}, \\ & \underline{q}_y \le \sum_{x\in\mathscr{X}_y} \pi_{xy} \le \bar{q}_y, \ \forall y \in \mathscr{Y}, \\ & \pi_{xy} \ge 0, \ \forall \{x,y\} \in \mathscr{E}, \end{aligned} \tag{1}$$

where $t_{xy} : \mathbb{R}_+ \to \mathbb{R}$ and $s_{xy} : \mathbb{R}_+ \to \mathbb{R}$ are utility functions for target node x and source node y, respectively. Furthermore, $\bar{p}_x \ge \underline{p}_x \ge 0$, $\forall x \in \mathscr{X}$ and $\bar{q}_y \ge \underline{q}_y \ge 0$, $\forall y \in \mathscr{Y}$. The constraints $\underline{p}_x \le \sum_{y\in\mathscr{Y}_x} \pi_{xy} \le \bar{p}_x$ and $\underline{q}_y \le \sum_{x\in\mathscr{X}_y} \pi_{xy} \le \bar{q}_y$ capture the limitations on the amount of requested and transferred resources at the target x and source y, respectively.

We have the following assumption on the utility functions.

Assumption 1. *The utility functions t_{xy} and s_{xy} are concave and monotonically increasing on π_{xy}, $\forall x \in \mathscr{X}, \forall y \in \mathscr{Y}$. Moreover, they are continuously differentiable with $t'_{xy} \le \rho$ and $s'_{xy} \le \rho$, where ρ is a positive constant.*

A rich class of functions satisfy the conditions in Assumption 1. For example, the utility functions t_{xy} and s_{xy} can be linear on π_{xy}, indicating a linear growth of benefits on the amount of transferred and consumed resources.

B. Distributed Optimal Transport

Next, we establish a distributed algorithm for computing the optimal transport strategy in (1). Our first step is to reformulate the optimization problem by introducing ancillary variables $\pi_{xy,t}$ and $\pi_{xy,s}$. The additional subscripts t and s indicate that the corresponding parameters belong to the target node or the source node, respectively. We then set $\pi_{xy} = \pi_{xy,t}$ and $\pi_{xy} = \pi_{xy,s}$, indicating that the solutions proposed by the targets and sources are consistent. This reformulation facilitates the design of a distributed algorithm which allows us to iterate through the process in obtaining the optimal transport plan. To this end, the reformulated optimal transport problem is presented as follows:

$$\begin{aligned} \min_{\Pi_t\in\mathscr{F}_t, \Pi_s\in\mathscr{F}_s, \Pi} \ & -\sum_{x\in\mathscr{X}} \sum_{y\in\mathscr{Y}_x} t_{xy}(\pi_{xy,t}) - \sum_{y\in\mathscr{Y}} \sum_{x\in\mathscr{X}_y} s_{xy}(\pi_{xy,s}) \\ \text{s.t.} \ & \pi_{xy,t} = \pi_{xy}, \ \forall \{x,y\} \in \mathscr{E}, \\ & \pi_{xy,s} = \pi_{xy}, \ \forall \{x,y\} \in \mathscr{E}, \end{aligned} \tag{2}$$

where $\Pi_t := \{\pi_{xy,t}\}_{x\in\mathscr{X}_y, y\in\mathscr{Y}}$, $\Pi_s := \{\pi_{xy,s}\}_{x\in\mathscr{X}, y\in\mathscr{Y}_x}$, $\mathscr{F}_t := \{\Pi_t | \pi_{xy,t} \ge 0, \underline{p}_x \le \sum_{y\in\mathscr{Y}_x} \pi_{xy,t} \le \bar{p}_x, \ \{x,y\} \in \mathscr{E}\}$, and $\mathscr{F}_s := \{\Pi_s | \pi_{xy,s} \ge 0, \underline{q}_y \le \sum_{x\in\mathscr{X}_y} \pi_{xy,s} \le \bar{q}_y, \ \{x,y\} \in \mathscr{E}\}$.

We resort to the alternating direction method of multipliers (ADMM) [14] to develop a distributed computational algorithm. First, let $\alpha_{xy,s}$ and $\alpha_{xy,t}$ be the Lagrangian multipliers associated with the constraint $\pi_{xy,s} = \pi_{xy}$ and $\pi_{xy,t} = \pi_{xy}$, respectively. The Lagrangian function associated with the optimization problem (2) can then be written as follows:

$$\begin{aligned} L(\Pi_t, \Pi_s, \Pi, \alpha_{xy,t}, \alpha_{xy,s}) = & -\sum_{x\in\mathscr{X}} \sum_{y\in\mathscr{Y}_x} t_{xy}(\pi_{xy,t}) \\ & - \sum_{y\in\mathscr{Y}} \sum_{x\in\mathscr{X}_y} s_{xy}(\pi_{xy,s}) + \sum_{x\in\mathscr{X}} \sum_{y\in\mathscr{Y}_x} \alpha_{xy,t}(\pi_{xy,t} - \pi_{xy}) \\ & + \sum_{y\in\mathscr{Y}} \sum_{x\in\mathscr{X}_y} \alpha_{xy,s}(\pi_{xy} - \pi_{xy,s}) + \frac{\eta}{2} \sum_{x\in\mathscr{X}} \sum_{y\in\mathscr{Y}_x} (\pi_{xy,t} - \pi_{xy})^2 \\ & + \frac{\eta}{2} \sum_{y\in\mathscr{Y}} \sum_{x\in\mathscr{X}_y} (\pi_{xy} - \pi_{xy,s})^2, \end{aligned} \tag{3}$$

where $\eta > 0$ is a positive scalar constant controlling the convergence rate in the algorithm designed below.

Note that in (3), the last two terms $\frac{\eta}{2}\sum_{x\in\mathscr{X}}\sum_{y\in\mathscr{Y}_x}(\pi_{xy,t} - \pi_{xy})^2$ and $\frac{\eta}{2}\sum_{y\in\mathscr{Y}}\sum_{x\in\mathscr{X}_y}(\pi_{xy} - \pi_{xy,s})^2$, acting as penalization, are quadratic. Hence, the Lagrangian function L is strictly convex, ensuring the existence of a unique optimal solution.

We next apply ADMM to the minimization problem in (2). The designed distributed algorithm is presented in the following proposition.

Proposition 1. *The iterative steps of applying ADMM to problem* (2) *are summarized as follows:*

$$\Pi_{x,t}(k+1) \in \arg\min_{\Pi_{x,t}\in\mathscr{F}_{x,t}} -\sum_{y\in\mathscr{Y}_x} t_{xy}(\pi_{xy,t}) + \sum_{y\in\mathscr{Y}_x} \alpha_{xy,t}(k)\pi_{xy,t} + \frac{\eta}{2}\sum_{y\in\mathscr{Y}_x}(\pi_{xy,t}-\pi_{xy}(k))^2, \tag{4}$$

$$\Pi_{y,s}(k+1) \in \arg\min_{\Pi_{y,s}\in\mathscr{F}_{y,s}} -\sum_{x\in\mathscr{X}_y} s_{xy}(\pi_{xy,s}) - \sum_{x\in\mathscr{X}_y} \alpha_{xy,s}(k)\pi_{xy,s} + \frac{\eta}{2}\sum_{x\in\mathscr{X}_y}(\pi_{xy}(k)-\pi_{xy,s})^2, \tag{5}$$

$$\pi_{xy}(k+1) = \arg\min_{\pi_{xy}} -\alpha_{xy,t}(k)\pi_{xy} + \alpha_{xy,s}(k)\pi_{xy} + \frac{\eta}{2}(\pi_{xy,t}(k+1)-\pi_{xy})^2 + \frac{\eta}{2}(\pi_{xy}-\pi_{xy,s}(k+1))^2, \tag{6}$$

$$\alpha_{xy,t}(k+1) = \alpha_{xy,t}(k) + \eta(\pi_{xy,t}(k+1)-\pi_{xy}(k+1))^2, \tag{7}$$

$$\alpha_{xy,s}(k+1) = \alpha_{xy,s}(k) + \eta(\pi_{xy}(k+1)-\pi_{xy,s}(k+1))^2, \tag{8}$$

where $\Pi_{\tilde{x},t} := \{\pi_{xy,t}\}_{y\in\mathscr{Y}_{\tilde{x}},x=\tilde{x}}$ *represents the solution at target node* $\tilde{x}\in\mathscr{X}$*, and* $\Pi_{\tilde{y},s} := \{\pi_{xy,s}\}_{x\in\mathscr{X}_{\tilde{y}},y=\tilde{y}}$ *represents the proposed solution at source node* $\tilde{y}\in\mathscr{Y}$*. In addition,* $\mathscr{F}_{x,t} := \{\Pi_{x,t}|\pi_{xy,t}\geq 0, y\in\mathscr{Y}_x, \underline{p}_x \leq \sum_{y\in\mathscr{Y}_x}\pi_{xy,t}\leq \bar{p}_x\}$*, and* $\mathscr{F}_{y,s} := \{\Pi_{y,s}|\pi_{xy,s}\geq 0, x\in\mathscr{X}_y, \underline{q}_y \leq \sum_{x\in\mathscr{X}_y}\pi_{xy,s}\leq \bar{q}_y\}$.

Proof. Let $\vec{x} = [\vec{\Pi}_{x,t}^T, \vec{\Pi}^T]^T$, $\vec{y} = [\vec{\Pi}^T, \vec{\Pi}_{y,s}^T]^T$, and $\alpha = [\{\alpha_{xy,t}\}^T, \{\alpha_{xy,s}\}^T]^T$, where $\vec{\cdot}$ denotes the vectorization operator. We note that these vectors are all $2|\mathscr{E}|\times 1$, where $|\mathscr{E}|$ denotes the number of connections between targets and sources. Now we can write the constraints in matrix form such that $A\vec{x}=\vec{y}$ where $A=[\mathbf{I},\mathbf{0},\mathbf{I},\mathbf{0}]$. Here $\mathbf{I}$ and $\mathbf{0}$ denote the identity and zero matrices respectively, both of which are $|\mathscr{E}|\times|\mathscr{E}|$. Next, we note that $\vec{x}\in\mathscr{F}_{\vec{x},t}$ and $\vec{y}\in\mathscr{F}_{\vec{y},s}$, where $\mathscr{F}_{\vec{x},t} = \{\vec{x}|\pi_{xy,t}\geq 0, \underline{p}_x \leq \sum_{y\in\mathscr{Y}_x}\pi_{xy,t}\leq \bar{p}_x, \{x,y\}\in\mathscr{E}\}$, $\mathscr{F}_{\vec{y},s} := \{\vec{y}|\pi_{xy,s}\geq 0, \underline{q}_y \leq \sum_{x\in\mathscr{X}_y}\pi_{xy,s}\leq \bar{q}_y, \{x,y\}\in\mathscr{E}\}$. In turn we can solve the minimization in (2) with the iterations: 1) $\vec{x}(k+1)\in\arg\min_{\vec{x}\in\mathscr{F}_{x,t}} L(\vec{x},\vec{y}(k),\alpha(k))$; 2) $\vec{y}(k+1)\in\arg\min_{\vec{y}\in\mathscr{F}_{y,s}} L(\vec{x}(k),\vec{y},\alpha(k))$; 3) $\alpha(k+1)=\alpha(k)+\eta(A\vec{x}(k+1)-\vec{y}(k+1))$, whose convergence is proved [14]. Because there is no coupling among $\Pi_{x,t},\Pi_{y,s},\pi_{xy},\alpha_{xy,t}$, and $\alpha_{xy,s}$, the above iterations can be decomposed to (4)-(8). ■

We can simplify steps (4)-(8) down to four steps, and the results are summarized below.

Proposition 2. *The iterations* (4)-(8) *can be simplified as follows:*

$$\Pi_{x,t}(k+1) \in \arg\min_{\Pi_{x,t}\in\mathscr{F}_{x,t}} -\sum_{y\in\mathscr{Y}_x} t_{xy}(\pi_{xy,t}) + \sum_{y\in\mathscr{Y}_x} \alpha_{xy}(k)\pi_{xy,t} + \frac{\eta}{2}\sum_{y\in\mathscr{Y}_x}(\pi_{xy,t}-\pi_{xy}(k))^2, \tag{9}$$

Algorithm 1 Distributed OT Algorithm

1: **while** $\Pi_{x,t}$ and $\Pi_{y,s}$ not converging **do**
2: Compute $\Pi_{x,t}(k+1)$ using (9), for all $x\in\mathscr{X}_y$
3: Compute $\Pi_{y,s}(k+1)$ using (10), for all $y\in\mathscr{Y}_x$
4: Compute $\pi_{xy}(k+1)$ using (11), for all $\{x,y\}\in\mathscr{E}$
5: Compute $\alpha_{xy}(k+1)$ using (12), for all $\{x,y\}\in\mathscr{E}$
6: **end while**
7: **return** $\pi_{xy}(k+1)$, for all $\{x,y\}\in\mathscr{E}$

$$\Pi_{y,s}(k+1) \in \arg\min_{\Pi_{y,s}\in\mathscr{F}_{y,s}} -\sum_{x\in\mathscr{X}_y} s_{xy}(\pi_{xy,s}) - \sum_{x\in\mathscr{X}_y} \alpha_{xy}(k)\pi_{xy,s} + \frac{\eta}{2}\sum_{x\in\mathscr{X}_y}(\pi_{xy}(k)-\pi_{xy,s})^2, \tag{10}$$

$$\pi_{xy}(k+1) = \frac{1}{2}(\pi_{xy,t}(k+1)+\pi_{xy,s}(k+1)), \tag{11}$$

$$\alpha_{xy}(k+1) = \alpha_{xy}(k) + \frac{\eta}{2}(\pi_{xy,t}(k+1)-\pi_{xy,s}(k+1)). \tag{12}$$

Proof. The simplification can be obtained straightforwardly by first characterizing the solution to (6) and then substituting it into (7) and (8). ■

For convenience, we summarize the distributed OT algorithm into Algorithm 1.

III. Differentially Private Distributed Optimal Transport

In this section, we first present the privacy concerns in the developed distributed OT in Section II. We then develop a differentially private distributed OT algorithm that promotes nodes' privacy explicitly during decision updates.

A. Privacy Concerns in the Distributed OT

In the previous distributed OT algorithm, the intermediate results are shared between connected nodes during updates. This sharing mechanism raises privacy concerns as an adversary that can access this result (e.g., through eavesdropping attack) has the ability to infer the participants' private information. Specifically, the adversary could leverage the compromised information $\Pi_{x,t}(k)$ and $\Pi_{y,s}(k)$ at each update step, k, to infer the node's private information including the sensitive preference parameters in the utility functions t_{xy} and s_{xy}. We denote the set of private preference information at node p by D_p, $p\in\mathscr{P}$.

We next use an example to further illustrate node's private information set. Specifically, we consider utility functions admitting a linear form for both the sender and receiver: $t_{xy}(\pi_{xy}) = \delta_{xy}\pi_{xy}$ and $s_{xy}(\pi_{xy}) = \gamma_{xy}\pi_{xy}$, where $\delta_{xy},\gamma_{xy}\in\mathbb{R}_+$. Then, for a target node $x\in\mathscr{X}$, we have set $D_x = \{\delta_{xy} : \forall y\in\mathscr{Y}_x\}$. Similarly for a source node $y\in\mathscr{Y}$, we have set $D_y = \{\gamma_{xy} : \forall x\in\mathscr{X}_y\}$. The information contained in D_p is crucial for developing optimal transport plans. Leakage of such private information is undesired in many resource allocation scenarios, especially those with societal impacts. For example, in the distribution of scarce vaccine resources, these preference parameters could indicate the severity of epidemics in different neighborhoods (modeled by nodes). It

is obvious that each participant does not want to leak this piece of information to other unauthorized parties.

To this end, we aim to protect the privacy of each node in the transport network using differential privacy [5]. Specifically, we propose to add randomness to the transport decisions communicated between each pair of source-target nodes during updates, preventing the adversary from learning the sensitive utility parameters of nodes simply based on the transport decisions. To achieve this goal, first, let D_p and D'_p be two information/data sets differ by one data point (utility parameter). In other words, their *Hamming Distance* is equal to 1, denoted by $H(D_p, D'_p) = 1$. Here, $H(D_p, D'_p) = \sum_{i=1}^{|D_p|} \mathbf{1}\{i : d_i \neq d'_i\}$, where d_i and d'_i denote the ith data point in the information set D_p and D'_p, respectively. Recall that the data points in these sets refer to the nodes' utility parameters which we aim to protect from leakage under the condition that the adversary intercepts the transport plans. The formal definition of differential privacy is presented below.

Definition 1 ($\beta_p(k)$-Differential Privacy)**.** *Consider the transport network $\mathcal{G} = \{\mathcal{P}, \mathcal{E}\}$, where $\mathcal{P}$ is composed of both source nodes and target nodes, and $\mathcal{E}$ is a set of edges connecting the nodes. At each node $p \in \mathcal{P}$, there is an information set D_p which is used to compute the resource transport plan. Let R be a randomized counterpart of Algorithm 1. Further, let $\beta(k) = \big(\beta_1(k), \beta_2(k), ..., \beta_{|\mathcal{P}|}(k)\big) \in \mathbb{R}_+^{|\mathcal{P}|}$, where $\beta_p(k) \in \mathbb{R}_+$ is the privacy parameter of node p at iteration k. Consider the outputs $\Pi_{x,t}(k)$ and $\Pi_{y,s}(k)$ at iteration k of Algorithm 1. Let D'_p be any information set such that $H(D'_p, D_p) = 1$ and $\widetilde{\Pi}^t_x(k)$ and $\widetilde{\Pi}^s_y(k)$ be the corresponding outputs of Algorithm 1 while using the information set D'_p. The algorithm R is $\beta_p(k)$-differentially private for any D'_p for all nodes $p \in \mathcal{P}$ and for all possible sets of outcome solutions S, if the following condition is satisfied at every iteration k:*

$$\Pr[\Pi_p(k) \in S] \leq \exp\big(\beta_p(k)\big) \cdot \Pr[\widetilde{\Pi}_p \in S], \tag{13}$$

where $\Pi_p(k) = \begin{cases} \Pi_{p,t}(k), \text{ if } p \in \mathcal{X}, \\ \Pi_{p,s}(k), \text{ if } p \in \mathcal{Y}, \end{cases}$ *and* $\widetilde{\Pi}_p(k) = \begin{cases} \widetilde{\Pi}_{p,t}(k), \text{ if } p \in \mathcal{X}, \\ \widetilde{\Pi}_{p,s}(k), \text{ if } p \in \mathcal{Y}. \end{cases}$

B. Output Variable Perturbation

In order to ensure that the sensitive preference information at each node remains private when transport plans are published over the network, we develop a differentially private algorithm based on output variable perturbation. This algorithm involves adding random noise to the output decision variables $\Pi_{x,t}(k+1)$ and $\Pi_{y,s}(k+1)$ during updates. More specifically, the random noise vectors, $\varepsilon_x(k+1) \in \mathbb{R}^{|\mathcal{Y}_x|}$ and $\varepsilon_y(k+1) \in \mathbb{R}^{|\mathcal{X}_y|}$ are added to the variables $\Pi_{x,t}(k+1)$ and $\Pi_{y,s}(k+1)$ obtained by (9) and (10), respectively.

Recall that $p \in \mathcal{P} = \mathcal{X} \cup \mathcal{Y}$ and thus $p = x$, $\forall x \in \mathcal{X}$, and $p = y$, $\forall y \in \mathcal{Y}$. The random noise vector $\varepsilon_p(k)$ is generated according to a distribution with density function $F_p(\varepsilon) \sim e^{-\xi_p(k)||\varepsilon||}$. Here, $\xi_p(k) = \frac{\rho}{\eta}\beta_p(k)$, where β_p is a privacy term at each node p.

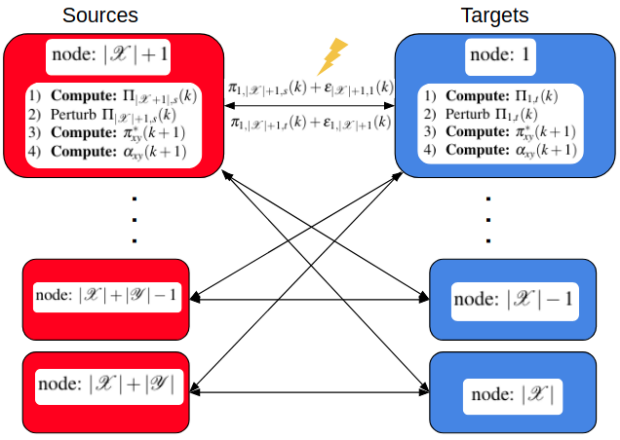


Fig. 1. Illustration of the differentially private distributed OT scheme. The information exchanged between nodes is susceptible to be intercepted by the adversary (e.g., by eavesdropping attack to the wireless channel). Hence, an appropriate random noise is added to the outputs at each update step.

Thus, the proposed solutions at the target node x and the source node y at step $k+1$ admit

$$\begin{aligned} \Pi^*_{x,t}(k+1) &= \Pi_{x,t}(k+1) + \varepsilon_x(k+1), \\ \Pi^*_{y,s}(k+1) &= \Pi_{y,s}(k+1) + \varepsilon_y(k+1), \end{aligned} \tag{14}$$

where $\Pi^*_{x,t}$ and $\Pi^*_{x,t}$ are perturbed solutions of Π^t_x and Π^t_x, respectively. The distributed OT algorithm with output perturbation includes the following steps:

$$\begin{aligned} \Pi_{x,t}(k+1) \in \arg \min_{\Pi^t_x \in \mathcal{F}^t_x} & - \sum_{y \in \mathcal{Y}_x} t_{xy}(\pi_{xy,t}) \\ & + \sum_{y \in \mathcal{Y}_x} \alpha_{xy}(k)\pi_{xy,t} + \frac{\eta}{2} \sum_{y \in \mathcal{Y}_x} \big(\pi_{xy,t} - \pi_{xy}(k)\big)^2, \end{aligned} \tag{15}$$

$$\Pi^*_{x,t}(k+1) = \Pi_{x,t}(k+1) + \varepsilon_x(k+1), \tag{16}$$

$$\begin{aligned} \Pi_{y,s}(k+1) \in \arg \min_{\Pi_{y,s} \in \mathcal{F}_{y,s}} & - \sum_{x \in \mathcal{X}_y} s_{xy}(\pi_{xy,s}) \\ & - \sum_{x \in \mathcal{X}_y} \alpha_{xy}(k)\pi_{xy,s} + \frac{\eta}{2} \sum_{x \in \mathcal{X}_y} \big(\pi_{xy}(k) - \pi_{xy,s}\big)^2, \end{aligned} \tag{17}$$

$$\Pi^*_{y,s}(k+1) = \Pi_{y,s}(k+1) + \varepsilon_y(k+1), \tag{18}$$

$$\pi^*_{xy}(k+1) = \frac{1}{2}\big(\pi^*_{xy,t}(k+1) + \pi^*_{xy,s}(k+1)\big), \tag{19}$$

$$\alpha_{xy}(k+1) = \alpha_{xy}(k) + \frac{\eta}{2}\big(\pi^*_{xy,t}(k+1) - \pi^*_{xy,s}(k+1)\big). \tag{20}$$

As a result of the perturbation in (16) and (18), $\Pi^*_{x,t}(k)$ and $\Pi^*_{y,s}(k)$ are randomized. Specifically, within each iteration, the node perturbs the output variable $\Pi_{x,t}(k)$ or $\Pi_{y,s}(k)$ respectively in order to obtain $\Pi^*_{x,t}(k)$ or $\Pi^*_{y,s}(k)$. The proposed scheme is further illustrated in Fig. 1. It is important to note that the information sets at each node, i.e., D_p containing sensitive utility parameters, remain untouched and not perturbed. Due to the random output perturbation, the transport strategy does not converge to a deterministic value compared with the distributed algorithm in Section II-B.

Instead, the algorithm converges approximately and oscillates within a bounded interval. The magnitude of the oscillation is directly related to the differential privacy parameter β_p chosen by each node $p \in \mathscr{P}$. When β_p becomes larger, $\forall p \in \mathscr{P}$, the differentially privacy algorithm tends to converge to the same solution yielded by Algorithm 1. Since noise is added to each output, the solution will oscillate around the optimal solution. To ensure convergence we check that the oscillation is within some threshold. For convenience, the differentially private distributed OT algorithm based on the output variable perturbation is summarized in Algorithm 2.

Algorithm 2 Differentially Private Distributed OT Algorithm With Output Variable Perturbation

for $k = 0, 1, 2, \ldots$ **do**
  for $x \in \mathscr{X}_y$ **do**
    Compute $\Pi_{x,t}(k+1)$ using (15)
    Compute $\Pi^*_{x,t}(k+1)$ using (16)
  end for
  for $y \in \mathscr{Y}_x$ **do**
    Compute $\Pi_{y,s}(k+1)$ using (17)
    Compute $\Pi^*_{y,s}(k+1)$ using (18)
  end for
  Compute $\pi^*_{xy}(k+1)$ using (19), for all $\{x,y\} \in \mathscr{E}$
  Compute $\alpha_{xy}(k+1)$ using (20), for all $\{x,y\} \in \mathscr{E}$
end for
return $\pi^*_{xy}(k+1)$, for all $\{x,y\} \in \mathscr{E}$

We further have the following Theorem 1 to theoretically guarantee the privacy-preserving property of Algorithm 2.

Theorem 1. *The proposed Algorithm 2 is β-differentially private with $\beta_p(k)$ for node p at iteration k. Let $Q(\Pi^*_{x,t}|D_x)$ and $Q(\Pi^*_{x,t}|D'_x)$ be the probability density functions for $\Pi^*_{x,t}$ given the information sets D_x and D'_x such that $H(D_x, D'_x) = 1$. The ratio of probability density of $\Pi^*_{x,t}$ is bounded:*

$$\frac{Q(\Pi^*_{x,t}(k)|D_x)}{Q(\Pi^*_{x,t}(k)|D'_x)} \leq e^{\beta_x(k)}. \tag{21}$$

It follows similarly for the probability density of $\Pi^*_{y,s}$, i.e.,

$$\frac{Q(\Pi^*_{y,s}(k)|D_y)}{Q(\Pi^*_{y,s}(k)|D'_y)} \leq e^{\beta_y(k)}. \tag{22}$$

Note that (21) *and* (22) *directly imply* $\frac{\Pr(\Pi^*_{x,t}(k)|D_x)}{\Pr(\Pi^*_{x,t}(k)|D'_x)} \leq e^{\beta_x(k)}$ *and* $\frac{\Pr(\Pi^*_{y,s}(k)|D_y)}{\Pr(\Pi^*_{y,s}(k)|D'_y)} \leq e^{\beta_y(k)}$*, respectively.*

Proof. We first show the bounded ratio in (21). We have $\frac{Q(\Pi^*_{x,t}(k)|D_x)}{Q(\Pi^*_{x,t}(k)|D'_x)} = \frac{F_x(\varepsilon_x(k))}{F_x(\varepsilon'_x(k))} = \frac{e^{-\xi_x(k)||\varepsilon_x(k)||}}{e^{-\xi_x(k)||\varepsilon'_x(k)||}}$. Our goal is to find a $\xi_x(k)$ such that the following inequality holds $\xi_x(k)(||\varepsilon_x(k)|| - ||\varepsilon'_x(k)||) \leq \beta_p(k)$. Let $W = \arg\min_{\Pi_{x,t}} f_x(k|D_x)$ and $W' = \arg\min_{\Pi_{x,t}} f_x(k|D'_x)$, where $f_x(k)$ is the objective function for the target node $x \in \mathscr{X}$ at iteration k, shown in (15). Also, let g and h be defined at each node $x \in \mathscr{X}$ such that $g(\Pi^*_{x,t}(k)) = f_x(k|D_x)$ and $h(\Pi^*_{x,t}(k)) = f_x(k|D'_x) - f_x(k|D_x)$.

Therefore, $h(\Pi^*_{x,t}(k)) = -\tilde{t}_{xy}(\pi_{xy,t}) + t_{xy}(\pi_{xy,t})$, where $\tilde{t}_{xy}$ refers to the altered utility function due to the difference between D'_x and D_x. Assumption 1 implies that $f_x(k|D_p) = g(\Pi^*_{x,t}(k))$ and $f_x(k|D'_x) = g(\Pi^*_{x,t}(k)) + h(\Pi^*_{x,t}(k))$ are both convex. We differentiate $h(\Pi^*_{x,t}(k))$ with respect to $\Pi^*_{x,t}(k)$ and get:

$$\nabla h(\Pi^*_{x,t}(k)) = -\tilde{t}'_{xy}(\pi_{xy,t}) + t'_{xy}(\pi_{xy,t}).$$

Assumption 1 further implies that $0 \leq t'_{xy} \leq \rho$. Thus, $||\nabla h(\Pi^*_{x,t})|| \leq \rho$. From the definitions of W and W', we have $\nabla g(W) = \nabla g(W') + \nabla h(W') = 0$. Based on Lemma 14 in [15] and knowing that $g(\cdot)$ is η-strongly convex, the following inequality holds: $\langle \nabla g(W) - g(W'), W - W' \rangle \geq \eta ||W - W'||^2$. Thus, by Cauchy-Schwartz inequality, we obtain

$$||W - W'|| \cdot ||\nabla h(W')|| \geq (W - W')^T \nabla h(W') =$$
$$\langle \nabla g(W) - g(W'), W - W' \rangle \geq \eta ||W - W'||^2.$$

Dividing both sides by $\eta||W - W'||$ yields $||W - W'|| \leq \frac{1}{\eta}||\nabla h(W')|| \leq \frac{\rho}{\eta}$. From (16), we have $||W - W'|| = ||\varepsilon_x(k) - \varepsilon'_x(k)|| \leq \frac{1}{\eta}||\nabla h(W')||$. Thus, we obtain

$$\xi_x(k)(||\varepsilon_x(k)|| - ||\varepsilon'_x(k)||) \leq \xi_x(k)(||\varepsilon_x(k) - \varepsilon'_x(k)||) \leq \frac{\rho}{\eta}\xi_x(k).$$

By choosing $\xi_x(k) = \frac{\eta}{\rho}\beta_p(k)$, the inequality $\xi_x(k)(||\varepsilon_x(k) - \varepsilon'_x(k)||) \leq \beta_p(k)$ holds. Thus, the output variable perturbation is β_p-differentially private for target node $x \in \mathscr{X}$. The proof follows identically for the perturbed output variable $\Pi^*_{y,s}(k)$ at the source node $y \in \mathscr{Y}$ and hence omitted. ■

In summary, the proposed Algorithm 2 guarantees the privacy of all participating nodes during their decision sharing.

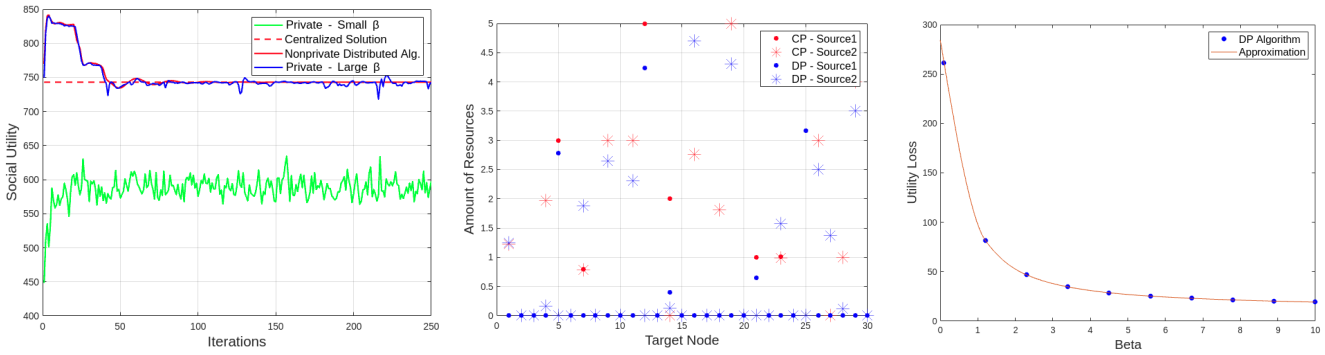
IV. Numerical Case Studies

In this section, we corroborate the effectiveness of the developed differentially private algorithm and show how the added privacy impacts the transport plan and its efficiency.

We construct a transport network with four source nodes and thirty target nodes in which every source node is connected to all target nodes, i.e., the network is complete. The upper bounds at the target nodes $\bar{p}_x$ are kept small (smaller than 5), while the upper bounds at the source nodes $\bar{q}_y$ are relatively larger (between 20 and 40). Such selection yields that the resources at the origin can be transported to heterogeneous target nodes. Additionally, we consider linear utility functions $t_{xy}(\pi_{xy}) = \delta_{xy}\pi_{xy}$, and $s_{xy}(\pi_{xy}) = \gamma_{xy}\pi_{xy}, \forall \{x,y\} \in \mathscr{E}$. The utility parameters δ_{xy} and γ_{xy} are randomly chosen integers between 1 and 5 for each pair of connection, $\forall \{x,y\} \in \mathscr{E}$.

In the following study, we investigate the impact of privacy parameter β_p on the transport utility. According to the definition, a smaller β_p yields a higher level of privacy. We compare the results for two sets of β_p. For the first one, we assign a value of 1 to β_p, $p \in \mathscr{P}$. For the larger value of β_p we use 1000. Furthermore, we select $\eta = 1$ and $\rho = 2$.

We leverage the developed algorithms, Algorithms 1 and 2, to compute the transport plans. The results are shown in Fig. 2. First, we observe that in Fig. 2(a), the trajectory of



(a) Social Utility (b) Transport Plan (c) Privacy and Transport Efficiency Tradeoff

Fig. 2. (a) shows the performance of the proposed algorithms. (b) depicts the optimal transport plans designed by the central planner (CP) and the solution given by the distributed differentially private (DP) algorithm. (c) shows an increase of the privacy level (smaller β_p) decreases the transport utility, reflecting the trade-off between privacy and transport efficiency.

transport plan yielded by the differentially private algorithm converges approximately to a certain value. The oscillation at the tail is due to the random noise added to the decision at each output perturbation step. We can also see that when β_p is small, the resulting social utility (i.e., transport efficiency), which is an aggregation of the utilities of all participating nodes, is relatively small. In comparison, when β_p is large, the social utility is close to the one returned by Algorithm 1 where differential privacy is not incorporated. Fig. 2(c) further shows this phenomenon and reveals the inherent trade-off between the amount of added privacy and the transport efficiency. Fig. 2(b) illustrates how the privacy factor affects the transport plan. The decreased optimality due to the privacy promotion indicates that the resource allocation is no longer taking full advantage of how much source nodes can provide or how much target nodes can request. For example, the target node 12 can request at most 5 units of resources, and does so when privacy is not added to the algorithm. When privacy is concerned, it only requests and receives 4.2 units of resources and hence the social utility is decreased.

V. Conclusion

This paper has developed a differentially private distributed optimal transport algorithm with a theoretical guarantee of achieved privacy. The algorithm protects the sensitive information at each node by perturbing the output of the transport schemes shared between connected nodes during updates. Under the designed mechanism, even if the transport decision is intercepted during its transmission, the adversary still cannot discover the underlying sensitive information used in the transport strategy design. The privacy level for each node can be determined appropriately by considering its trade-off with the resulting transport efficiency. Future work includes extending the current model-based distributed optimal transport framework to data-driven learning-based optimal transport while considering data privacy in the learning process.

References

[1] R. Zhang and Q. Zhu, "Consensus-based distributed discrete optimal transport for decentralized resource matching," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 5, no. 3, pp. 511–524, 2019.

[2] J. Hughes and J. Chen, "Fair and distributed dynamic optimal transport for resource allocation over networks," in *55th Annual Conference on Information Sciences and Systems (CISS)*, 2021.

[3] ——, "Resilient and distributed discrete optimal transport with deceptive adversary: A game-theoretic approach," in *IEEE Control System Letters*, 2022, pp. 1166–1171.

[4] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Theory of cryptography conference*. Springer, 2006, pp. 265–284.

[5] C. Dwork, A. Roth *et al.*, "The algorithmic foundations of differential privacy." *Foundations and Trends in Theoretical Computer Science*, vol. 9, no. 3-4, pp. 211–407, 2014.

[6] T. Zhang and Q. Zhu, "Dynamic differential privacy for ADMM-based distributed classification learning," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 1, pp. 172–187, 2017.

[7] X. Zhang, M. M. Khalili, and M. Liu, "Improving the privacy and accuracy of ADMM-based distributed algorithms," in *International Conference on Machine Learning*. PMLR, 2018, pp. 5796–5805.

[8] J. Pawlick and Q. Zhu, "A mean-field stackelberg game approach for obfuscation adoption in empirical risk minimization," in *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 2017, pp. 518–522.

[9] K. Chaudhuri, C. Monteleoni, and A. D. Sarwate, "Differentially private empirical risk minimization," *Journal of Machine Learning Research*, vol. 12, no. 29, pp. 1069–1109, 2011.

[10] Y. Zhang, Z. Hao, and S. Wang, "A differential privacy support vector machine classifier based on dual variable perturbation," *IEEE Access*, vol. 7, pp. 98 238–98 251, 2019.

[11] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang, "Deep learning with differential privacy," *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security*, 2016.

[12] M. Du, K. Wang, X. Liu, S. Guo, and Y. Zhang, "A differential privacy-based query model for sustainable fog data centers," *IEEE Transactions on Sustainable computing*, vol. 4, no. 2, pp. 145–155, 2017.

[13] T. Zhang and Q. Zhu, "Distributed privacy-preserving collaborative intrusion detection systems for VANETs," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 1, pp. 148–161, 2018.

[14] S. Boyd, N. Parikh, and E. Chu, *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*. Now Publishers, 2011.

[15] S. Shalev-Shwartz, "Online learning: Theory, algorithms, and applications," *PhD Dissertation*, 2007.