Cost Bounds for Pickup and Delivery Problems with Application to Large-Scale Transportation Systems

Kyle Treleaven, Marco Pavone, Emilio Frazzoli

***Abstract*— Demand-responsive transport (DRT) systems, where users generate requests for transportation from a pickup point to a delivery point, are expected to increase in usage dramatically as the inconvenience of privately-owned cars in metropolitan areas becomes excessive. However, despite the increasing role of DRT systems, there are very few rigorous results characterizing achievable performance (in terms, e.g., of stability conditions). In this paper, our aim is to bridge this gap for a rather general model of DRT systems, which takes the form of a generalized Dynamic Pickup and Delivery Problem. The key strategy is to develop analytical bounds for the optimal cost of the Euclidean Stacker Crane Problem (ESCP), which represents a general *static* model for DRT systems. By leveraging such bounds, we characterize a necessary and sufficient condition for the stability of DRT systems; the condition depends only on the workspace geometry, the stochastic distributions of pickup and delivery points, customers' arrival rate, and the number of vehicles. Our results exhibit some surprising features that are absent in traditional spatially-distributed queueing systems.**

I. Introduction

Private automobiles have dramatically changed the concept of *personal urban mobility* by enabling fast and anytime point-to-point travel within large cities. However, the fact that the urban population is projected to jump from the current 3.5 billion to more than 6 billion in the next 30 years [1], coupled with the fact that the availability of urban land for road and parking is bound to decrease, indicates that private automobiles are an unsustainable solution for the future of personal mobility in dense urban environments. Hence, a paradigm shift is emerging whereby the outdated concept of personal urban mobility based on private cars is being replaced by the concept of *large-scale one-way vehicle sharing*.

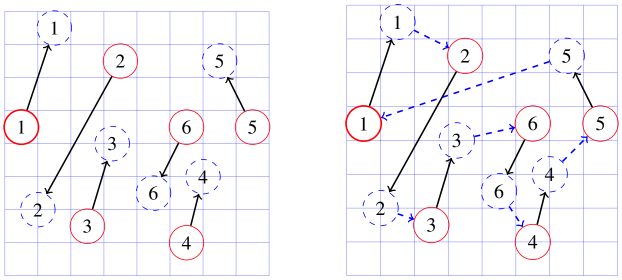
One of the main paradigms for one-way vehicle sharing is represented by demand-responsive transport (DRT) systems, where people drive (or are driven by) shared vehicles from a pickup point to a delivery point. Even though the previous considerations suggest that DRT systems will likely become the backbone for personal urban mobility in large metropolitan areas, there are very few rigorous results about the characterization of their achievable performance (in terms, e.g., of stability conditions). In this paper, our aim is to bridge this gap for a rather general model of DRT system. Specifically, we model a DRT system as a *Dynamic* Pickup and Delivery Problem (DPDP), where pickup requests arrive according to a renewal process and are randomly located according to a general probability distribution; corresponding delivery locations are also randomly distributed according to a general probability distribution (possibly *different* from the pickup distribution), and a number of unit-capacity vehicles must transport demands from their pickup locations to their delivery locations. The objective is to derive a necessary and sufficient condition for the stability of the system, in the sense that customers' waiting times stay bounded at all times. The key strategy and contribution is the development of analytical bounds for the optimal cost of the Euclidean Stacker Crane Problem (ESCP), which represents a *static* model for DRT systems: in the ESCP, each customer is associated with a pickup location and a delivery location, and the objective is to find a minimum-length tour visiting all locations with the constraint that each pickup location and its associated delivery location are visited in consecutive order. By leveraging these bounds, we derive a necessary and sufficient stability condition for our DRT model that only depends on the workspace geometry, the stochastic distributions of pickup and delivery points, customers' arrival rate, and the number of vehicles.

Literature overview. DPDPs with unit-capacity vehicles, which are arguably a rather general model for DRT systems, are generally treated as a sequence of static subproblems and their performance characteristics, such as stability conditions, are generally not characterized. Excellent surveys on heuristics, metaheuristics and online algorithms for DPDPs modeling DRT systems can be found in [2] and [3]. Even though these algorithms are quite effective in addressing DPDPs, alone they do not give any indication of fundamental limits of performance. To the best of our knowledge, the only *analytical* studies for DPDPs modeling DRT systems are [4] and [5]. Specifically, in [4] the authors study the single vehicle case of the problem under the constraint that pickup and delivery distributions are *identical*; in [5] the authors derive bounds for the more general case of multiple vehicles, however under the quite unrealistic assumption of three-dimensional workspaces and equal distributions of pickup and delivery sites. On the contrary, in this paper we derive a necessary and sufficient stability condition for the more realistic case of multiple servicing vehicles and possibly *different* distributions for pickup and delivery sites. We will show that when such distributions are different, our stability condition shows an additional term compared to stability conditions for traditional spatially-distributed queueing systems. Our

Kyle Treleaven and Emilio Frazzoli are with the Laboratory for Information and Decision Systems, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139 {ktreleav, frazzoli}@mit.edu.

Marco Pavone is with the Department of Aeronautics and Astronautics, Stanford University, Stanford, CA 94305 pavone@stanford.edu.

This research was supported in part by the Future Urban Mobility project of the Singapore-MIT Alliance for Research and Technology (SMART) Center, with funding from Singapore's National Research Foundation.



(a) Six pickup/delivery pairs are generated in the Euclidean plane.

(b) Dashed arrows *combined* with the solid arrows represent a stacker crane tour.

Fig. 1. Example of Euclidean Stacker Crane Problem in two dimensions. Solid and dashed circles denote pickup and delivery points, respectively; solid arrows denote the routes from pickup points to their delivery points.

stability condition, by giving structural insights into DRT systems, would provide a system designer with essential information to build business and strategic planning models regarding, e.g., fleet sizing.

Structure of the paper. This paper is structured as follows. In Section II we provide some background on the Euclidean Stacker Crane Problem (which represents a static approximation for DPDPs), on the Euclidean Bipartite Matching Problem, and on some relevent notions in probability theory. Then, in Section III we introduce the DPDP we wish to study and we formally state the objective of the paper, i.e., to derive a necessary and sufficient condition for such a model of DRT systems. In Section IV we present lower and upper bounds on the optimal cost of the ESCP, which will be instrumental, in Section VI, to derive the desired stability condition. Section VII presents our concluding discussion and directions for future work.

II. Background Material

In this section we summarize the background material used in the paper. Specifically, we review the stochastic Euclidean Stacker Crane Problem, the stochastic Euclidean Bipartite Matching Problem (EBMP), a related concept in transportation theory, and a generalized Law of Large Numbers.

A. The Euclidean Stacker Crane Problem

Let $\mathcal{X} = \{x_1, \dots, x_n\}$ and $\mathcal{Y} = \{y_1, \dots, y_n\}$ be two sets of points in the d-dimensional Euclidean space $\mathbb{R}^d$, where $d \geq 1$. The Euclidean Stacker Crane Problem (ESCP) is to find a minimum-length tour through the points in $\mathcal{X} \cup \mathcal{Y}$ with the property that each point x_i (which we call the ith *pickup*) is immediately followed by the point y_i (the ith *delivery*); in other words, the pair (x_i, y_i) must be visited in consecutive order (see Figure 1). The length of a tour is the sum of all Euclidean distances along the tour. We will refer to such a tour as an optimal stacker crane tour, and to a tour that is not minimum-length but still satisfies the pickup-to-delivery constraints as a stacker crane tour. Note that the ESCP is a *constrained* version of the well-known Euclidean Traveling Salesman Problem.

In this paper we focus on a *stochastic* version of the ESCP. Let $\mathcal{X} = \{X_1, \dots, X_n\}$ be a set of points that are independent and identically distributed (i.i.d.) in a compact set $\Omega \subset \mathbb{R}^d$ and distributed according to a density $\varphi_{\mathrm{P}} : \Omega \to \mathbb{R}_{\geq 0}$; let $\mathcal{Y} = \{Y_1, \dots, Y_n\}$ be a set of points that are i.i.d. in Ω and distributed according to a density $\varphi_{\mathrm{D}} : \Omega \to \mathbb{R}_{\geq 0}$. As before, we interpret each pair (X_i, Y_i) as the pickup and delivery sites, respectively, of some transportation request, and we seek to determine the cost of an optimal stacker crane tour through all points. We will refer to this stochastic version of ECSP as ESCP$(n, \varphi_{\mathrm{P}}, \varphi_{\mathrm{D}})$, and we will write $\mathcal{X}, \mathcal{Y} \sim$ ESCP$(n, \varphi_{\mathrm{P}}, \varphi_{\mathrm{D}})$ to mean that $\mathcal{X}$ contains n pickup sites i.i.d. with density φ_{P}, and $\mathcal{Y}$ contains n delivery sites i.i.d. with density φ_{D}. An important contribution of this paper will be to characterize the behavior of the optimal stacker crane tour through $\mathcal{X}$ and $\mathcal{Y}$ as a function of parameters n, φ_{P}, and φ_{D}; throughout the paper we assume that distributions are absolutely continuous with densities φ_{P} and φ_{D}.

B. The Euclidean Bipartite Matching Problem

Let $\mathcal{X} = \{x_1, \dots, x_n\}$ and $\mathcal{Y} = \{y_1, \dots, y_n\}$ be two sets of points in $\mathbb{R}^d$. The Euclidean Bipartite Matching Problem is to find a permutation $\sigma^* : \{1, \dots, n\} \mapsto \{1, \dots, n\}$ (not necessarily unique), such that the sum of the Euclidean distances between the matched pairs $\{(y_i, x_{\sigma^*(i)})$ for $i = 1, \dots, n\}$ is minimized, i.e.:

$$\sum_{i=1}^{n} \|x_{\sigma^*(i)} - y_i\| = \min_{\sigma \in \Pi_n} \sum_{i=1}^{n} \|x_{\sigma(i)} - y_i\|,$$

where $\|\cdot\|$ denotes the Euclidean norm and Π_n denotes the set of all permutations over n elements. Let $Q := (\mathcal{X}, \mathcal{Y})$; we refer to the left-hand side in the above equation as the optimal bipartite matching cost $L_{\mathrm{M}}(Q)$; we refer to $l_{\mathrm{M}}(Q) := L_{\mathrm{M}}(Q)/n$ as the *average* match cost.

The EBMP over *random* sets of points enjoys some remarkable properties. Specifically, let $\mathcal{X} = \{X_1, \dots, X_n\}$ be a set of n points that are i.i.d. in a compact set $\Omega \subset \mathbb{R}^d$, $d \geq 3$, and distributed according to a density $\varphi : \Omega \to \mathbb{R}_{\geq 0}$; let $\mathcal{Y} = \{Y_1, \dots, Y_n\}$ be a set of n points that are i.i.d. in Ω and distributed according to the same density $\varphi : \Omega \to \mathbb{R}_{\geq 0}$. Let $Q = (\mathcal{X}, \mathcal{Y})$. In [6] it is shown that there exists a constant $\beta_{\mathrm{M},d}$ such that the optimal bipartite matching cost $L_{\mathrm{M}}(Q) = \min_{\sigma \in \Pi_n} \sum_{i=1}^{n} \|X_{\sigma(i)} - Y_i\|$ has limit

$$\lim_{n \to +\infty} \frac{L_{\mathrm{M}}(Q)}{n^{1-1/d}} = \beta_{\mathrm{M},d} \int_{\Omega} \bar{\varphi}(x)^{1-1/d} \, dx, \qquad (1)$$

almost surely, where $\bar{\varphi}$ is the density of the absolutely continuous part of the point distribution. The constant $\beta_{\mathrm{M},3}$ has been estimated numerically as $\beta_{\mathrm{M},3} \approx 0.7080 \pm 0.0002$ [7].

In the case $d = 2$ (i.e. the planar case) the following (weaker) result holds with high probability [8]

$$L_{\mathrm{M}}(Q)/(n \log n)^{1/2} \leq \gamma \qquad (2)$$

for some positive constant γ (if φ is uniform, it also holds with high probability that $L_{\mathrm{M}}(Q)/(n \log n)^{1/2}$ is bounded below by a positive constant [9]).

To the best of our knowledge, there are no similar results in the literature that apply when the distribution of $\mathcal{X}$ points is different from the distribution of $\mathcal{Y}$ points (which is the typical case for transportation systems); moreover, despite a close relation between the stochastic ESCP and the stochastic EBMP, the asymptotic cost of a stochastic ESCP has not been characterized to date.

C. Euclidean Wasserstein distance

A significant component of the paper will be to generalize the results in [6] to address the case that the distribution of $\mathcal{X}$ points is different from that of $\mathcal{Y}$ points. The following notion of transportation complexity will prove useful.

Let φ_{P} and φ_{D} be two probability densities over $\Omega \subset \mathbb{R}^d$. The *Euclidean Wasserstein distance*, between φ_{P} and φ_{D}, is defined as

$$W(\varphi_{\mathrm{P}}, \varphi_{\mathrm{D}}) = \inf_{\gamma \in \Gamma(\varphi_{\mathrm{P}}, \varphi_{\mathrm{D}})} \int_{x,y \in \Omega} \|y - x\| \; d\gamma(x, y), \quad (3)$$

where $\Gamma(\varphi_{\mathrm{P}}, \varphi_{\mathrm{D}})$ denotes the set of measures over product-space $\Omega \times \Omega$ having marginal densities φ_{P} and φ_{D} respectively. The Euclidean Wasserstein distance is a continuous version of the Earth Mover's distance; properties of the generalized version are discussed in [10].

D. An Asymptotically Optimal Polynomial-Time Algorithm for the Stochastic ESCP

In [11] the authors have introduced a polynomial-time algorithm for the Euclidean Stacker Crane Problem, called SPLICE, that was shown to be asymptotically optimal for ESCP($n, \varphi_{\mathrm{P}}, \varphi_{\mathrm{D}}$), with $d \geq 2$; optimality is in the sense that $\lim_{n\to+\infty} L_{\mathrm{SPLICE}}(n)/L^*(n) = 1$, almost surely, where $L_{\mathrm{SPLICE}}(n)$ is the length of the output of SPLICE, and $L^*(n)$ is the optimal tour length. The behavior of this algorithm will be instrumental to prove bounds for the length of the optimal stacker crane tour; hence, we briefly describe its logic:

Algorithm SPLICE

Input: a set of *demands* $\mathcal{S} = \{(x_1, y_1), \dots, (x_n, y_n)\}$, $n > 1$.

Output: a stacker crane tour through $\mathcal{S}$.

1: **initialize** $\sigma \leftarrow$ solution to Euclidean bipartite matching problem between sets $\mathcal{X} = \{x_1, \dots, x_n\}$ and $\mathcal{Y} = \{y_1, \dots, y_n\}$ computed by using a bipartite matching algorithm $\mathcal{M}$.

2: Add the n pickup-to-delivery links $x_i \to y_i$, $i = 1, \dots, n$.

3: Add the n matching links $y_i \to x_{\sigma(i)}$, $i = 1, \dots, n$.

4: Apply a *rewiring* heuristic to connect subtours.

The key idea behind SPLICE is to make connections from delivery sites back to pickup sites in accordance with an optimal bipartite matching between the sets of pickup and delivery sites. Unfortunately, this procedure is likely to generate two or more *disconnected subtours*, and so, in general, the result is not a stacker crane tour. *Fortunately*, it was shown that the number of disconnected subtours is "asymptotically small", and so any "rewiring" heuristic can be used to form a feasible stacker crane tour from the relatively small number of disconnected subtours. The asymptotic optimality of SPLICE relies crucially on the growth order for the number of subtours generated by SPLICE with respect to n, the size of the problem instance. Since we will need a similar analysis for the bounds derived in Section IV, we present the result governing the growth order here.

Lemma 2.1 (Asymptotic number of subtours [11]): Let $\mathcal{X}, \mathcal{Y} \sim$ ESCP($n, \varphi_{\mathrm{P}}, \varphi_{\mathrm{D}}$) be a random ESCP instance of size n, in a compact d-dimensional Euclidean environment $\Omega \subset \mathbb{R}^d$, where $d \geq 2$. Let N_n be the number of subtours generated by the SPLICE algorithm on inputs $\mathcal{X}$ and $\mathcal{Y}$. Then $\lim_{n\to+\infty} N_n/n = 0$, almost surely.

Remark 2.2: One can prove, using the same arguments of the proof, that $\lim_{n\to+\infty} N_n/n^{1-1/d} = 0$ almost surely, for any integer $d \geq 3$.

E. The strong law of absolute differences

The last bit of background is a slightly more general *law of large numbers* result. Let $X_1, \dots, X_n$ be a sequence of scalar random variables that are i.i.d. with mean $\mathbb{E}X$ and finite variance. Then the sequence of cumulative sums $S_n = \sum_{i=1}^n X_i$ has the property (discussed, e.g., in [12]) that

$$\lim_{n\to\infty} \frac{S_n - \mathbb{E}S_n}{n^\alpha} = 0, \qquad \text{almost surely,}$$

for any $\alpha > 1/2$. Note that the well-known Strong Law of Large Numbers (SLLN) is the special case where $\alpha = 1$.

III. Problem statement

We consider the following model for DRT systems, which is known in the literature as DPDP: a total of m vehicles travel at unit velocity within a workspace Ω; the vehicles have unlimited range and *unit* capacity (i.e., they can transport one demand at a time). Demands are generated according to a time-invariant renewal process, with time intensity $\lambda \in \mathbb{R}_{>0}$. A newly arrived demand has an associated pickup location which is independent and identically distributed in Ω according to a density φ_{P}. Each demand must be transported from its pickup location to its delivery location, at which time it is removed from the system. The delivery locations are also i.i.d. in Ω according to a density φ_{D}. A policy for routing the vehicles is said to be *stabilizing* if the expected number of demands in the system remains uniformly bounded at all times; in this paper, we consider policies that are *causal*, but for which no additional information restrictions apply (i.e. the policy could depend on any information about past or outstanding requests). The objective of the paper is as follows: *find a necessary and sufficient condition for the existence of stabilizing policies as a function of the system's parameters, i.e., λ, m, φ_{P}, φ_{D}, Ω.*

This problem has been studied in [5] under the restrictive assumptions $\varphi_{\mathrm{D}} = \varphi_{\mathrm{P}} := \varphi$ and $d \geq 3$; in that paper, it has been shown that if one defines the "load factor" as

$$\varrho \doteq \lambda \, \mathbb{E}_\varphi \|Y - X\| / m,$$

where Y and X are two random points in Ω distributed according to the distribution φ, then the condition $\varrho < 1$ is necessary and sufficient for a stabilizing policy to exist. However, that analysis—and indeed the result itself—is no longer valid if $\varphi_{\mathrm{D}} \neq \varphi_{\mathrm{P}}$. This paper will show how the definition of load factor has to be modified for the more realistic case $\varphi_{\mathrm{D}} \neq \varphi_{\mathrm{P}}$. Pivotal in our approach is to characterize, with almost sure analytical bounds, the scaling of the optimal solution of ESCP$(n, \varphi_{\rm P}, \varphi_{\rm D})$ with respect to the problem size.

IV. Analytical Bounds on the Cost of the ESCP

In this section we derive analytical bounds on the cost of the optimal stacker crane tour. The resulting bounds are useful for two reasons: (i) they give further insight into the ESCP (and the EBMP), and (ii) they will allow us to find a necessary and sufficient stability condition for our model of DRT systems (i.e., for the 1-DPDP).

The development of these bounds follows from an analysis of the growth order, with respect to the instance size n, of the EBMP matching on $Q_n = (\mathcal{X}_n, \mathcal{Y}_n)$, where $\mathcal{X}_n, \mathcal{Y}_n \sim$ ESCP$(n, \varphi_{\rm P}, \varphi_{\rm D})$. The main technical challenge is to extend the results in [6], about the length of the matching to the case where $\varphi_{\rm P}$ and $\varphi_{\rm D}$ are not identical. We first derive in Section IV-A a lower bound on the length of the EBMP matching for the case $\varphi_{\rm P} \neq \varphi_{\rm D}$ (and resulting lower bound for the ESCP); then in Section IV-B we find the corresponding upper bounds.

A. A Lower Bound on the Length of the ESCP

In the rest of the paper, we let $\mathcal{C} = \{C^1, \ldots, C^{|\mathcal{C}|}\}$ denote some finite partition of Euclidean environment Ω into $|\mathcal{C}|$ *cells*. We denote by $\varphi_{\rm P}(C^i) := \int_{x \in C^i} \varphi_{\rm P}(x)dx$ the measure of cell C^i under the pickup distribution (with density $\varphi_{\rm P}$), i.e., the probability that a particular pickup X is in the ith cell. Similarly, we denote by $\varphi_{\rm D}(C^i) := \int_{y \in C^i} \varphi_{\rm D}(y)dy$ the cell's measure under the delivery distribution (with density $\varphi_{\rm D}$), i.e., the probability that a particular delivery Y is in the ith cell. Most of the results of the paper are valid for arbitrary partitions of the environment; however, for some of the more delicate analysis we will refer to the following particular construction. Without loss of generality, we assume that the environment $\Omega \subset \mathbb{R}^d$ is a hyper-cube with side-length L. For some integer $r \geq 1$, we construct a partition $\mathcal{C}_r$ of Ω by slicing the hyper-cube into a grid of r^d smaller cubes, each length L/r on a side; inclusion of subscript r in our notation will make the construction explicit. The ordering of cells in $\mathcal{C}_r$ is arbitrary.

Our first result bounds the average length of a match in the optimal bipartite matching, $l_{\rm M}(Q_n)$, asymptotically from below. In preparation for this result we present Problem 1, a linear optimization problem whose solution maps partitions to real numbers.

Problem 1 (Optimistic "rebalancing"):

$$
\begin{aligned}
\underset{\{\alpha_{ij} \geq 0\}_{i,j \in \{1,\ldots,r^d\}}}{\text{Minimize}} \quad & \sum_{ij} \alpha_{ij} \min_{y \in C^i, x \in C^j} \|x - y\| \\
\text{subject to} \quad & \sum_j \alpha_{ij} = \varphi_{\rm D}(C^i) \quad \text{for all } C^i \in \mathcal{C}, \\
& \sum_i \alpha_{ij} = \varphi_{\rm P}(C^j) \quad \text{for all } C^j \in \mathcal{C}.
\end{aligned}
$$

We denote by $\mathcal{T}(\mathcal{C})$ the feasible set of Problem 1, and we refer to a feasible solution $A(\mathcal{C}) := [\alpha_{ij}]$ as a *transportation matrix*. We denote by $\underline{A}(\mathcal{C}) := [\underline{\alpha}_{ij}]$ the *optimal* solution of Problem 1, and we denote by $\underline{l}(\mathcal{C})$ the cost of the optimal solution.

Lemma 4.1 (Lower bound on the cost of EBMP): Let $\mathcal{X}_n, \mathcal{Y}_n \sim$ ESCP$(n, \varphi_{\rm P}, \varphi_{\rm D})$, and let $Q_n = (\mathcal{X}_n, \mathcal{Y}_n)$. For any finite partition $\mathcal{C}$ of Ω, $\liminf_{n \to \infty} l_{\rm M}(Q_n) \geq \underline{l}(\mathcal{C})$ almost surely, where $\underline{l}(\mathcal{C})$ denotes the value of Problem 1.

Proof: Let σ denote the optimal bipartite matching of Q_n. For a particular partition $\mathcal{C}$, we define random variables $\hat{\alpha}_{ij} := \left|\left\{k : Y_k \in C^i, X_{\sigma(k)} \in C^j\right\}\right|/n$ for every pair (C^i, C^j) of cells; that is, $\hat{\alpha}_{ij}$ denotes the fraction of matches under σ whose $\mathcal{Y}$-endpoints are in C^i and whose $\mathcal{X}$-endpoints are in C^j. Let $\hat{\mathcal{T}}_n$ be the set of matrices with entries $\{\alpha_{ij} \geq 0\}_{i,j=1,\ldots,|\mathcal{C}|}$, such that $\sum_i \alpha_{ij} = \left|\mathcal{X}_n \cap C^j\right|/n$ for all $C^j \in \mathcal{C}$ and $\sum_j \alpha_{ij} = \left|\mathcal{Y}_n \cap C^i\right|/n$ for all $C^i \in \mathcal{C}$; note $\{\hat{\alpha}_{ij}\}$ itself is an element of $\hat{\mathcal{T}}_n$. Then the average match length $l_{\rm M}(Q_n)$ is bounded below by

$$
\begin{aligned}
l_{\rm M}(Q_n) = \frac{1}{n} \sum_{k=1}^{n} \left\|X_{\sigma(k)} - Y_k\right\| &\geq \sum_{ij} \hat{\alpha}_{ij} \min_{y \in C^i, x \in C^j} \|x - y\| \\
&\geq \min_{A \in \hat{\mathcal{T}}_n} \sum_{ij} \alpha_{ij} \min_{y \in C^i, x \in C^j} \|x - y\|.
\end{aligned}
$$

The key observation is that $\lim_{n\to\infty} \left|\left\{\mathcal{X}_n \cap C^j\right\}\right|/n = \varphi_{\rm P}(C^j)$, and $\lim_{n\to\infty} \left|\left\{\mathcal{Y}_n \cap C^i\right\}\right|/n = \varphi_{\rm D}(C^i)$, almost surely. Applying standard sensitivity analysis (see Chapter 5 of [13]), it can be shown that the final expression converges almost surely to $\underline{l}(\mathcal{C})$ as $n \to +\infty$; thus, we obtain the lemma. A version of this proof with a detailed sensitivity analysis appears in [14]. ■

We are interested in the tightest possible lower bound, and so we define $\underline{l} := \sup_{\mathcal{C}} \underline{l}(\mathcal{C})$. Remarkably, the supremum lower bound $\underline{l}$ is equivalent to the Wasserstein distance between $\varphi_{\rm D}$ and $\varphi_{\rm P}$, and so we can refine Lemma 4.1 as follows.

Lemma 4.2 (Best lower bound on the cost of EBMP): Let $\mathcal{X}_n, \mathcal{Y}_n \sim$ ESCP$(n, \varphi_{\rm P}, \varphi_{\rm D})$, and let $Q_n = (\mathcal{X}_n, \mathcal{Y}_n)$. Then

$$
\liminf_{n\to\infty} l_{\rm M}(Q_n) \geq W(\varphi_{\rm D}, \varphi_{\rm P}), \qquad \text{almost surely.} \tag{4}
$$

Proof: [Proof (Sketch)] The lemma is proved by showing that $\sup_{\mathcal{C}} \underline{l}(\mathcal{C}) = W(\varphi_{\rm D}, \varphi_{\rm P})$. By construction, Problem 1 is a discrete approximation and lower bound of (3); moreover, it can be shown that $\lim_{r\to+\infty} \underline{l}(\mathcal{C}_r) - W(\varphi_{\rm D}, \varphi_{\rm P}) \to 0^-$, where $\mathcal{C}_r$ is the grid partition of d^r cubes. Applying Lemma 4.1 to this sequence of partitions obtains the lemma. A complete proof of the lemma appears in [14], deriving both the approximation bound and the limit of the sequence. ■

Henceforth in the paper, we will abandon the notation $\underline{l}$ in favor of $W(\varphi_{\rm D}, \varphi_{\rm P})$ to denote this lower bound. This connection to the Wasserstein distance yields the following notable result.

Proposition 4.3: The lower bound $W(\varphi_{\rm D}, \varphi_{\rm P})$ of (4) is equal to zero if and only if $\varphi_{\rm D} = \varphi_{\rm P}$.

Proof: The proposition follows immediately from the fact that the Wasserstein distance is known to satisfy the axioms of a metric on $\Gamma(\varphi_{\rm D}, \varphi_{\rm P})$. ■

An intuitive explanation of the proposition is that if some fixed area $\mathcal{A}$ in the environment has unequal proportions of $\mathcal{X}$ points versus $\mathcal{Y}$ points, then a positive fraction of the matches associated with $\mathcal{A}$ (a positive fraction of all matches) must have endpoints *outside* of $\mathcal{A}$, i.e., at positive distance. Such an area can be identified whenever $\varphi_{\mathrm{P}} \neq \varphi_{\mathrm{D}}$.

The implication of Lemma 4.2 is that the average match length is asymptotically no less than some constant which depends *only* on the workspace geometry and the spatial distribution of pickup and delivery points; moreover, that constant is generally *non-zero*. We are now in a position to state the main result of this section.

Theorem 4.4 (Lower bound on the cost of ESCP): Let $L^*(n)$ be the length of the optimal stacker crane tour through $\mathcal{X}_n, \mathcal{Y}_n \sim \mathrm{ESCP}(n, \varphi_{\mathrm{P}}, \varphi_{\mathrm{D}})$, for compact $\Omega \in \mathbb{R}^d$, where $d \geq 2$. Then

$$\liminf_{n \to +\infty} L^*(n)/n \geq \mathbb{E}_{\varphi_{\mathrm{P}}\varphi_{\mathrm{D}}} \|Y - X\| + W(\varphi_{\mathrm{D}}, \varphi_{\mathrm{P}}), \quad (5)$$

almost surely.

Proof: A stacker crane tour is composed of pickup-to-delivery links and delivery-to-pickup links. The latter describe *some* bipartite matching having cost no less than the optimal cost for the EBMP. Thus, one can write $L^*(n)/n \geq \frac{1}{n}\sum_{i=1}^{n} \|Y_i - X_i\| + \frac{1}{n} L_{\mathrm{M}}(Q_n)$. The first term of the last expression goes to $\mathbb{E}_{\varphi_{\mathrm{P}}\varphi_{\mathrm{D}}} \|Y - X\|$ almost surely. By Lemma 4.2, the second term is bounded below asymptotically, almost surely, by $W(\varphi_{\mathrm{D}}, \varphi_{\mathrm{P}})$. ■

B. An Upper Bound on the Length of the ESCP

In this section we produce a sequence that bounds $L_{\mathrm{M}}(Q_n)$ asymptotically from above, and matches the linear scaling of (5). The bound relies on the performance of Algorithm 2, a *randomized* algorithm for the stochastic EBMP. The idea of Algorithm 2 is that each point $y \in \mathcal{Y}$ randomly generates an associated *shadow site* X', so that the collection $\mathcal{X}'$ of shadow sites "looks like" the set of actual pickup sites. An *optimal* matching is produced between $\mathcal{X}'$ and $\mathcal{X}$ which assists in the matching between $\mathcal{Y}$ and $\mathcal{X}$; specifically, if $x \in \mathcal{X}$ is the point matched to X', then the matching produced by Algorithm 2 contains (y, x). An illustrative diagram can be found in Figure 2.

Algorithm 2 is specifically designed to have two important properties for random sets Q_n: First, $\mathbb{E}\|X' - Y\|$ is predictably controlled by "tuning" inputs—a partition $\mathcal{C}$ of the environment and "policy matrix" $A(\mathcal{C})$—chosen as a function of n; second, $L_{\mathrm{M}}((\mathcal{X}', \mathcal{X}))/n \to 0^+$ as $n \to +\infty$. Later we will show that $\mathcal{C}$ and $A(\mathcal{C})$ can be chosen so that $\mathbb{E}\|X' - Y\| \to W(\varphi_{\mathrm{D}}, \varphi_{\mathrm{P}})$ (as $n \to +\infty$) leading to a bipartite matching algorithm whose performance matches the lower bound of (2).

Lemma 4.5 (Similarity of $\mathcal{X}'$ to $\mathcal{X}$): Let $X_1, \ldots, X_n$ be a set of points that are i.i.d. with density φ_{P}; let $Y_1, \ldots, Y_n$ be a set of points that are i.i.d. with density φ_{D}. Then Algorithm 2 generates shadow sites $X'_1, \ldots, X'_n$, which are (i) jointly independent of $X_1, \ldots, X_n$, and (ii) mutually i.i.d., with density φ_{P}.

Proof: Lemma 4.5 relies on basic laws of probability, and the proof is omitted in the interest of brevity. A complete proof of the lemma is provided in [14]. ■

The importance of this lemma is that it allows us to apply equation (1) of Section II-B to characterize $L_{\mathrm{M}}((\mathcal{X}', \mathcal{X}))$.

Algorithm 2 Randomized EBMP (parameterized)

Input: *pickup* points $\mathcal{X} = \{x_1, \ldots, x_n\}$, *delivery* points $\mathcal{Y} = \{y_1, \ldots, y_n\}$, probability densities $\varphi_{\mathrm{P}}(\cdot)$ and $\varphi_{\mathrm{D}}(\cdot)$, partition $\mathcal{C}$ of the workspace, and matrix $A(\mathcal{C}) \in \mathcal{T}(\mathcal{C})$.

Output: a bi-partite matching between $\mathcal{Y}$ and $\mathcal{X}$.

1: **initialize** $\mathcal{X}' \leftarrow \emptyset$.
2: **initialize** matchings $\overline{M} \leftarrow \emptyset$; $\hat{M} \leftarrow \emptyset$; $M \leftarrow \emptyset$.
3: // *generate "shadow pickups"*
4: **for** $y \in \mathcal{Y}$ **do**
5: Let C^i be the cell containing y.
6: Sample J; $J = j$ with probability $\alpha_{ij}/\varphi_{\mathrm{D}}(C^i)$.
7: Sample X' with pdf $\varphi_{\mathrm{P}}(\,\cdot\,| X' \in C^J)$.
8: Insert X' into $\mathcal{X}'$ and (y, X') into $\overline{M}$.
9: **end for**
10: $\hat{M} \leftarrow$ an optimal EBMP between $\mathcal{X}'$ and $\mathcal{X}$.
11: // *construct the matching*
12: **for** $X' \in \mathcal{X}'$ **do**
13: Let (y, X') and (X', x) be the matches in $\overline{M}$ and $\hat{M}$, respectively, whose $\mathcal{X}'$-endpoints are X'.
14: Insert (y, x) into M.
15: **end for**
16: **return** M

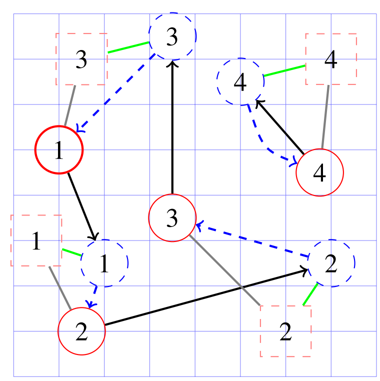


Fig. 2. Algorithm 2: Demands are labeled with integers. Pickup and delivery sites are represented by solid and dashed circles, respectively. Pickup-to-delivery links are shown as black arrows. Shadow pickups are shown as dashed squares, with undirected links to their generators (delivery sites); also shown are optimal matching links between shadows and pickups. Dashed arrows show the resulting induced matching. Note, this solution produces two disconnected subtours $(1, 2, 3)$ and (4).

Lemma 4.6 (Delivery-to-Pickup Lengths): Let Y be a random point with probability density φ_{D}; let X' be the *shadow site* of $y = Y$ generated by lines 5-7 of Algorithm 2, running with inputs $\mathcal{C}$ and $A(\mathcal{C})$. Then $\mathbb{E}\|X' - Y\| \leq \sum_{ij} \alpha_{ij} \max_{y \in C^i, x \in C^j} \|x - y\|$.

Proof: First, we observe that $\mathbb{E}\|X' - Y\| \leq \sum_{ij} \mathbb{P}[Y \in C^i, X' \in C^j] \max_{y \in C_i, x \in C_j} \|x - y\|$. By chain rule, we can write $\mathbb{P}[Y \in C^i, X' \in C^j] = \mathbb{P}[Y \in C^i] \times \mathbb{P}[X' \in C^j | Y \in C^i]$. Finally, noting that $\mathbb{P}[Y \in C^i] = \varphi_{\mathrm{D}}(C^i)$, and $\mathbb{P}[X' \in C^j | Y \in C^i] = \mathbb{P}[J = j | Y \in C^i] = \alpha_{ij}/\varphi_{\mathrm{D}}(C^i)$, we obtain the result. ■

Given a finite partition $\mathcal{C}$, it should be desirable to choose $A(\mathcal{C})$ in order to optimize the performance of Algorithm 2; that is, minimize the expected length of the matching produced. We can minimize at least the bound of Lemma 4.6 using the solution of Problem 2 (shown below); we denote the optimal solution $\overline{A}(\mathcal{C})$.

Problem 2 (Pessimistic "rebalancing"):

$$
\begin{aligned}
\underset{\{\alpha_{ij}\geq 0\}_{i,j\in\{1,\ldots,r^d\}}}{\text{Minimize}} \quad & \sum_{ij} \alpha_{ij} \max_{y\in C^i, x\in C^j} \|x-y\| \\
\text{subject to} \quad & \sum_{j} \alpha_{ij} = \varphi_{\mathrm{D}}(C^i) \quad \text{for all } C^i \in \mathcal{C}, \\
& \sum_{i} \alpha_{ij} = \varphi_{\mathrm{P}}(C^j) \quad \text{for all } C^j \in \mathcal{C}.
\end{aligned}
$$

Now we present Algorithm 3, described in pseudo-code, which computes[1] a specific partition $\mathcal{C}$, and then invokes Algorithm 2 with inputs $\mathcal{C}$ and $\overline{A}(\mathcal{C})$.

Algorithm 3 Randomized EBMP

Input: *pickup* points $\mathcal{X} = \{x_1, \ldots, x_n\}$, *delivery* points $\mathcal{Y} = \{y_1, \ldots, y_n\}$, and prob. densities $\varphi_{\mathrm{P}}(\cdot)$ and $\varphi_{\mathrm{D}}(\cdot)$.
Output: a bi-partite matching between $\mathcal{Y}$ and $\mathcal{X}$.
Require: an arbitrary *resolution* function $\mathrm{res}(n) \in \omega(n^{1/d})$, where d is the dimension of the space.

```
r ← res(n).
C ← grid partition C_r, of r^d cubes.
A ← Ā(C), the solution of Problem 2.
Run Algorithm 2 on (X, Y, φ_P, φ_D, C, A), producing matching M.
return M
```

Lemma 4.7 (Granularity of Algorithm 3): Let r be the resolution parameter, and $\mathcal{C}_r$ the resulting grid-based partition, used by Algorithm 3. Let Y be a random variable with probability density φ_{D}, and let X' be the *shadow site* of $y = Y$ generated by lines 5-7 of Algorithm 2, running under Algorithm 3. Then $\mathbb{E}\,\|X' - Y\| - W(\varphi_{\mathrm{D}}, \varphi_{\mathrm{P}}) \leq 2L\sqrt{d}/r$.

Proof: [Proof (Sketch)] Using Lemma 4.6, we can bound the difference $\mathbb{E}\,\|X' - Y\| - W(\varphi_{\mathrm{D}}, \varphi_{\mathrm{P}})$ by the difference between Problems 1 and 2. Problems 1 and 2 are discrete lower- and upper- approximations, respectively, of equation (3), which converge at a rate $2L\sqrt{d}/r$. A complete proof of the lemma appears in [14]. ■

We are now in a position to present an upper bound on the cost of the optimal EBMP matching that holds in the general case when $\varphi_{\mathrm{P}} \neq \varphi_{\mathrm{D}}$.

Lemma 4.8 (Upper bound on the cost of EBMP): Let $\mathcal{X}_n, \mathcal{Y}_n \sim \mathrm{ESCP}(n, \varphi_{\mathrm{P}}, \varphi_{\mathrm{D}})$, and let $Q_n = (\mathcal{X}_n, \mathcal{Y}_n)$. For $d \geq 3$,

$$\limsup_{n\to+\infty} \frac{L_{\mathrm{M}}(Q_n) - nW(\varphi_{\mathrm{D}}, \varphi_{\mathrm{P}})}{n^{1-1/d}} \leq \kappa(\varphi_{\mathrm{P}}, \varphi_{\mathrm{D}}), \tag{6}$$

almost surely, where

$$\kappa(\varphi_{\mathrm{P}}, \varphi_{\mathrm{D}}) := \min_{\phi\in\{\varphi_{\mathrm{P}},\varphi_{\mathrm{D}}\}} \left\{ \beta_{\mathrm{M},d} \int_{\Omega} \phi(x)^{1-1/d}\, dx \right\}. \tag{7}$$

For $d = 2$,

$$\frac{L_{\mathrm{M}}(Q_n) - nW(\varphi_{\mathrm{D}}, \varphi_{\mathrm{P}})}{\sqrt{n \log n}} \leq \gamma, \tag{8}$$

with high probability as $n \to +\infty$, for a positive constant γ.

Proof: We first focus on the case $d \geq 3$. The proof relies on the characterization of the length of the bipartite matching produced by Algorithm 3 (which also bounds the length of the *optimal* matching). By the triangle inequality, the length $\tilde{L}_{\mathrm{M}}(Q_n)$ of its matching is at most the sum of the matches between $\mathcal{X}$ and $\mathcal{X}'$, plus the distances from the sites in $\mathcal{Y}$ to their shadows in $\mathcal{X}'$, i.e.

$$\tilde{L}_{\mathrm{M}}(Q_n) \leq L_{\mathrm{M}}((\mathcal{X}', \mathcal{X})) + L_{\mathcal{Y}\mathcal{X}'}, \tag{9}$$

where $L_{\mathcal{Y}\mathcal{X}'} = \sum_{(Y,X')\in\overline{M}} \|X' - Y\|$. By subtracting on both sides of equation (9) the term $nW(\varphi_{\mathrm{D}}, \varphi_{\mathrm{P}})$, and dividing by $n^{1-1/d}$, we obtain

$$
\begin{aligned}
\frac{\tilde{L}_{\mathrm{M}}(Q_n) - nW(\varphi_{\mathrm{D}}, \varphi_{\mathrm{P}})}{n^{1-1/d}}
&\leq \frac{L_{\mathrm{M}}((\mathcal{X}', \mathcal{X}))}{n^{1-1/d}} + \frac{L_{\mathcal{Y}\mathcal{X}'} - nW(\varphi_{\mathrm{D}}, \varphi_{\mathrm{P}})}{n^{1-1/d}} \\
&= \frac{L_{\mathrm{M}}((\mathcal{X}', \mathcal{X}))}{n^{1-1/d}} + \frac{L_{\mathcal{Y}\mathcal{X}'} - n\,\mathbb{E}\,\|X' - Y\|}{n^{1-1/d}} \\
&\qquad + O\left(\frac{n^{1/d}}{r}\right),
\end{aligned}
$$

where the last equality follows from Lemma 4.7. Lemma 4.5 allows us to apply equation (1) to $L_{\mathrm{M}}((\mathcal{X}', \mathcal{X}))$, and so the limit of the first term is

$$\lim_{n\to+\infty} \frac{L_{\mathrm{M}}((\mathcal{X}', \mathcal{X}))}{n^{1-1/d}} = \beta_{\mathrm{M},d} \int_{\Omega} \varphi_{\mathrm{P}}(x)^{1-1/d}\, dx$$

almost surely. We observe that $n\,\mathbb{E}\,\|X' - Y\|$ is the expectation of $L_{\mathcal{Y}\mathcal{X}'}$, and so the second term goes to zero almost surely (absolute differences law, Section II-E). The resolution function of Algorithm 3 ensures that the third term vanishes. Collecting these results, we obtain the inequality in (6) with $\phi = \varphi_{\mathrm{P}}$. To complete the proof for the case $d \geq 3$, we observe that Algorithm 2 could be alternatively defined as follows: the points in $\mathcal{X}$ generate a set $\mathcal{Y}'$ of shadow sites; the intermediate matching is now between $\mathcal{Y}$ and $\mathcal{Y}'$. One can then prove results congruent with the results in Lemmas 4.5, 4.6, and 4.7. By following the same line of reasoning, one can finally prove the inequality in (6) with $\phi = \varphi_{\mathrm{D}}$. This concludes the proof for the case $d \geq 3$. The proof for the case $d = 2$ follows the same logic and is omitted in the interest of brevity. ■

We can leverage this result to derive the main result of this section, which is an asymptotic upper bound for the optimal cost of the ESCP. In addition to having the same linear scaling as our lower bound, the bound also includes "next-order" terms.

Theorem 4.9 (Upper bound on the cost of ESCP): Let $\mathcal{X}_n, \mathcal{Y}_n \sim \mathrm{ESCP}(n, \varphi_{\mathrm{P}}, \varphi_{\mathrm{D}})$ be a random instance of the ESCP, for compact $\Omega \in \mathbb{R}^d$, where $d \geq 2$. Let $L^*(n)$ be the length of the optimal stacker crane tour through $\mathcal{X}_n \cup \mathcal{Y}_n$. Then, for $d \geq 3$,

$$\limsup_{n\to+\infty} \frac{L^*(n) - n\Big[\mathbb{E}_{\varphi_{\mathrm{P}}\varphi_{\mathrm{D}}}\,\|Y - X\| + W(\varphi_{\mathrm{D}}, \varphi_{\mathrm{P}})\Big]}{\kappa(\varphi_{\mathrm{P}}, \varphi_{\mathrm{D}})\, n^{1-1/d}} \leq 1, \tag{10}$$

[1] In the definition of the algorithm, we use the "small omega" notation, where $f(\cdot) \in \omega(g(\cdot))$ implies $\lim_{n\to\infty} f(n)/g(n) = \infty$.

almost surely. For $d = 2$,

$$\frac{L^*(n) - n\Big[\mathbb{E}_{\varphi_P\varphi_D}\|Y - X\| + W(\varphi_D, \varphi_P)\Big]}{\gamma\sqrt{n\log n}} \leq 1, \quad (11)$$

with high probability as $n \to +\infty$, for a positive constant γ.

Proof: We first consider the case $d \geq 3$. Let $L_{\text{SPLICE}}(n)$ be the length of the SCP tour through $\mathcal{X}_n, \mathcal{Y}_n$ generated by SPLICE. Let $Q_n = (\mathcal{X}_n, \mathcal{Y}_n)$. One can write

$$\begin{aligned} L_{\text{SPLICE}}(n) &\leq \sum_{i=1}^{n}\|Y_i - X_i\| + L_{\text{M}}(Q_n) + \max_{x,y\in\Omega}\|x-y\|\,N_n \\ &= \left(\sum_{i=1}^{n}\|Y_i - X_i\| - n\mathbb{E}_{\varphi_P\varphi_D}\|Y-X\|\right) \\ &\quad + \Big(L_{\text{M}}(Q_n) - nW(\varphi_D,\varphi_P)\Big) \\ &\quad + n\Big[\mathbb{E}_{\varphi_P\varphi_D}\|Y-X\| + W(\varphi_D,\varphi_P)\Big] \\ &\quad + \max_{x,y\in\Omega}\|x-y\|\,N_n. \end{aligned}$$

The following results hold almost surely: The first term of the last expression is $o(n^{1-1/d})$ (absolute differences); by Lemma 4.8, the second term is $\kappa(\varphi_P, \varphi_D)n^{1-1/d} + o(n^{1-1/d})$; finally, by Remark 2.2, one has $\lim_{n\to+\infty} N_n/n^{1-1/d} = 0$. Collecting these results, dividing on both sides by $\kappa(\varphi_P, \varphi_D)n^{1-1/d}$, and noting that by definition $L^*(n) \leq L_{\text{SPLICE}}(n)$, one obtains the claim. The proof for the case $d = 2$ is almost identical and is omitted. ∎

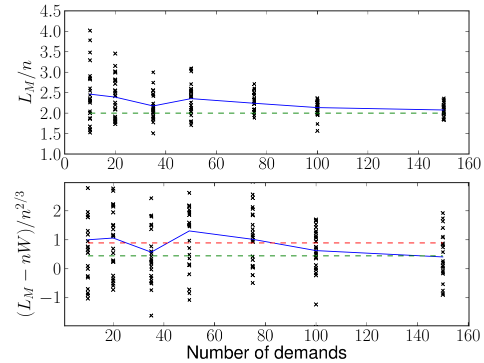
V. Simulation Results

In this section, we compare the observed scaling of the length of the EBMP, as a function of instance size, with what is predicted by equations (4) and (6). The implications for the scaling of the ESCP—equations (5) and (10)—are immediate. We focus on two examples of pickup/delivery distributions (φ_P, φ_D):

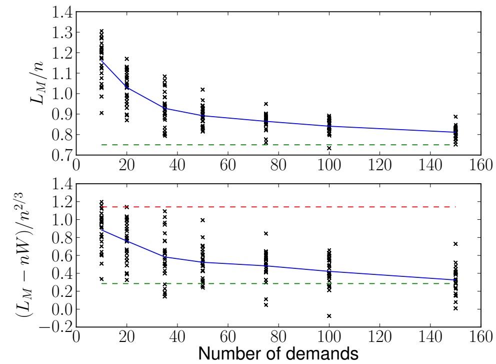
Case I—Unit Cube Arrangement: In the first case, the pickup site density φ_P places one-half of the probability uniformly over a unit cube centered along the x-axis at $x = -4$, and the other half uniformly over the unit cube centered at $x = -2$. The delivery site density φ_D places one-half of the probability uniformly over the cube at $x = -4$ and the other half over a new unit cube centered at $x = 2$.

Case II—Co-centric Sphere Arrangement: In the second case, pickup sites are uniformly distributed over a sphere of radius $R = 2$, and delivery sites are uniformly distributed over a sphere of radius $r = 1$; both spheres are centered at the origin.

Figure 3(a) (top) shows a scatter plot of $(n, L_{\text{M}}/n)$ with one point for each of twenty-five (25) *Case I* trials per size-category; that is, the x-axis denotes the size of the instance, and the y-axes denotes the average length of a match in the optimal matching solution. Additionally, the plot shows a curve (solid line) through the empirical mean in each size category, and a dashed line showing the Wasserstein distance between φ_D and φ_P, i.e. the predicted asymptotic limit to which the sequence should converge. Figure 3(b) (top) is analogous to Figure 3(a) (top), but for the *Case II* trials. Both plots exhibit the predicted approach of L_{M}/n to the constant $W(\varphi_D, \varphi_P) > 0$; the convergence in Figure 3(b) (top) appears slower because W is smaller. Figure 3(a) (bottom) shows a scatter plot of $\big(n, (L_{\text{M}} - W)/n^{2/3}\big)$ from the same data, with another solid curve through the empirical mean. Also shown are constants κ and $\tilde{\kappa}$ (dashed lines): κ is the asymptotic upper bound of equation (6); $\tilde{\kappa}$ is a smaller constant that results from bringing the $\min$ inside the integral in equation (7). Figures 3(b) (bottom) is again analogous to Figures 3(a) (bottom), and both plots indicate asymptotic convergence to a constant no larger than $\kappa(\varphi_D, \varphi_P)$. In fact, these cases give some credit to a developing conjecture of the authors: that the minimization in (7) can be moved inside the integral to provide a smaller (often much smaller) constant factor.



(a)



(b)

Fig. 3. Scatter plots of $(n, L_{\text{M}}/n)$ (top) and $\big(n, (L_{\text{M}} - W)/n^{2/3}\big)$ (bottom), with one point for each of twenty-five trials per size category. Figure 3(a) shows results for random samples under the distribution of *Case I*; Figure 3(b) shows results for random samples under the distribution of *Case II*.

VI. Stability Condition for DRT Systems

In this section we present a necessary and sufficient condition for the stability of DRT systems, modeled as DPDPs. Specifically, let us define the load factor as

$$\varrho := \lambda\,[\mathbb{E}_{\varphi_P\varphi_D}\|Y - X\| + W(\varphi_D, \varphi_P)]/m. \quad (12)$$

Note that when $\varphi_{\mathrm{D}} = \varphi_{\mathrm{P}}$, one has $W(\varphi_{\mathrm{D}}, \varphi_{\mathrm{P}}) = 0$ (by Lemma 4.2), and the above definition reduces to the definition of load factor given in [5] (valid for $d \geq 3$ and $\varphi_{\mathrm{D}} = \varphi_{\mathrm{P}}$).

The following theorem is the main result of the paper.

Theorem 6.1 (Stability condition for DRT systems): Consider the DPDP defined in Section III, which serves as a model of DRT systems. Then, the condition

$$\varrho = \lambda \left[\mathbb{E}_{\varphi_{\mathrm{P}}\varphi_{\mathrm{D}}} \|Y - X\| + W(\varphi_{\mathrm{D}}, \varphi_{\mathrm{P}})\right]/m < 1$$

is necessary and sufficient for the existence of stabilizing policies.

Proof: [Proof (Sketch)] Here we only present a sketch of the proof. A complete proof of the theorem appears in [14]. Let us first consider necessity. By Theorem 4.4 each demand requires, on average, a service time strictly greater than $\mathbb{E}_{\varphi_{\mathrm{P}}\varphi_{\mathrm{D}}} \|Y - X\| + W(\varphi_{\mathrm{D}}, \varphi_{\mathrm{P}})$. Since demands arrive, on average, at a rate of λ, and there are m vehicles, the fraction of time spent by the vehicles to provide service is, on average, $\lambda \left[\mathbb{E}_{\varphi_{\mathrm{P}}\varphi_{\mathrm{D}}} \|Y - X\| + W(\varphi_{\mathrm{D}}, \varphi_{\mathrm{P}})\right]/m = \varrho$. Hence, to have any hope of stability, such fraction has to be less than one, i.e., $\varrho < 1$.

Let us now consider sufficiency. The proof is constructive in the sense that a particular *gated* policy is stabilizing; the policy repeatedly applies algorithm SPLICE to determine tours through the outstanding demands, splits the tour into m equal length fragments, and assigns a fragment to each vehicle. The proof is based on a recursive relation bounding the expected number of demands in the system at the times when new tours are computed. The details of the proof can be found in [14]. ■

The stability condition in Theorem 6.1 only depends on the workspace geometry, the stochastic distributions of pickup and delivery points, customers' arrival rate, and the number of vehicles, and makes explicit the roles of the different parameters in affecting the performance of the overall system. We believe that this characterization would be instrumental for a system designer of DRT systems to build business and strategic planning models regarding, e.g., fleet sizing.

VII. Conclusion

In this paper we have derived a necessary and sufficient stability condition for demand-responsive transport systems, modeled as DPDPs with general and possibly different pickup and delivery distributions. Central to our approach has been the development of asymptotic bounds for the cost of the optimal solution to the stochastic ESCP. Our bounds prove that the ratio between the ESCP optimal cost and the number of demands goes to a constant almost surely (or with high probability in two dimensions), and that, perhaps surprisingly, this constant is generally *larger* than the expected pickup-to-delivery distance; in other words, the optimal tour is characterized by *non-vanishing* delivery-to-pickup links, whose length we have related to a known measure of transportation complexity, the Wasserstein distance. These bounds have allowed the derivation of a stability condition for our model of DRT systems that only depends on the workspace geometry, the spatial distributions of pickup and delivery points, customers' arrival rate, and the number of vehicles. Compared to traditional vehicle routing problems and spatially-distributed queueing systems, the stability condition presents an extra term due to the effect of non-vanishing delivery-to-pickup links.

This paper leaves numerous important extensions open for further research. First, in this paper we have restricted our analysis to stability conditions: in future work we plan to characterize *quality-of-service*, e.g., in terms of customer waiting time or time-to-service, and develop optimal (or near optimal) routing policies. Second, it would be of interest to address the even more general case where the pickup and delivery locations are statistically correlated. Finally, devising similar stability conditions and results for a *network model* of DRT systems (i.e., where a graph better describes the environment than Euclidean space) would be of both theoretical and practical importance.

Acknowledgments

The authors would like to thank Prof. Javed Aslam for suggesting the important connection between our lower bounds and the Wasserstain distance.

References

[1] UN. World urbanization prospects: The 2007 revision population database. Technical report, United Nations, 2007.

[2] G. Berbeglia, J. F. Cordeau, and G. Laporte. Dynamic pickup and delivery problems. *European Journal of Operational Research*, 202(1):8 – 15, 2010.

[3] S. N. Parragh, K. F. Doerner, and R. F. Hartl. A survey on pickup and delivery problems. *Journal fur Betriebswirtschaft*, 58(2):81 – 117, 2008.

[4] M. R. Swihart and J. D. Papastavrou. A stochastic and dynamic model for the single-vehicle pick-up and delivery problem. *European Journal of Operational Research*, 114(3):447–464, 1999.

[5] M. Pavone, K. Treleaven, and E. Frazzoli. Fundamental performance limits and efficient polices for Transportation-On-Demand systems. In *Proc. IEEE Conf. on Decision and Control*, pages 5622–5629, December 2010.

[6] V. Dobric and J. E. Yukich. Asymptotics for transportation cost in high dimensions. *Journal of Theoretical Probability*, 8(1):97–118, 1995.

[7] J. Houdayer, J.H. Boutet de Monvel, and O.C. Martin. Comparing mean field and euclidean matching problems. *The European Physical Journal B - Condensed Matter and Complex Systems*, 6(3):383–393, 1998.

[8] M. Talagrand. The Ajtai-Komlós-Tusnády matching theorem for general measures. In *Probability in Banach spaces*, volume 30. Birkhaüser Boston, Boston, MA, 1992.

[9] M. Ajtai, J. Komlós, and G. Tusnády. On optimal matchings. *Combinatorica*, 4(4):259–264, December 1984.

[10] L. Ruschendorf. The Wasserstein distance and approximation theorems. *Probability Theory and Related Fields*, 70:117–129, 1985. 10.1007/BF00532240.

[11] K. Treleaven, M. Pavone, and E. Frazzoli. An asymptotically optimal algorithm for pickup and delivery problems. In *Proc. IEEE Conf. on Decision and Control*, December 2011.

[12] L. E Baum and M. Katz. Convergence rates in the law of large numbers. *Transactions of the American Mathematical Society*, 120(1):108–123, 1965.

[13] D. Bertsimas and J. N. Tsitsiklis. *Introduction to Linear Optimization*. Athena Scientific, February 1997.

[14] K. Treleaven, M. Pavone, and E. Frazzoli. Asymptotically optimal algorithms for pickup and delivery problems with application to Large-Scale transportation systems. *arXiv:1202.1327*, February 2012.