Decentralized Algorithms for Stochastic and Dynamic Vehicle Routing with General Demand Distribution

Marco Pavone, Emilio Frazzoli, Francesco Bullo

Abstract—**We present decentralized algorithms for a class of stochastic and dynamic vehicle routing problems, known as the multiple-vehicle Dynamic Traveling Repairperson Problem (m-DTRP), in which demands arrive randomly over time and their locations have an arbitrary distribution, and the objective is to minimize the average waiting time between the appearance of a demand and the time it is visited by a vehicle. The best previously known control algorithms rely on *centralized, a-priori* task assignment, and are therefore of limited applicability in scenarios involving large ad-hoc networks of autonomous vehicles. By combining results from geometric probability and locational optimization, we provide a policy that solves, providing a constant-factor approximation to the optimal achievable performance, the decentralized version of the m-DTRP; such policy (i) does not rely on centralized and a priori task assignment, (ii) is spatially distributed, scalable to large networks, and adaptive to network changes. Simulation results are presented and discussed.**

I. INTRODUCTION

Dynamic vehicle routing problems occur when it is desired that one or more vehicles visit customers (demands) that arrive sequentially over a period of time. The objective is to schedule these visits in a way that aims at optimizing the quality of service, as perceived by the customers (e.g., minimize the wait for delivery/service) [1]. As a canonical example of an application with these characteristics, consider surveillance missions with Unmanned Aerial Vehicles (UAVs). In these missions, the UAVs must ensure continued coverage of a certain area; as events occur, i.e., as new targets are detected by on-board sensors or other assets (e.g., intelligence, high-altitude or orbiting platforms, etc.), UAVs must proceed to the location of the new event and provide close-range information. Variations of problems falling in this class have been studied in a number of papers in the recent past, e.g., [2]–[4]. In most of these papers, the problem is set up in such a way that the location of demands is fixed and known a priori; a strategy is computed that attempts to optimize the cost of servicing the known demands. In the present work, we wish to address the case in which new demands are generated continuously by a stochastic process: we will provide algorithms for minimizing the expected waiting time between the appearance of a demand and the time it is serviced by one of the vehicles.

In [5], restricting the attention to the case in which demands' locations are uniformly distributed in the workspace, we introduced and partially studied a decentralized receding horizon policy for the m-DTRP problem that is spatially distributed, scalable to large networks, and adaptive to network changes. The contribution of this paper is threefold. First, we focus on the single-agent case, and we characterize the behavior of the aforementioned receding horizon policy in the general setting in which demands' locations follow a general distribution. Second, using concepts from locational optimization, we derive algorithms that allow a team of agent to reach, in a decentralized way, an equitable partition of a convex set on the plane. Finally, combining the results of the first part of the paper with those of the second part, we provide a policy that solves, providing a constant-factor approximation to the optimal achievable performance, the decentralized version of the m-DTRP.

We believe that our decentralized algorithms for computing equitable partitions of planar convex sets could find application in many other coverage problems. Simulation results are presented and discussed.

II. PRELIMINARIES

In this section, we briefly describe some known concepts from probability and locational optimization, on which we will rely extensively later in the paper.

A. Variation of an Integral Function due to a Domain Change.

The following result is related to classic divergence theorems. Let $\Omega = \Omega(y) \subset \mathbb{R}^d$ be a region that depends smoothly on a real parameter $y \in \mathbb{R}$ and that has a well-defined boundary $\partial\Omega(y)$ for all y. Let h be a density function over $\mathbb{R}^d$. Then

$$\frac{d}{dy}\int_{\Omega(y)} h(x)\,dx = \int_{\partial\Omega(y)}\left(\frac{dx}{dy}\cdot n(x)\right)h(x)\,dx, \quad (1)$$

where $v\cdot w$ denotes the scalar product between vectors v and w, n is the unit outward normal to $\partial\Omega(y)$, and dx/dy denotes the derivative of the boundary points with respect to y.

B. Voronoi Diagrams

Let $G = (g_1,\cdots,g_m)$ be a vector of m points in a metric space Q. Let the Voronoi region $V_i = V_i(G)$ be the set of all points $q \in Q$ such that $\|q - g_i\| \leqslant \|q - g_j\|$ for all $i \neq j$. If Q is a finite-dimensional Euclidean space, the boundary of each V_i is a *convex polygon*. The set of regions $\{V_1,\cdots,V_m\}$ is called the Voronoi diagram for the generators $(g_1,\cdots,g_m)$. When the two Voronoi regions V_i and V_j are adjacent (i.e., they share an edge), g_i is called a *Voronoi neighbor* of g_j (and

vice-versa). We also define the (i, j)-face as $\Delta_{ij} = V_i \bigcap V_j$. We will shortly refer to the vertices of a face Δ_{ij} as the set $\{u_l\}$, without any additional subscript: the hidden subscripts will be clear from the context.

Voronoi diagrams enjoy the Perpendicular Bisector Property: for any set of $\{\lambda_l\}, \lambda_l \geqslant 0$ with $\sum_l \lambda_l = 1$, such that $\sum_l \lambda_l u_l \in \Delta_{ij}$, we have

$$\left(\sum_l \lambda_l u_l - \frac{g_i + g_j}{2}\right) \cdot (g_j - g_i) = 0, \qquad (2)$$

in words, the face Δ_{ij} bisects the line segment joining g_j and g_i and that line segment is perpendicular to the face.

We introduce the following notation (for Voronoi regions in $\mathbb{R}^2$): $O_{ij}^g = (g_j + g_i)/2$; $O_{ij}^u = (u_1 + u_2)/2$; $\gamma_{ij} = \|g_j - g_i\|$; $\delta_{ij} = \|u_2 - u_1\|$.

Finally, we define a Voronoi diagram equitable with respect to an integrable density function $g : Q \to \mathbb{R}_+$, if the integral weighted by g over a Voronoi cell is identical for all cells.

C. Asymptotic and Worst-Case Properties of the Traveling Salesperson Problem in the Euclidean Plane

The Euclidean Traveling Salesperson Problem (TSP) is formulated as follows: given a set D of n points in R^d, find the minimum-length tour (i.e., closed path) through all points in D. Let $\text{TSP}(D)$ denote the minimum length of a tour through all the points in D; by convention, $\text{TSP}(\emptyset) = 0$. Assume that the locations of the n points are random variables independently and identically distributed in a compact set Q; in [6] it is shown that there exists a constant $\beta_{\text{TSP},d}$ such that, almost surely,

$$\lim_{n \to +\infty} \frac{\text{TSP}(D)}{n^{1-1/d}} = \beta_{\text{TSP},d} \int_Q \bar{f}(q)^{1-1/d} \, dq , \qquad (3)$$

where $\bar{f}$ is the density of the absolutely continuous part of the distribution of the points; the current best estimate of the constant in Eq. (3) for the case $d = 2$ is $\beta_{\text{TSP},2} \simeq 0.7120$ [7]. According to [8], if Q is a "fairly compact and fairly convex" set in the plane, then Eq. (3) provides an adequate estimate of the optimal TSP tour length for values of n as low as 15.

III. NOTATION AND PROBLEM FORMULATION

The basic version of the problem we wish to study in this paper is known as the Dynamic Traveling Repairperson Problem (DTRP), and was studied by Bertsimas and van Ryzin in [9]; in the basic setting, a single vehicle and a uniform demand distribution are considered. The same authors later extended the problem to the multiple vehicle case [10], and to the general demand and interarrival time distribution case [1]. In this section, we define the DTRP problem and its necessary elements.

Let the workspace $Q \subset \mathbb{R}^d$ be a convex polygon, and let $\|\cdot\|$ denote the Euclidean norm in $\mathbb{R}^d$. For simplicity, we will consider the planar case, i.e., $d = 2$, with the understanding that extensions to higher dimensions are possible; we will use the shorthand $\beta = \beta_{\text{TSP},2}$. Consider m holonomic vehicles, modeled as point masses. The vehicles are free to move,

with bounded velocity, within the environment Q; without loss of generality, we will assume that the maximum velocity magnitude is unitary. The vehicles are identical, and have unlimited fuel and demand servicing capacity.

Demands are generated according to a homogeneous (i.e., time-invariant) spatio-temporal Poisson process, with time intensity λ, and spatial density $f(x)$ supported on Q. Demand locations $\{X_j; \ j \geqslant 1\}$ are i.i.d. (i.e., *independent* and *identically distributed*) and distributed according to density f. They become known (are realized) at a demand's arrival epoch. Thus, at time t we know with *certainty* the locations of demands that arrived prior to time t, but future demand locations form an i.i.d. sequence. The notation $f(x)$ is short for $f(x_1, x_2)$. The density $f(x)$ satisfies:

$$\mathbb{P}[X_j \in S] = \int_S f(x) \, dx \quad \forall S \subseteq Q, \quad \text{and} \int_Q f(x) \, dx = 1.$$

We assume that the number of initial outstanding demands is a random variable with finite first moment.

A demand is considered serviced when one of the agents reaches its location. For simplicity, agents are not required to stop or to loiter in proximity of demands: the extension to the case with additional on-site servicing time is straightforward, but the notation is more cumbersome. Information on outstanding demands at time t is summarized as a finite set of demand positions $D(t)$. Demands are inserted in D as soon as they are generated, while they are removed from D upon servicing.

Denote by T^j the time elapsed from the time the jth demand is generated to the time it is serviced. The steady-state system time is defined by $T \triangleq \lim_{j \to \infty} \mathbb{E}[T^j]$. The objective of the m-DTRP is to find a causal policy for servicing demands that minimizes T; the infimum of the system time over all causal stabilizing policies is denoted by T^*.

Finally, we need two definitions; in these definitions, X is the location of a randomly chosen demand and T is its waiting time.

Definition 3.1: A policy π is called *spatially unbiased* if: $\mathbb{E}[T|X \in S] = \mathbb{E}[T], \quad \forall S \subseteq Q$; a policy π is called *spatially biased* if there exist sets $S_1, S_2 \subseteq Q$ such that: $\mathbb{E}[T|X \in S_1] > \mathbb{E}[T|X \in S_2]$.

In the following, we are interested in designing control policies that provide constant-factor approximations of the optimal achievable performance. A policy π, yielding a steady-state system time T_π, is said to provide a constant-factor approximation if there exist a factor $\theta > 1$ such that $T_\pi \leq \theta T^*$. Moreover, we are interested in decentralized, scalable, adaptive control policies, that rely only on the local exchange of information between neighboring vehicles, and do not require explicit knowledge of the global structure of the network.

IV. EXISTING RESULTS FOR THE m-DTRP

As mentioned before, the main reference on dynamic vehicle routing problems to date is the work of Bertsimas and van Ryzin [1], [9], [10]. In these works, lower bounds for the optimal system time are derived in the light load case

(i.e., $\lambda \to 0^+$), and in the heavy load case (i.e., $\lambda \to +\infty$). Subsequently, policies are designed for the two cases, and their performance is compared to the lower bounds. These results are obtained through techniques drawn from combinatorial optimization, queueing theory, and geometric probability.

A. Lower Bounds on the Optimal System Time in Light and Heavy Load

For the light-load case, a tight lower bound on the system time is derived in [10]. This lower bound is strongly related to the optimization of the continuous m-median function (or Weber function) H_m. (See [11] and references therein for an introduction to the optimization of the Weber function.) Let H_m^* be the global minimum of H_m, then

$$T^* \geq H_m^* \triangleq \min_{P \in Q^m} \int_Q f(x) \left(\min_{i \in \{1, \cdots, m\}} \|p_i - x\| \right) \, dx, \quad (4)$$

where $P = (p_1, \cdots, p_m)$ is a vector of m distinct points in Q. This bound actually holds in every load condition, but becomes tight as $\lambda \to 0^+$. Notice that, for $m > 1$, the optimization of H_m is a non-convex optimization problem.

In heavy load, we have the following lower bound [1]

$$\lim_{\lambda \to +\infty} \frac{T^*}{\lambda} \geqslant \frac{\kappa^2}{m^2} \left[\int_Q f^\alpha(x) dx \right]^{\frac{1}{1-\alpha}}, \quad (5)$$

where $\kappa \geqslant \frac{2}{3\sqrt{\pi}}$, and $\alpha = 1/2$ within the class of spatially unbiased policies, while $\alpha = 2/3$ within the class of spatially biased policies. It is shown in [1] that the lower bound for biased policies (i.e., with $\alpha = 2/3$) is always lower than or equal to the lower bound for unbiased policies ($\alpha = 1/2$) for all densities f. These bounds are not known to be tight; moreover, they differ from the previous light load bound in that they are asymptotic bounds (i.e., valid for $\lambda \to +\infty$). It is possible to show that allowing biased service will result in a strict reduction of the optimal mean system time for any nonuniform density f.

B. Available Policies for the Light and Heavy Load Case

For the light-load case, an optimal policy, presented in [10], is the "m-Stochastic Queue Median Policy." This policy, although optimal, has several drawbacks: (i) it does not stabilize the system as the load increases, (ii) it relies on the solution of the continuous m-median problem that is *NP-hard* for $m > 1$, (iii) it relies on *a priori* computation and, thus, is not adaptive to changes in the environment. Moreover, it requires a *centralized* assignment for the agent locations.

For the heavy-load case, optimal policies are not known. In [1], provably good policies, both in the unbiased and biased classes, are provided. Although these policies are provably good, they share two major requirements, namely: (i) a centralized data structure, (ii) intensive off-line computation.

V. THE SINGLE-AGENT RECEDING HORIZON POLICY

As discussed in the above section, the policies presented in [1], [10] are not well suited for the decentralized version of the m-DTRP. The following *conjecture*, presented in [1], suggests a strategy for the task of finding a provably good decentralized policy: let μ^* denote an optimal single-agent DTRP policy in heavy load; then, dividing the region Q into m subregions of equal size, assigning one vehicle to each subregion, and having vehicles follow the single-agent policy μ^* in each subregion is an optimal m-agent policy in heavy load. With this conjecture in mind, we consider the following strategy: (i) we first design, for the single agent case, a policy that does not require off-line computation; (ii) we then extend the single agent policy to the multi agent case via the concept of *virtual generators*: the virtual generators are a means to achieve in a decentralized way an equitable partition of Q.

For the single agent case, we propose the following receding horizon policy (a similar version was first proposed by the authors in [5] in the setting of uniform demand distribution). In what follows, given a tour T of D, a fragment of T is a connected subset of T. Moreover, let P_1^* be the median of Q (i.e., $P_1^* = \operatorname{argmin} H_1$), and p be the current agent's position.

Single-Vehicle Receding Horizon (sRH) policy —
Given $\eta \in (0, 1]$:
- while the set of demands D is empty, move at unit speed toward P_1^* if $p \neq P_1^*$, otherwise stop;
- while D is not empty, do the following: (i) compute the TSP tour through all demands in D, (ii) find the fragment of length $\eta \operatorname{TSP}(D)$ that maximizes the number of reached demands, (iii) service the optimal fragment starting from the endpoint closest to the current position (if the entire tour is selected, start from the closest demand). Repeat.

In other words, if $D \neq \emptyset$, the vehicle looks for a maximum-reward TSP fragment.

A. Analysis of the sRH

We first consider the light-load case (i.e., $\lambda \to 0^+$) and present a result whose proof follows the treatment in [5].

Theorem 5.1: The sRH policy is asymptotically optimal in the light-load case, that is, $T_{\text{sRH}} \to T^*$ as $\lambda \to 0^+$.

The following result characterizes the behavior of the sRH policy in heavy load. Its derivation is based on ideas outlined in [5]. For lack of space, its proof is omitted and will be added in a forthcoming paper.

Theorem 5.2: An upper bound on the cost of the sRH policy in *heavy load* is:

$$\lim_{\lambda \to \infty} \frac{T_{\text{sRH}}}{\lambda} \leqslant \frac{\beta^2}{2 - \eta} \left(\int_Q f^{1/2}(x) dx \right)^2. \quad (6)$$

Comparing this upper bound with the lower bound in Eq. (5) with $\alpha = 1/2$, we obtain:

$$T_{\text{sRH}} \leq \frac{\beta_{\text{TSP},2}^2}{(2 - \eta)\kappa^2} T^*, \qquad \text{as} \quad \lambda \to +\infty.$$

As $\eta \to 0^+$ we obtain the same factor (i.e., 80%) as the best known unbiased policy [1]. We conclude that, in heavy load, the performance of the sRH policy has the same performance guarantee as the best known unbiased policy.

VI. Simulation results for the Heavy Load Case

In this section, through simulations, we check the correctness of the upper bound for the sRH policy in heavy-load and we investigate in-depth its *distributional* behavior. The algorithms described in the paper were implemented in Matlab®6.1, with external calls to the program `linkern` (freely available for academic research use at www.math.princeton.edu/tsp/concorde.html).

Roughly speaking, the sRH policy is increasingly more biased as the horizon η decreases. Thus, we expect the upper bound in Eq. (6) (showing a dependency on f peculiar to unbiased policies) to be increasingly less tight as η decreases.

In the simulation experiments, we consider a square region $\mathcal{A}_2$ with area $1 - \varepsilon$, and inside $\mathcal{A}_2$ a significantly smaller square region $\mathcal{A}_1$ with area ε. Region $\mathcal{A}_1$ will represent a *peak* of demand concentration. Within each region demands are uniformly distributed. Points fall in region $\mathcal{A}_1$ with probability $1 - \delta$ and in region $\mathcal{A}_2$ with probability δ. Thus, the density is piecewise uniform with:

$$ f(x) = \begin{cases} \frac{1-\delta}{\varepsilon} & x \in \mathcal{A}_1 \\ \frac{\delta}{1-\varepsilon} & x \in \mathcal{A}_2 \end{cases} $$

In all simulations, we have considered, as the initial condition, a number of outstanding demands 20% higher than the steady state expected value.

The upper bound for the heavy load case has been successfully tested for various values of δ and λ. In all simulations, we set $\varepsilon = 10^{-4}$. Clearly, the greater λ, the more accurate the bound is. The simulation results reported in Fig. 1 refer to the case $\delta = 0.19$ and $\lambda = 200$. As suggested by the dependence
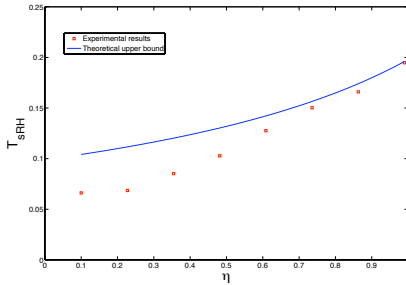

Fig. 1. Dependence of the system time over η ($\delta = 0.19$).

of the upper bound over η, the system time decreases as η is decreased. Moreover, the upper bound is less and less tight as η decreases.

To estimate the actual dependence of the system time of the the sRH policy on $f(x)$ (i.e., the distributional behavior), we use the approach outlined in [1]. We consider the following representation of the dependence of the system time on the density $f(x)$ for a particular policy μ:

$$ T_\mu = \gamma_\mu \Xi(\alpha), \quad \Xi(\alpha) = \left(\int_Q f^\alpha(x) dx \right)^{\frac{1}{1-\alpha}}, \quad \gamma_\mu \text{ constant.} $$

As seen before, in the unbiased case $\alpha = 1/2$ and in the optimal biased case $\alpha = 2/3$. For the particular density $f(x)$ given before and for ε small it is possible to show that: $\log(T_\mu) = \frac{\alpha}{1-\alpha} \log(\delta) + c_\mu$, where c_μ depends on the policy and the system parameters. Thus, by plotting $\log(T_\mu)$ (or $\log(N_\mu)$) against $\log(\delta)$ and performing a linear regression, one can estimate α and hence the distributional dependence of the policy μ. Note that since $\log(\cdot)$ is increasing and $\log(\delta) < 0$, higher values of α imply lower system times.

We set $\varepsilon = 10^{-4}$ and $\lambda = 200$. Then, a different simulation run is performed for eleven values of δ in the range 0.05 to 0.9999 (this last value corresponds to a uniform demand distribution) and eight values of η in the range 0.01 to 1. The observed system time is then recorded for each δ and η.

We would expect a value of α equal to 0.5 for $\eta = 1$ (when all the TSP is serviced at each epoch); moreover, we would expect a lower system time as η is decreased, since the policy becomes more reactive. Fig. 2 shows the functional dependence of α over η. As expected, when $\eta = 1$, $\alpha \simeq 0.5$. Moreover, α increases as η is decreased; therefore, we conclude that decreasing η improves *both* the scaling factor in the system time (i.e., the factor $\frac{1}{2-\eta}$), *and* its distributional behavior.
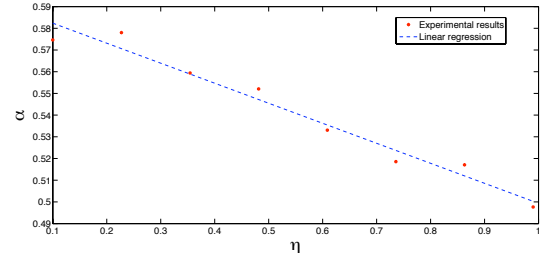

Fig. 2. Functional dependence of α over η.

VII. The Multi-Agent Receding Horizon Policy

Here we design a decentralized policy for the DTRP problem applicable to multiple-vehicle systems. The basic procedure to extend the single-vehicle policy to the multiple-vehicle case consists in partitioning Q into *regions of dominance* and in letting each agent to execute the sRH policy within *its own* region of dominance. In order to achieve an efficient partition of Q in a decentralized way, we introduce certain points, that we call *virtual generators*, associated to each agent; we denote the virtual generators as $G = (g_1, g_2, \ldots, g_m) \in \mathbb{R}^{2m}$. We denote the ith Voronoi cell in the Voronoi partition induced by G as $V_i \triangleq V_i(G)$. We define the region of dominance for agent i as the convex polygon $D_i = V_i \bigcap Q$; we call the partition into regions of dominance induced by G as $D(G)$. Notice that the virtual generators are not physical points, but are rather artificial variables, in general not coinciding with the vehicles' positions.

We shall assume that each vehicle has sufficient information available to determine: (1) its Voronoi cell, and (2) the locations of all outstanding events in its region of dominance.

A control policy that relies on information (1) and (2) is *Voronoi-distributed*, in the sense that the behavior of each vehicle depends only on the location of the other agents with contiguous Voronoi cells. In [12, Section 2.3], it is shown that the number of Voronoi neighbors of each generator is on average less than or equal to 6. Accordingly, Voronoi-distributed policies are *spatially distributed* and *scalable* in the number of agents. A spatially distributed algorithm for the local computation and maintenance of Voronoi cells is provided in [13].

The key idea is to enable the virtual generators to move along the gradient of a suitable locational optimization function in such a way that an effective partition $D(G^*)$ is achieved. Following the arguments provided in [14], such partition should guarantee the same workload for each agent. Here, we restrict our attention to the heavy load case; the light load case is significantly simpler and has been already discussed in [5]. In the following, we first find a suitable locational optimization function and then we compute its gradient.

A. The Relevant Locational Optimization Function

As discussed before, each agent executes the sRH policy in its own region of dominance D_i. We want now to find an upper bound on the system time when a particular partition $D(G)$ is adopted. The conditional demand density in any given region D_i is $f(x)/\int_{D_i} f(x)\,dx$. Moreover, the arrival rate in each region D_i is $\lambda \int_{D_i} f(x)\,dx$. Therefore, the system time for a demand X that lands in region D_i is asymptotically upper bounded by: $\mathbb{E}\left(T|X \subseteq D_i\right) \leqslant \lambda\beta^2 \left(\int_{D_i} f^{1/2}(x)dx\right)^2 \big/ (2 - \eta)$. Since the probability that a demand lands in the region D_i is $\int_{D_i} f(x)\,dx$, an asymptotic upper bound for the system time in the multi agent case is:

$$T \leqslant \frac{\lambda\beta^2}{2 - \eta} \sum_{i=1}^{m} \left(\int_{D_i} f(x)\,dx\right)\left(\int_{D_i} f^{1/2}(x)\,dx\right)^2 \leqslant$$
$$\frac{\lambda\beta^2}{2 - \eta} \sum_{i=1}^{m} \left(\int_{V_i} f^{1/2}(x)\,dx\right)^2 \triangleq \frac{\lambda\beta^2}{2 - \eta} L_m(G). \quad (7)$$

In the second inequality, recall that the support of f is Q. As it will be clearer later, $L_m(G)$ is an appropriate locational optimization function; it is fairly easy to show that $L_m(G)$ is non-convex.

B. The Differential of the Locational Optimization Function

The differential of the locational optimization function L_m is presented in the following theorem. We point out that this gradient can be computed in a *distributed* way, since it depends only on the location of the other agents with contiguous Voronoi cells. In the following, let be $A_k \triangleq \int_{V_k} f^{1/2}(x)\,dx \quad k = 1, 2, \cdots, m$.

Theorem 7.1: The partial derivative of the locational opti-

mization function is:

$$\frac{\partial L_m(G)}{\partial g_i} = 2 \sum_{j \in N(i)} \delta_{ij}\left(A_i - A_j\right)\left[\frac{1}{2}\int_0^1 f^{1/2}(x(\lambda))\,d\lambda\,n_{ij} + \right.$$
$$\left. + \frac{1}{\gamma_{ij}}\int_0^1 \left(f^{1/2}(x(\lambda))\,(x(\lambda) - O_{ij}^g)\right)d\lambda\right], \quad (8)$$

Proof: Since the motion of a generator g_i affects the Voronoi region V_i and its neighboring regions V_j for $j \in \{j_1, \cdots, j_{k_i}\} \triangleq N(i)$, we have

$$\frac{\partial L_m(G)}{\partial g_i} = 2A_i\frac{\partial A_i}{\partial g_i} + 2 \sum_{j \in N(i)} A_j\frac{\partial A_j}{\partial g_i} \quad (9)$$

Now, the result in Eq. (1) provides the means to analyze the variation of an integral function due to a domain change. Since the boundary of V_i satisfies $\partial V_i = \cup_j \Delta_{ij}$, where $\Delta_{ij} = \Delta_{ji}$ is the edge between V_i and V_j, we have

$$\frac{\partial}{\partial g_i^k}\int_{V_i} f^{1/2}(x)\,dx = \sum_{j \in N(i)} \int_{\Delta_{ij}} \left(\frac{dx}{dg_i^k}\cdot n_{ij}(x)\right)f^{1/2}(x)\,dx$$
$$\frac{\partial}{\partial g_i^k}\int_{V_j} f^{1/2}(x)\,dx = \int_{\Delta_{ij}} \left(\frac{dx}{dg_i^k}\cdot n_{ji}(x)\right)f^{1/2}(x)\,dx, \quad (10)$$

where g_i^k is the kth coordinate of generator g_j ($k \in \{1, 2\}$) and we define n_{ij} as the unit normal to Δ_{ij} outward of V_i (therefore we have $n_{ji} = -n_{ij}$).

The first step to compute the derivative of the boundary points in Eq. (10) is to introduce a parametrization for the boundary segment Δ_{ij}. Let u_1 and u_2 (as discussed in Section II we avoid any further subscript) be the endpoints of segment Δ_{ij}. Recall that the workspace is assumed to be bounded, therefore both vertices are finite. We introduce the natural parametrization: $x(\lambda) = u_1\lambda + u_2(1 - \lambda)$, with $\lambda \in [0, 1]$. Let δ_{ij} be the length of border Δ_{ij}. Clearly, we have that $\|x'(\lambda)\| = \|u_1 - u_2\| = \delta_{ij}$.

To evaluate the scalar product between the boundary points and the unit outward normal to the border in Eq. (10), we differentiate the Perpendicular Bisector Property (Eq. (2)) with respect to g_i^k for any point $x(\lambda) \in \Delta_{ij}$; we get

$$\frac{\partial x(\lambda)}{\partial g_i^k}\cdot(g_j - g_i) = \frac{1}{2}e^k\cdot(g_j - g_i) + e^k\cdot\left(x(\lambda) - \frac{g_i + g_j}{2}\right), \quad (11)$$

where e^k is the kth canonical vector ($k \in \{1, 2\}$) in $\mathbb{R}^2$. By the Perpendicular Bisector Property, we have that $n_{ij} = (g_j - g_i)/\|g_j - g_i\|$ and the desired explicit expressions for the scalar products in Eq. (10) follow immediately (recalling that $n_{ji} = -n_{ij}$). In what follows, let γ_{ij}, O_{ij}^g and O_{ij}^u be defined as in Section II. By using the definition of line integral, it is now easy to show (in compact form):

$$\frac{\partial}{\partial g_i}\int_{V_i} f^{1/2}(x)\,dx = \sum_{j \in N(i)} \left[\frac{\delta_{ij}}{2}\int_0^1 f^{1/2}(x(\lambda))\,d\lambda\,n_{ij} + \right.$$
$$\left. + \frac{\delta_{ij}}{\gamma_{ij}}\int_0^1 \left(f^{1/2}(x(\lambda))\,(x(\lambda) - O_{ij}^g)\right)d\lambda\right],$$

$$\frac{\partial}{\partial g_i} \int_{V_j} f^{1/2}(x)\,dx = -\frac{\delta_{ij}}{2} \int_0^1 f^{1/2}(x(\lambda))\,d\lambda\, n_{ij} +$$

$$-\frac{\delta_{ij}}{\gamma_{ij}} \int_0^1 \left(f^{1/2}(x(\lambda))\,(x(\lambda) - O_{ij}^g) \right) d\lambda.$$

Collecting these results, we get the theorem. ∎

The expression in Eq. (8) simplifies considerably when the demands follow a uniform distribution (i.e., $f(x) = 1/A$, where A is the area of the workspace). In such case

$$\frac{\partial L_m(G)}{\partial g_i} = 2 \sum_{j \in N(i)} \frac{\delta_{ij}}{A^{1/2}} (A_i - A_j) \left[\frac{1}{2} n_{ij} + \frac{1}{\gamma_{ij}} (O_{ij}^u - O_{ij}^g) \right],$$
(12)

Assuming that there exists a Voronoi diagram equitable with respect to $f^{1/2}(x)$ (henceforth we assume this condition holds, unless otherwise stated), from the expression of the gradient of $L_m(G)$ we see that one of the critical point of $L_m(G)$ will satisfy $A_i = A_j \quad \forall i, j$. This point correspond to the global minimum of $L_m(G)$, as it can be easily verified by Lagrange multiplier arguments.

C. A Gradient Descent Law for the Virtual Generators

Assume the virtual generators obey a first order dynamical behavior described by $\dot{g}_i = -\frac{\partial L_m(G)}{\partial g_i}$. For the closed-loop system induced by such gradient descent law, the virtual generators converge asymptotically to the set of critical points of $L_m(G)$. If the the virtual generators converge to the global minimum of $L_m(G)$, we have $\int_{V_i} f^{1/2}(x)\,dx = \int_Q f^{1/2}(x)\,dx/m$. Therefore, the system time (Eq. (7)) reduces to

$$T \leqslant \frac{\lambda \beta^2 \left(\int_Q f^{1/2}(x)dx \right)^2}{m^2(2 - \eta)}.$$
(13)

Some remarks are in order. First, in this section we have postulated that there exists a Voronoi diagram equitable with respect to the function $f^{1/2}(x)$. Although this assumption is reasonable in most cases, to date it is not possible to state that for every density function $g(x)$ there always exists a corresponding equitable Voronoi diagram. Second, the proposed gradient descent law is clearly not guaranteed to find a global minimum of $L_m(G)$. Therefore, the result in Eq. (13) holds only *locally*. On the other hand, local optimality is the common price to pay in change of decentralization.

In Fig. (3) we test the dependence of the system time over m. We consider $\lambda = 130$, a uniform demand distribution and 5 values for m, namely $m = 2, 4, 6, 8, 10$. We notice that the system time decreases quadratically in m, as expected. Remember that the upper bound on the system time holds asymptotically in λ.

VIII. Conclusion

In this paper we discuss decentralized algorithms for vehicle routing in a stochastic time-varying environment with general demand distribution. Our control policy is spatially distributed in the sense that the behavior of each vehicle depends only
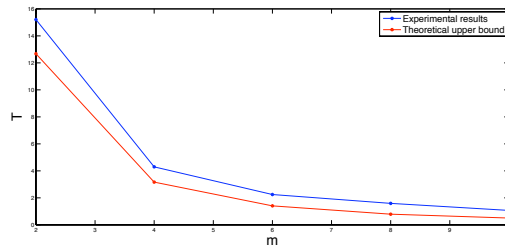

Fig. 3. Dependence of the system time over m.

on the location of all other vehicles with contiguous Voronoi cells. We conclude by mentioning some limitations of our approach. We considered omni-directional vehicles with first order dynamics: non-holonomic constraints will have to be taken into account for application to UAVs or other systems. In this paper, all demands are removed from the demand queue only upon service; in some scenarios, demands might disappear before being serviced. These issues are under current study; first results for the latter problem can be found in [15].

References

[1] D. J. Bertsimas and G. J. van Ryzin. Stochastic and dynamic vehicle routing in the Euclidean plane with multiple capacitated vehicles. *Advances in Applied Probability*, 25(4):947–978, 1993.

[2] P. Chandler, S. Rasmussen, and M. Pachter. UAV cooperative path planning. In *Proc. AIAA Conf. on Guidance, Navigation, and Control*, 2000.

[3] R. W. Beard, T. W. McLain, M. A. Goodrich, and E. P. Anderson. Coordinated target assignment and intercept for unmanned air vehicles. *IEEE Trans. Robotics and Automation*, 18(6):911–922, 2002.

[4] W. Li and C. Cassandras. A cooperative receding horizon controller for multivehicle uncertain environments. *IEEE Trans. on Automatic Control*, 51(2):242–257, 2006.

[5] E. Frazzoli and F. Bullo. Decentralized algorithms for vehicle routing in a stochastic time-varying environment. In *Proc. IEEE Conf. on Decision and Control*, pages 3357–3363, 2004.

[6] J. M. Steele. Probabilistic and worst case analyses of classical problems of combinatorial optimization in Euclidean space. *Mathematics of Operations Research*, 15(4):749, 1990.

[7] G. Percus and O. C. Martin. Finite size and dimensional dependence of the Euclidean traveling salesman problem. *Physical Review Letters*, 76(8):1188–1191, 1996.

[8] R. C. Larson and A. R. Odoni. *Urban Operations Research*. Prentice-Hall, Englewood Cliffs, NJ, 1981.

[9] D. J. Bertsimas and G. J. van Ryzin. A stochastic and dynamic vehicle routing problem in the Euclidean plane. *Operations Research*, 39:601–615, 1991.

[10] D. J. Bertsimas and G. J. van Ryzin. Stochastic and dynamic vehicle routing with general demand and interarrival time distributions. *Operations Research*, 41(1):60–76, 1993.

[11] Z. Drezner, editor. *Facility Location: A Survey of Applications and Methods*. Springer Series in Operations Research. Springer Verlag, New York, 1995.

[12] K. Sugihara A. Okabe, B. Boots and S. N. Chiu. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. John Wiley & Sons, New York, NY, 2000.

[13] J. Cortés, S. Martínez, T. Karatas, and F. Bullo. Coverage control for mobile sensing networks. *IEEE Transactions On Robotics and Automation*, 20(2):243–255, 2004.

[14] D. J. Bertsimas and G. J. van Ryzin. Stochastic and dynamic vehicle routing with general interarrival and service time distributions. *Advances in Applied Probability*, 25:947–978, 1993.

[15] M. Pavone, N. Bisnik, E. Frazzoli, and V. Isler. Decentralized vehicle routing in a stochastic and dynamic environment with customer impatience. In *Proc. Robocomm Conf.*, Athens, Greece, October 2007.