Adaptive information collection by robotic sensor networks for spatial estimation

Rishi Graham Jorge Cortés

Abstract

This work deals with trajectory optimization for a robotic sensor network sampling a spatio-temporal random field. We examine the optimal sampling problem of minimizing the maximum predictive variance of the estimator over the space of network trajectories. This is a high-dimensional, multi-modal, nonsmooth optimization problem, known to be NP-hard even for static fields and discrete design spaces. Under an asymptotic regime of near-independence between distinct sample locations, we show that the solutions to a novel generalized disk-covering problem are solutions to the optimal sampling problem. This result effectively transforms the search for the optimal trajectories into a geometric optimization problem. Constrained versions of the latter are also of interest as they can accommodate trajectories that satisfy a maximum velocity restriction on the robots. We characterize the solution for the unconstrained and constrained versions of the geometric optimization problem as generalized multicircumcenter trajectories, and provide algorithms which enable the network to find them in a distributed fashion. Several simulations illustrate our results.

I. Introduction

Intelligent information collection is an exciting field with many scientific, industrial, and security applications. Path planning, either a priori or online, is an important part of any data collection mission. When the underlying process being studied is modeled as random, special attention should be given to the choice of sample locations in order to minimize uncertainty in the resulting estimation. Our aim in this paper is to characterize the optimal trajectories for

Submitted on April 2, 2010. An early version of this work was submitted to the 2010 IEEE CDC as [1].

Rishi Graham is with the Department of Applied Mathematics and Statistics, University of California, Santa Cruz, rishig@ams.ucsc.edu

Jorge Cortés is with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, cortes@ucsd.edu

sampling a spatio-temporal random field modeled as a Gaussian field, and design distributed coordination algorithms that help robotic sensor networks determine them. We assume that the mean and covariance of the field are known, and concentrate on minimizing the maximum predictive variance. Because of our interest in online and adaptive operation, we consider a fairly general optimization problem where some of the samples in the network trajectory might be fixed in the optimization. This allows us to consider dynamic situations in which the network composition changes because of agents' arrivals and departures, or information is received from the environment or a human operator about changing conditions.

Literature review: There is a rich literature on the use of model uncertainty to drive the placement of sensing devices, e.g., [2], [3], [4]. Most of this research has focused on choosing from discrete sets of hypothetical sampling locations, and until recently all of it has made use of centralized computational techniques. Even choosing a fixed number of sampling locations from a discrete set has been shown to be NP-hard [5]. In cooperative control, various works consider mobile sensor networks performing spatial estimation tasks. [6], [7], [8] consider deterministic models with a stochastic measurement error term. [9] addresses the multiple robot path planning problem by greedily choosing way points from a discrete set of possible sensing locations. [10] considers a robotic sensor network with centralized control estimating a static field from samples with both sensing and localization error. In [11], a deterministic model is used, where the random elements come as unknown model parameters, and localization error is included. The work [12] uses a Gaussian process model where all information is globally available via all-to-all communication. Given the difficulty of optimizing within the whole set of network trajectories, [13] restricts the optimization problem to a subset of possible paths described by a finite set of parameters. [14] considers a single snapshot scenario (when agents only take one round of measurements) over a discrete sampling space in the limit of near-independence. Our previous work [15] has built on this setup to characterize the optimal network configurations in continuous sampling spaces and established the connection with Voronoi partitions [16] and geometric optimization [17], [18].

Statement of contributions: Our first contribution pertains to the characterization of the solutions of the optimal sampling problem for minimizing the prediction variance. We introduce a weighted distance metric called the correlation distance and define a novel generalized disk-covering function based on it. We show that minimizing this function is equivalent to minimizing

the maximum prediction variance in the limit of near-independence, thus turning the optimization problem into a geometric one. Our next contributions all pertain to the solution of this geometric problem. We first introduce a form of generalized Voronoi partition based on the maximal correlation between a given predictive location and the samples. Assuming a fixed network trajectory, we show that this partition minimizes the maximal correlation distance over all partitions of the predictive space. We next define multicircumcenter trajectories, which minimize the maximal correlation distance over all trajectories for a fixed partition. The combination of these two results gives rise to the optimal trajectories for the correlation distance disk-covering problem. The final stage of our solution is to define an extension of the maximal correlation partition which takes into account the positions of consecutive samples taken by the same robotic agent. Over this extended set, we define a notion of centering which ensures that the distance between such consecutive samples does not exceed a maximum distance. We show that these constrained multicenter trajectories optimize the correlation distance disk-covering problem over the set of distance-constrained trajectories. Finally, using the duality between optimal trajectory and optimal partition, we design a Lloyd-type algorithm which enables the network to arrive at locally optimal trajectories. At any step of the experiment, our strategy is capable of optimizing the remainder of the trajectories as new information arrives.

Organization: Section II introduces some preliminary mathematical notions. Section III discusses the robotic network model and introduces the statistical setup. We introduce the notion of near-independence in Section IV, and make the connection to the correlation distance disk covering problem. Section V introduces the maximal correlation partition. In Sections VI and VII the optimal trajectories are constructed for a fixed partition in the unconstrained and constrained cases, respectively. Section VIII presents the distributed coordination algorithms. Finally, Section IX gathers our conclusions and ideas for future work. For clarity of exposition, the proofs are presented in the appendix.

II. PRELIMINARIES

We start with some notation for standard geometric objects. Let $\mathbb{R}$, $\mathbb{R}_{>0}$ and $\mathbb{R}_{\geq 0}$ denote the set of reals, positive reals and nonnegative reals, respectively. Let $\mathrm{vrs}(p)$ denote the unit vector in direction p, i.e., $\mathrm{vrs}(p) = p/\|p\|$. For $p \in \mathbb{R}^d$ and $r \in \mathbb{R}_{>0}$, we let $\overline{B}(p, r)$ denote the *closed ball* of radius r centered at p. For a set W, we denote by $|W|$, $\mathrm{bnd}(W)$, $\mathrm{int}(W)$, and $\mathrm{co}(W)$ its

cardinality, the boundary, the interior, and the convex hull, respectively. A set $W \subset \mathbb{R}^d$ is *convex*, respectively *strictly convex* if, for every $s_1, s_2 \in W$ and $\alpha \in (0,1)$, we have $\alpha s_1 + (1-\alpha)s_2 \in W$, respectively, $\alpha s_1 + (1-\alpha)s_2 \in \mathrm{int}(W)$. For a bounded set $W \subset \mathbb{R}^d$, we let $\mathrm{CC}(W)$ and $\mathrm{CR}(W)$ denote the *circumcenter* and *circumradius* of W, respectively, that is, the center and radius of the smallest-radius d-sphere enclosing W.

Let $\mathfrak{P}(W)$ (respectively $\mathbb{F}(W)$) denote the collection of subsets (respectively, finite subsets) of W. Let $i_{\mathbb{F}} : (\mathbb{R}^d)^n \to \mathbb{F}(\mathbb{R}^d)$ be the natural immersion, i.e., $i_{\mathbb{F}}(P)$ contains only the distinct points in $P = (p_1, \ldots, p_n) \in (\mathbb{R}^d)^n$. Let $\|\cdot\|$ denote the Euclidean distance function on $\mathbb{R}^d$. We are interested in distances between points and subsets of $\mathbb{R}^d$. Let $\mathrm{d}_{\max} : \mathbb{R}^d \times \mathfrak{P}(\mathbb{R}^d) \to \mathbb{R}$ denote the maximum distance between a point and set, i.e., $\mathrm{d}_{\max}(s, W) = \sup_{s \in W}\{\|s - s\|\}$. For a vector, $S = (s_1, \ldots, s_n)^T \in (\mathbb{R}^d)^n$, let $\pi_k : (\mathbb{R}^d)^n \to \mathbb{R}^d$ denote the canonical projection onto the kth factor, i.e. $\pi_k(S) = s_k$. For a function $f : \mathbb{R}^d \to \mathbb{R}$ and $c \in \mathbb{R}$, let

$$S_{\text{lvl}}(f, c) = \{s \in \mathbb{R}^d \mid f(s) = c\}, \quad S_{\text{sublvl}}(f, c) = \{s \in \mathbb{R}^d \mid f(s) \leq c\},$$

denote the c-level and c-sublevel sets of f, respectively.

A. Nonsmooth analysis

Here we present some useful notions from nonsmooth analysis following [19]. A function $f : \mathbb{R}^d \to \mathbb{R}$ is *locally Lipschitz at* $s \in \mathbb{R}^d$ if there exist positive constants L_s and ϵ such that $|f(y) - f(y')| \leq L_s \|y - y'\|$ for all $y, y' \in \overline{B}(s, \epsilon)$. The function f is *locally Lipschitz on* $W \subseteq \mathbb{R}^d$ if it is locally Lipschitz at s, for all $s \in W$. A function $f : \mathbb{R}^d \to \mathbb{R}$ is *regular at* $s \in \mathbb{R}^d$ if for all $v \in \mathbb{R}^d$, the right and generalized directional derivatives of f at s in the direction of v, coincide. The interested reader is referred to [19] for the precise definition of these directional derivatives. The *generalized gradient* of a locally Lipschitz function f is

$$\partial f(s) = \mathrm{co}\{\lim_{i \to +\infty} df(s_i) \mid s_i \to s\,,\ s_i \notin W \cup \Omega_f\},$$

where $\Omega_f \subset \mathbb{R}^d$ denotes the set of points at which f fails to be differentiable, and W denotes any other set of measure zero. Note that this definition coincides with $df(s)$ if f is continuously differentiable at s. A point $s \in \mathbb{R}^d$ which satisfies that $0 \in \partial f(s)$ is called a *critical point of* f.

For a given closed, convex set $G \subset \mathbb{R}^d$, let $N_G : G \to \mathfrak{P}(\mathbb{R}^d)$ and $T_G : G \to \mathfrak{P}(\mathbb{R}^d)$ map locations in G to the normal cone and the tangent cone of G, respectively. Specifically, we have

$$N_G(x) = \{y \in \mathbb{R}^d \mid y^T(x - z) \geq 0,\ \forall z \in G\}, \quad T_G(x) = \{y \in \mathbb{R}^d \mid y^T z \leq 0,\ \forall z \in N_G(x)\}.$$

B. Spatio-temporal simple kriging

Here we describe the spatial interpolation process known as kriging following [20], adapted to a spatio-temporal context. Let Z denote a spatio-temporal process of interest taking values on a compact and convex region $\mathcal{D} \subset \mathbb{R}^d$. We assume the form

$$Z(s,t) = \mu(s,t) + \omega(s,t), \quad (s,t) \in \mathcal{D} \times \mathbb{R}_{\geq 0}, \tag{1}$$

where μ is a known function mapping space-time location to the mean value, and ω is a zero mean random space-time process with known covariance. We assume that ω has a separable covariance structure, which exhibits second-order stationarity and isotropy in the spatial dimensions, i.e.,

$$\mathrm{Cov}[\omega(s_i,t_i), \omega(s_j,t_j)] = g_0\, g_s(\|s_i - s_j\|) g_t(t_i,t_j), \tag{2}$$

for correlation functions $g_s : \mathbb{R}_{\geq 0} \to (0,1]$, and $g_t : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to [0,1]$, and constant $g_0 \in \mathbb{R}_{>0}$. We assume that g_s is strictly decreasing and continuously differentiable with nonzero derivative except possibly at 0 (i.e., $g_s'(d) < 0$ for $d > 0$). Note the assumption that the image of the spatial correlation function is strictly nonzero. These assumptions include the popular exponential, Gaussian, and Matérn correlation functions [21].

Assume that $n \in \mathbb{Z}_{>0}$ sensing agents take samples at each of a sequence of discrete timesteps $\{1, \ldots, k_{\max}\}$, with $k_{\max} \in \mathbb{Z}_{>0}$. Let $S_i = (s_i^{(1)}, \ldots, s_i^{(k_{\max})})^T \in \mathcal{D}^{k_{\max}}$ denote the spatial locations of samples taken over the course of the experiment by the ith agent, and let $S = (S_1^T, \ldots, S_n^T)^T \in (\mathcal{D}^{k_{\max}})^n$ denote the locations of all samples taken by the network. We use $I_{\text{samp}} = \{1, \ldots, n\} \times \{1, \ldots, k_{\max}\}$ to denote the set of index pairs into the sample vector. We refer often to vectors of elements indexed by both agent and timestep, such as the elements of S. To save space, we use the shorthand notation $(a_1^{(1)}, \ldots, a_n^{(k_{\max})}) = (a_1^{(1)}, \ldots, a_1^{(k_{\max})}, \ldots, a_n^{(1)}, \ldots, a_n^{(k_{\max})})$. Let $Y = (y_1^{(1)}, \ldots, y_n^{(k_{\max})})^T \in (\mathbb{R}^{k_{\max}})^n$ denote the values of all samples taken at locations S. We assume that the data are corrupted with a measurement error so that,

$$y_i^{(k)} = Z(s_i^{(k)}, k) + \epsilon_i, \qquad \epsilon_i \overset{\text{iid}}{\sim} \text{Normal}\left(0, \tau^2\right), \tag{3}$$

where $\tau^2 > 0$, and "iid" denotes independent and identically distributed. Although it is not a technical requirement, the assumption that the sampling noise is iid simplifies notation and exposition. The covariance between $y_i^{(k)}$ and $y_j^{(l)}$ is given by

$$\mathrm{Cov}[y_i^{(k)}, y_j^{(l)}] = \begin{cases} g_0\, g_s(0) g_t(k,k) + \tau^2, & \text{if } (i,k) = (j,l), \\ g_0\, g_s(\|s_i - s_j\|) g_t(k,l), & \text{otherwise.} \end{cases}$$

Let $\mathbf{\Sigma} = \Sigma(S)$ denote the covariance matrix of Y. Where disambiguation is not required, we use bold face to denote explicit dependence on S.

The *simple kriging predictor* at $(s,t) \in \mathcal{D} \times \mathbb{R}_{\geq 0}$ minimizes the error variance $\sigma^2(s,t;S) = \mathrm{Var}(Z(s,t) - p(s,t;Y))$ among all unbiased predictors of the form $p(s,t;Y) = \sum_{i=1}^{n} \sum_{k=1}^{k_{\max}} l_i^{(k)} y_i^{(k)} + a,\ a \in \mathbb{R}$. The simple kriging predictor at $(s,t) \in \mathcal{D} \times \mathbb{R}_{\geq 0}$ corresponds then to the Linear Unbiased Minimum Variance Estimator (LUMVE),

$$\hat{p}_{\mathrm{SK}}(s,t;Y) = \mu(s,t) + \boldsymbol{c}^T \mathbf{\Sigma}^{-1}(Y - \boldsymbol{\mu}), \tag{4}$$

with $\boldsymbol{\mu} = (\mu(s_1^{(1)},1), \ldots, \mu(x_n^{(k_{\max})}, k_{\max}))^T$, $\boldsymbol{c} = \mathrm{Cov}[Z(s,t), Y] \in (\mathbb{R}^{k_{\max}})^n$, and error variance,

$$\sigma^2(s,t;S) = g_0 g_s(0) g_t(t,t) - \boldsymbol{c}^T \mathbf{\Sigma}^{-1} \boldsymbol{c}. \tag{5}$$

Note that the function σ^2 only depends on the location of the samples and is invariant under permutations of the space-time sample locations.

III. Problem statement

Here we describe the model for the robotic network and provide the objective function for optimal sampling.

A. Robotic network model

Consider a group $\{R_1, \ldots, R_n\}$ of $n \in \mathbb{Z}_{>0}$ robotic sensing agents taking measurements of a spatio-temporal process of interest over $\mathcal{D} \subset \mathbb{R}^d$. The position of robot $i \in \{1, \ldots, n\}$ at time $t \in \mathbb{R}$ is denoted by $p_i(t) \in \mathcal{D}$, and $P(t) = (p_1(t), \ldots, p_n(t))^T \in \mathcal{D}^n$ denotes the vector of all positions. The robots take point measurements of the random process at their location at discrete instants of time in $\mathbb{Z}_{>0}$. The results of the paper are independent of the particular robot dynamics, so long as each agent is able to move up to any point within a distance $u_{\max} \in \mathbb{R}_{>0}$ between consecutive sampling times. We assume that the agents have a tunable communication radius that allows them to transmit prospective sample locations to nearby agents.

B. Objective function for spatial estimation

We consider the scenario where the robotic network is given a time frame $[1, k_{\max}]$, with $k_{\max} \in \mathbb{Z}_{>0}$, to sample the spatio-temporal process Z. A natural objective is to design sampling

trajectories in such a way as to minimize the uncertainty of an estimate of the field at time $k_{\max}$ generated from samples taken up to that time. Here, we consider an objective function inspired by the notion of G-optimality from optimal design [20], [3]. The *maximum error variance* $\mathcal{M} : (\mathcal{D}^{k_{\max}})^n \to \mathbb{R}$ of estimates made at time $k_{\max}$ over the region $\mathcal{D}$ is

$$\mathcal{M}(S) = \max_{s \in \mathcal{D}} \sigma^2((s, k_{\max}); S) = g_0 g_s(0) g_t(k_{\max}, k_{\max}) - \min_{s \in \mathcal{D}} \left\{ \boldsymbol{c}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{c} \right\}. \tag{6}$$

Note that $\mathcal{M}$ corresponds to a "worst-case scenario," where we consider locations in the domain at which the error variance of the LUMVE is maximal. Under the assumption of noisy measurements, i.e., $\tau^2 > 0$, the function σ^2 is well-defined for any $s \in \mathcal{D}$ and $S \in (\mathcal{D}^{k_{\max}})^n$. Indeed, the dependence of σ^2 on the sample locations is continuous, and hence $\mathcal{M}$ is also well-defined. Our goal is to find the sampling trajectories $S \in (\mathcal{D}^{k_{\max}})^n$ that minimize the objective function $\mathcal{M}$. Note that the simplified case of $k_{\max} = 1$ corresponds to optimal sampling locations for a single snapshot of a static field, where, under near-independence [15], multicircumcenter Voronoi configurations are optimal. The problem of trajectory optimization treated here is considerably more complex. We should also note that all of our results hold for predictions of the field made at times other than $k_{\max}$.

IV. Optimal solutions under near-independence

The objective function $\mathcal{M}$ is not convex and nonsmooth. The problem of finding an explicit characterization for its optimizers is especially hard: even for $k_{\max} = 1$, the optimization of $\mathcal{M}$ is known to be NP-hard over discrete spaces [5]. In this section we consider instead the optimization of $\mathcal{M}$ when the correlation function is raised to the power $\alpha \in \mathbb{R}_{>0}$. This is equivalent to considering, instead of (2), the spatio-temporal covariance

$$\mathrm{Cov}[\omega(s_i, t_i), \omega(s_j, t_j)] = g_0 \, g_s^\alpha(\|s_i - s_j\|) g_t^\alpha(t_i, t_j).$$

The meaning of this modified problem is as follows: as α grows, the correlation between distinct space-time locations vanishes. This asymptotic regime of increasingly smaller correlation between distinct points is known as *near-independence*, see [14]. Note that the correlation function $(g_s g_t)^\alpha$ retains much of the shape of the original correlation function (e.g., smoothness, range, etc), so this analysis is helpful in determining the properties of the original problem as well. To ease the exposition, we denote by $\boldsymbol{c}^{\{\alpha\}}$, resp. $\boldsymbol{\Sigma}^{\{\alpha\}}$, the vector $\boldsymbol{c}$, resp. the matrix $\boldsymbol{\Sigma}$,

with the correlation in each element raised to the power α. Similarly, let $\mathcal{M}^{\{\alpha\}} : (\mathcal{D}^{k_{\max}})^n \to \mathbb{R}$ be defined as

$$\mathcal{M}^{\{\alpha\}}(S) = g_0\big(g_s(0)g_t(k_{\max}, k_{\max})\big)^{\alpha} - \min_{s\in\mathcal{D}} \big\{(\boldsymbol{c}^{\{\alpha\}})^T(\boldsymbol{\Sigma}^{\{\alpha\}})^{-1}\boldsymbol{c}^{\{\alpha\}}\big\}.$$

Therefore, our objective is to characterize the asymptotic minimizers of this function. To do so, we need to introduce a family of weighted distance measures based on correlation. Define $\phi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and $w : \{1, \ldots, k_{\max}\} \to \mathbb{R}_{\geq 0}$ by,

$$\phi(d) = -\log(g_s(d)), \qquad w(k) = -\log(g_t(k_{\max}, k)).$$

The function w gives a weight which depends on the temporal correlation between sample time k and predictive time $k_{\max}$. The function ϕ is strictly increasing and continuously differentiable with strictly positive derivative except possibly at zero. It therefore admits an inverse, $\phi^{-1} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$. The correlation between a sample at step k and prediction at step $k_{\max}$ induces the weighted distance function, $\delta_k : \mathcal{D} \times \mathcal{D} \to \mathbb{R}_{\geq 0}$,

$$\delta_k(s_1, s_2) = \phi(\|s_1 - s_2\|) + w(k). \tag{7}$$

We refer to δ_k as the *correlation distance* associated with sample time k, and note that $\delta_k(s, s_i^{(k)}) = -\log\big(g_s(\|s - s_i^{(k)}\|)g_t(k_{\max}, k)\big)$. The following result classifies its level sets.

Lemma IV.1 (Correlation level sets) *For each $k \in \{1, \ldots, k_{max}\}$, $s \in \mathcal{D}$ and $c \in \mathbb{R}$, one has $S_{lvl}(s' \mapsto \delta_k(s', s), c) = \mathrm{bnd}\big(\overline{B}(s, r_k(c))\big)$, where $r_k : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, defined by*

$$r_k(c) = \begin{cases} \phi^{-1}(c - w(k)) & \text{if } c \geq w(k), \\ 0 & \text{otherwise}, \end{cases}$$

is strictly increasing and continuously differentiable on the interval $(w(k), \infty)$, with derivative $r_k{}'(c) = \frac{1}{\phi'(r_k(c))}$.

We are interested in those samples with smallest correlation distance to a given predictive location. Note that this is equivalent to the samples with highest correlation to the predictive location. We must therefore consider the possibility of samples with identical correlation to all predictive locations. Let S_{unique} be the following set of possible trajectories, which ensures the

spatio-temporal uniqueness of any samples that achieve the maximal correlation distance from any predictive location,

$$S_{\text{unique}} = \Big\{ S = (s_1^{(1)}, \ldots, s_n^{(k_{\max})})^T \in (\mathcal{D}^{k_{\max}})^n \mid \nexists (i,k) \neq (j,l) \in I_{\text{samp}} \text{ and } s \in \mathcal{D}, \text{ s.t.} \\ \delta_k(s, s_i^{(k)}) = \min_{(i',k') \in I_{\text{samp}}} \delta_k(s, s_{i'}^{(k')}) \text{ and } \delta_k(s', s_i^{(k)}) = \delta_l(s', s_j^{(l)}),\ \forall s' \in \mathcal{D} \Big\}.$$

Note that for samples $s_i^{(k)}$ and $s_j^{(l)}$ to have identical correlation distance to all predictive locations requires that $s_i^{(k)} = s_j^{(l)}$ *and* $g_t(k_{\max}, k) = g_t(k_{\max}, l)$. We are now ready to characterize the minimizers of $\mathcal{M}^{\{\alpha\}}$ as α grows.

Theorem IV.2 (Global minimizers of $\mathcal{M}$ under near-independence) *Let $\mathcal{H} : (\mathcal{D}^{k_{max}})^n \to \mathbb{R}$ denote the* correlation distance disk-covering *function, defined by*

$$\mathcal{H}(S) = \max_{s \in \mathcal{D}} \big\{ \min_{(i,k) \in I_{\text{samp}}} \{\delta_k(s, s_i^{(k)})\} \big\}. \tag{8}$$

For $\Omega \subset (\mathcal{D}^{k_{max}})^n$ compact, let $S_{mcc} \in \Omega$ be a global minimizer of the correlation disk-covering function $\mathcal{H}$ over Ω. Further assume that $S_{mcc} \in S_{unique}$. Then, as $\alpha \to \infty$, S_{mcc} asymptotically globally optimizes $\mathcal{M}^{\{\alpha\}}$ over Ω, that is, $\mathcal{M}^{\{\alpha\}}(S_{mcc})$ approaches a global minimum over Ω.

The proof of the theorem can be reproduced for local minimizers of $\mathcal{H}$ over Ω to arrive at the following result.

Corollary IV.3 (Local minimizers of $\mathcal{M}$ under near-independence) *For $\Omega \subset (\mathcal{D}^{k_{max}})^n$ compact, let $S_{mcc} \in \Omega$ be a local minimizer of the correlation disk-covering function $\mathcal{H}$ over Ω. Then, as $\alpha \to \infty$, S_{mcc} asymptotically locally optimizes $\mathcal{M}^{\{\alpha\}}$ over Ω, that is, $\mathcal{M}^{\{\alpha\}}(S_{mcc})$ approaches a local minimum over Ω.*

The generality of the subspace Ω in Theorem IV.2 and Corollary IV.3 also allows us to apply the result to two situations of particular importance. First, we may restrict the samples to feasible trajectories based on vehicular movement limitations, and the initial positions of the vehicles, which we will call *anchor points*. This amounts to a restriction on each agent trajectory, and we define the range-based constraint set, $\Omega_{\text{Rg}} \subset (\mathcal{D}^{k_{\max}})^n$ as, $\Omega_{\text{Rg}} = \prod_{i=1}^n \Omega_{\text{Rg}_i}$, where

$$\Omega_{\text{Rg}_i} = \big\{ (s_i^{(1)}, \ldots, s_i^{(k_{\max})})^T \in \mathcal{D}^{k_{\max}} \;\big|\; \|s_i^{(1)} - p_i(0)\| \leq u_{\max} \text{ and } \\ \|s_i^{(k)} - s_i^{(k-1)}\| \leq u_{\max},\ \forall k \in \{2, \ldots, k_{\max}\} \big\}. \tag{9}$$

Our results also hold for a more general problem, optimizing over all $P(0) \in \mathcal{D}^n$, however this setup is directed at online path planning where the benefits of distributed implementation shine. Second, a change in mission parameters at time $k-1$, $k \in \{2, \ldots, k_{\max}\}$, might prompt optimization over just those locations not yet sampled, i.e., $\Omega_{\text{Rg}}^{(\geq k)} = \prod_{i=1}^{n} \Omega_{\text{Rg}_i}^{(\geq k)}$, where

$$\Omega_{\text{Rg}_i}^{(\geq k)} = \big\{ (s_i^{(k)}, \ldots, s_i^{(k_{\max})})^T \in \mathcal{D}^{k_{\max}-k+1} \;\big|\; \|s_i^{(k)} - p(k-1)\| \leq u_{\max} \text{ and } \|s_i^{(k')} - s_i^{(k'-1)}\| \leq u_{\max},\ \forall k' \in \{k+1, \ldots, k_{\max}\} \big\}. \quad (10)$$

For ease of notation, we assume that these decisions and path adjustments are made at sample time instants, and thus the anchor points for optimization over $\Omega_{\text{Rg}_i}^{(\geq k)}$ are the sample locations at step $k-1$, but the process is easily extensible to optimization between sample times.

Theorem IV.2 shows that the optimization of the maximum error variance is equivalent to a geometric optimization problem in the near-independence range. This remarkable result allows us to turn the search for the optimizers of $\mathcal{M}^{\{\alpha\}}$ into the search for the optimizers of the correlation disk-covering function $\mathcal{H}$ defined in (8). This is what we tackle in the following sections.

V. Maximal correlation partition

In this section, we introduce the maximal correlation partition associated to a network trajectory. This partition will be instrumental in determining the optimizers of $\mathcal{H}$. In the context of this work, a partition of $\mathcal{D}$ is a collection of compact subsets, $\mathcal{W} = \{W_1^{(1)}, \ldots, W_n^{(k_{\max})}\}$ with disjoint interiors whose union is $\mathcal{D}$. For any $S \in S_{\text{unique}}$, let $\mathcal{MC}(S) = (\text{MC}_1^{(1)}(S), \ldots, \text{MC}_n^{(k_{\max})}(S))$ denote the *maximal correlation partition* defined by

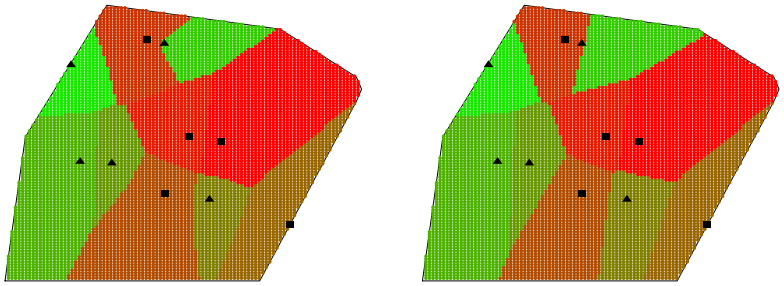
$$\text{MC}_i^{(k)}(S) = \big\{ s \in \mathcal{D} \;\big|\; \delta_k(s, s_i^{(k)}) \leq \delta_l(s, s_j^{(l)}),\ \forall (j,l) \neq (i,k) \big\}. \quad (11)$$

This partition corresponds to a generalized Voronoi partition [16] for distance measure ϕ and weights given by w. In general, the maximal correlation regions are neither convex nor star-shaped. Note that, depending on the weights and locations, $\text{MC}_i^{(k)}(S)$ might be empty for some i. Let $\text{I} : \mathfrak{P}(\mathcal{D}) \to \{1, \ldots, n * k_{\max}\}$ map a partition to the number of nonempty cells it contains, which we term the *index* of the partition. The following lemma gives some special cases in which $\mathcal{MC}$ is equal to distance-based partitions known in the literature, see e.g., [16], [22].

Lemma V.1 (Special cases of $\mathcal{MC}$) *The maximal correlation partition $\mathcal{MC}(S)$ corresponds to*

- *the Voronoi partition of $\mathcal{D}$ with generators S, if all weights are equal,*
- *the power diagram, if the spatial correlation is the Gaussian, $g_s(d) = e^{-\alpha d^2}$, with $\alpha \in \mathbb{R}_{>0}$,*
- *the additively weighted Voronoi partition, if the spatial correlation is the exponential, $g_s(d) = e^{-\alpha d}$, with $\alpha \in \mathbb{R}_{>0}$.*

Figure 1 illustrates the latter two types of partitions. For $S \in S_{\text{unique}}$, the correlation distance



(a) (b)

Fig. 1. Examples of maximal correlation partition in which each cell is defined by the predictive locations with highest (a) exponential correlation and (b) Gaussian correlation to a given (generating) sample. In both cases, two timesteps are shown. Samples taken at step 1 are shown as filled triangles, those taken at step 2 are shown as filled boxes.

disk-covering function can be restated in terms of the maximal correlation partition as,

$$\mathcal{H}(S) = \max_{(i,k)\in I_{\text{samp}}} \left\{ \max_{s\in \text{MC}_i^{(k)}(S)} \{\delta_k(s, s_i^{(k)})\} \right\}. \tag{12}$$

This expression is important because it clearly shows how $\mathcal{H}$ has a double dependence on the network trajectory S: through the value of the correlation distance and through the maximal correlation partition. This motivates us to define an extension of $\mathcal{H}$ as follows: for a given sample vector $S \in (\mathcal{D}^{k_{\max}})^n$ and a partition $W = \{W_1^{(1)}, \ldots, W_n^{(k_{\max})}\} \subset \mathfrak{P}(\mathcal{D})$ of the predictive space, define $\mathcal{H}_{\mathcal{W}} : (\mathcal{D}^{k_{\max}})^n \to \mathbb{R}$ by

$$\mathcal{H}_{\mathcal{W}}(S) = \max_{\substack{(i,k)\in I_{\text{samp}} \\ W_i^{(k)}\neq\emptyset}} \left\{ \max_{s\in W_i^{(k)}} \{\delta_k(s, s_i^{(k)})\} \right\}. \tag{13}$$

Note that if $S \in S_{\text{unique}}$, then $\mathcal{H}(S) = \mathcal{H}_{\mathcal{MC}(S)}(S)$. This function is particularly useful in our search for the optimizers of $\mathcal{H}$ because it allows us to decouple the two dependencies of this function on the network trajectory. The following result characterizes the maximal correlation partition as the optimal partition for $\mathcal{H}_{\mathcal{W}}$ given a fixed network trajectory.

Proposition V.2 ($\mathcal{H}$-optimality of the maximal correlation partition) *For any $S \in S_{unique}$ and any partition $\mathcal{W} \subset \mathfrak{P}(\mathcal{D})$ of $\mathcal{D}$ with $\mathrm{I}(\mathcal{W}) \leq \mathrm{I}(\mathcal{MC}(S))$,*

$$\mathcal{H}(S) \leq \mathcal{H}_{\mathcal{W}}(S), \tag{14}$$

that is, the maximal correlation partition $\mathcal{MC}(S)$ is optimal for $\mathcal{H}$ among all partitions of $\mathcal{D}$ of less than or equal index.

Proposition V.2 implies that, in order to fully characterize the optimizers of $\mathcal{H}$, it is sufficient to characterize the optimizers of $\mathcal{H}_{\mathcal{W}}$ for a fixed arbitrary partition. The latter formulation is advantageous because of the single dependence of the value of $\mathcal{H}_{\mathcal{W}}$ on the network trajectory.

VI. Unconstrained optimal trajectories for a given partition

In this section, our objective is to characterize the optimal network trajectories of $\mathcal{H}_{\mathcal{W}}$ for a fixed partition $\mathcal{W} = \{W_1^{(1)}, \ldots, W_n^{(k_{\max})}\} \subset \mathfrak{P}(\mathcal{D})$ of $\mathcal{D}$. We will find it useful to start our analysis with the simplified problem of locating a single sample to minimize the maximum correlation distance to a single predictive region. We will then build on this analysis to tackle the more complex multiple sample problem.

A. Single sample unconstrained problem

For $(i,k) \in I_{\text{samp}}$ with $W_i^{(k)} \neq \emptyset$, consider the task of choosing where R_i should take the sample at time k. Let $\mathrm{MCD}_i^{(k)} : \mathcal{D} \to \mathbb{R}_{>0}$ be defined as,

$$\mathrm{MCD}_i^{(k)}(s) = \max_{s' \in W_i^{(k)}} \delta_k(s', s). \tag{15}$$

Note that $\mathrm{MCD}_i^{(k)}$ corresponds to $\mathcal{H}_{\mathcal{W}}$ for a single agent and single sample at timestep k. For any $s \in \mathcal{D}$, it is important to note that the maximum correlation distance, $\mathrm{MCD}_i^{(k)}(s)$ is attained at the same locations in $W_i^{(k)}$ as the maximum Euclidean distance, i.e.,

$$\operatorname*{argmax}_{s' \in W_i^{(k)}} \delta_k(s', s) = \operatorname*{argmax}_{s' \in W_i^{(k)}} \|s' - s\|.$$

In the next result, which follows from Lemma IV.1, we characterize the sublevel sets of $\mathrm{MCD}_i^{(k)}$.

Lemma VI.1 (Sublevel sets of MCD) *For any $c \in \mathbb{R}_{\geq 0}$, the set $S_{sublvl}(\mathrm{MCD}_i^{(k)}, c)$ is closed, bounded, and strictly convex.*

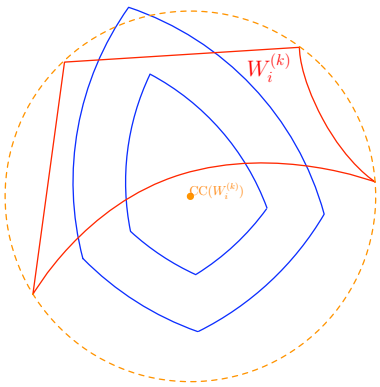


Fig. 2. A two-dimensional example of the level sets of $\mathrm{MCD}_i^{(k)}$. The dashed circle is the circumcircle. The closed curves around the circumcenter represent two different level sets of $\mathrm{MCD}_i^{(k)}$.

Figure 2 shows a two-dimensional example of the level sets of $\mathrm{MCD}_i^{(k)}$. The following result on the generalized gradient of the maximum correlation distance function makes use of [23, Theorem 2.1] and [19, Theorem 2.3.9].

Lemma VI.2 (Smoothness of $\mathrm{MCD}_i^{(k)}$) *The* $\mathrm{MCD}_i^{(k)}$ *is locally Lipschitz and regular, and its generalized gradient takes the form*

$$\partial \mathrm{MCD}_i^{(k)}(s) = \mathrm{co}\{\phi'(\mathrm{d}_{\max}(s, W_i^{(k)}))\, \mathrm{vrs}(s - s') \mid s' \in \operatorname*{argmax}_{s^* \in W_i^{(k)}} \delta_k(s^*, s)\}.$$

We next characterize the minimizers of $\mathrm{MCD}_i^{(k)}$.

Proposition VI.3 ($\mathrm{CC}(W_i^{(k)})$ minimizes $\mathrm{MCD}_i^{(k)}$) *The function* $\mathrm{MCD}_i^{(k)}$ *has a global minimum at* $\mathrm{CC}(W_i^{(k)})$ *and no other critical points.*

Remark VI.4 Note that Proposition VI.3 implies that the circumcenter minimizes the maximum Euclidean distance to an arbitrary set. •

B. Multiple sample unconstrained problem

Here, we use the results of Section VI-A to tackle the multiple sample problem, i.e., the characterization of the optima of the network objective $\mathcal{H}_{\mathcal{W}}$. We can equivalently write (13) as

$$\mathcal{H}_{\mathcal{W}}(S) = \max_{\substack{(i,k) \in I_{\mathrm{samp}} \\ W_i^{(k)} \neq \emptyset}} \mathrm{MCD}_i^{(k)}(s_i^{(k)}).$$

The following result on the generalized gradient of $\mathcal{H}_{\mathcal{W}}$ follows from using Lemma VI.2 and [19, Proposition 2.3.12] on this expression.

Lemma VI.5 (Smoothness of $\mathcal{H}_{\mathcal{W}}$) *The function $\mathcal{H}_{\mathcal{W}}$ is locally Lipschitz and regular, and its generalized gradient takes the form*

$$\partial\mathcal{H}_{\mathcal{W}}(S) = \operatorname{co}\left\{\partial \mathrm{MCD}_i^{(k)}(S) \mid (i,k) \in I_{\text{samp}} \text{ s.t. } \mathrm{MCD}_i^{(k)}(S) = \mathcal{H}_{\mathcal{W}}(S)\right\},$$

where, with a slight abuse of notation, we use $\mathrm{MCD}_i^{(k)}(S)$ *to denote the map* $S \mapsto \mathrm{MCD}_i^{(k)}(s_i^{(k)})$.

In order to extend Proposition VI.3 to the multiple sample case, we first need to introduce a piece of notation to account for the possibility of empty regions in the maximal correlation partition. Let $\overline{\mathrm{CC}} : \mathfrak{P}(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ be defined by

$$\overline{\mathrm{CC}}(W, s) = \begin{cases} \mathrm{CC}(W) & \text{if } W \neq \emptyset, \\ s & \text{otherwise.} \end{cases}$$

Let $\overline{\mathrm{CC}}(\mathcal{W}, S) = \left(\overline{\mathrm{CC}}(W_1^{(1)}, s_1^{(1)}), \ldots, \overline{\mathrm{CC}}(W_n^{(k_{\max})}, s_n^{(k_{\max})})\right)^T$ denote a vector of such circumcenter locations. We are now ready to state a generalization of Proposition VI.3.

Proposition VI.6 ($\mathcal{H}_{\mathcal{W}}$-optimal trajectories) *For any $S = (s_1^{(1)}, \ldots, s_n^{(k_{max})})^T \in S_{unique}$, any partition $\mathcal{W} = \{W_1^{(1)}, \ldots, W_n^{(k_{max})}\} \subset \mathfrak{P}(\mathcal{D})$ of $\mathcal{D}$, and any $\tilde{S} = (\tilde{s}_1^{(1)}, \ldots, \tilde{s}_n^{(k_{max})})^T \in (\mathcal{D}^{k_{max}})^n$,*

$$\mathcal{H}_{\mathcal{W}}\left(\overline{\mathrm{CC}}(\mathcal{W}, \tilde{S})\right) \leq \mathcal{H}_{\mathcal{W}}(S), \tag{16}$$

that is, the circumcenter locations $\overline{\mathrm{CC}}(\mathcal{W}, \tilde{S})$ are optimal for $\mathcal{H}_{\mathcal{W}}$ among all network trajectories.

Note the duality between the results in Proposition V.2 (for a fixed network configuration, the maximal correlation partition is optimal) and Proposition VI.6 (for a fixed partition, the circumcenter locations are optimal). The combination of these two results allow us provide the following characterization of the optimizers of the correlation disk-covering function $\mathcal{H}$.

Proposition VI.7 (Generalized multicircumcenter trajectories optimize $\mathcal{H}$) *Consider $S = (s_1^{(1)}, \ldots, s_n^{(k_{max})})^T \in (\mathcal{D}^{k_{max}})^n$ such that $s_i^{(k)} = \mathrm{CC}\left(MC_i^{(k)}(S)\right)$ for each $(i,k) \in I_{\text{samp}}$ with $MC_i^{(k)}(S) \neq \emptyset$. Then S is a local minimizer of $\mathcal{H}$ over $(\mathcal{D}^{k_{max}})^n$. We call such a network trajectory a* generalized multicircumcenter trajectory. *Furthermore, if $\mathrm{I}(\mathcal{MC}(S)) = n * k_{max}$, then S is a global minimizer of $\mathcal{H}$ over $(\mathcal{D}^{k_{max}})^n$.*

VII. Range-constrained optimal trajectories for a given partition

In this section, our objective is to characterize the optimizers of $\mathcal{H}_{\mathcal{W}}$ over Ω_{Rg} for a fixed partition $\mathcal{W}$. We begin our discussion by providing a useful alternative expression for $\mathcal{H}_{\mathcal{W}}$. Let $\mathcal{W}_i = \{W_i^{(1)}, \ldots, W_i^{(k_{\max})}\}$ denote the elements of the partition $\mathcal{W}$ assigned to the samples in the trajectory of R_i. With a slight abuse of notation, we may write

$$\mathcal{H}_{\mathcal{W}}(S) = \max_{\substack{i \in \{1,\ldots,n\} \\ \mathcal{W}_i \neq \emptyset}} \mathcal{H}_{\mathcal{W}_i}(S_i), \quad \text{where} \quad \mathcal{H}_{\mathcal{W}_i}(S_i) = \max_{\substack{k \in \{1,\ldots,k_{\max}\} \\ W_i^{(k)} \neq \emptyset}} \left\{ \text{MCD}_i^{(k)}(s_i^{(k)}) \right\}.$$

The condition $\mathcal{W}_i \neq \emptyset$ indicates that there is at least one nonempty $W_i^{(k)} \in \mathcal{W}_i$. The above expression clearly shows that, for a fixed partition, minimizing $\mathcal{H}_{\mathcal{W}}$ over the space of network trajectories is equivalent to (independently) minimizing each of the functions $\mathcal{H}_{\mathcal{W}_i}$ over the space of trajectories of the robot R_i. As a consequence, we structure our discussion in three parts. First, we deal with the single sample problem. Then, we build on this discussion to address the problem of finding an optimal sampling trajectory for a *single* agent. Finally, we combine individual agent trajectories into a network trajectory to find the constrained optimizers of $\mathcal{H}_{\mathcal{W}}$.

A. Single sample constrained problem

Proposition VI.3 allows a simple, geometric interpretation of the minimizer of $\text{MCD}_i^{(k)}$. Our objective here is to obtain a similar characterization for the range-constrained problem. We first consider the single sample problem over a general closed convex constraint set.

Proposition VII.1 (Constrained minimizers of $\text{MCD}_i^{(k)}$) *Assume that $W_i^{(k)} \neq \emptyset$. Let $\Gamma \subset \mathbb{R}^d$ be closed and convex. Then a point $s^* \in \Gamma$ is the unique minimizer of $\text{MCD}_i^{(k)}$ over Γ if and only if $\mathbf{0} \in \partial \text{MCD}_i^{(k)}(s^*) + N_\Gamma(s^*)$.*

Let us now specify the range based constraint set for $s_i^{(k)}$. The *set of constraining locations* of $(i,k) \in I_{\text{samp}}$ are the locations of robot R_i at sample times $k-1$ and $k+1$,

$$S_{\text{cs}}(k, S_i) = \{p(k') \mid k' \in K_{\text{cs}}(k)\}, \text{ where } K_{\text{cs}}(k) = \{k-1, k+1\} \cap \{0, \ldots, k_{\max}\}.$$

Note that in all but the initial anchor point, this set corresponds to the sample locations immediately preceding and following the (i,k)th sample. Let $\Gamma^{(k)} : \mathcal{D}^{k_{\max}} \to \mathfrak{P}(\mathbb{R}^d)$ map a network

trajectory to the intersection of $u_{\max}$-balls centered at locations in the set of constraining locations of (i,k), i.e.,

$$\Gamma^{(k)}(S_i) = \bigcap_{s \in S_{\text{cs}}(k,S_i)} \overline{B}(s, u_{\max}). \tag{17}$$

The set $\Gamma^{(k)}(S_i)$ corresponds to Ω_{Rg} with all other samples fixed in space. Restricting $S_i^{(k)}$ to $\Gamma^{(k)}(S_i)$ ensures that R_i does not violate the maximum distance requirement $u_{\max}$.

In order to state the main result of this section, we will find it useful to introduce an extension of the predictive set $W_i^{(k)}$ which incorporates the position of sample (i,k) relative to $\Gamma^{(k)}(S_i)$. To that end, define $\text{EPt}^{(k:k')} : \mathcal{D}^{k_{\max}} \to \mathbb{R}^d,\ (i,k) \in I_{\text{samp}},\ k' \in K_{\text{cs}}(k)$ by

$$\text{EPt}^{(k:k')}(S_i) = s_i^{(k)} + r_k(\mathcal{H}_{\mathcal{W}_i}(S_i)) \frac{s_i^{(k')} - s_i^{(k)}}{u_{\max}}, \tag{18}$$

The reason for the use of $\mathcal{H}_{\mathcal{W}_i}(S_i)$ will be made apparent in Section VII-B. For now, it is only important that $\mathcal{H}_{\mathcal{W}}(S_i) \geq \text{MCD}_i^{(k)}(s_i^{(k)})$. The location $\text{EPt}^{(k:k')}(S_i)$ can be seen as the projection of $s_i^{(k')}$ onto the surface of $\overline{B}(s_i^{(k)}, r_k(\mathcal{H}_{\mathcal{W}_i}(S_i)) \frac{\|s_i^{(k')} - s_i^{(k)}\|}{u_{\max}})$. Then, we extend the predictive set by the extended constraint points as follows. Let $\widetilde{W}_i^{(k)} : \mathcal{D}^{k_{\max}} \to \mathfrak{P}(\mathbb{R}^d),\ (i,k) \in I_{\text{samp}}$ be the *constraint extended predictive set*,

$$\widetilde{W}_i^{(k)}(S_i) = \text{co}\left(W_i^{(k)},\ \{\text{EPt}^{(k:k')}(S_i) \mid k' \in K_{\text{cs}}(k)\}\right).$$

A point $s \in \widetilde{W}_i^{(k)}(S_i)$ is *active in centering* if there is no neighborhood of s which might be added to $\widetilde{W}_i^{(k)}(S_i)$ without changing the circumcenter. It can be seen from (18) that $\text{EPt}^{(k:k')}(S_i)$ is active in centering if and only if $r_k(\mathcal{H}_{\mathcal{W}_i}(S_i)) \frac{\|s_i^{(k)} - s_i^{(k')}\|}{u_{\max}} \geq r_k\big(\text{MCD}_i^{(k)}(s_i^{(k)})\big)$. Figure 3 shows an example of the extended predictive set.

The next result gives a geometric interpretation of the constrained optimum in terms of $\widetilde{W}$.

Proposition VII.2 (Extended circumcenter minimizes $\text{MCD}_i^{(k)}$ over $\Gamma^{(k)}(S_i)$) *Assume that $\Gamma^{(k)}(S_i)$ and $W_i^{(k)}$ are nonempty. Further assume that the scaling factor for the extended constraints satisfies $\mathcal{H}_{\mathcal{W}_i}(S_i) = \text{MCD}_i^{(k)}(s_i^{(k)})$. Then $s_i^{(k)}$ is the unique minimizer of $\text{MCD}_i^{(k)}$ over $\Gamma^{(k)}(S_i)$ iff $s_i^{(k)} = \text{CC}\left(\widetilde{W}_i^{(k)}(S_i)\right)$.*

B. Multiple sample single agent constrained problem

Here we extend the constrained solution above to a single agent optimizing its own trajectory. and characterize the optima of $\mathcal{H}_{\mathcal{W}_i}$ over the constraint set Ω_{Rg_i} defined in (9) in terms of *centered*

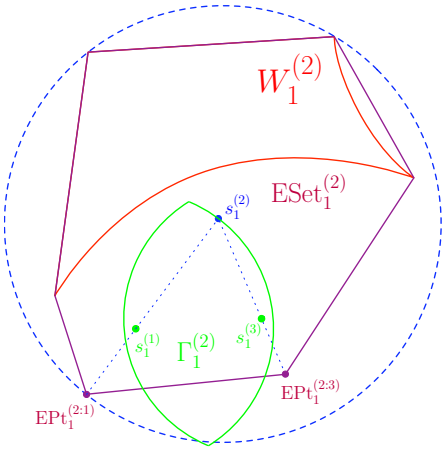


Fig. 3. A two-dimensional example of the extended center representation of a critical point of the constrained problem. The dashed circle is the circumcircle of $\widetilde{W}_1^{(2)}$, with circumcenter $s_1^{(2)}$. Note that $s_1^{(2)}$ is on the boundary of $\Gamma^{(2)}$ formed by $s_1^{(1)}$, and thus $\mathrm{EPt}^{(2:1)}$ is active in centering.

sub-sequences. In order to facilitate discussion of generalized gradients, let $\mathrm{d}^{(k:k')} : \mathcal{D}^{k_{\max}} \to \mathbb{R}_{\geq 0}$, $k, k' \in \{1, \ldots, k_{\max}\}$ be defined as $\mathrm{d}^{(k:k')}(S_i) = \|s_i^{(k)} - s_i^{(k')}\|$, and let $\mathrm{d}^{(1:0)}(S_i) = \mathrm{d}\,01(S_i) = \|s_i^{(1)} - p_i(0)\|$. With a slight abuse of notation, we use

$$\widetilde{W}_i^{(k)}(S_i; K_C) = \mathrm{co}\left(W_i^{(k)}, \{\mathrm{EPt}^{(k:k')}(S_i) \mid k' \in K_{\mathrm{cs}}(k) \cap K_C\}\right).$$

to denote constraint extended sets as calculated with a subset of the constraint points.

Lemma VII.3 (Centered sequences satisfy range constraint) *Let $S_i \in \mathcal{D}^{k_{max}}$, and let $K_C \subseteq \{1, \ldots, k_{max}\}$ define a sequence of consecutive samples from S_i such that each is at the circumcenter of the extended set formed by consecutive neighbors in the sequence, i.e.,*

$$s_i^{(k)} = \mathrm{CC}\left(\widetilde{W}_i^{(k)}(S_i; \{0\} \cup K_C)\right), \quad \textit{for all } k \in K_C,$$

Then $\mathrm{d}^{(k:k')}(S_i) \leq u_{max}$, for all $k \in K_C$ and $k' \in (\{0\} \cup K_C) \cap K_{\mathrm{cs}}(k)$. We call such a sequence centered.

Figure 4 shows an example of a centered sequence.

In the unconstrained case, optimizing $\mathcal{H}_{\mathcal{W}}$ takes the form of centering each sample within its predictive region, which may be characterized in terms of the generalized gradient of MCD. Given our discussion for the single sample constrained problem, in particular Proposition VII.2, we next characterize the gradient of the maximum correlation distance to the *extended* predictive

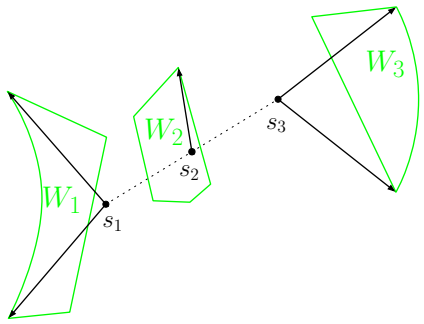


Fig. 4. Two-dimensional three sample example of a centered sequence. The solid arrows show the directions from the sample to the farthest points in the associated predictive region. For illustrative purposes, we have used a correlation distance equivalent to Euclidean distance.

region, $\widetilde{W}$, and thereby the optimal agent trajectories in terms of centered sequences. We begin with a result on the effect of the trajectory on the constraint extended predictive sets.

Lemma VII.4 (Correlation distance to extended constraints) *Let $(i,k) \in I_{\text{samp}}$ and $k' \in K_{\text{cs}}(k)$, and let $S_i \in \mathcal{D}^{k_{max}}$ such that $s_i^{(k)} \neq s_i^{(k')}$. Let $\text{CDE}_i^{(k:k')} : \mathcal{D}^{k_{max}} \to \mathbb{R}$ be defined by*

$$\text{CDE}_i^{(k:k')}(S_i) = \delta_k(s_i^{(k)}, \text{EPt}^{(k:k')}(S_i)).$$

The function $\text{CDE}_i^{(k:k')}$ is locally Lipschitz and regular near S_i, and its generalized gradient at S_i takes the form

$$\partial \text{CDE}_i^{(k:k')}(S_i) = \frac{\phi'(\|\text{EPt}^{(k:k')}(S_i) - s_i^{(k)}\|)}{u_{max}} \times \\ \times \left(r_k(\mathcal{H}_{\mathcal{W}_i}(S_i)) \partial\, \mathrm{d}^{(k:k')}(S_i) + \frac{\mathrm{d}^{(k:k')}(S_i)}{\phi'(r_k(\mathcal{H}_{\mathcal{W}_i}(S_i)))} \partial \mathcal{H}_{\mathcal{W}_i}(S_i) \right).$$

In the expression of the gradient of $\text{CDE}_i^{(k:k')}$ where $k' \neq 0$, note that since $s_i^{(k)} \neq s_i^{(k')}$, the set $\partial\, \mathrm{d}^{(k:k')}(S_i)$ consists of a single vector whose only nonzero components are the kth and k'th entries. Likewise $\partial\, \mathrm{d}^{(1:0)}(S_i)$ is nonzero only in the first entry.

We next characterize the function which maps the maximum correlation from a sample to any point in its constraint extended predictive set.

Lemma VII.5 (Extended set correlation distance) *Let $(i,k) \in I_{\text{samp}}$. Let $\text{MCD}_{\widetilde{W}}^{(k)} : \mathcal{D}^{k_{max}} \to \mathbb{R}$ map the ith trajectory to the maximum correlation distance from $s_i^{(k)}$ to the corresponding*

constraint extended predictive set, i.e.,

$$\mathrm{MCD}^{(k)}_{\widetilde{W}}(S_i) = \max_{s \in \widetilde{W}^{(k)}_i(S_i)} \delta_k\big(s, s^{(k)}_i\big).$$

Further assume that either $W^{(k)}_i \neq \emptyset$, or *there is an* $s \in S_{\mathrm{cs}}(k, S_i)$ *such that* $s^{(k)}_i \neq s$. *Then* $\mathrm{MCD}^{(k)}_{\widetilde{W}}$ *is locally Lipschitz and regular, and the generalized gradient takes the form*

$$\partial \mathrm{MCD}^{(k)}_{\widetilde{W}}(S_i) = \begin{cases} \partial \mathrm{MCD}^{(k)}_i(S_i) & \text{if } \mathrm{MCD}^{(k)}_i(S_i) > \mathrm{CDE}^{(k)}_{\max}(S_i), \\ \partial \mathrm{CDE}^{(k)}_{\max}(S_i) & \text{if } \mathrm{MCD}^{(k)}_i(S_i) < \mathrm{CDE}^{(k)}_{\max}(S_i), \\ \mathrm{co}\left\{\partial \mathrm{MCD}^{(k)}_i(S_i), \partial \mathrm{CDE}^{(k)}_{\max}(S_i)\right\} & \text{if } \mathrm{MCD}^{(k)}_i(S_i) = \mathrm{CDE}^{(k)}_{\max}(S_i), \end{cases}$$

where $\mathrm{MCD}^{(k)}_i(S_i)$ *denotes the map* $S_i \mapsto \mathrm{MCD}^{(k)}_i(s^{(k)}_i)$, $\mathrm{CDE}^{(k)}_{\max}(S_i) = \max_{l \in K_{\mathrm{cs}}(k)} \mathrm{CDE}^{(k:l)}_i(S_i)$, *and* $\partial \mathrm{CDE}^{(k)}_{\max}(S_i) = \mathrm{co}\{\partial \mathrm{CDE}^{(k:k')}_i(S_i) \mid k' \in \mathrm{argmax}_{l \in K_{\mathrm{cs}}(k)} \mathrm{CDE}^{(k:l)}_i(S_i)\}$.

The constrained objective function for a single agent may be defined as

$$\mathcal{H}_{\widetilde{W}_i}(S_i) = \max_{k \in \{1,\ldots,k_{\max}\}} \mathrm{MCD}^{(k)}_{\widetilde{W}}(S_i).$$

Note that this function may be calculated entirely by R_i. The following proposition describes the smoothness of the per-agent constrained objective function.

Proposition VII.6 (Extended maximum correlation distance) *Let* $i \in \{1, \ldots, n\}$ *and assume that the set* $\mathcal{W}_i$ *contains at least one nonempty element. The function* $\mathcal{H}_{\widetilde{W}_i}$ *is locally Lipschitz and regular and its gradient takes the form*

$$\partial \mathcal{H}_{\widetilde{W}_i}(S_i) = \mathrm{co}\left\{\partial \mathrm{MCD}^{(k)}_{\widetilde{W}}(S_i), k \in \{1, \ldots, k_{max}\} \mid \mathrm{MCD}^{(k)}_{\widetilde{W}}(S_i) = \mathcal{H}_{\widetilde{W}_i}(S_i)\right\}. \tag{19}$$

Lemma VII.7 (Equality of $\mathcal{H}_{\widetilde{W}_i}$ and $\mathcal{H}_{\mathcal{W}_i}$ over $\Omega_{\mathbf{Rg}_i}$) *Let* $i \in \{1, \ldots, n\}$ *and* $S_i \in \Omega_{\mathrm{Rg}_i}$. *Then* $\mathcal{H}_{\widetilde{W}_i}(S_i) = \mathcal{H}_{\mathcal{W}_i}(S_i)$.

We next characterize the critical points of $\mathcal{H}_{\widetilde{W}_i}$ in terms of a special case of centered sequences.

Lemma VII.8 (Maximal elements define sub-sequences within centered sequences) *Let* $K_C \subseteq \{1, \ldots, k_{max}\}$ *define a centered sequence of samples in* S_i *with* $\max_{k \in K_C} \mathrm{MCD}^{(k)}_i(s^{(k)}_i) = \mathcal{H}_{\mathcal{W}_i}(S_i)$. *Then there is a sub-sequence,* $K_{MC} \subseteq K_C$ *which is centered and such that every* $k \in K_{MC}$ *satisfies* $\mathrm{MCD}^{(k)}_{\widetilde{W}}(s^{(k)}_i) = \mathcal{H}_{\mathcal{W}_i}(S_i)$. *We refer to a sequence such as* K_{MC} *as* maximally centered.

Proposition VII.9 (Global minimizers of $\mathcal{H}_{\widetilde{W}_i}$ on $\Omega_{\mathbf{Rg}_i}$ contain maximally centered sequences) *A trajectory $S_i \in \Omega_{\mathrm{Rg}_i}$ is a critical point of $\mathcal{H}_{\widetilde{W}_i}$ if and only if it contains at least one maximally centered sequence of samples. Furthermore, any such critical point globally minimizes $\mathcal{H}_{\mathcal{W}_i}$ on Ω_{Rg_i}.*

C. Multiple agent constrained problem

Finally, we combine agent trajectories into a network trajectory to find the constrained optimizers of $\mathcal{H}_{\mathcal{W}}$. First, define $\mathcal{H}_{\widetilde{\mathcal{W}}} : (\mathcal{D}^{k_{\max}})^n \to \mathbb{R}$ by

$$\mathcal{H}_{\widetilde{\mathcal{W}}}(S) = \max_{i \in \{1,\ldots,n\}} \mathcal{H}_{\widetilde{W}_i}(S_i). \tag{20}$$

The following result extends Lemma VII.7 to the network.

Lemma VII.10 (Equality of $\mathcal{H}_{\widetilde{\mathcal{W}}}$ and $\mathcal{H}_{\mathcal{W}}$ over $\Omega_{\mathbf{Rg}}$) *Let $S \in \Omega_{\mathrm{Rg}}$. Then $\mathcal{H}_{\widetilde{\mathcal{W}}}(S) = \mathcal{H}_{\mathcal{W}}(S)$.*

The critical points of the extended network objective function may now be characterized. The proof of this result follows from Proposition VII.9.

Proposition VII.11 (Global minima of $\mathcal{H}_{\widetilde{\mathcal{W}}}$ on $\Omega_{\mathbf{Rg}}$ contain maximally centered sequences) *A trajectory $S \in \Omega_{\mathrm{Rg}}$ is a critical point of $\mathcal{H}_{\widetilde{\mathcal{W}}}$ if and only if there is at least one $i \in \operatorname{argmax}_{i \in \{1,\ldots,n\}} \mathcal{H}_{\mathcal{W}_i}(S_i)$ such that S_i contains at least one maximally centered sequence. Furthermore, any such critical point is a global minimum of $\mathcal{H}_{\mathcal{W}}$ over Ω_{Rg}*

Proposition VII.11 allows us to think of the optimization of $\mathcal{H}_{\mathcal{W}}$ independently for each agent. If each agent optimizes their own trajectory (cf. Proposition VII.9), then the resulting network trajectory is optimal. Along with Proposition V.2, this allows the following result on the optimal trajectories of the correlation disk-covering function $\mathcal{H}$ over Ω_{Rg}.

Proposition VII.12 (Range-constrained generalized multicircumcenter trajectory) *Let $S = (S_1^T, \ldots, S_n^T) \in (\mathcal{D}^{k_{max}})^n$ such that each S_i contains at least one maximally centered sequence with respect to the partition $\mathcal{W} = \mathcal{MC}(S)$. Then S is a local minimizer of $\mathcal{H}$ over Ω_{Rg}. We call such a network trajectory a* range-constrained generalized multicircumcenter trajectory. *Furthermore, if $\mathrm{I}(\mathcal{MC}(S)) = n * k_{max}$, then S is a global minimizer of $\mathcal{H}$ over Ω_{Rg}.*

Remark VII.13 Note that if each S_i is centered, then it must contain a maximally centered sequence, and thus S is a range-constrained generalized multicircumcenter trajectory. •

The following proposition allows for partial optimization of trajectories which are already under way, based on minimizing the maximum error *over the remainder of the experiment*. The proof is a direct result of Proposition VII.9, where the samples being optimized over are anchored by the last sample already taken.

Proposition VII.14 (Partially fixed range-constrained generalized multicircumcenter trajectory) *Let $k^* \in \{2, \dots, k_{max}\}$, and assume that samples $\{1, \dots, k^* - 1\}$ have been taken (thus the locations are now fixed). Let $S = (S_1^T, \dots, S_n^T) \in (\mathcal{D}^{k_{max}})^n$ such that, for each $i \in \{1, \dots, n\}$, $\exists\, K_i \subseteq \{k^*, \dots, k_{max}\}$ which defines a maximal sequence of samples in S_i, with anchor point $p_i(k^* - 1)$. Then S is a local minimizer of the map $(s_1^{(k^*)}, \dots, s_n^{(k_{max})}) \mapsto \mathcal{H}(S)$ over $\Omega_{\text{Rg}}^{(\geq k^*)}$. Furthermore, if $\mathrm{I}(\mathcal{MC}(S)) = n * k_{max}$, then S is a global minimum of the constrained problem.*

VIII. The Generalized Multicircumcenter Algorithm

Given our discussion in the previous sections, here we synthesize coordination algorithms to find the optimal trajectories of the correlation disk-covering $\mathcal{H}$ with and without range-constraints. The design of these strategies is based on the characterizations stated in Propositions VI.7 and VII.12 for the unconstrained and the constrained cases, respectively.

Table I presents the Generalized Multicircumcenter Algorithm, based on the well-known Lloyd algorithm for data clustering, by which the network may find a minimizer of $\mathcal{H}$ over $\Omega_{\text{Rg}}^{(\geq k^*)}$ for some $k^* \in \{1, \dots, k_{\max}\}$. With slight adjustments, the same algorithm works for the unconstrained case.

Figure 5 shows results of a simulation of the Generalized Multicircumcenter Algorithm, leaving out the initial anchor points to illustrate optimization over the set of all initial positions. The convergence properties of the algorithm are characterized in the following result.

Proposition VIII.1 (Convergence of the Generalized Multicircumcenter Algorithm) *The* Generalized Multicircumcenter Algorithm *is distributed over the partition $\mathcal{MC}(S^{\{j\}})$, meaning that at step $j + 1$, R_i need only communicate with $R_{i'}$ for each $i' \in \{1, \dots, n\}$ such that $MC_i^{(k)}(S^{\{j\}})$ adjacent to $MC_{i'}^{(k')}(S^{\{j\}})$ for some k, k'. Furthermore, $S^{\{j\}} \in \Omega_{\text{Rg}}^{(\geq k^*)}$, for all $j \in \mathbb{Z}_{>0}$. As $j \to \infty$, $S^{\{j\}}$ approaches a $S^* \in (\mathcal{D}^{k_{max}})^n$, and if $S^* \notin S_{unique}$, then S^* is a minimizer of $\mathcal{H}$ over $\Omega_{\text{Rg}}^{(\geq k^*)}$.*

Goal:	Find a minimum of $\mathcal{H}$ over $\Omega_{\text{Rg}}^{(\geq k^*)}$
Input:	(i) Sample interval $[k^*, k_{\max}]$
	(ii) Anchor points, $p_i(k^*-1),\ i \in \{1,\dots,n\}$
	(ii) Initial trajectory, $S^{\{0\}} = (S_1^{\{0\}},\dots,S_n^{\{0\}})^T \in \Omega_{\text{Rg}}^{(\geq k^*)}$, with $S_i^{\{0\}}$ the ith *agent* trajectory
Assume:	(i) R_i has a communication radius, $R_{\text{com}} \in \mathbb{R}_{>0}$ which is large enough to communicate its trajectory to any other agents whose samples are neighbors in $\mathcal{MC}$
	(ii) If $k^* > 1$, R_i knows the locations of all past samples which neighbor any future samples of R_i in $\mathcal{MC}$

For $j \in \mathbb{Z}_{>0}$, each robot R_i, $i \in \{1,\dots,n\}$ executes synchronously

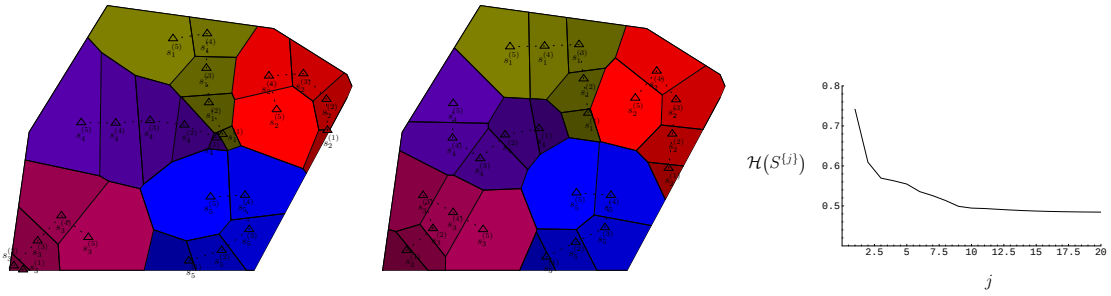
1: send all future elements of $S_i^{\{j-1\}}$ to robots within a distance of R_{com}

2: calculate $\text{MC}_i^{(k)}(S^{\{j-1\}})$ for $k \in \{k^*,\dots,k_{\max}\}$

3: run gradient descent of $\mathcal{H}_{\widetilde{W}_i}$ *on future samples only* to find a centered agent trajectory, $S_i^{\{j\}} \in \Omega_{\text{Rg}_i}^{(\geq k^*)}$

TABLE I

GENERALIZED MULTICIRCUMCENTER ALGORITHM.



(a) (b) (c)

Fig. 5. Simulation of 20 iterations of the GENERALIZED MULTICIRCUMCENTER ALGORITHM with no initial anchor points. (a) Shows the initial trajectory $S^{\{0\}}$. (b) Shows the final trajectory $S^{\{20\}}$. In each case, the associated maximal correlation partition is drawn, with the different colors representing different agents and different intensities of each color representing the timestep at which the given sample is to be taken (more intense colors represent later timesteps). The dashed lines show the path each agent will take. (c) Shows the value of $\mathcal{H}(S^{\{j\}})$ as a function of j.

Remark VIII.2 We suspect that the limit points of the GENERALIZED MULTICIRCUMCENTER ALGORITHM are in S_{unique} except for initial conditions in a set of measure zero, but establishing this fact is challenging because of the delicate interplay between the objective function and the constraints. Extensive simulations have reinforced our idea that this intuition is correct. •

We next turn our attention to an adaptive approach to optimal path planning. Before moving to

take the kth sample, an intelligent network of robotic sensors might receive updated information from an external source (a change in the environment or network composition, or even human input). One or more of the agents may switch from sensing mode to actuation mode, or back. The GENERALIZED MULTICIRCUMCENTER ALGORITHM directly applies to such a situation, because it optimizes over only those sample locations *not yet fixed*. The network will arrive at a trajectory which minimizes the maximum error variance over all trajectories feasible to the network moving forward. Table II describes the SEQUENTIAL GENERALIZED MULTICIRCUMCENTER ALGORITHM for performing this sequential optimization. The convergence of the SEQUENTIAL GENERALIZED MULTICIRCUMCENTER ALGORITHM follows from Proposition VIII.1, and Figure 6 depicts an illustrative example.

Goal:	Sequentially updated optimization.
Input:	(i) Initial trajectory, $S^{\{0\}} = (S_1^{\{0\}}, \ldots, S_n^{\{0\}})^T \in \Omega_{\text{Rg}}$, with $S_i^{\{0\}}$ the ith *agent* trajectory (ii) Status information about correlation structure, domain boundaries, and network composition

Initialization

1: network calculates optimal trajectory, S, via GENERALIZED MULTICIRCUMCENTER ALGORITHM

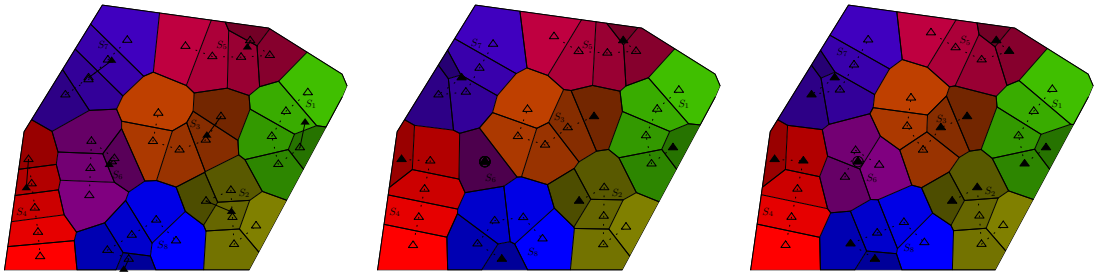
For $k \in \{1, \ldots, k_{\text{max}}\}$

1: move to kth location in optimal trajectory and take kth sample
2: **if** status input changed since previous optimization **then**
3: run the GENERALIZED MULTICIRCUMCENTER ALGORITHM to calculate a new optimal network trajectory over $\Omega_{\text{Rg}}^{(k+1)}$, holding the sample locations at steps $1, \ldots, k$ fixed

TABLE II
SEQUENTIAL GENERALIZED MULTICIRCUMCENTER ALGORITHM

IX. CONCLUSIONS

We have considered a robotic sensor network taking samples of a spatio-temporal process. As criteria for optimization we have taken the maximum error variance of the prediction made at the end of the experiment. Under the asymptotic regime of near-independence, we have shown that minimizing this error is equivalent to minimizing the correlation distance disk-covering function, thus allowing geometric solutions. We have introduced the maximal correlation partition and showed that it is the optimal partition of the predictive space for the disk-covering function given



(a) (b) (c)

Fig. 6. Evolution of three steps of the SEQUENTIAL GENERALIZED MULTICIRCUMCENTER ALGORITHM with $n = 8$ robots, $k_{\max} = 5$ steps, and Gaussian correlation. In (a), the initial trajectory is calculated from the initial anchor points $p_i(0)$. In (b), the first set of samples have been taken, and R_6 has dropped out to perform another task (for this simulation, R_6 remains stationary during this task). The figure shows the result of the GENERALIZED MULTICIRCUMCENTER ALGORITHM as run by the remaining 7 agents over timesteps $\{2, \dots, k_{\max}\}$. In (c), after the second set of samples have been taken, R_6 joins the network again. The figure shows the result of optimizing over steps $\{3, \dots, k_{\max}\}$ with all agents. In all three plots, the anchor points and any past samples are shown as solid triangles, with solid lines connecting the initial anchors to the first samples, the optimized samples at steps $\{k^*, \dots, k_{\max}\}$ are empty triangles, with dashed lines connecting each agent trajectory. The last sample location of the dropped agent is circled. In each case, the associated maximal correlation partition is drawn, with the different colors representing different agents and different intensities of each color representing the timestep at which the given sample is to be taken (more intense colors represent later timesteps).

a fixed network trajectory. We have introduced the novel notion of multicircumcenter trajectories and established their optimality with regards to the disk-covering function given a fixed partition. We have also defined a notion of extended sets which encodes a maximum movement restriction into a form of geometric centering, yielding the constrained multicircumcenter trajectory which is optimal over the set of all range-constrained trajectories. On the design front, we have synthesized distributed strategies that allow the network to calculate an optimal trajectory. In an ongoing experiment, the optimization can be executed online to recalculate the remaining sample locations in the face of changes in the environment, network structure, or human input. Future work will include the study of more complex predictive regions and of alternative optimality criteria.

REFERENCES

[1] R. Graham and J. Cortés, "Generalized multicircumcenter trajectories for optimal design under near-independance," in *IEEE Conf. on Decision and Control*, (Atlanta, Georgia), Dec. 2010. Submitted.

[2] K. Chaloner and I. Verdinelli, "Bayesian experimental design, a review," *Statistical Science*, vol. 10, no. 3, pp. 273–304, 1995.

[3] F. Pukelsheim, *Optimal Design of Experiments*, vol. 50 of *Classics in Applied Mathematics*. Philadelphia, PA: SIAM, 2006.

[4] E. P. Liski, N. K. Mandal, K. R. Shah, and B. K. Sinha, *Topics in Optimal Design*, vol. 163 of *Lecture Notes in Statistics*. New York: Springer, 2002.

[5] C.-W. Ko, J. Lee, and M. Queyranne, "An exact algorithm for maximum entropy sampling," *Operations Research*, vol. 43, no. 4, pp. 684–691, 1995.

[6] P. Ögren, E. Fiorelli, and N. E. Leonard, "Cooperative control of mobile sensor networks: Adaptive gradient climbing in a distributed environment," *IEEE Transactions on Automatic Control*, vol. 49, no. 8, pp. 1292–1302, 2004.

[7] K. M. Lynch, I. B. Schwartz, P. Yang, and R. A. Freeman, "Decentralized environmental modeling by mobile sensor networks," *IEEE Transactions on Robotics*, vol. 24, no. 3, pp. 710–724, 2008.

[8] S. Martínez, "Distributed interpolation schemes for field estimation by mobile sensor networks," *IEEE Transactions on Control Systems Technology*, vol. 18, no. 2, pp. 491–500, 2010.

[9] A. Singh, A. Krause, C. Guestrin, and W. J. Kaiser, "Efficient informative sensing using multiple robots," *Journal of Artificial Intelligence Research*, vol. 34, pp. 707–755, 2009.

[10] D. O. Popa, K. Sreenath, and F. L. Lewis, "Robotic deployment for environmental sampling applications," in *International Conference on Control and Automation*, (Budapest, Hungary), pp. 197–202, June 2005.

[11] M. F. Mysorewala, *Simultaneous robot localization and mapping of parameterized spatio-temporal fields using multi-scale adaptive sampling*. PhD thesis, University of Texas at Arlington, 2008.

[12] J. Choi, J. Lee, and S. Oh, "Biologically-inspired navigation strategies for swarm intelligence using spatial Gaussian processes," in *IFAC World Congress*, (Seoul, Korea), July 2008.

[13] N. E. Leonard, D. Paley, F. Lekien, R. Sepulchre, D. M. Fratantoni, and R. Davis, "Collective motion, sensor networks and ocean sampling," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 48–74, 2007.

[14] M. E. Johnson, L. M. Moore, and D. Ylvisaker, "Minimax and maximin distance designs," *Journal of Statistical Planning and Inference*, vol. 26, pp. 131–148, 1990.

[15] R. Graham and J. Cortés, "Asymptotic optimality of multicenter Voronoi configurations for random field estimation," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 153–158, 2009.

[16] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. Wiley Series in Probability and Statistics, Wiley, 2 ed., 2000.

[17] P. K. Agarwal and M. Sharir, "Efficient algorithms for geometric optimization," *ACM Computing Surveys*, vol. 30, no. 4, pp. 412–458, 1998.

[18] Z. Drezner and H. W. Hamacher, eds., *Facility Location: Applications and Theory*. Springer, 2001.

[19] F. H. Clarke, *Optimization and Nonsmooth Analysis*. Canadian Mathematical Society Series of Monographs and Advanced Texts, Wiley, 1983.

[20] N. A. C. Cressie, *Statistics for Spatial Data*. New York: Wiley, 1993. revised edition.

[21] P. Abrahamsen, "A review of Gaussian random fields and correlation functions," Technical Report 917, Norwegian Computing Center, Oslo, Norway, 1997. Electronically available at http://publications.nr.no/917_Rapport.pdf.

[22] M. de Berg, M. van Kreveld, M. Overmars, and O. Schwarzkopf, *Computational Geometry: Algorithms and Applications*. Springer, 2 ed., 2000.

[23] F. H. Clarke, "Generalized gradients and applications," *Transactions of the American Mathematical Society*, vol. 205, pp. 247–262, 1975.

[24] J. Cortés and F. Bullo, "Coordination and geometric optimization via distributed dynamical systems," *SIAM Journal on Control and Optimization*, vol. 44, no. 5, pp. 1543–1574, 2005.

[25] J. P. LaSalle, *The Stability and Control of Discrete Processes*, vol. 62 of *Applied Mathematical Sciences*. Springer, 1986.

APPENDIX

PROOFS AND SUPPORTING RESULTS FROM SECTION IV

We begin with some notation and preliminary results. Let $\mathrm{mcds} : \mathcal{D} \times (\mathcal{D}^{k_{\max}})^n \to \mathbb{F}(I_{\text{samp}})$ denote the *minimal correlation distance set* (MCDS) defined as,

$$\mathrm{mcds}(s, S) = \underset{(i,k)\in I_{\text{samp}}}{\operatorname{argmin}} \left\{\delta_k(s, s_i^{(k)})\right\}.$$

Note that mcds defines the set of samples in S with the highest correlation to s. Let $g_{\max} : \mathcal{D} \times (\mathcal{D}^{k_{\max}})^n \to \mathbb{R}$ map location and trajectory to this maximal correlation value, i.e.,

$$g_{\max}(s, S) = g_s(\|s - s_i^{(k)}\|) g_t(k_{\max}, k), \quad \forall\, (i,k) \in \mathrm{mcds}(s, S).$$

The following result describes a useful result on the dimensionality of the intersection of any two correlation distance surfaces.

Lemma A.1 (Equidistant sets are at most $d-1$ dimensional surfaces) *Assume that $S \in S_{\text{unique}}$, and let $(i,k),(j,l) \in I_{\text{samp}}$. Define $\gamma = \{s \in \mathbb{R}^d \mid \delta_k(s, s_i^{(k)}) = \delta_l(s, s_j^{(l)})\} \subset \mathbb{R}^d$. Then $\gamma = \mathbb{R}^d$ if and only if $(i,k) = (j,l)$. Otherwise, if $\gamma \neq \emptyset$, then it describes a surface in $\mathbb{R}^d$ which is at most $d-1$ dimensional.*

Proof: First, consider the shape of the correlation distance surfaces $s \mapsto \delta_k(s, s_i^{(k)})$ and $s \mapsto \delta_l(s, s_j^{(l)})$ in $\mathbb{R}^{d+1}$. From (7), it can be seen that the two surfaces differ only by a translation which is a result of both the spatial and temporal locations of the sample. The assumption that $S \in S_{\text{unique}}$ implies that $\gamma = \mathbb{R}^d$ if and only if $(i,k) = (j,l)$. Next, assume $\gamma \neq \mathbb{R}^d$ and $\gamma \neq \emptyset$. It can be shown that either the two correlation distance surfaces are tangent and that the tangent surface is contained within a one-dimensional line, or the gradient of the function $s \mapsto \delta_k(s, s_i^{(k)}) - \delta_l(s, s_j^{(l)})$ over $\gamma \setminus \{s_i^{(k)}, s_j^{(l)}\}$ is nonzero, implying that the dimension of γ is at most $d-1$. ■

The above lemma allows the following result on the cardinality of the MCDS.

Proposition A.2 (Cardinality of MCDS) *Assume that $S \in S_{unique}$. Then,*

$$\min_{s\in\mathcal{D}} \left\{ g_{\max}(s,S) \,|\, \mathrm{mcds}(s,S)| \right\} = \min_{s\in\mathcal{D}} \left\{ g_{\max}(s,S) \right\}.$$

Proof: We proceed by contradiction. If the statement is false, then there exists $s^\dagger \in \mathcal{D}$ such that $s^\dagger \in \mathrm{argmin}_{s\in\mathcal{D}} \left\{ g_{\max}(s,S) \,|\, \mathrm{mcds}(s,S)| \right\}$, and $|\mathrm{mcds}(s^\dagger,S)| > 1$. Define $\Gamma \subset \mathcal{D}$ by $\Gamma = \{ s \in \mathcal{D} \mid |\mathrm{mcds}(s,S)| > 1 \}$. Note that $s^\dagger \in \Gamma$, and $\Gamma \subseteq \bigcup_{i\neq j} \gamma_{ij}$. Lemma A.1 shows that Γ is the union of a finite number of surfaces of dimension at most $d-1$ embedded in $\mathbb{R}^d$. For any $\epsilon \in \mathbb{R}_{>0}$, there is a location $s^* \in \mathcal{D} \setminus \Gamma$ which satisfies $\|s^\dagger - s^*\| < \epsilon$. Thus $|\mathrm{mcds}(s^*,S)| = 1$. Since $g_{\max}(s,S)$ changes continuously with s, for ϵ small enough we have, $g_{\max}(s^*,S)|\mathrm{mcds}(s^*,S)| < g_{\max}(s^\dagger,S)|\mathrm{mcds}(s^\dagger,S)|$, which is a contradiction. ■

We are now ready to prove the main result.

Proof of Theorem IV.2: Note that minimizing $\mathcal{M}^{\{\alpha\}}$ on Ω_{Rg} is equivalent to maximizing the function $L^{\{\alpha\}} : \Omega_{\mathrm{Rg}} \to \mathbb{R}$ defined by $L^{\{\alpha\}}(S) = \min_{s\in\mathcal{D}} \left\{ (c^{\{\alpha\}})^T \times (\mathbf{\Sigma}^{\{\alpha\}})^{-1} (c^{\{\alpha\}}) \right\}$. Let $\lambda_{\min}$ and $\lambda_{\max} : \Omega_{\mathrm{Rg}} \times \mathbb{R} \to \mathbb{R}$ be such that $\lambda_{\min}(S,\alpha)$, $\lambda_{\max}(S,\alpha)$ denote, respectively, the minimum and the maximum eigenvalue of $\mathbf{\Sigma}^{\{\alpha\}}$. Note that with $\tau^2 \neq 0$, we have $0 < \lambda_{\min}(S,\alpha) \leq \lambda_{\max}(S,\alpha)$. Gershgorin circles and Proposition A.2 yield the asymptotic bounds,

$$\frac{g_0^2}{\lambda_{\max}(S,\alpha)} \min_{s\in\mathcal{D}} \{ g_{\max}(s,S)^{2\alpha}(1+o(1)) \} \leq L^{\{\alpha\}}(S) \leq \frac{g_0^2}{\lambda_{\min}(S,\alpha)} \min_{s\in\mathcal{D}} \{ g_{\max}(s,S)^{2\alpha}(1+o(1)) \}.$$

Consider, then, comparing an arbitrary sampling trajectory $S^* \in \Omega_{\mathrm{Rg}}$ against a global minimizer of $\mathcal{H}$ on Ω_{Rg}, say S_{mcc}. We can write,

$$\frac{L^{\{\alpha\}}(S^*)}{L^{\{\alpha\}}(S_{mcc})} \leq \frac{\frac{1}{\lambda_{\max}(S^*,\alpha)} \min_{s\in\mathcal{D}} \left\{ g_{\max}(s,S^*)^{2\alpha}(1+o(1)) \right\}}{\frac{1}{\lambda_{\min}(S_{mcc},\alpha)} \min_{s\in\mathcal{D}} \left\{ g_{\max}(s,S_{mcc})^{2\alpha}(1+o(1)) \right\}}. \tag{21}$$

Next we take a closer look at the eigenvalues. Note that the covariance matrix, $\mathbf{\Sigma}^{\{\alpha\}}$ becomes diagonal for large α. This gives us $\lim_{\alpha\to\infty} 1/(g_0+\tau^2)\mathbf{\Sigma}^{\{\alpha\}} = \boldsymbol{I}_{nk_{\max}}$, and it can be seen that $\lambda_{\max}(S,\alpha)/(g_0+\tau^2)$ and $\lambda_{\min}(S,\alpha)/(g_0+\tau^2)$ tend to 1 for any sample trajectory $S \in \Omega_{\mathrm{Rg}}$. Finally, since S_{mcc} minimizes the maximum over s of the minimum over (i,k) of $\delta_k(s, s_i^{(k)}) = \phi(\|s - s_i^{(k)}\|) - w(k)$, it equivalently maximizes the minimum value of $g_{\max}(s,S)$. For any $S \in \Omega_{\mathrm{Rg}}$, $\min_{s\in\mathcal{D}} \{ g_{\max}(s,S)^{2\alpha} \} \leq \min_{s\in\mathcal{D}} \{ g_{\max}(s,S_{mcc})^{2\alpha} \}$. Thus the ratio (21) is bounded by $1+o(1)$. Therefore, in the limit as $\alpha \to \infty$, minimizing $\mathcal{M}^{\{\alpha\}}$ over Ω_{Rg} is equivalent to minimizing the maximum covariance disk-covering function, $\mathcal{H}$ on Ω_{Rg}. ■

PROOFS AND SUPPORTING RESULTS FROM SECTION V

Proof of Proposition V.2: Let $(i,k) \in I_{\text{samp}}$ and $s_* \in \mathcal{D}$ be such that $\mathcal{H}(S) = \delta_k(s_*, s_i^{(k)})$. By definition, given a partition $\mathcal{W} = \{W_1^{(1)}, \ldots, W_n^{(k_{\max})}\}$ of $\mathcal{D}$, there exists a pair, $(j,l) \in I_{\text{samp}}$, such that $s_* \in W_j^{(l)}$. The definition of $\mathcal{MC}$ and the assumption that $\mathrm{I}(\mathcal{W}) \leq \mathrm{I}(\mathcal{MC}(S))$ leads to the implication chain, $\mathcal{H}(S) = \delta_k(s_*, s_i^{(k)}) \leq \delta_l(s_*, s_j^{(l)}) \leq \max_{s \in W_j^{(l)}} \delta_l(s, s_j^{(l)}) \leq \mathcal{H}_{\mathcal{W}}(S)$. ■

PROOFS AND SUPPORTING RESULTS FROM SECTION VI

Proof of Lemma VI.1: For $c \leq w(k)$, we have $S_{\text{sublvl}}(\text{MCD}_i^{(k)}, c) = \emptyset$. Otherwise, it is the intersection of an infinite set of closed d-spheres, which is a strictly convex set. ■

Proof of Proposition VI.3: First, note that $\text{MCD}_i^{(k)}$ and the map $s \mapsto \mathrm{d}_{\max}(s, W_i^{(k)})$ have the same extrema. In [24] it is shown that the latter function has a unique global minimum at $\mathrm{CC}(W_i^{(k)})$, when $W_i^{(k)}$ is taken to be a convex polygon. Identical reasoning yields the same result for any closed, bounded and nonempty $W_i^{(k)}$. Thus $\mathrm{CC}(W_i^{(k)})$ is a global minimum of $\text{MCD}_i^{(k)}$. The requirement that $\phi'(d) > 0$ for all $d > 0$ suffices to ensure that $\text{MCD}_i^{(k)}$ does not have any critical points which are not critical points of the Euclidean maximum distance function. Since that function has no critical points other than $\mathrm{CC}(W_i^{(k)})$, the result follows. ■

Proof of Proposition VI.6: For each $(i,k) \in I_{\text{samp}}$ with $W_i^{(k)} \neq \emptyset$, we can write,

$$\max_{s \in W_i^{(k)}} \big\{\delta_k\big(s, \overline{\mathrm{CC}}(W_i^{(k)}, \tilde{s}_i^{(k)})\big)\big\} = \phi\big(\max_{s \in W_i^{(k)}} \|s - \overline{\mathrm{CC}}(W_i^{(k)}, \tilde{s}_i^{(k)}))\|\big) + w(k) \leq$$

$$\leq \phi\big(\max_{s \in W_i^{(k)}} \|s - s_i^{(k)}\|\big) + w(k) = \max_{s \in W_i^{(k)}} \big\{\delta_k(s, s_i^{(k)})\big\}.$$

Taking the maximum over all nodes implies (16). ■

PROOFS AND SUPPORTING RESULTS FROM SECTION VII-A

We begin with this supporting result on strictly convex sets.

Lemma A.3 (Strict convexity) *Let $G \subset \mathbb{R}^d$ be closed, bounded, and strictly convex. For any $s_1, s_2 \in G$ and $v \in N_G(s_2) \setminus \{\mathbf{0}\}$, $v^T \operatorname{vrs}(s_1 - s_2) < 0$. Equivalently,* $\operatorname{vrs}(s_1 - s_2) \in \operatorname{int}(T_G(s_2))$.

Proof of Proposition VII.1: Necessity is a result of [19, Corollary to Proposition 2.4.3]. To show sufficiency, assume that $\mathbf{0} \in \partial\text{MCD}_i^{(k)}(s^*) + N_\Gamma(s^*)$, and we consider two cases. If $\mathrm{CC}(W_i^{(k)}) \in \Gamma$, the result follows by Proposition VI.3. We proceed by contradiction. Assume that

$s^* \neq \mathrm{CC}(W_i^{(k)})$, and $\mathbf{0} \in \partial\mathrm{MCD}_i^{(k)}(s^*) + N_\Gamma(s^*)$, but s^* is not a unique minimizer. Then $\exists s^\dagger \in \Gamma$ such that $\mathrm{MCD}_i^{(k)}(s^\dagger) \leq \mathrm{MCD}_i^{(k)}(s^*)$. By Proposition VI.3, s^* is not a critical point of $\mathrm{MCD}_i^{(k)}$. It follows that there is at least one nonzero vector, $v_G \in \partial\mathrm{MCD}_i^{(k)}(s^*)$ with $-v_G \in N_\Gamma(s^*)$, which implies $v_G^T \,\mathrm{vrs}(s^\dagger - s^*) \geq 0$. We know that $s^\dagger \in S_{\text{sublvl}}(\mathrm{MCD}_i^{(k)}, \mathrm{MCD}_i^{(k)}(s^*))$, and by Lemma VI.1, $S_{\text{sublvl}}(\mathrm{MCD}_i^{(k)}, \mathrm{MCD}_i^{(k)}(s^*))$ is strictly convex. By [19, Theorem 2.4.7 Corollary 1], $v_G \in N_{S_{\text{sublvl}}(\mathrm{MCD}_i^{(k)}, \mathrm{MCD}_i^{(k)}(s^*))}(s^*)$. Lemma A.3 yields, $v_G^T \,\mathrm{vrs}(s^\dagger - s^*) < 0$, a contradiction. Therefore s^* is the unique global minimizer of $\mathrm{MCD}_i^{(k)}$ over Γ. ■

We will need this supporting result on the circumcenter of the extended set.

Lemma A.4 ($s = \mathrm{CC}(\widetilde{W}_i^{(k)}(S_i))$ **implies** $s \in \Gamma^{(k)}(S_i)$**)** *Assume that* $W_i^{(k)} \neq \emptyset$. *Let* $S_i \in \mathcal{D}^{k_{max}}$ *such that* $\Gamma^{(k)}(S_i) \neq \emptyset$. *If* $s_i^{(k)} = \mathrm{CC}\left(\widetilde{W}_i^{(k)}(S_i)\right)$ *then* $s_i^{(k)} \in \Gamma^{(k)}(S_i) \bigcap \mathcal{D}$.

Proof: Assume that $s_i^{(k)} = \mathrm{CC}\left(\widetilde{W}_i^{(k)}(S_i)\right)$. Equation (18) and the fact that $s_i^{(k)} \in \mathrm{co}\left(\widetilde{W}_i^{(k)}(S_i)\right)$ imply that $s_i^{(k)} \in \mathcal{D}$. That $s_i^{(k)} \in \Gamma^{(k)}(S_i)$ follows by contradiction from the fact that $s_i^{(k)} \notin \Gamma^{(k)}(S_i)$ implies that $s_i^{(k)} = \mathrm{CC}(\mathrm{co}\{S_{\text{cs}}(k, S_i)\})$, and the fact that $\Gamma^{(k)}(S_i)$ is the nonempty intersection of d-spheres of equal radii centered at points in $S_{\text{cs}}(k, S_i)$. ■

Proof of Proposition VII.2: As a result of Lemma A.4, $s_i^{(k)} = \mathrm{CC}\left(\widetilde{W}_i^{(k)}(S_i)\right)$ implies that $s_i^{(k)} \in \Gamma^{(k)}(S_i)$. We may therefore assume $s_i^{(k)} \in \Gamma^{(k)}(S_i)$. Note that since $s_i^{(k)} \in \Gamma^{(k)}(S_i)$, we may write, $\mathrm{d}_{\max}(s_i^{(k)}, \widetilde{W}_i^{(k)}(S_i)) = r_k(\mathcal{H}_{\mathcal{W}_i}(S_i)) = r_k\big(\mathrm{MCD}_i^{(k)}(s_i^{(k)})\big)$. If, in addition, $\mathrm{d}^{(k:k')}(S_i) = u_{\max}$ for some $k' \in K_{\text{cs}}(k)$, then we also have, $r_k(\mathcal{H}_{\mathcal{W}_i}(S_i)) = \|s_i^{(k)} - \mathrm{EPt}^{(k:k')}(S_i)\|$. Let $\xi_{\text{EPt}} \subset \mathbb{R}^d$, respectively $\xi_W \subset \mathbb{R}^d$ denote the sets of unit vectors pointing from $s_i^{(k)}$ to the extended constraint points at a distance of $r_k(\mathcal{H}_{\mathcal{W}_i}(S_i))$, respectively to the points in $W_i^{(k)}$ at a distance of $r_k(\mathcal{H}_{\mathcal{W}_i}(S_i))$, i.e.,

$$\xi_{\text{EPt}} = \left\{ \mathrm{vrs}(s_i^{(k')} - s_i^{(k)}) \middle| k' \in K_{\text{cs}}(k) \text{ s.t. } \|s_i^{(k)} - \mathrm{EPt}^{(k:k')}(S_i)\| = r_k(\mathcal{H}_{\mathcal{W}_i}(S_i)) \right\}$$

$$\xi_W = \{\mathrm{vrs}(s - s_i^{(k)}) \mid s \in W_i^{(k)} \text{ s.t. } \|s_i^{(k)} - s\| = r_k(\mathcal{H}_{\mathcal{W}_i}(S_i))\}.$$

It can be deduced from Equation (18) that the set $\{\mathbf{0} \bigcup \xi_{\text{EPt}}\}$ spans $N_{\Gamma^{(k)}(S_i)}(s_i^{(k)})$. By extension of Proposition VI.3, we may conclude that $s_i^{(k)} = \mathrm{CC}\left(\widetilde{W}_i^{(k)}(S_i)\right)$ if and only if $\mathbf{0} \in \mathrm{co}\{\xi_W \bigcup \xi_{\text{EPt}}\}$. It can be seen that $\mathbf{0} \in \mathrm{co}\{\xi_W \bigcup \xi_{\text{EPt}}\}$ if and only if $\mathbf{0} \in \partial\mathrm{MCD}_i^{(k)}(s_i^{(k)}) + N_{\Gamma^{(k)}(S_i)}(s_i^{(k)})$. By Proposition VII.1, we have our result. ■

PROOFS AND SUPPORTING RESULTS FROM SECTION VII-B

Proof of Lemma VII.3: The result follows by simple contradiction from two observations for any $k' \in K_{\text{cs}}(k) \cap K_C$. First, if $\mathrm{d}^{(k:k')} > u_{\max}$, then $s_i^{(k)} = \mathrm{CC}(\widetilde{W}_i^{(k)}(S_i; K_C))$ would imply that $K_{\text{cs}}(k) \cap K_C = \{k-1, k+1\}$ and $s_i^{(k)} = \frac{s_i^{(k-1)} + s_i^{(k+1)}}{2}$. Second, the first and last samples in the sequence must satisfy $\delta_k\big(\mathrm{EPt}^{(k:k')}(S_i), s_i^{(k)}\big) \leq \mathrm{MCD}_i^{(k)}(S_i)$. ■

Proof of Lemma VII.4: From Equation (18), we can write,

$$\|\mathrm{EPt}^{(k:k')}(S_i) - s_i^{(k)}\| = \frac{r_k(\mathcal{H}_{\mathcal{W}_i}(S_i))}{u_{\max}}\, \mathrm{d}^{(k:k')}(S_i).$$

It has been established that $\mathcal{H}_{\mathcal{W}}$ is locally Lipschitz and regular, as is $\mathrm{d}^{(k:k')}$. The gradient is derived from [19, Proposition 2.3.13] and a special case of [19, Theorem 2.3.9]. ■

The following result characterizes critical points of $\partial\mathrm{CDE}_i^{(k:k')} \subset \partial\mathcal{H}_{\widetilde{W}_i}(S_i)$.

Corollary A.5 (Critical points of ∂CDE$_i^{(k:k')}(S_i)$) *Let $S_i \in \Omega_{\mathrm{Rg}_i}$, and let $k, k' \in \{1, \ldots, k_{max}\}$. If $\mathbf{0} \in \partial\mathrm{CDE}_i^{(k:k')}(S_i) \subset \partial\mathcal{H}_{\widetilde{W}_i}(S_i)$ then* all *of the following hold,*

$$\mathcal{H}_{\mathcal{W}_i}(S_i) = \mathrm{MCD}_i^{(k)}(S_i) = \mathrm{MCD}_i^{(k')}(S_i) \tag{22a}$$

$$\mathbf{0} \in \mathrm{co}\{\partial\mathrm{MCD}_i^{(k)}(s_i^{(k)}), \partial\mathrm{MCD}_i^{(k')}(s_i^{(k')})\} \tag{22b}$$

$$s_i^{(k)} = \mathrm{CC}\left(\widetilde{W}_i^{(k)}(S_i; \{k'\})\right). \tag{22c}$$

Proof: First, note that since $S_i \in \Omega_{\mathrm{Rg}_i}$, we have $\partial\mathrm{CDE}_i^{(k:k')}(S_i) \subset \partial\mathcal{H}_{\widetilde{W}_i}(S_i)$ if and only if $\mathrm{d}^{(k:k')}(S_i) = u_{\max}$. From Lemma VII.4 it can be seen that $\partial\mathrm{CDE}_i^{(k:k')}(S_i)$ is proportional to the sum of two vector sets, one of which consists of a single vector which is nonzero only in the kth and k'th components, and the other is $\partial\mathcal{H}_{\mathcal{W}_i}(S_i)$. Any vector in $\partial\mathcal{H}_{\mathcal{W}_i}(S_i)$ is zero everywhere except (possibly) the element corresponding to a single timestep. Thus $\mathbf{0} \in \partial\mathrm{CDE}_i^{(k:k')}(S_i)$ only if Equation (22a) holds. Solving the two simultaneous equations $\mathbf{0} \in \pi_k(\partial\mathrm{CDE}_i^{(k:k')})$ and $\mathbf{0} \in \pi_{k'}(\partial\mathrm{CDE}_i^{(k:k')})$ yields the other results. ■

Proof of Lemma VII.5: Note that $\mathrm{MCD}_{\widetilde{W}}^{(k)}(S_i) = \max\left\{\mathrm{MCD}_i^{(k)}(S_i), \mathrm{CDE}_{\max}^{(k)}(S_i)\right\}$. By Lemmas VII.4 and VI.2, $\mathrm{MCD}_{\widetilde{W}}^{(k)}(S_i)$ can be seen to be the maximum of locally Lipschitz and regular functions, and therefore locally Lipschitz and regular itself. The form of the gradient follows from application of [19, Proposition 2.3.12]. ■

Proof of Proposition VII.6: The $\mathrm{MCD}_{\widetilde{W}}^{(k)}$ is locally Lipschitz and regular for all $k \in \mathrm{argmax}_{k \in \{1,\ldots,k_{\max}\}} \mathrm{MCD}_{\widetilde{W}}^{(k)}(S_i)$. Since $\mathcal{H}_{\widetilde{W}_i}$ is the maximum of locally Lipschitz and regular

functions, it is locally Lipschitz and regular itself. The form of the gradient follows from application of [19, Proposition 2.3.12]. ■

Proof of Lemma VII.7: For any $k \in \{1, \dots, k_{\max}\}$ and $k' \in K_{\text{cs}}(k)$, $S_i \in \Omega_{\text{Rg}}$ implies that $\text{CDE}_i^{(k:k')}(S_i) \leq \mathcal{H}_{\mathcal{W}_i}(S_i)$. By definition, we also have $\text{MCD}_i^{(k)}(S_i) \leq \mathcal{H}_{\mathcal{W}_i}(S_i)$, with equality for at least one k. We may then write, $\mathcal{H}_{\widetilde{W}_i}(S_i) = \max_{k \in \{1,\dots,k_{\max}\}} \text{MCD}_{\widetilde{W}}^{(k)}(S_i) = \mathcal{H}_{\mathcal{W}_i}(S_i)$. ■

Proof of Lemma VII.8: First, note that since $S_i \in \Omega_{\text{Rg}_i}$, for any $k \in K_C$ and $k' \in K_{\text{cs}}(k)$, we have $\text{CDE}_i^{(k:k')}(S_i) \leq \mathcal{H}_{\mathcal{W}_i}(S_i)$, with equality if and only if $\text{d}^{(k:k')}(S_i) = u_{\max}$. If this condition is not met, then sample $s_i^{(k')}$ is not active in the centering of $s_i^{(k)}$. Furthermore, if $\text{MCD}_{\widetilde{W}}^{(k')}(S_i) < \mathcal{H}_{\mathcal{W}_i}(S_i)$, then $\text{CDE}_i^{(k':k)}(S_i) < \mathcal{H}_{\mathcal{W}_i}(S_i)$. Thus any sample which does not have maximal distance to its *extended* set can not be active in the centering of a sample which does. If k is maximal, and k' is not, then the sub-sequence which includes k but not k' is also centered. Thus a maximally centered sequence may be constructed around any maximal sample in K_C. ■

Proposition A.6 (Maximally centered trajectories are optimal) *Let $\mathcal{W}_i \subset \mathfrak{P}(\mathcal{D})$ and $S_i \in \Omega_{\text{Rg}_i}$ such that the entire sequence, S_i is maximally centered. Then S_i is the unique strict global minimizer of $\mathcal{H}_{\widetilde{W}_i}$ over Ω_{Rg_i}.*

Proof: Let $\tilde{S}_i = (\tilde{s}_i^{(1)}, \dots, \tilde{s}_i^{(k_{\max})})^T \in \Omega_{\text{Rg}_i}$ such that $\mathcal{H}_{\widetilde{W}_i}(\tilde{S}_i) \leq \mathcal{H}_{\widetilde{W}_i}(S_i)$. By Lemma VI.1, the set $G_{\text{Sub}}^{(k)} = S_{\text{sublvl}}(\text{MCD}_i^{(k)}, \mathcal{H}_{\widetilde{W}_i}(S_i))$ is convex for any $k \in \{1, \dots, k_{\max}\}$. Let $G_{\text{CSub}}^{(0)} = \{p_i(0)\}$, and let $G_{\text{CSub}}^{(k)} = \{s \in \mathbb{R}^d \mid \exists k' \in K_{\text{cs}}(k),\ s' \in G_{\text{Sub}}^{(k')} \text{ with } \|s - s'\| \leq u_{\max}\}$, also a convex set. Since $\tilde{S}_i \in \Omega_{\text{Rg}_i}$, $\tilde{s}_i^{(k)} \in G_{\text{CSub}}^{(k)}$ for each $k \in \{1, \dots, k_{\max}\}$, and since $\mathcal{H}_{\widetilde{W}_i}(\tilde{S}_i) \leq \mathcal{H}_{\widetilde{W}_i}(S_i)$, $\tilde{s}_i^{(k)} \in G_{\text{Sub}}^{(k)}$. Making use of the similarity between the extended set formulation and the Lagrangian of the constrained one-center problem, it can be shown that $G_{\text{ESub}}^{(k)} \cap G_{\text{Sub}}^{(k)} = \{s_i^{(k)}\}$. Thus $\tilde{S}_i = S_i$ is the unique global minimum of $\mathcal{H}_{\widetilde{W}_i}$ over Ω_{Rg_i}. ■

Proof of Proposition VII.9: We begin with the critical point result. We consider three separate cases inspired by Lemma VII.5 and Proposition VII.6. First, if there is a $k \in \{1, \dots, k_{\max}\}$ with $\mathbf{0} \in \partial \text{MCD}_i^{(k)}(S_i) \subset \partial \mathcal{H}_{\widetilde{W}_i}(S_i)$, then $\{k\}$ defines a maximally centered sequence in S_i.

Second, assume that $\mathbf{0} \notin \mathcal{H}_{\mathcal{W}_i}(S_i)$, but that $\exists k \in \text{argmax}_{k' \in \{1,\dots,k_{\max}\}} \text{MCD}_{\widetilde{W}}^{(k')}(S_i)$ with $\mathbf{0} \in \partial \text{CDE}_{\max}^{(k)}(S_i)$. From Corollary A.5, it can be deduced that $\exists k' \in \text{argmax}_{l \in K_{\text{cs}}(k)} \text{CDE}_i^{(k:l)}(S_i)$ such that $\{k, k'\}$ is a maximally centered sequence.

Finally, assume that $\mathbf{0} \notin \mathcal{H}_{\mathcal{W}_i}(S_i)$ and there is no k with $\mathbf{0} \in \partial \text{CDE}_{\max}^{(k)}(S_i) \subset \partial \mathcal{H}_{\widetilde{W}_i}(S_i)$. With a slight abuse of notation, let $\text{MCD}_{\widetilde{W}}^{(k)}(S_i; K) = \max_{s \in \widetilde{W}_i^{(k)}(S_i;K)} \delta_k\big(s, s_i^{(k)}\big)$. In this case it

can be shown that $\mathbf{0} \in \partial\mathcal{H}_{\widetilde{W}_i}(S_i)$ if and only if there is a sequence $K^* \subseteq \{1, \dots, k_{\max}\}$ of two or more consecutive samples which satisfies $\mathbf{0} \in \operatorname{co}\{\text{MCD}_i^{(k)}(S_i) \mid k \in K^*\}$, and for all $k \in K^*$, $\mathbf{0} \in \pi_k(\partial\text{MCD}_{\widetilde{W}}^{(k)}(S_i; K^*))$ and $\partial\text{MCD}_{\widetilde{W}}^{(k)}(S_i; K^*) \subset \mathcal{H}_{\widetilde{W}_i}(S_i)$. It can be shown that the first two conditions are satisfied if and only if K^* defines a centered sequence, while the last requires that it be maximal.

This proves that S_i is a critical point if and only if it contains at least one maximally centered sequence. That any critical point is a global minimum follows by applying Proposition A.6 to any maximally centered sequence in S_i. ■

PROOFS AND SUPPORTING RESULTS FROM SECTION VIII

Proof of Proposition VIII.1: We use the discrete-time LaSalle invariance principle [25] to show convergence. Let $\text{T} : (\mathcal{D}^{k_{\max}})^n \to (\mathcal{D}^{k_{\max}})^n$ denote the evolution map of the GENERALIZED MULTICIRCUMCENTER ALGORITHM, i.e., $S^{\{j\}} = \text{T}(S^{\{j-1\}})$. Note that Ω is positively invariant with respect to T, and that $\mathcal{H}$ is nonincreasing along T on Ω. Since Ω is bounded, any evolution is bounded. The maps T and $\mathcal{H}$ are both continuous on Ω. By the discrete time LaSalle invariance principle, any evolution with initial condition $S^{\{0\}} \in \Omega$ must converge to M, the largest invariant set with respect to T contained in $Z = \{S \in \Omega \mid \mathcal{H}(\text{T}(S)) = \mathcal{H}(S)\} \subset \Omega$.

Now, let $M_{\min}$ denote the set of all global minimizers of $\mathcal{H}$ on Ω, and note that $M_{\min} \subseteq M$. We reason by contradiction to show that $M_{\min} = M$. Assume that there is a trajectory, $S^{\{0\}} \in M \setminus M_{\min}$. Since $M \subset Z$, we have $\mathcal{H}(S^{\{1\}}) = \mathcal{H}(S^{\{0\}})$. Consider the fixed-partition optimization at step 0. Let $\mathcal{W} = \mathcal{MC}(S^{\{0\}})$, and let $i \in \operatorname{argmax}_{i' \in \{1,\dots,n\}} \mathcal{H}_{\mathcal{W}_i}(S_i^{\{0\}})$. Since $S^{\{0\}}$ is not a global minimizer of $\mathcal{H}$ over Ω, it is not a global minimizer of $\mathcal{H}_{\widetilde{\mathcal{W}}}$ over Ω, thus $S_i^{\{0\}}$ is not a global minimizer of $\mathcal{H}_{\widetilde{W}_i}$ over Ω_i. On the other hand, $S_i^{\{1\}}$ is a global minimizer of $\mathcal{H}_{\widetilde{W}_i}$, and we have $\mathcal{H}_{\widetilde{W}_i}(S_i^{\{1\}}) < \mathcal{H}_{\widetilde{W}_i}(S_i^{\{0\}})$. This is true for all such i, thus $\mathcal{H}_{\widetilde{\mathcal{W}}}(S^{\{1\}}) < \mathcal{H}_{\widetilde{\mathcal{W}}}(S^{\{0\}})$. By Lemma VII.10 and Proposition V.2, we can write, $\mathcal{H}_{\mathcal{W}}(S^{\{0\}}) > \mathcal{H}_{\mathcal{W}}(S^{\{1\}}) \geq \mathcal{H}(S^{\{1\}})$. Thus $\mathcal{H}(S^{\{0\}}) > \mathcal{H}(S^{\{1\}})$, which contradicts the assumption that $S^{\{0\}} \in Z$. Therefore $M_{\min} = M$, and the result follows. ■