# Generalized multicircumcenter trajectories for optimal design under near-independence

Rishi Graham Jorge Cortés

*Abstract*— **This work deals with trajectory optimization for a network of robotic sensors sampling a spatio-temporal random field. We examine the problem of minimizing over the space of network trajectories the maximum predictive variance of the estimator. 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 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 problem as generalized multicircumcenter trajectories, and provide distributed algorithms to find them.**

## I. Introduction

Intelligent data collection is an exciting field with many scientific, industrial, and safety applications. Path planning, either a priori or online, is an important part of any data collection mission. In this paper, we examine optimal trajectories for sampling a spatio-temporal random field modeled as a Gaussian process. We assume that the mean and covariance of the field are known, and concentrate on minimizing the maximum predictive variance.

*Literature review:* There is a rich literature on the use of model uncertainty to drive the placement of sensing devices, e.g., [1], [2], [3]. 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 [4]. In cooperative control, various works consider mobile sensor networks performing spatial estimation tasks. [5], [6] consider deterministic models with a stochastic measurement error term. [7] addresses the multiple robot path planning problem by choosing way points from a discrete set of possible sensing locations. In [8], a deterministic model is used, where the random elements come as unknown model parameters, and localization error is included. The work [9] uses a Gaussian process model where all information is globally available via all-to-all communication. [10] considers optimal sampling trajectories from a parameterized set of paths. [11] discusses the tracking of level curves in a noisy scalar field.

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`

*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 its minimization 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. 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. We show that these constrained multicenter trajectories optimize the correlation distance disk-covering problem over the set of distance-constrained trajectories. Finally, we present a version of Lloyd's algorithm which enables the network to arrive at locally optimal trajectories. This may be performed at any step of the experiment to optimize the remainder of the trajectories as new information arrives. For reasons of space, all proofs are omitted and will appear elsewhere.

## II. Preliminaries

We present here some useful notation. Let $\mathbb{R}$, $\mathbb{R}_{>0}$ and $\mathbb{R}_{\geq 0}$ denote the set of reals, positive reals and nonnegative reals, respectively. Given $\mathcal{D} \subset \mathbb{R}^d$, $d \in \mathbb{N}$, we use the shorthand notation $\mathcal{D}_e = \mathcal{D} \times \mathbb{R}_{\geq 0}$. For $p \in \mathbb{R}^d$ and $r \in \mathbb{R}_{>0}$, let $\overline{B}(p, r)$ denote the *closed ball* of radius $r$ centered at $p$. For a set $W$, we denote by $|W|$, $\operatorname{bnd}(W)$, $\operatorname{int}(W)$, and $\operatorname{co}(W)$ its cardinality, boundary, interior, and convex hull, respectively. A *convex polytope* is the convex hull of a finite point set. For a bounded set $W \subset \mathbb{R}^d$, $\operatorname{CC}(W)$ denotes the *circumcenter* of $W$, i.e., the center of the smallest-radius $d$-sphere enclosing $W$. Finally, $\mathfrak{P}(W)$ denotes the collection of subsets of $W$.

### A. Nonsmooth analysis

Here we present some useful notions from nonsmooth analysis following [12]. 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 $\epsilon$

such that $|f(y)-f(y')| \leq L_s\|y-y'\|$ for all $y, y' \in \overline{B}(s,\epsilon)$. $f$ is *locally Lipschitz on* $W \subseteq \mathbb{R}^d$ if it is locally Lipschitz at $s$, for all $s \in W$. $f : \mathbb{R}^d \to \mathbb{R}$ 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. For a given closed, convex set $G \subset \mathbb{R}^d$, let $N_G(x) = \{y \in \mathbb{R}^d \mid \ < y, x-z > \geq 0,\ \forall z \in G\}$ be the normal cone of $G$ at $x$.

### B. Spatio-temporal simple kriging

Let $Z$ denote a spatio-temporal process taking values on a convex polytope $\mathcal{D} \subset \mathbb{R}^d$ of the form

$$Z(s,t) = \mu(s,t) + \omega(s,t), \quad (s,t) \in \mathcal{D}_e, \tag{1}$$

where $\mu$ is a known mean value function, and $\omega$ is a zero mean random space-time process with known separable covariance, which exhibits second-order stationarity and isotropy in the spatial dimensions, i.e.,

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

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 $0$.

Let $n \in \mathbb{Z}_{>0}$ sensing agents take samples at a sequence of discrete timesteps $\{1, \dots, k_{\text{max}}\}$, $k_{\text{max}} \in \mathbb{Z}_{>0}$. Let $S_i = (s_i^{(1)}, \dots, s_i^{(k_{\text{max}})})^T \in \mathcal{D}^{k_{\text{max}}}$ denote the spatial locations of samples taken over the course of the experiment by the $i$th agent, and let $S = (S_1^T, \dots, S_n^T)^T \in (\mathcal{D}^{k_{\text{max}}})^n$ denote the locations of all samples taken by the network. Let $I_{\text{samp}} = \{1, \dots, n\} \times \{1, \dots, k_{\text{max}}\}$. 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 notation $(a_1^{(1)}, \dots, a_n^{(k_{\text{max}})}) = (a_1^{(1)}, \dots, a_1^{(k_{\text{max}})}, \dots, a_n^{(1)}, \dots, a_n^{(k_{\text{max}})})$. Let $Y = (y_1^{(1)}, \dots, y_n^{(k_{\text{max}})})^T \in (\mathbb{R}^{k_{\text{max}}})^n$ denote the values of all samples taken at locations $S$. We assume that the data are corrupted with a measurement error so that,

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

where $\tau^2 > 0$. The covariance between $y_i^{(k)}$ and $y_j^{(l)}$ is

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

Let $\mathbf{\Sigma} = \Sigma(S)$ denote the covariance matrix of $Y$, where bold face is used to denote explicit dependence on $S$.

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

$$\hat{p}_{\text{SK}}(s,t;Y) = \mu(s,t) + \boldsymbol{c}^T \mathbf{\Sigma}^{-1}(Y - \boldsymbol{\mu}), \tag{3}$$

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

$$\sigma^2(s,t;S) = g_0 g_s(0) g_t(t,t) - \boldsymbol{c}^T \mathbf{\Sigma}^{-1} \boldsymbol{c}. \tag{4}$$

$\sigma^2$ 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 $\{R_1, \dots, R_n\}$ of $n \in \mathbb{Z}_{>0}$ robotic sensing agents taking measurements of a spatio-temporal process of interest over a convex polytope $\mathcal{D} \subset \mathbb{R}^d$, for $d > 1$. The robots take point measurements of the random process at discrete instants of time in $\mathbb{Z}_{>0}$. Our results below are independent of the particular robot dynamics, so long as each agent is able to move up to a distance $u_{\text{max}} \in \mathbb{R}_{>0}$ between consecutive sampling times.

### B. Objective function for spatial estimation

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 $k_{\text{max}}$ generated from samples taken up to that time. Here, we consider an objective function inspired by the notion of G-optimality from optimal design [13], [2]. The *maximum error variance* $\mathcal{M} : (\mathcal{D}^{k_{\text{max}}})^n \to \mathbb{R}$ of estimates made at time $k_{\text{max}}$ over the region $\mathcal{D}$ is

$$\begin{aligned} \mathcal{M}(S) &= \max_{s \in \mathcal{D}} \sigma^2((s, k_{\text{max}}); S) \\ &= g_0 g_s(0) g_t(k_{\text{max}}, k_{\text{max}}) - \min_{s \in \mathcal{D}} \left\{ \boldsymbol{c}^T \mathbf{\Sigma}^{-1} \boldsymbol{c} \right\}. \end{aligned} \tag{5}$$

Note that $\mathcal{M}$ corresponds to a "worst-case scenario," where we deal with locations at which the error variance of the LUMVE is maximal. Our goal is to find the sampling trajectories $S \in (\mathcal{D}^{k_{\text{max}}})^n$ that minimize this objective function.

## 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 $k_{\text{max}} = 1$, the optimization of $\mathcal{M}$ is known to be NP-hard over discrete spaces [4]. In this section we consider instead the optimization of $\mathcal{M}$ when the correlation function is raised to the power $\alpha \in \mathbb{R}_{>0}$, i.e., $(g_s g_t)^\alpha$. This correlation function retains much of the shape of the original correlation (e.g., smoothness, range, etc). This asymptotic regime of increasingly smaller correlation between distinct points as $\alpha$ grows is known as *near-independence*, see [14]. To ease the exposition, we denote by $\boldsymbol{c}^{\{\alpha\}}$, resp. $\mathbf{\Sigma}^{\{\alpha\}}$, the vector $\boldsymbol{c}$, resp. the matrix $\mathbf{\Sigma}$, with the correlation in each element raised to the power $\alpha$. Similarly, let $\mathcal{M}^{\{\alpha\}} : (\mathcal{D}^{k_{\text{max}}})^n \to \mathbb{R}$ be defined as

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

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 $w : \{1, \dots, k_{\text{max}}\} \to \mathbb{R}_{\geq 0}$ by,

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

The function $\phi$ 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, $\delta_k : \mathcal{D} \times \mathcal{D} \to \mathbb{R}_{\geq 0}$,

$$\delta_k(s_1, s_2) = \phi(\|s_1 - s_2\|) + w(k). \tag{6}$$

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

Let $S_{\text{unique}}$ 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,

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

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 $\alpha$ grows.

**Theorem IV.1 (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*

$$\mathcal{H}(S) = \max_{s \in \mathcal{D}} \Big\{ \min_{(i,k) \in I_{\text{samp}}} \{\delta_k(s, s_i^{(k)})\} \Big\}. \tag{7}$$

*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 $\Omega$. Further assume that $S_{mcc} \in S_{unique}$. Then, as $\alpha \to \infty$, $S_{mcc}$ asymptotically globally optimizes $\mathcal{M}^{\{\alpha\}}$ over $\Omega$, that is, $\mathcal{M}^{\{\alpha\}}(S_{mcc})$ approaches a global minimum over $\Omega$.*

The generality of $\Omega$ in Theorem IV.1 allows us to apply the result to two situations of particular importance. First, we may restrict the samples to feasible trajectories based on limitations on the agents' motion, and their initial positions, which we call *anchor points*. 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

$$\Omega_{\text{Rg}_i} = \big\{(s_i^{(1)}, \dots, s_i^{(k_{\max})})^T \in \mathcal{D}^{k_{\max}} \mid \|s_i^{(1)} - p_i(0)\| \leq u_{\max} \text{ and } \|s_i^{(k)} - s_i^{(k-1)}\| \leq u_{\max},\ \forall k \in \{2, \dots, k_{\max}\}\big\}. \tag{8}$$

Second, a change in mission parameters at time $k-1$, $k \in \{2, \dots, k_{\max}\}$, 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

$$\Omega_{\text{Rg}_i}^{(\geq k)} = \big\{(s_i^{(k)}, \dots, s_i^{(k_{\max})})^T \in \mathcal{D}^{k_{\max}-k+1} \mid \|s_i^{(k)} - p(k-1)\| \leq u_{\max} \text{ and } \|s_i^{(k')} - s_i^{(k'-1)}\| \leq u_{\max},\ \forall k' \in \{k+1, \dots, k_{\max}\}\big\}. \tag{9}$$

Theorem IV.1 shows that the optimization of the maximum error variance is equivalent to a geometric optimization problem in the near-independence range.

## V. Maximal correlation partition

Here, we introduce the maximal correlation partition associated to a network trajectory. A partition of $\mathcal{D}$ is a collection of compact subsets, $\mathcal{W} = \{W_1^{(1)}, \dots, W_n^{(k_{\max})}\}$ with disjoint interiors whose union is $\mathcal{D}$. For any $S \in S_{\text{unique}}$, let $\mathcal{MC}(S) = (\text{MC}_1^{(1)}(S), \dots, \text{MC}_n^{(k_{\max})}(S))$ denote the *maximal correlation partition* defined by

$$\text{MC}_i^{(k)}(S) = \big\{s \in \mathcal{D} \mid \delta_k(s, s_i^{(k)}) \leq \delta_l(s, s_j^{(l)}),\ \forall (j,l) \neq (i,k)\big\}. \tag{10}$$

This partition corresponds to a generalized Voronoi partition [15] for distance measure $\phi$ and weights given by $w$. In general, the maximal correlation regions are neither convex nor star-shaped. Depending on the weights and locations, $\text{MC}_i^{(k)}(S)$ might be empty for some $i$. Let $\text{I} : \mathfrak{P}(\mathcal{D}) \to \{1, \dots, n * k_{\max}\}$ map a partition to the number of nonempty cells it contains, which we term the *index* of the partition.

For $S \in S_{\text{unique}}$, the correlation distance disk-covering function can be restated as,

$$\mathcal{H}(S) = \max_{(i,k) \in I_{\text{samp}}} \Big\{ \max_{s \in \text{MC}_i^{(k)}(S)} \{\delta_k(s, s_i^{(k)})\} \Big\}. \tag{11}$$

This expression is important because it 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 $W = \{W_1^{(1)}, \dots, W_n^{(k_{\max})}\} \subset \mathfrak{P}(\mathcal{D})$ of the predictive space, define $\mathcal{H}_{\mathcal{W}} : (\mathcal{D}^{k_{\max}})^n \to \mathbb{R}$ by

$$\mathcal{H}_{\mathcal{W}}(S) = \max_{\substack{(i,k) \in I_{\text{samp}} \\ W_i^{(k)} \neq \emptyset}} \Big\{ \max_{s \in W_i^{(k)}} \{\delta_k(s, s_i^{(k)})\} \Big\}. \tag{12}$$

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.

**Proposition V.1 ($\mathcal{H}$-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 $\text{I}(\mathcal{W}) \leq \text{I}(\mathcal{MC}(S))$, one has that $\mathcal{H}(S) \leq \mathcal{H}_{\mathcal{W}}(S)$, 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.1 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

Our objective is to characterize the optimal network trajectories of $\mathcal{H}_{\mathcal{W}}$ for a fixed partition $\mathcal{W}$ of $\mathcal{D}$. Rewrite (12) as

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

In the following result, let $\overline{\mathrm{CC}}(W, s) = \mathrm{CC}(W)$ if $W \neq \emptyset$, and $\overline{\mathrm{CC}}(W, s) = s$ otherwise, and let $\overline{\mathrm{CC}}(\mathcal{W}, S) = \big(\overline{\mathrm{CC}}(W_1^{(1)}, s_1^{(1)}), \ldots, \overline{\mathrm{CC}}(W_n^{(k_{\max})}, s_n^{(k_{\max})})\big)^T$ denote a vector of such circumcenter locations.

**Proposition VI.1 ($\mathcal{H}_{\mathcal{W}}$-optimal trajectories)** *For $S = (s_1^{(1)}, \ldots, s_n^{(k_{max})})^T \in S_{unique}$, a partition $\mathcal{W} = \{W_1^{(1)}, \ldots, W_n^{(k_{max})}\} \subset \mathfrak{P}(\mathcal{D})$ of $\mathcal{D}$, and $\tilde{S} = (\tilde{s}_1^{(1)}, \ldots, \tilde{s}_n^{(k_{max})})^T \in (\mathcal{D}^{k_{max}})^n$, one has that $\mathcal{H}_{\mathcal{W}}\big(\overline{\mathrm{CC}}(\mathcal{W}, \tilde{S})\big) \leq \mathcal{H}_{\mathcal{W}}(S)$, i.e., the circumcenter locations $\overline{\mathrm{CC}}(\mathcal{W}, \tilde{S})$ are optimal for $\mathcal{H}_{\mathcal{W}}$.*

The combination of Propositions V.1 and VI.1 allows us provide the following characterization of the optimizers of $\mathcal{H}$.

**Proposition VI.2 (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 $(i, k) \in I_{\mathrm{samp}}$ 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. *Moreover, 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

Here, our objective is to characterize the optimizers of $\mathcal{H}_{\mathcal{W}}$ over $\Omega_{\mathrm{Rg}}$ for a fixed partition $\mathcal{W}$. Let $\mathcal{W}_i = \{W_i^{(1)}, \ldots, W_i^{(k_{\max})}\}$ denote the elements of $\mathcal{W}$ assigned to the samples in the trajectory of $R_i$. We may write

$$\mathcal{H}_{\mathcal{W}}(S) = \max_{i \in \{1,\ldots,n\},\ \mathcal{W}_i \neq \emptyset} \mathcal{H}_{\mathcal{W}_i}(S_i),$$

where $\mathcal{H}_{\mathcal{W}_i}(S_i) = \max_{k \in \{1,\ldots,k_{\max}\},\ W_i^{(k)} \neq \emptyset} \left\{\mathrm{MCD}_i^{(k)}(s_i^{(k)})\right\}$. 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 shows that, for a fixed partition, minimizing $\mathcal{H}_{\mathcal{W}}$ over the space of network trajectories is equivalent to (independently) minimizing each $\mathcal{H}_{\mathcal{W}_i}$ over the space of trajectories of the robot $R_i$. Hence, we structure our discussion in three parts. First, we deal with the single sample problem. Then, we find an optimal sampling trajectory for a *single* agent. Finally, we combine individual agent trajectories into a network trajectory.

### A. Single sample constrained problem

We consider the single sample problem over a general closed convex constraint set.

**Proposition VII.1 (Constrained minimizers of $\mathrm{MCD}_i^{(k)}$)** *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 $\mathrm{MCD}_i^{(k)}$ over $\Gamma$ iff $\mathbf{0} \in \partial\mathrm{MCD}_i^{(k)}(s^*) + N_\Gamma(s^*)$.*

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

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

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. Define $\Gamma^{(k)} : \mathcal{D}^{k_{\max}} \to \mathfrak{P}(\mathbb{R}^d)$,

$$\Gamma^{(k)}(S_i) = \bigcap_{s \in S_{\mathrm{cs}}(k, S_i)} \overline{B}(s, u_{\max}). \tag{13}$$

The set $\Gamma^{(k)}(S_i)$ corresponds to $\Omega_{\mathrm{Rg}}$ 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 $u_{\max}$.

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

$$\mathrm{EPt}^{(k:k')}(S_i) = s_i^{(k)} + r_k(\mathcal{H}_{\mathcal{W}_i}(S_i)) \frac{s_i^{(k')} - s_i^{(k)}}{u_{\max}}, \tag{14}$$

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 $\mathcal{H}_{\mathcal{W}}(S_i) \geq \mathrm{MCD}_i^{(k)}(s_i^{(k)})$. The location $\mathrm{EPt}^{(k:k')}(S_i)$ 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 $\widetilde{W}_i^{(k)} : \mathcal{D}^{k_{\max}} \to \mathfrak{P}(\mathbb{R}^d)$, $(i, k) \in I_{\mathrm{samp}}$ be the *constraint extended predictive set*,

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

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. Figure 1 shows an example of the extended predictive set.

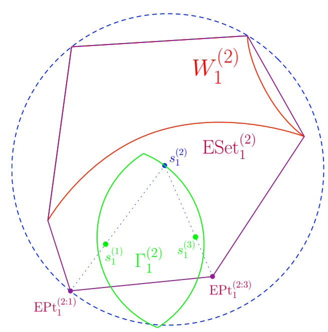



Fig. 1. A two-dimensional example of the extended center representation of a critical point of the constrained problem. The dashed circle is the circumcircle of $\widetilde{W}_1^{(2)}$, with circumcenter $s_1^{(2)}$. Note that $s_1^{(2)}$ is on the boundary of $\Gamma^{(2)}$ formed by $s_1^{(1)}$, and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

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

**Proposition VII.2 (Extended circumcenter minimizes $\mathrm{MCD}_i^{(k)}$ 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 $\mathcal{H}_{\mathcal{W}_i}(S_i) = \mathrm{MCD}_i^{(k)}(s_i^{(k)})$. Then $s_i^{(k)}$ is the unique minimizer of $\mathrm{MCD}_i^{(k)}$ over $\Gamma^{(k)}(S_i)$ iff $s_i^{(k)} = \mathrm{CC}\left(\widetilde{W}_i^{(k)}(S_i)\right)$.*

*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 $\Omega_{\mathrm{Rg}_i}$ in (8) in terms of *centered* sub-sequences. To ease the exposition, let $\mathrm{d}^{(k:k')} : \mathcal{D}^{k_{\max}} \to \mathbb{R}_{\geq 0}$, $\mathrm{d}^{(k:k')}(S_i) = \|s_i^{(k)} - s_i^{(k')}\|$. We use $\widetilde{W}_i^{(k)}(S_i; K_C) = \mathrm{co}\big(W_i^{(k)}, \{\mathrm{EPt}^{(k:k')}(S_i) \mid k' \in K_{\mathrm{cs}}(k) \cap K_C\}\big)$ 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 $K_C \subseteq \{1, \dots, k_{max}\}$ 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.,*

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

*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.

Define the constrained objective function for an agent,

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

where $\mathrm{MCD}_{\widetilde{W}}^{(k)}(S_i) = \max_{s \in \widetilde{W}_i^{(k)}(S_i)} \delta_k\big(s, s_i^{(k)}\big)$. Note that $\mathcal{H}_{\widetilde{W}_i}$ may be calculated entirely by $R_i$. Moreover, for $S_i \in \Omega_{\mathrm{Rg}_i}$, it holds that $\mathcal{H}_{\widetilde{W}_i}(S_i) = \mathcal{H}_{\mathcal{W}_i}(S_i)$.

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

**Lemma VII.4 (Maximal elements define sub-sequences within centered sequences)** *Let $K_C \subseteq \{1, \dots, k_{max}\}$ define a centered sequence of samples in $S_i$ with $\max_{k \in K_C} \mathrm{MCD}_i^{(k)}(s_i^{(k)}) = \mathcal{H}_{\mathcal{W}_i}(S_i)$. Then there is a sub-sequence, $K_{MC} \subseteq K_C$ which is centered and such that every $k \in K_{MC}$ satisfies $\mathrm{MCD}_{\widetilde{W}}^{(k)}(s_i^{(k)}) = \mathcal{H}_{\mathcal{W}_i}(S_i)$. We refer to a sequence such as $K_{MC}$ as* maximally centered.

**Proposition VII.5 (Global minimizers of $\mathcal{H}_{\widetilde{W}_i}$ on $\Omega_{\mathrm{Rg}_i}$ contain maximally centered sequences)** *A trajectory $S_i \in \Omega_{\mathrm{Rg}_i}$ is a critical point of $\mathcal{H}_{\widetilde{W}_i}$ iff it contains at least one maximally centered sequence of samples. Furthermore, any such critical point globally minimizes $\mathcal{H}_{\mathcal{W}_i}$ on $\Omega_{\mathrm{Rg}_i}$.*

*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

$$\mathcal{H}_{\widetilde{\mathcal{W}}}(S) = \max_{i \in \{1,\dots,n\}} \mathcal{H}_{\widetilde{W}_i}(S_i). \tag{15}$$

Note that $\mathcal{H}_{\widetilde{\mathcal{W}}}(S) = \mathcal{H}_{\mathcal{W}}(S)$ for $S \in \Omega_{\mathrm{Rg}}$. Next, we characterize the critical points of $\mathcal{H}_{\widetilde{\mathcal{W}}}$.

**Proposition VII.6 (Global minima of $\mathcal{H}_{\widetilde{\mathcal{W}}}$ on $\Omega_{\mathrm{Rg}}$ 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 $i \in \mathrm{argmax}_{i \in \{1,\dots,n\}} \mathcal{H}_{\mathcal{W}_i}(S_i)$ 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 $\Omega_{\mathrm{Rg}}$*

Proposition VII.6 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.5), then the resulting network trajectory is optimal. Along with Proposition V.1, this allows the following result on the optimal trajectories of the correlation disk-covering function $\mathcal{H}$ over $\Omega_{\mathrm{Rg}}$.

**Proposition VII.7 (Range-constrained generalized multicircumcenter trajectory)** *Let $S = (S_1^T, \dots, S_n^T) \in (\mathcal{D}^{k_{max}})^n$ 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 $\Omega_{\mathrm{Rg}}$. 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 $\Omega_{\mathrm{Rg}}$.*

The following results allows for partial optimization of trajectories which are already under way, based on minimizing the maximum error *over the remainder of the experiment*.

**Proposition VII.8 (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 $\Omega_{\mathrm{Rg}}^{(\geq k^*)}$. 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. 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 $\Omega_{\mathrm{Rg}}^{(\geq k^*)}$ for some $k^* \in \{1, \dots, k_{\max}\}$. With slight adjustments, the same algorithm works for the unconstrained case. Figure 2 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.

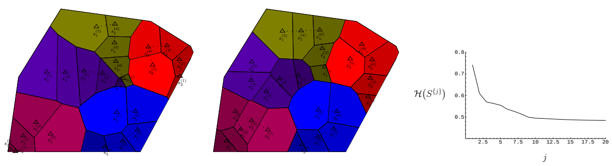
**Proposition VIII.1** *The* Generalized Multicircumcenter Algorithm *is distributed over the partition $\mathcal{MC}(S^{\{j\}})$, meaning that at step $j+1$, $R_i$ need only communicate with $R_{i'}$ for each $i' \in \{1, \dots, n\}$ such that $MC_i^{(k)}(S^{\{j\}})$ adjacent to $MC_{i'}^{(k')}(S^{\{j\}})$ for some $k, k'$. Furthermore, $S^{\{j\}} \in \Omega_{\mathrm{Rg}}^{(\geq k^*)}$, 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 $\Omega_{\mathrm{Rg}}^{(\geq k^*)}$.*

**Goal:** Find a minimum of $\mathcal{H}$ over $\Omega_{\text{Rg}}^{(\geq k^*)}$
**Input:** (i) Sample interval $[k^*, k_{\text{max}}]$
(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 $i$th *agent* trajectory

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



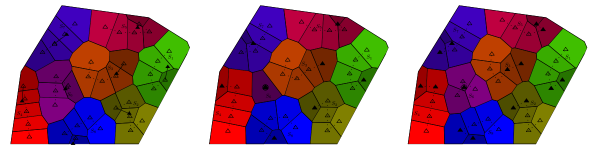
(a) (b) (c)

Fig. 2. 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 $j$.

We finish by discussing an adaptive approach to optimal path planning. Before moving to take the $k$th sample, the network might receive new 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. Figure 3 depicts an illustrative example of this procedure.

## IX. CONCLUSIONS

We have considered a robotic sensor network taking samples of a spatio-temporal process. The criteria for optimization has been 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 established its optimality properties with respect to the disk-covering function. We have introduced the novel notion of multicircumcenter trajectories and established their optimality for unconstrained and constrained versions of the disk-covering optimization problem. On the design front, we have synthesized distributed strategies that allow the network to calculate an optimal trajectory. Future work will include the study of more complex predictive regions and of alternative optimality criteria.



(a) (b) (c)

Fig. 3. Sequential implementation of GENERALIZED MULTICIRCUMCENTER ALGORITHM with $n = 8$ robots, $k_{\text{max}} = 5$ steps, and Gaussian correlation. In (a), the trajectory is calculated from the initial anchor points. 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 plot shows the result of the GENERALIZED MULTICIRCUMCENTER ALGORITHM as run by the remaining 7 agents over timesteps $\{2,\dots,k_{\text{max}}\}$. 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 $\{3,\dots,k_{\text{max}}\}$ 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 $\{k^*,\dots,k_{\text{max}}\}$ are empty triangles, with dashed lines connecting each agent trajectory. The last sample location of the dropped agent is circled. The color convention is the same as in Figure 2.

## ACKNOWLEDGEMENTS

This work was partially supported by NSF CAREER award ECS-0546871 and NSF award CCF-0917166.

## REFERENCES

[1] K. Chaloner and I. Verdinelli, “Bayesian experimental design, a review,” *Statistical Science*, vol. 10, no. 3, pp. 273–304, 1995.
[2] F. Pukelsheim, *Optimal Design of Experiments*, vol. 50 of *Classics in Applied Mathematics*. Philadelphia, PA: SIAM, 2006.
[3] 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.
[4] 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.
[5] 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.
[6] 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.
[7] 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.
[8] 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.
[9] 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.
[10] 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.
[11] F. Zhang and N. E. Leonard, “Cooperative filters and control for cooperative exploration,” *IEEE Transactions on Automatic Control*, vol. 55, pp. 650–663, March 2010.
[12] F. H. Clarke, *Optimization and Nonsmooth Analysis*. Canadian Mathematical Society Series of Monographs and Advanced Texts, Wiley, 1983.
[13] N. A. C. Cressie, *Statistics for Spatial Data*. New York: Wiley, 1993. revised edition.
[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] 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.