Asymptotic optimality of multicenter Voronoi configurations for random field estimation

Rishi Graham Jorge Cortés

***Abstract*—This paper deals with multi-agent networks performing estimation tasks. Consider a network of mobile agents with sensors that can take measurements of a spatial stochastic process. Using a statistical technique known as kriging, a field estimate may be calculated over the environment, with an associated error variance at each point. We study a single-snapshot scenario, in which the spatial process mean is known and each agent can only take one measurement. We consider two optimization problems with respect to the measurement locations, using as objective functions the maximum error variance and the extended prediction variance. We show that, as the correlation between distinct locations vanishes, circumcenter and incenter Voronoi configurations become network configurations that optimize the maximum error variance and the extended prediction variance, respectively. We also present distributed coordination algorithms that steer the network towards these configurations. The technical approach draws on tools from geostatistics, computational geometry, linear algebra, and dynamical systems.**

I. INTRODUCTION

Mobile sensor networks are envisioned to perform distributed sensing and data fusion tasks in a wide range of scenarios, including environmental monitoring, oceanographic research, and distributed surveillance. This paper considers sensor networks taking measurements of physical processes modeled as spatial random fields. Standard interpolation techniques produce estimates of the field at each point of the environment of interest, along with a measure of the accuracy of the estimate. In this paper, we consider the problem of where to place the agents when a single measurement is to be taken by each. Addressing this problem is an initial step towards the more ambitious goal of characterizing optimal coordinated agent trajectories when multiple measurements are possible. We assume that the mean of the process is known, and we study the limiting case of near independence between distinct locations. The assumption of near independence has been suggested as a first step in gathering data in a relatively large space [1]. Our results show that the solution of the single-snapshot scenario is both elegant and technically challenging. We make no assertion that a correlated spatial field is accurately modeled by near-independence. This asymptotic assumption merely provides an analytical framework that justifies the intuitive notion of space filling design, which is surprisingly difficult to prove optimal in general.

This version: September, 2008

Rishi Graham is with the Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA 95064, USA, rishig@ams.ucsc.edu

Jorge Cortés is with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA 92093, USA, cortes@ucsd.edu

Literature review: Kriging [2], [3] is a standard geostatistical technique that produces estimates of spatial processes based on data collected at a finite number of locations. An advantage of kriging over other spatial interpolation methods is that it provides a measure of the uncertainty associated to the estimator. The optimal design literature [4], [5] deals with the problem of designing experiments to optimize the resulting statistical estimation. Of particular interest are the notions of G-optimality, minimizing the maximum error variance, and D-optimality, minimizing the generalized variance of the estimator. The work [6] introduces performance metrics for optimal estimation in oceanographic research. The works [7], [8] propose distributed optimal estimation strategies for deterministic fields, when the sensor measurements taken by individual agents are uncorrelated. The only source of uncertainty is the stochastic measurement errors. In [9], the emphasis is on finding optimal agent trajectories over a given interval of time among a parameterized set of trajectories. Here, instead, we focus on optimal network configurations for the estimation of the random field at a single snapshot. In our technical approach, we have been inspired by [1], which considers the problem of minimizing the maximum uncertainty over a discrete space and shows that minimax configurations are asymptotically optimal as the correlation between any two distinct points vanishes. Minimax configurations minimize the maximum distance to the nearest agent from any point in space. We make the connection to Voronoi partitions of continuous spaces, which are a classical notion in computational geometry [10]. The work [11] defines circumcenter and incenter Voronoi configurations and proposes coordination algorithms which steer the network to these configurations.

Statement of contributions: In this paper, we consider two performance metrics for optimal placement of sensor networks based on kriging. Kriging produces a Linear Unbiased Minimum Variance Estimator (LUMVE) of the random field at any location. We first characterize the continuity properties of the error variance of the estimator as a function of the network configuration. In the case of zero measurement error, this is not trivial. Previous results in the optimal design literature have avoided this problem by optimizing over a discrete set of possible configurations, while we consider the continuous space of all agent locations within the region. Next, we define our first optimality criterion, the maximum error variance of the estimator as a function of network configuration. This gives a measure of the worst-case estimate over the region. We study its critical points asymptotically, as the correlation between any two distinct points vanishes. We define a second optimality criterion, the extended prediction variance of the estimator as a novel form of D-optimality. This criterion gives a measure of the overall information provided by the estimator. We introduce a method for applying this criterion to a bounded region. We study the critical points of this function within the same asymptotic framework as the first. Our main results show that circumcenter, respectively incenter, Voronoi configurations are asymptotically optimal for the maximum error variance over the environment, respectively the extended prediction variance. In general, these objective functions pose nonconvex and high-dimensional optimization problems. In addition, the first criterion is nonsmooth. For these reasons, it is difficult to obtain exactly the configurations that optimize them. Our results are relevant to the extent that they guarantee that, for scenarios with small enough correlation between distinct points, circumcenter and incenter Voronoi configurations are optimal for appropriate measures of uncertainty. The network can achieve these configurations by executing simple distributed dynamical systems.

Organization: Section II introduces basic computational geometric notions and presents an overview of kriging. Section III states the problem of interest. We present our main results in Section IV on the optimality of circumcenter and incenter Voronoi configurations. Section V presents simulations to illustrate our results. Finally, Section VI gathers our conclusions and ideas for future work. Some proofs have been omitted for brevity, and can be found in [12].

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. We are concerned with operations on a compact and connected set $\mathcal{D}$ of Euclidean space $\mathbb{R}^d$, $d \in \mathbb{N}$. We denote by $\mathcal{D}^n$ the Cartesian product of n copies of $\mathcal{D}$. For $p, q \in \mathbb{R}^d$, we let $]p, q[= \{\lambda p + (1-\lambda)q \mid \lambda \in]0, 1[\}$ denote the *open segment* with extreme points p and q. 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. We denote by $|S|$ and ∂S the cardinality and the boundary of a set S, respectively. A *convex polytope* is the convex hull of a finite point set. For a bounded set $S \subset \mathbb{R}^d$, we let $\mathrm{CC}(S)$ and $\mathrm{CR}(S)$ denote the *circumcenter* and *circumradius* of S, respectively, that is, the center and radius of the smallest-radius d-sphere enclosing S. The incenter set of S, denoted $\mathrm{IC}(S)$, is the set of the centers of maximum-radius d-spheres contained in S. The inradius of S, denoted by $\mathrm{IR}(S)$, is the common radius of these spheres.

We consider tuples or ordered sets of possibly coincident points, $P = (p_1, \dots, p_n) \in (\mathbb{R}^d)^n$. We refer to such an element as a *configuration*. Let $\mathfrak{P}(S)$ (respectively $\mathbb{F}(S)$) denote the collection of subsets (respectively, finite subsets) of S. We denote an element of $\mathbb{F}(\mathbb{R}^d)$ by $\mathcal{P} = \{p_1, \dots, p_n\} \subset \mathbb{R}^d$, where $p_1, \dots, p_n$ are distinct points in $\mathbb{R}^d$. 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 $P = (p_1, \dots, p_n)$. Let S_{coinc} be the set of all tuples in $(\mathbb{R}^d)^n$ which contain at least one coincident pair of points,

$$S_{\text{coinc}} = \{(p_1, \dots, p_n) \in (\mathbb{R}^d)^n \mid p_i = p_j \text{ for some } i, j \in \{1, \dots, n\}, i \neq j\}.$$

Let $\|\cdot\|$ denote the Euclidean distance function on $\mathbb{R}^d$. Define the distance $\mathrm{d} : \mathbb{R}^d \times \mathfrak{P}(\mathcal{D}) \to \mathbb{R}$ from a point in $\mathbb{R}^d$ to a set of points in $\mathcal{D}$ by $\mathrm{d}(s, \mathcal{P}) = \inf_{p \in \mathcal{P}} \{\|s - p\|\}$, and let $\mathrm{mds} : \mathbb{R}^d \times \mathfrak{P}(\mathcal{D}) \to \mathfrak{P}(\mathcal{D})$ be the *minimum distance set* map, $\mathrm{mds}(s, \mathcal{P}) = \{p \in \mathcal{P} \mid \|s - p\| = \mathrm{d}(s, P)\}$.

A. Voronoi partitions and multicenter problems

Here we present some relevant concepts on Voronoi diagrams [10], [13]. A *partition* of $\mathcal{D}$ is a collection of n polygons $\mathcal{W} = \{W_1, \dots, W_n\}$ with disjoint interiors whose union is $\mathcal{D}$. The *Voronoi partition* $\mathcal{V}(P) = (V_1(P), \dots, V_n(P))$ of $\mathcal{D}$ generated by $P = (p_1, \dots, p_n)$ is defined by

$$V_i(P) = \{q \in \mathcal{D} \mid \|q - p_i\| \leq \|q - p_j\|, \ \forall j \neq i\}.$$

We say that P is a *circumcenter Voronoi configuration* if $p_i = \mathrm{CC}(V_i(P))$, for all $i \in \{1, \dots, n\}$, and that P is an *incenter Voronoi configuration* if $p_i \in \mathrm{IC}(V_i(P))$, for all $i \in \{1, \dots, n\}$. An incenter Voronoi configuration is *isolated* if it has a neighborhood in $\mathcal{D}^n$ which does not contain any other incenter Voronoi configuration. Figure 1 shows examples of these configurations.

Consider the *disk-covering* and *sphere-packing multicenter* functions defined by

$$\mathcal{H}_{\text{DC}}(P) = \max_{s \in \mathcal{D}} \{\mathrm{d}(s, i_{\mathbb{F}}(P))\},$$

$$\mathcal{H}_{\text{SP}}(P) = \min_{i \neq j \in \{1, \dots, n\}} \left\{\frac{1}{2}\|p_i - p_j\|, \mathrm{d}(p_i, \partial\mathcal{D})\right\}.$$

We are interested in the configurations that optimize these multicenter functions. The minimization of $\mathcal{H}_{\text{DC}}$ corresponds to minimizing the largest possible distance of any point in $\mathcal{D}$ to one of the agents' locations given by $p_1, \dots, p_n$. We refer to it as the as the *multi-circumcenter problem*. The maximization of $\mathcal{H}_{\text{SP}}$ corresponds to the situation where we are interested in maximizing the coverage of the area $\mathcal{D}$ in such a way that the radius of the generators do not overlap (in order not to interfere with each other) or leave the environment. We refer to it as the *multi-incenter problem*. It is useful to define the *index function* $N : \mathcal{D}^n \to \mathbb{N}$ as

$$N(P) = \left| \operatorname*{argmin}_{p_i \neq p_j} \left\{\frac{1}{2}\|p_i - p_j\|, \mathrm{d}(p_i, \partial\mathcal{D})\right\}\right|.$$

B. Spatial prediction via simple kriging

This section reviews the geostatistical kriging procedure for the estimation of spatial processes, see e.g., [2], [14]. A random process Z is *second-order stationary* and isotropic if it has constant mean, $E(Z(s)) = \mu$, and its covariance is of the form $\mathrm{Cov}(Z(p_1), Z(p_2)) = g(\|p_1 - p_2\|)$, for some decreasing function $g : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$. The covariance matrix of the set of points $p_1, \dots, p_n \in \mathcal{D}$ is $\boldsymbol{\Sigma} = \Sigma(P) = [g(\|p_1 - p_2\|)]_{i,j=1}^n \in \mathbb{R}^{n \times n}$. When it is clear from the context, we use bold face to denote explicit dependence on P. We define $c : \mathcal{D} \times \mathcal{D}^n \to \mathbb{R}^n$ to be the vector of covariances between a point $s \in \mathcal{D}$ and the locations in P, i.e., $\boldsymbol{c} = c(s, P) = (g(\|s - p_1\|), \dots, g(\|s - p_n\|))^T$. The associated correlation function $\rho : \mathbb{R}^d \times \mathbb{R}^d \to [0, 1]$ is $\rho(p_1, p_2) = \frac{g(\|p_1 - p_2\|)}{g(0)}$. Throughout the paper, we make the following assumptions on the model for the spatial random process Z of interest. We assume that Z is of the form

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

and that the mean function μ is known. Also, δ is a zero-mean second-order stationary random process with a known decreasing isotropic covariance function, g. We further assume that g is everywhere differentiable. Some examples of such functions are the exponential, cubic, spherical, modified Bessel, and rational quadratic covariance functions.

Assume measurement data $\boldsymbol{y} = (Y(p_1), \dots, Y(p_n))^T$ are corrupted with error such that

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

where $\tau^2 \geq 0$, and "iid" denotes independent and identically distributed. The assumption that the errors ϵ_i, $i \in \{1, \dots, n\}$ are independent and identically distributed corresponds to the fact that the robotic network is equipped with distinct identical sensors. In the error case, the covariance between $Y(p_i)$ and $Y(p_j)$ is given by

$$\text{Cov}(Y(p_i), Y(p_j)) = \begin{cases} g(\|p_i - p_j\|) + \tau^2, & \text{if } i = j, \\ g(\|p_i - p_j\|), & \text{if } i \neq j. \end{cases}$$

Note that the covariance matrix of P with respect to the noisy process Y may be written $\boldsymbol{\Sigma}_\tau = \Sigma_\tau(P) = \boldsymbol{\Sigma} + \tau^2 I_n$, where I_n denotes the $n \times n$ identity matrix.

The *simple kriging predictor* at $s \in \mathcal{D}$ minimizes the error variance $\sigma^2(s; p_1, \dots, p_n) = \text{Var}(Z(s) - p(s; Y(p_1), \dots, Y(p_n)))$ among all unbiased predictors of the form $p(s; Y(p_1), \dots, Y(p_n)) = \sum_{i=1}^n l_i Y(p_i) + k$. The simple kriging predictor at $s \in \mathcal{D}$ corresponds then to the LUMVE,

$$\hat{p}_{\text{SK}}(s; Y(p_1), \dots, Y(p_n)) = \mu(s) + \boldsymbol{c}^T \boldsymbol{\Sigma}_\tau^{-1}(\boldsymbol{y} - \boldsymbol{\mu}), \tag{3}$$

where $\boldsymbol{\mu} = (\mu(p_1), \dots, \mu(p_n))^T$. The error variance of $\hat{p}_{\text{SK}}$ at $s \in \mathcal{D}$ is

$$\sigma^2(s; p_1, \dots, p_n) = g(0) - \boldsymbol{c}^T \boldsymbol{\Sigma}_\tau^{-1} \boldsymbol{c}. \tag{4}$$

Note that σ^2 is invariant under permutations of $p_1, \dots, p_n$.

III. Objective functions

Consider a network of n agents in a convex polytope $\mathcal{D} \subset \mathbb{R}^d$. Assume each agent has a sensor and can take a noisy measurement $Y(p_i)$ as in (2) of the spatial process Z at its position p_i. A natural objective is to select locations to take measurements in such a way as to minimize the uncertainty in the estimate of the spatial process. Here, we consider two objective functions inspired by the notions of G- and D-optimality from optimal design [2], [4].

The *maximum error variance* is

$$\begin{aligned} \mathcal{M}(p_1, \dots, p_n) &= \max_{s \in \mathcal{D}} \sigma^2(s; p_1, \dots, p_n) \\ &= g(0) - \min_{s \in \mathcal{D}} \{\boldsymbol{c}^T \boldsymbol{\Sigma}_\tau^{-1} \boldsymbol{c}\}. \end{aligned} \tag{5a}$$

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. Let us make an important observation about the well-posedness of $\mathcal{M}$. Under noisy measurements, i.e., $\tau^2 > 0$, the function σ^2 is well-defined for any $s \in \mathcal{D}$ and $(p_1, \dots, p_n) \in \mathcal{D}^n$. Indeed, the dependence of σ^2 on the network configuration is continuous, and hence, $\mathcal{M}$ is also well-defined. However, when no measurement noise is present, i.e., $\tau^2 = 0$, then the matrix $\boldsymbol{\Sigma}_\tau = \boldsymbol{\Sigma}$ in (4) is not invertible for configurations that belong to S_{coinc}, and therefore, it is not clear what the value of σ^2 is. Proposition IV.2 below states that, in the no measurement noise case, σ^2 is a continuous function of the configuration under suitable technical conditions on the covariance structure of the spatial field.

Our second objective function requires some background. The generalized variance [15] of the LUMVE is defined as $|\boldsymbol{\Sigma}_\tau^{-1}|$, where $|\cdot|$ denotes the determinant. Minimizing $|\boldsymbol{\Sigma}_\tau^{-1}|$ is equivalent to minimizing $-|\boldsymbol{\Sigma}_\tau|$. For discrete state spaces, it can be shown [1] that configurations which maximize the minimum distance between agents asymptotically minimize $-|\boldsymbol{\Sigma}_\tau|$ in the limit of near independence. This tends to place agents on the boundary of $\mathcal{D}$. Since we are only interested in predictions over $\mathcal{D}$, we would like a notion of optimality which penalizes agents too close to the boundary as it does agents too close to each other. To this end, let $\gamma : \mathcal{D} \to \mathbb{R}^d$ map a point in $\mathcal{D}$ to its mirror image reflected across the nearest boundary of $\mathcal{D}$. Formally, $\gamma(s) \in s + 2\left(\operatorname{argmin}_{s^* \in \partial\mathcal{D}} \{\|s^* - s\|\} - s\right)$. Note that $\gamma(s)$ is in general not unique, and is not a smooth function of s. However, $\|s - \gamma(s)\|$ is smooth, and is the same for all values of $\gamma(s)$. Now consider minimizing the determinant of the estimator which would result if we had data from all agents as well as their reflections. The *extended prediction variance* is then

$$\mathcal{E}(p_1, \dots, p_n) = -\left|\Sigma_\tau(p_1, \dots, p_n, \gamma(p_1), \dots, \gamma(p_n))\right|. \tag{5b}$$

Since $\mathcal{E}$ does not require inversion of the covariance matrix, it is always well-posed.

Our goal is to find the network configurations $(p_1, \dots, p_n) \in \mathcal{D}^n$ that minimize the objective functions $\mathcal{M} : \mathcal{D}^n \to \mathbb{R}$ and $\mathcal{E} : \mathcal{D}^n \to \mathbb{R}$.

IV. Optimal configurations for spatial prediction

In this section, we provide several results that characterize the optimal network configurations for the objective functions $\mathcal{M}$ and $\mathcal{E}$. In Section IV-A, we show that minima of $\mathcal{M}$ cannot be in S_{coinc}. This fact is useful in Section IV-B where we show that circumcenter and incenter Voronoi configurations are asymptotically optimal for $\mathcal{M}$ and $\mathcal{E}$, respectively.

A. Coincident configurations are not minima of the maximum error variance

In this section, we examine the effect of the location of a subset of agents on the error variance terms. In particular we are interested in comparing $\sigma^2(s; P)$ against $\sigma^2(s; i_{\mathbb{F}}(P))$ for configurations $P \in S_{\text{coinc}}$. The following lemma provides a useful decomposition of σ^2.

Lemma IV.1 *The estimation error variance function may be written in the form*

$$\sigma^2(s; P) = \sigma^2(s; \overline{P}) - \frac{\left(\mathcal{N}(s, p_1; \overline{P})\right)^2}{\sigma^2(p_1; \overline{P}) + \tau^2}, \tag{6}$$

with $\mathcal{N}(s,p_1;\overline{P}) = g(\|s-p_1\|) - c^T(s,\overline{P})\Sigma_\tau(\overline{P})^{-1}c(p_1,\overline{P})$ *and* $\overline{P}=(p_2,\dots,p_n)\in\mathcal{D}^{n-1}$.

This fact may be proved using [16, Proposition 8.2.4] for the inverse of a partitioned symmetric matrix. Equation (6) may be applied repeatedly to isolate the effects of any subset of locations in P. In the following proposition we consider the behavior of $\mathcal{M}$ as agents move around $\mathcal{D}$. It can be proved by showing that, under the given assumptions, for $P\in\mathcal{D}\setminus S_{\text{coinc}}$ and $P'\in\mathcal{D}\bigcap S_{\text{coinc}}$, $\lim_{P\to P'}\sigma^2(s;P)=\sigma^2(s;i_{\mathbb{F}}(P'))$.

Proposition IV.2 (Continuity of estimation error variance) *Let Z be second-order stationary with isotropic covariance function,* $\operatorname{Cov}[Z(p_1),Z(p_2)] = g(\|p_1-p_2\|)$, *with* $g:\mathbb{R}_{\geq 0}\to\mathbb{R}_{\geq 0}$ *differentiable. Assume* $g'(0)\neq 0$ *and* $\tau^2=0$. *Then, for* $s\in\mathcal{D}$, *the error variance,* $(p_1,\dots,p_n)\mapsto\sigma^2(s;p_1,\dots,p_n)$ *is continuous. Moreover,* $\sigma^2(s;P)=\sigma^2(s;i_{\mathbb{F}}(P))$ *for* $P\in S_{coinc}$.

Under the assumptions of Proposition IV.2, we can extend the mean-squared error function by continuity to include configurations in S_{coinc}. With a slight abuse of notation, in the case of no measurement error, we use $\sigma^2(s;P)$ to denote $\sigma^2(s;i_{\mathbb{F}}(P))$ for $P\in S_{\text{coinc}}$.

Proposition IV.3 (Minima of $\mathcal{M}$ are not in S_{coinc}) *Let* $P^\dagger\in\mathcal{D}^n$ *be a strict local minimum of the map* $P\mapsto\mathcal{M}(P)$. *Under the assumptions of Proposition IV.2,* $P^\dagger\notin S_{coinc}$.

Proof: We proceed by contradiction. Assume $P^\dagger\in S_{\text{coinc}}$. Consider a configuration $P\in\mathcal{D}^n\setminus S_{\text{coinc}}$ in a neighborhood of $P^\dagger$ such that $i_{\mathbb{F}}(P^\dagger)\subset i_{\mathbb{F}}(P)$. Let $s,s^\dagger\in\mathcal{D}$ such that $\mathcal{M}(P)=\sigma^2(s;P)$ and $\mathcal{M}(P^\dagger)=\sigma^2(s^\dagger;P^\dagger)$. Using Lemma IV.1 and Proposition IV.2, one can deduce that $\sigma^2(s;P^\dagger)\geq\sigma^2(s;P)$. By the definition of $\mathcal{M}$, $\sigma^2(s^\dagger;P^\dagger)\geq\sigma^2(s;P^\dagger)$. Therefore $\mathcal{M}(P^\dagger)=\sigma^2(s^\dagger;P^\dagger)\geq\sigma^2(s;P^\dagger)\geq\sigma^2(s;P)=\mathcal{M}(P)$, which is a contradiction. ■

B. Multicenter Voronoi configurations are asymptotically optimal

Let us consider the objective functions $\mathcal{M}$ and $\mathcal{E}$ introduced in Section III but with covariance function C^k, $k\in\mathbb{N}$. This is equivalent to considering the correlation, ρ^k. As k grows, the correlation between distinct points in $\mathcal{D}$ vanishes. Note that ρ^k 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 $\boldsymbol{c}^{(k)}$, respectively $\boldsymbol{\Sigma}_\tau^{(k)}$, the vector $\boldsymbol{c}$, respectively the matrix $\boldsymbol{\Sigma}_\tau$, with each element raised to the kth power. Similarly, let $\mathcal{M}^{(k)},\mathcal{E}^{(k)}:D^n\to\mathbb{R}$ be defined as

$$\mathcal{M}^{(k)}(p_1,\dots,p_n) = g^k(0) - \min_{s\in\mathcal{D}}\{(\boldsymbol{c}^{(k)})^T(\boldsymbol{\Sigma}_\tau^{(k)})^{-1}\boldsymbol{c}^{(k)}\},$$

$$\mathcal{E}^{(k)}(p_1,\dots,p_n) = -\left|\Sigma_\tau^{(k)}\left(p_1,\dots,p_n,\gamma(p_1),\dots,\gamma(p_n)\right)\right|.$$

First we establish a result on the cardinality of the minimum distance set. Let $C_{\text{mds}}:\mathbb{R}^d\times\mathcal{D}^n\to\mathbb{R}$ such that $C_{\text{mds}}(s,P) = g(\|s-p\|)$, for any $p\in\operatorname{mds}(s,P)$. Note that C_{mds} is well-defined.

Proposition IV.4 (Cardinality of minimum distance set) *Let the covariance function C be continuous. For* $P\in\mathcal{D}^n\setminus S_{coinc}$, *one has* $\min_{s\in\mathcal{D}}\{C_{\text{mds}}(s,P)\,|\operatorname{mds}(s,P)|\} = \min_{s\in\mathcal{D}}\{C_{\text{mds}}(s,P)\}$.

We are now ready to prove one of the main results of the paper. The proof follows a similar line of reasoning to [1].

Theorem IV.5 (Minima of $\mathcal{M}$ under near independence) *Let* $P_{mcc}\in\mathcal{D}^n$ *be a global minimizer of the multi-circumcenter problem. Then, as* $k\to\infty$, P_{mcc} *asymptotically globally optimizes* $\mathcal{M}^{(k)}$, *that is,* $\mathcal{M}^{(k)}(P_{\text{mcc}})$ *approaches a global minimum.*

Proof: Note that minimizing $\mathcal{M}^{(k)}$ is equivalent to finding the tuples P which maximize the function $L^{(k)}:\mathcal{D}^n\to\mathbb{R}$ defined as $L^{(k)}(P)=\min_{s\in\mathcal{D}}\left\{(c^{(k)}(s,P))^T(\Sigma_\tau^{(k)}(P))^{-1}(c^{(k)}(s,P))\right\}$. Let $\lambda_{\min}$ and $\lambda_{\max}:\mathcal{D}^n\times\mathbb{R}\to\mathbb{R}$ be such that $\lambda_{\min}(P,k)$, $\lambda_{\max}(P,k)$ denote, respectively, the minimum and the maximum eigenvalue of $\Sigma_\tau^{(k)}(P)$. We can see that $L^{(k)}(P)$ is bounded above by $\lambda_{\max}(P,k)\sum_{p\in P}g(\|s-p\|)^{2k}$ and below by $\lambda_{\min}(P,k)\sum_{p\in P}g(\|s-p\|)^{2k}$. For a given s, in terms of the minimum distance set we can write

$$\sum_{p\in P}g(\|s-p\|)^{2k} = \sum_{p\in\operatorname{mds}(s,P)}g(\|s-p\|)^{2k} + \sum_{p\in P\setminus\operatorname{mds}(s,P)}g(\|s-p\|)^{2k}$$
$$= |\operatorname{mds}(s,P)|\,C_{\text{mds}}(s,P)^{2k} + \sum_{p\in P\setminus\operatorname{mds}(s,P)}g(\|s-p\|)^{2k}.$$

As $k\to\infty$ the elements in the minimum distance set dominate, so we are left with

$$\sum_{p\in P}g(\|s-p\|)^{2k} = |\operatorname{mds}(s,P)|\,C_{\text{mds}}(s,P)^{2k} + o(C_{\text{mds}}(s,P)^{2k}).$$

From Proposition IV.4,

$$\min_{s\in\mathcal{D}}\{|\operatorname{mds}(s,P)|\,C_{\text{mds}}(s,P)\} = \min_{s\in\mathcal{D}}\{C_{\text{mds}}(s,P)\},$$

so we can write

$$\min_{s\in\mathcal{D}}\Big\{\sum_{p\in P}g(\|s-p\|)^{2k}\Big\} = \min_{s\in\mathcal{D}}\left\{C_{\text{mds}}(s,P)^{2k}\left(1+o(1)\right)\right\}.$$

Consider, then, comparing an arbitrary configuration P^* against a global minimizer of $\mathcal{H}_{\text{DC}}$, say P_{mcc}. In the zero measurement error case, by Proposition IV.3, we can assume without loss of generality that $P^*\notin S_{\text{coinc}}$. Therefore, no matter what the value of τ is, we can safely use the eigenvalues of $(\boldsymbol{\Sigma}_\tau^{(k)})^{-1}$ to provide bounds. Specifically,

$$\frac{L^{(k)}(P^*)}{L^{(k)}(P_{\text{mcc}})} \leq \frac{\lambda_{\max}(P^*,k)\min_{s\in\mathcal{D}}\left\{C_{\text{mds}}(s,P^*)^{2k}\left(1+o(1)\right)\right\}}{\lambda_{\min}(P_{\text{mcc}},k)\min_{s\in\mathcal{D}}\left\{C_{\text{mds}}(s,P_{\text{mcc}})^{2k}\left(1+o(1)\right)\right\}}. \tag{7}$$

Next we take a closer look at the eigenvalues. Note that if we divide $\Sigma_\tau^{(k)}(P)$ by the common factor of $(g(0)+\tau^2)^k$, the resulting correlation matrix becomes diagonal for large k. This gives us $\lim_{k\to\infty}1/(g(0)+\tau^2)^k\Sigma_\tau^{(k)}(P)=I_n$, and it can be seen that $\lambda_{\min}(P,k)/(g(0)+\tau^2)^k$ and $\lambda_{\max}(P,k)/(g(0)+\tau^2)^k$ both tend to 1 for any configuration P. Finally, since P_{mcc} minimizes the maximum distance to any point $s \in \mathcal{D}$, it maximizes the minimum covariance, so for any $P \in \mathcal{D}^n$, $\min_{s\in\mathcal{D}} C_{\mathrm{mds}}(s,P) \le \min_{s\in\mathcal{D}} C_{\mathrm{mds}}(s,P_{\mathrm{mcc}})$. Thus the ratio (7) is bounded by $1+o(1)$. Therefore, in the limit as $k\to\infty$, minimizing $\mathcal{M}^{(k)}$ is equivalent to solving the multi-circumcenter problem. ∎

The proof of the theorem can be reproduced for local minimizers of the multi-circumcenter problem to arrive at the following result.

Corollary IV.6 *Let $P_{\mathrm{mcc}} \in \mathcal{D}^n$ be a local minimizer of the multi-circumcenter problem. Then, as $k\to\infty$, P_{mcc} asymptotically optimizes $\mathcal{M}^{(k)}$, that is, $\mathcal{M}^{(k)}(P_{\mathrm{mcc}})$ approaches a minimum.*

According to [11], under certain technical conditions, solutions to the multi-circumcenter problem are circumcenter Voronoi configurations. Next, let us present a similar asymptotic result for the extended prediction variance.

Theorem IV.7 (Minima of $\mathcal{E}$ under near independence) *Let $P_{\mathrm{mic}} \in \mathcal{D}^n$ be a global maximizer of the multi-incenter problem with lowest index. Then, as $k\to\infty$, P_{mic} asymptotically globally optimizes $\mathcal{E}^{(k)}$, that is, $\mathcal{E}^{(k)}(P_{\mathrm{mic}})$ approaches a global minimum.*

Proof: Expanding the objective function for asymptotically dominant terms, we may write

$$\mathcal{E}^{(k)}(P) = -(g(0)^k+\tau^2)^{2n} + \left(g(0)^k+\tau^2\right)^{2n-2} J^{(k)}(P) + \\ + o\left(\left(g(0)^k+\tau^2\right)^{2n-2} J^{(k)}(P)\right),$$

where

$$J^{(k)}(P) = \sum_{i\neq j} g(\|p_i-p_j\|)^{2k} + \sum_{i,j=1}^{n} g(\|p_i-\gamma(p_j)\|)^{2k} + \\ + \sum_{i\neq j} g(\|\gamma(p_i)-\gamma(p_j)\|)^{2k}.$$

Asymptotically all but the largest terms in $J^{(k)}(P)$ will drop out, and minimizing $\mathcal{E}^{(k)}(P)$ becomes equivalent to minimizing those terms. The largest terms in $J^{(k)}(P)$ correspond to the shortest distance between the locations of either the agents or their reflected images. For any two agent locations, $p_i, p_j \in \mathcal{D}$, and any of their reflections $\gamma(p_i), \gamma(p_j)$ the minimum distance between any two of the four points can be reduced to $\min\{\|p_i-p_j\|, \|p_i-\gamma(p_i)\|, \|p_j-\gamma(p_j)\|\}$ (note that this is not in general true for non-convex domains). Thus the shortest distance between agents in P and their reflections may be expressed as $2\mathcal{H}_{\mathrm{SP}}(P)$, though the index of P might be larger than 1. Therefore we have $J^{(k)}(P) = N(P)\left(g(2\mathcal{H}_{\mathrm{SP}}(P))^{2k}\right)(1+o(1))$. Consider comparing an arbitrary configuration, $P^* \in \mathcal{D}^n$ against P_{mic}. We have

$$\frac{J^{(k)}(P_{\mathrm{mic}})}{J^{(k)}(P^*)} = \frac{N(P_{\mathrm{mic}})\left(g(2\mathcal{H}_{\mathrm{SP}}(P_{\mathrm{mic}}))^{2k}\right)(1+o(1))}{N(P^*)\left(g(2\mathcal{H}_{\mathrm{SP}}(P^*))^{2k}\right)(1+o(1))}.$$

If P^* is not a global solution of the multi-incenter problem, we have $\mathcal{H}_{\mathrm{SP}}(P_{\mathrm{mic}}) > \mathcal{H}_{\mathrm{SP}}(P^*)$, and since $g(\cdot)$ is decreasing this gives us

$$\lim_{k\to\infty} \frac{J^{(k)}(P_{\mathrm{mic}})}{J^{(k)}(P^*)} = 0.$$

If, on the other hand, P^* is a global solution of the multi-incenter problem, then, using the fact that P_{mic} has the lowest index among all of them, we deduce $\frac{J^{(k)}(P_{\mathrm{mic}})}{J^{(k)}(P^*)} \le 1+o(1)$. ∎

The proof of the theorem can be reproduced for isolated local maximizers of the multi-incenter problem to arrive at the following result.

Corollary IV.8 *Let $P_{\mathrm{mic}} \in \mathcal{D}^n$ be an isolated local maximizer of the multi-incenter problem. Then, as $k\to\infty$, P_{mic} asymptotically optimizes $\mathcal{E}^{(k)}$, that is, $\mathcal{E}^{(k)}(P_{\mathrm{mic}})$ approaches a minimum.*

According to [11], under certain technical conditions, solutions to the multi-incenter problem are incenter Voronoi configurations.

C. Distributed coordination algorithms

Given the results in Theorems IV.5 and IV.7, it is of interest to design coordination algorithms that steer a network of mobile agents towards circumcenter and incenter Voronoi configurations. We do this following the exposition in [11]. In light of the results in Section IV-B, this enables the network to perform a spatial prediction which is asymptotically optimal as $k\to\infty$. Note that these algorithms are not intended to provide optimal trajectories for multiple sequential measurements. That problem is left for future work.

Let us assume each agent can move according to a first-order dynamical model $\dot{p}_i = u_i$, $i \in \{1,\dots,n\}$. Consider the following coordination algorithms

$$\dot{p}_i = \mathrm{CC}(V_i(P)) - p_i, \tag{8a}$$

$$\dot{p}_i \in \mathrm{IC}(V_i(P)) - p_i, \tag{8b}$$

for each $i \in \{1,\dots,n\}$. Note that (8b) is a differential inclusion. We understand its solutions in the Filippov sense [17]. Both coordination algorithms are Voronoi distributed, meaning that each agent only requires information from its Voronoi neighbors in order to execute its control law. The equilibrium points of the flow (8a) are the circumcenter Voronoi configurations, whereas the equilibrium points of the flow (8b) are incenter Voronoi configurations. Furthermore, the evolution of $\mathcal{H}_{\mathrm{DC}}$ along (8a) is monotonically decreasing, while the evolution of $\mathcal{H}_{\mathrm{SP}}$ along (8b) is monotonically increasing. The convergence properties of these coordination algorithms, as well as alternative flows with similar distributed properties that can also be used to steer the network to center Voronoi configurations, are studied in [11].

V. SIMULATIONS

With the aim of illustrating the results presented in Section IV, we performed simulations for both objective functions $\mathcal{M}$ and $\mathcal{E}$ with $n = 5$ agents. In our simulations, we used as domain $\mathcal{D}$ the convex polygon with vertices $\{(0, 0.1), (2.5, 0.1), (3.45, 1.6), (3.5, 1.7), (3.45, 1.8), (2.7, 2.2), (1.0, 2.4), (0.2, 1.3)\}$ and as isotropic covariance the one defined via $g : \mathbb{R} \to \mathbb{R}$, $g(\|s_1 - s_2\|) = e^{-\frac{1}{5}\|s_1-s_2\|}$. Note that the mean function, μ, does not play a role in determining the optimal network configurations. Figure 1 shows the multicenter configurations obtained with the flows (8).

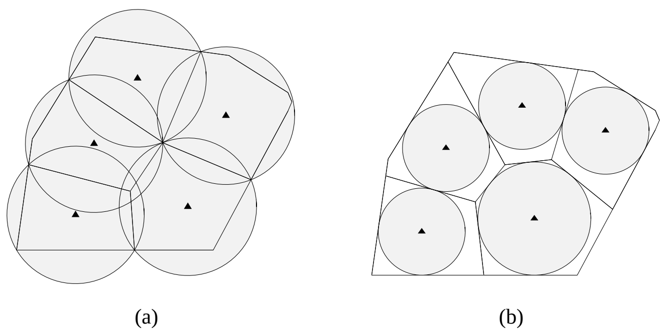

Fig. 1. Multicenter configurations found from different random starting positions using (a) the flow (8a) and (b) the flow (8b).

A. Analysis of simulations for $\mathcal{M}^{(k)}$

Using $\mathcal{M}^{(1)}$ we ran over 1000 random trials, each time running a gradient descent algorithm, and chose the local minimum configuration with the smallest value of $\mathcal{M}^{(1)}$ to be our approximation of a global minimum. From this configuration P_*, we generated a multi-circumcenter configuration using (8a), depicted in Figure 1(a). For increasing values of k, we ran a gradient descent of $\mathcal{M}^{(k)}$ to find the best local configuration near P_*. We plotted $\mathcal{M}^{(k)}$ as calculated with this new configuration against $\mathcal{M}^{(k)}$ as calculated with the multi-circumcenter configuration. For comparison, we also plotted the performance of a random (static) configuration which was not a local minimum. Figure 2 illustrates the result in Theorem IV.5. We halt the experiment at around

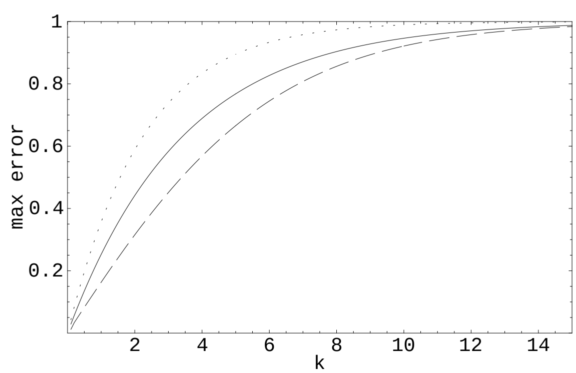

Fig. 2. Value of $\mathcal{M}^{(k)}$ for multi-circumcenter (solid), approximated global minimum (dashed) arrived at by gradient descent for each value of k, and random (dotted) configurations of 5 agents for increasing k. The covariance function is exponential.

$k = 15$ because the performance of the circumcenter Voronoi configuration becomes impossible to distinguish from the one of the minimizer of $\mathcal{M}^{(k)}$ at this resolution.

B. Analysis of simulations for $\mathcal{E}^{(k)}$

Using $\mathcal{E}^{(1)}$ we ran over 1000 random trials, each time running a gradient descent algorithm, and chose the local minimum configuration with the smallest value of $\mathcal{E}^{(1)}$ to be our approximation of a global minimum. From this configuration P_* we generated the multi-incenter configuration using (8b), depicted in Figure 1(b). For increasing values of k, we ran a gradient descent of $\mathcal{E}^{(k)}$ to find the best local configuration near P_*. We plotted $\mathcal{E}^{(k)}$ as calculated with this new configuration against $\mathcal{E}^{(k)}$ as calculated with the multi-incenter configuration. For comparison, we also plotted the performance of a random (static) configuration which was not a local minimum. Figure 3 illustrates the result stated in Theorem IV.7.

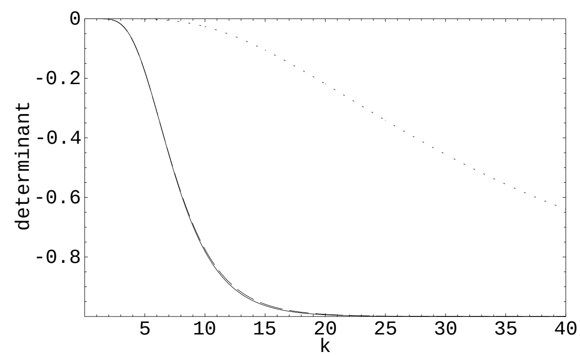

Fig. 3. Value of $\mathcal{E}^{(k)}$ for multi-incenter (solid), approximated global minimum (dashed) arrived at by gradient descent for each value of k, and random (dotted) configurations of 5 agents for increasing k. The covariance function is exponential. The performance of the global and multi-incenter configurations looks identical even though configurations are different at each k.

Remarkably, the performance of the incenter Voronoi configuration and the minimizer of $\mathcal{E}^{(k)}$ are almost identical, even for low values of k. The numerical simulations suggest that multi-incenter Voronoi configurations are near-optimal for the extended prediction criterion.

VI. Conclusions

We have used the maximum error variance and the extended variance of the LUMVE as metrics for optimal placement of mobile sensor networks estimating random fields. We have shown that under the assumption of near independence, circumcenter configurations minimize the maximum error variance and incenter configurations minimize the extended variance of the estimator. Under limited time or energy resources, or as a starting point for further exploration, a group of robotic sensors can begin by moving toward these configurations to start the estimation procedure.

Future work will explore: (i) regarding the asymptotic analysis, the determination of lower and upper bounds on the parameter k that guarantee that multicenter Voronoi configurations achieve a given a desired level of performance. In particular, we would like to determine the near-optimality in general of incenter Voronoi configurations for the extended variance criterion; (ii) the extension of the results to similar error metrics for the universal kriging predictor, where the mean function is unknown; and (iii) the characterization of the trajectories (rather than configurations) that provide optimal estimates of the random field as agents reconfigure and take successive measurements over time. Consideration will also be given to fields which change over time, and to distributed methods for estimation and data fusion.

ACKNOWLEDGMENTS

This research was supported in part by NSF CAREER Award ECS-0546871.

REFERENCES

[1] 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.

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

[3] M. L. Stein, *Interpolation of Spatial Data. Some Theory for Kriging*. Springer Series in Statistics, New York: Springer, 1999.

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

[5] 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.

[6] J. S. Willcox, J. G. Bellingham, Y. Zhang, and A. B. Baggeroer, "Performance metrics for oceanographic surveys with autonomous underwater vehicles," *IEEE Journal of Oceanic Engineering*, vol. 26, no. 4, pp. 711–725, 2001.

[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. Meyn, *Control Techniques for Complex Networks*. Cambridge University Press, 2007.

[9] 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.

[10] 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.

[11] 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.

[12] R. Graham and J. Cortés, "Asymptotic optimality of multicenter Voronoi configurations for random field estimation," Tech. Rep. ams2007-10, University of California, Santa Cruz, 2007. Electronically available at http:www.ams.ucsc.edu/research/report?ID=480.

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

[14] 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.

[15] S. S. Wilks, "Certain generalizations in the analysis of variance," *Biometrika*, vol. 24, no. 3/4, pp. 471–494, 1932.

[16] D. S. Bernstein, *Matrix Mathematics*. Princeton, NJ: Princeton University Press, 2005.

[17] A. F. Filippov, *Differential Equations with Discontinuous Righthand Sides*, vol. 18 of *Mathematics and Its Applications*. Kluwer Academic Publishers, 1988.