Statistical Coverage Control of Mobile Sensor Networks

Ömür Arslan, *Member, IEEE*

***Abstract*—A fundamental challenge of mobile sensor networks is automated active reconfiguration of sensors in response to environmental stimuli in order to maximize their total sensing quality (or minimize their total sensing cost) of events occurring over an environment. In this paper, given an event distribution over a convex environment, we consider mobile isotropic sensors with adjustable sensing range and propose a new family of provably correct reactive coverage control algorithms for both continuous- and discrete-time sensor dynamics. The proposed coverage control algorithms constantly (re)configure sensor positions and sensing ranges in order to minimize a statistical distance, in particular, an f-divergence, between the event distribution over the environment and the overall event detection probability of sensors. We show that the standard Voronoi-based coverage control law of homogeneous mobile sensor networks is a special case of our framework where the event detection probability of each sensor has a Gaussian form, the statistical distance is set to be the Kullback-Leibler (KL) divergence and sensor allocation is performed based on Voronoi diagrams. To increase the practicality of our framework, we also present its integration with a Voronoi-based collision avoidance strategy for disk-shaped sensor bodies and its extension to differential drive sensor dynamics, while retaining the stability properties.**

***Index Terms*—Sensor networks, coverage control, statistical distance, KL divergence, Voronoi diagrams, soft sensor allocation.**

I. Introduction

Over their static counterparts, mobile sensor networks offer increased flexibility and functionality in active information gathering for applications such as situational awareness, security, surveillance, search and rescue, and environmental monitoring. However, given an estimate of some event or activity distribution over an environment, sensors are usually needed to be actively reconfigured in order to achieve better monitoring of environmental conditions. In this paper, assuming a given event/activity distribution over a known convex environment, we consider the automated reconfiguration problem of a group of mobile isotropic sensors with adjustable sensing range. We propose a new family of simple greedy coverage control algorithms that reactively (re)configure sensor positions and sensing ranges, as illustrated in Fig. 1, in order to maximize their total collective coverage quality which is measured in terms of a statistical distance between the event distribution over the environment and the overall event detection probability of the sensors. Potential applications of the proposed coverage control algorithms are envisioned to include surveillance and environmental monitoring of large regions using adjustable resolution sensors, such as aerial surveillance with downward facing cameras [1] and underwater monitoring with acoustic sensors [2], at varying resolution for increased accuracy.

Ömür Arslan is with the Autonomous Motion Department, Max Planck Institute for Intelligent Systems, Tübingen, 72076 Germany. E-mail: omur.arslan@tuebingen.mpg.de

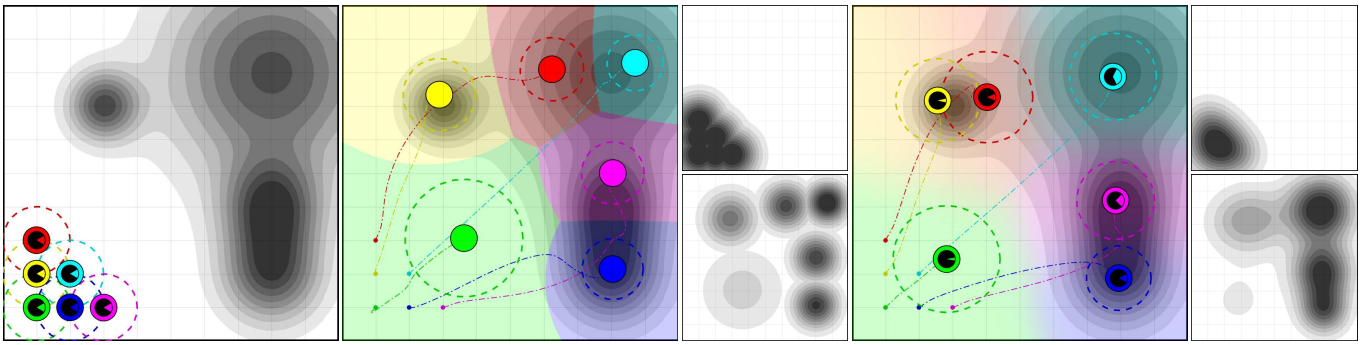
A. Motivation and Prior Literature

Finding a globally optimal coverage configuration of mobile sensor networks is usually very hard. For example, many related facility localization (e.g., p-center and p-median) problems [3] are known to be NP hard. This complexity is generally mitigated by utilizing greedy gradient methods that can efficiently find a locally optimal sensing configuration using gradient descent or ascent along a coverage objective.[1]

A widely used coverage objective is the expected sensing cost of event locations, where the spatial sensing cost of sensors is usually assumed to be a (monotone increasing) function of the Euclidean distance between sensor position and event location [8]. Such geometric approaches also leverage Voronoi diagrams to decouple the problems of optimal sensor allocation and sensor configuration optimization, as well as to model the heterogeneity of sensors in their sensing capabilities via additively and multiplicatively weighted Voronoi diagrams [9]–[11]. For example, the standard Voronoi-based coverage control algorithm of mobile sensor networks continuously moves each sensor towards its associated Voronoi centroid. This simple coverage control strategy is known to asymptotically converge to a locally optimal coverage configuration where each sensor is located at the respective Voronoi centroid and therefore minimizes the expected squared Euclidean distance to events occuring within its Voronoi cell [9]–[11]. Another common coverage objective is the joint event detection probability of sensors, where each sensor is modelled by means of its event detection probability (e.g., a Gaussian distribution) that is parametrized according to the sensor's sensing capabilities [12]–[14]. Moreover, artificial potential fields, constructed as the gradient of a repulsive potential function between sensors, are also employed for minimizing sensor overlap and so for maximizing the total coverage area [15]. In fact, the common functional form of these coverage objectives of geometric, probabilistic and potential field approaches can be unified using a "mixing function" to describe a general family of coverage cost functions [16].

Alternatively, we propose a new perspective for quantifying coverage quality of a sensor network by comparing the overall event detection probability of sensors with the (given or

[1]The coverage quality of such locally optimal configurations strongly depends on the choice of initial sensor network configurations. Using proper initialization [4] and incremental deployment [5], one can obtain a locally optimal coverage configuration that is as "good" as a globally optimal coverage configuration. One can also validate the quality of a locally optimal coverage configuration by checking the stability of sensor allocation [6], [7].



(a) (b) (c) (d) (e)

Fig. 1. Statistical coverage control of mobile isotropic sensor networks with variable sensing range aims to continuously adjust sensor positions and sensing ranges in order to asymptotically minimize an f-divergence between the (given or estimated) event distribution over an environment and the overall event detection probability of sensors. (a) The level sets (gray) of a given event distribution in (98) over the environment and an initial sensor network coverage configuration, where the dashed circles depict sensors' sensing ranges and the pie-charts attached to disk-shaped colored sensor bodies illustrate sensors' mixture weights for soft sensor allocation. (b) A locally optimal Voronoi-centroidal coverage configuration and the associated Voronoi-based sensor allocation where each Voronoi cell is coloured in accordance with the respective sensor body color. (c) The maximum of event detection probabilities of sensors at the initial (top) and the final (bottom) coverage configurations. (d) A locally optimal soft centroidal coverage configuration and the associated soft sensor allocation where spatial colouring is in accordance with sensor body colors and proportional with soft assignment weights. (e) The weighted average of event detection probabilities of sensors at the initial (top) and the final (bottom) coverage configurations.[12] Here, the reverse Kullback-Leibler (KL) divergence is used as a statistical coverage quality measure, and the locally optimal coverage configurations in (b) and (d) are, respectively, obtained by following the continuous-time move-to-Voronoi-centroidal configuration law in (37) and the continuous-time move-to-soft-centroidal configuration law in (33).[13]

estimated) event distribution over the environment using a statistical distance, in particular, an f-divergence, where each sensor is modelled using its event detection probability that is parametrized by sensor position and sensing range, see Fig. 1. We show that the standard Voronoi-based coverage control law of homogeneous isotropic mobile sensors in [9] is a special case of our approach, where each sensor is modelled using a Gaussian event detection probability, the statistical distance is set to be the Kullback-Leibler (KL) divergence and sensor allocation is performed based on the standard Voronoi partition of the environment. Thus, the proposed approach offers an extension of Voronoi-based coverage control algorithms with similar qualitative properties and interpretation (i.e., we introduce new "move-to-centroidal-configuration" laws) for mobile isotropic sensors with adjustable sensing range, whereas most existing work on coverage control of mobile sensor networks assumes fixed (and possibly heterogeneous) sensing ranges [9]–[11]. Our interest in sensors with adjustable sensing range is motivated by the increased availability of low-cost controllable sensors (e.g., cameras) for accurate environment monitoring at varying resolution [17].

Adjustable sensing range plays a critical role in sensor (especially camera) networks due to the trade-off between sensitivity and sensing region volume [17]. Among the others, the most closely related approaches [18], [19] use Gaussian mixture models to represent the overall sensing quality of a pan/tilt/zoom camera network and design a discrete-time coverage control strategy that reconfigures camera orientation and zoom levels based on the iterations of the standard Expectation-Maximization (EM) algorithm [20], which is known to decrease the KL divergence between the overall sensing quality of the network and the event distribution. In this paper, in order to generate smooth coverage motion, in addition to its continuous-time generalization, we describe a discrete-time generalization, with adjustable step size, of the EM algorithm using the MM (Majorize-Minimization/Minorize-Maximization) principle [21] that decreases an f-divergence between the overall event detection probability of a mobile sensor network and the event distribution over the environment. In sensor networks, Gaussian mixture models also find applications in event distribution estimation in unknown environments [22]–[25] and in interactive multirobot control using time-varying density function [26].

Finally, it is worth mentioning that sensors with adjustable sensing range are usually studied in (randomly deployed, very large) static sensor networks in order to design efficient scheduling algorithms to prolong network lifetime by minimizing energy consumption [27]–[29]. Moreover, adaptive Voronoi diagrams with adjustable additive and/or multiplicative weights are used for load balancing among sensors and area-constrained coverage control [30]–[32], but without making an explicit connection between Voronoi weights and sensor capabilities. We here introduce a new notion of additively and multiplicatively weighted generalized Voronoi diagrams, for isotropic sensors with adjustable sensing range, whose weights are a function of sensors' sensing ranges.

B. Contributions and Organization of the Paper

In this paper, we consider the reactive coverage control problem of mobile isotropic sensor networks with adjustable sensing range. Following a probabilistic approach, in Section II-A, we model the event detection probability of each sensor by a Gaussian distribution parametrized by sensor position and sensing range, which also yields a natural spatial sensing cost function. In Section II-B we show a novel use of this probabilistic sensing model to define a new additively and multiplicatively weighted Voronoi diagram for determining optimal sensor allocation in mobile sensor networks with adjustable sensing range. To quantify the event coverage performance of sensors, in Section II-C we introduce a new statistical coverage quality measure based on an f-divergence between the event distribution over the environment and the overall event detection probability of the sensors that is determined by either a weighted average or the maximum of individual event

detection probabilities of the sensors, defining, respectively, a *soft* or a *hard* sensor allocation scheme. In Section II-E we provide an explicit characterization of locally optimal coverage configurations for both soft and hard sensor allocation. In Section III we propose a new family of simple reactive coverage control algorithms, for both continuous- and discrete-time sensor dynamics, that asymptotically (re)configure the sensors towards a locally optimal coverage configuration. In Section III-C, we present three practical extensions of our framework for discrete event distributions, for collision avoidance of disk-shaped finite-size sensors and for nonholonomically constrained kinematic differential drive dynamics, while maintaining the stability properties. In Section IV, we provide numerical simulations demonstrating the resulting coverage motion and locally optimal coverage configurations for both the KL divergence and the Hellinger distance. In Section V, we conclude with a summary of our contributions and future research directions.

In summary, the main contributions of the present paper over the existing literature on coverage control of mobile sensor networks are:

- a new probabilistic spatial sensing model for mobile sensors with adjustable sensing range,
- a fresh perspective on soft and hard sensor allocation in mobile sensor networks,
- a novel statistical coverage quality measure based on an f-divergence between the event detection probability of sensors and the event distribution over the environment,
- a new family of provably correct, simple reactive coverage control algorithms for continuous- and discrete-time sensor dynamics.

In addition to these completely new contributions, we also present practical extensions for differential-drive sensor dynamics and collision avoidance based on our prior work [33]. Moreover, we discuss in details how our proposed statistical coverage control methods generalize the standard Voronoi-based coverage control approaches [9].

II. Quantifying Coverage via Statistical Distances

In this section, we first introduce a probabilistic spatial sensing model for isotropic mobile sensors with adjustable sensing range and describe its use for constructing generalized Voronoi diagrams that determine optimal sensor allocation in such mobile sensor networks. Then, we propose a new coverage quality measure based on a statistical distance between a given event distribution over the environment and the overall event detection probability of sensors, and present an explicit characterization of locally optimal coverage configurations.

A. Mobile Sensor Networks with Adjustable Sensing Range

For ease of exposition, we consider a convex environment $\mathcal{W} \subset \mathbb{R}^d$ in the d-dimensional Euclidean space $\mathbb{R}^d$ that is monitored by a group of mobile isotropic heterogeneous sensors, located at positions $\mathbf{p} := (\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_n) \in \mathcal{W}^n$ and with adjustable sensing ranges[2] $\boldsymbol{r} := (r_1, r_2, \ldots, r_n) \in (\mathbb{R}_{>0})^n$, each of whose event detection probability $q_{\mathrm{s}_i}(\mathrm{x})$ at event location $\mathrm{x} \in \mathbb{R}^d$ has the following normalized Gaussian form[3]

$$q_{\mathrm{s}_i}(\mathrm{x}) := \frac{1}{\left(2\pi r_i^2\right)^{\frac{d}{2}}} \exp\left(-\frac{\|\mathrm{x}-\mathrm{p}_i\|^2}{2r_i^2}\right), \qquad (1)$$

where $\|.\|$ denotes the standard Euclidean norm. For ease of presentation, we find it convenient to denote the state of a mobile sensor network with adjustable sensing range by $\mathbf{s} := (\mathrm{s}_1, \mathrm{s}_2, \ldots, \mathrm{s}_n) \in (\mathcal{W} \times \mathbb{R}_{>0})^n$, where each sensor's state $\mathrm{s}_i := (\mathrm{p}_i, r_i) \in \mathcal{W} \times \mathbb{R}_{>0}$ is comprised of its position $\mathrm{p}_i \in \mathcal{W}$ and adjustable desired sensing range $r_i \in \mathbb{R}_{>0}$.

Note that, by applying the "two-" or 'three-sigma" rules of thumbs, one can consider the effective sensing range of a sensor to be $2r_i$ or $3r_i$ depending on the desired confidence level of 95% or 99.7%, respectively. Also observe that the normalized form of the event detection probability q_{s_i} in (1) ensures that the event detection probability of a sensor decreases with increasing sensing range, which is consistent with the trade-off between the sensitivity and the sensing range of many physical finite-resolution sensors, such as cameras, radars, acoustic and ultrasonic range sensors. Therefore, the adjustable sensing range parameter r_i intuitively corresponds to zoom control of cameras [17] or signal power control of dynamic range radars [34].

Finally, we find it useful to emphasize that the negative log-likelihood of the event detection probability q_{s_i} defines a natural spatial sensing cost c_{s_i} for event location $\mathrm{x} \in \mathbb{R}^d$ as

$$c_{\mathrm{s}_i}(\mathrm{x}) := -\log q_{\mathrm{s}_i}(\mathrm{x}) = \frac{\|\mathrm{x}-\mathrm{p}_i\|^2}{2r_i^2} + \frac{d}{2}\log r_i^2 + \frac{d}{2}\log 2\pi, \quad (2)$$

which is an additively and multiplicatively weighted power distance based on the standard Euclidean distance [35].

B. Optimal Sensor Allocation

To determine which part of the environment each sensor should monitor given other sensors' configurations, we follow a standard optimal resource allocation strategy widely used in facility localization [8], [36], quantization [37], [38] and sensor networks [9]; and we assign each event location to the sensor with maximal event detection probability (resp. minimal sensing cost) so that the total event detection probability of sensors is maximized (resp. the total sensing cost

[2]Here, $\mathbb{R}_{>0}$ ($\mathbb{R}_{\geq 0}$) denotes the set of positive (nonnegative) real numbers.

[3]Depending on the application setting, especially the physical nature of sensors, one can consider other kernel functions to define a distance between sensor position and event location, and alternative probability distribution functions. For example, if the event detection probability $q_{\mathrm{s}_i}(\mathrm{x})$ of a sensor is defined, using a kernel function $\mathcal{K} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_{\geq 0}$, as

$$q_{\mathrm{s}_i}(\mathrm{x}) := \frac{1}{\left(2\pi r_i^2\right)^{\frac{d}{2}}} \exp\Big(-\frac{\mathcal{K}(\mathrm{x}, \mathrm{p}_i)}{2r_i^2}\Big),$$

then the sensor's spatial sensing cost c_{s_i} becomes

$$c_{\mathrm{s}_i}(\mathrm{x}) := -\log q_{\mathrm{s}_i}(\mathrm{x}) = \frac{\mathcal{K}(\mathrm{x}, \mathrm{p}_i)}{2r_i^2} + \frac{d}{2}\log r_i^2 + \frac{d}{2}\log 2\pi,$$

which has the same general form as the standard sensing cost function used in coverage control of homogeneous mobile sensor networks [9]. In this paper, in order to design simple intuitive coverage control algorithms with an explicit closed form, we limit our attention to $\mathcal{K}(\mathrm{x}, \mathrm{p}_i) = \|\mathrm{x} - \mathrm{p}_i\|^2$. In future papers, in addition to other kernel functions, we plan to extend our results to the exponential family of distributions such as Gamma distributions for modelling sensor directionality and Beta distributions for modelling limited range sensing.

is minimized). For any given sensor network configuration $\mathbf{s}\in(\mathcal{W}\times\mathbb{R}_{>0})^n$, this yields an additively and multiplicatively weighted Voronoi diagram, $\mathcal{V}(\mathbf{s}):=\{V_1,V_2,\dots,V_n\}$, of the environment $\mathcal{W}$, whose blocks are given by

$$V_i := \Big\{ \mathrm{x}\in\mathcal{W} \Big| q_{\mathrm{s}_i}(\mathrm{x}) \geq q_{\mathrm{s}_j}(\mathrm{x}),\ \ \forall j\neq i \Big\}, \tag{3a}$$

$$= \Big\{ \mathrm{x}\in\mathcal{W} \Big| c_{\mathrm{s}_i}(\mathrm{x}) \leq c_{\mathrm{s}_j}(\mathrm{x}),\ \ \forall j\neq i \Big\}, \tag{3b}$$

$$= \Big\{ \mathrm{x}\in\mathcal{W} \Big| \tfrac{\|\mathrm{x}-\mathrm{p}_i\|^2}{d\,r_i^2}+\log r_i^2 \leq \tfrac{\|\mathrm{x}-\mathrm{p}_j\|^2}{d\,r_j^2}+\log r_j^2,\ \forall j\neq i \Big\}, \tag{3c}$$

and have quadratic (i.e., circular or planar) boundary segments between adjacent Voronoi cells [8], see Fig. 2(a)-(b). It is useful to observe that if sensors share identical sensing range parameters, i.e., $r_i=r_j$ for all $i\neq j$, then $\mathcal{V}(\mathbf{s})$ is equivalent to the standard Voronoi diagram of $\mathcal{W}$, generated based on sensor positions $\mathbf{p}=(\mathrm{p}_1,\mathrm{p}_2,\dots,\mathrm{p}_n)$ and the Euclidean distance, i.e., $V_i=\{\mathrm{x}\in\mathcal{W}\,|\,\|\mathrm{x}-\mathrm{p}_i\|\leq\|\mathrm{x}-\mathrm{p}_j\|,\,\forall j\neq i\}$ [8].

Despite their flexibility in modelling different phenomena, generalized Voronoi diagrams might suffer from some form of degeneracy such as empty Voronoi cells or Voronoi cells excluding generating points [8]. In our case in (3), one can observe both situations: the Voronoi cell of a sensor can be empty, especially if its sensing range is very large compared to the others' sensing ranges and the diameter of the environment $\mathcal{W}$, denoted by $\mathrm{diam}(\mathcal{W}):=\max_{\mathrm{x},\mathrm{y}\in\mathcal{W}}\|\mathrm{x}-\mathrm{y}\|$; and, likewise, nonempty Voronoi cells might exclude the associated sensor positions. However, one can always ensure that Voronoi cells contain the respective sensor positions by bounding the sensors' difference in sensing range from above by their spatial separation distances. Hence, we find it convenient to summarize these observations as follows:

Proposition 1 *Let $\mathcal{V}(\mathbf{s})$ be the generalized Voronoi diagram of $\mathcal{W}$ in (3) associated with a sensor network configuration $\mathbf{s}\in(\mathcal{W}\times R_{>0})^n$. Then, one has the following properties:*

(i) $V_i=\varnothing$ *if* $\log r_i^2 > \log r_j^2+\frac{\mathrm{diam}(\mathcal{W})^2}{d\,r_j^2}$ *for all* $j\neq i$.

(ii) $\mathrm{p}_i\in V_i$ *iff* $\log r_i^2 \leq \log r_j^2+\frac{\|\mathrm{p}_i-\mathrm{p}_j\|^2}{d\,r_j^2}$ *for all* $j\neq i$.

(iii) $\mathrm{p}_i\in V_i$ *if* $\|\mathrm{p}_i-\mathrm{p}_j\|^2 \geq d\big|r_i^2-r_j^2\big|$ *for all* $j\neq i$.[4]

Proof. The result (i) directly follows from (3c), because $0\leq\|\mathrm{x}-\mathrm{p}_j\|\leq\mathrm{diam}(\mathcal{W})$ for any $\mathrm{x},\mathrm{p}_j\in\mathcal{W}$. Similarly, one can obtain the result (ii) by substituting $\mathrm{x}=\mathrm{p}_i$ in the predicate of (3c).

Finally, if $r_i\leq r_j$, then we have $\log r_i^2\leq\log r_j^2+\frac{\|\mathrm{p}_i-\mathrm{p}_j\|^2}{d\,r_j^2}$. Otherwise ($r_i>r_j$), $\|\mathrm{p}_i-\mathrm{p}_j\|^2\geq d\big(r_i^2-r_j^2\big)$ implies that $\frac{\|\mathrm{p}_i-\mathrm{p}_j\|^2}{d\,r_j^2}\geq\frac{r_i^2}{r_j^2}-1\geq\log\frac{r_i^2}{r_j^2}$, because $\log x\leq x^2-1$ for $x\geq 1$. Thus, the result (iii) is evident from the result (ii). ■

C. Quantifying Coverage Quality via Statistical Distances

Inspired by the use of Gaussian mixture models in learning approximate representations of complex probability distribution functions [39], we define the overall event detection probability of a sensor network, at a coverage configuration $\mathbf{s}\in(\mathcal{W}\times R_{>0})^n$, as a weighted average of the individual event detection probabilities of the sensors as

$$\varphi_{\boldsymbol{\theta}}(\mathrm{x}) := \sum_{i=1}^{n}\pi_i q_{\mathrm{s}_i}(\mathrm{x}), \tag{4}$$

where $\boldsymbol{\pi}:=(\pi_1,\pi_2,\dots,\pi_n)\in[0,1]^n$ denotes a list of normalized mixture weights that satisfy $\sum_{i=1}^n\pi_i=1$. For the sake of notational simplicity, we shall denote the extended *soft* state of a sensor network by $\boldsymbol{\theta}:=(\mathbf{s},\boldsymbol{\pi})\in(\mathcal{W}\times\mathbb{R}_{>0})^n\times\Omega(n)$, where $\Omega(n)$ is the set of normalized mixture weights,

$$\Omega(n) := \Big\{ \boldsymbol{\pi}\in[0,1]^n \Big| \sum\nolimits_{i=1}^{n}\pi_i=1 \Big\}. \tag{5}$$

The notion of soft sensor network configurations will become clear in Section II-E. Here, one can intuitively interpret mixture weights as task weights determining the role of each sensor in the coverage task. For example, if the sensors are desired to participate equally in the coverage task, their mixture weights can be set to be equal, i.e., $\pi_i=\frac{1}{n}$ for all i. Hence, (small) mixture weights can also be utilized to identify redundant sensors in the coverage task.

Note that $\varphi_{\boldsymbol{\theta}}(\mathrm{x})$ is a valid probability distribution over $\mathbb{R}^d$ as are all the individual event detection probabilities q_{s_i}. Hence, given an event distribution $\phi:\mathcal{W}\to\mathbb{R}_{>0}$ over $\mathcal{W}$, in order to determine the event coverage performance of a soft sensor network configuration, $\boldsymbol{\theta}\in(\mathcal{W}\times\mathbb{R}_{>0})^n\times\Omega(n)$, we propose to use an f-divergence [40], [41] that measures the statistical distance between the event distribution ϕ and the average event detection probability $\varphi_{\boldsymbol{\theta}}$ as[5]

$$D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi) := \int_{\mathcal{W}} f\left(\frac{\varphi_{\boldsymbol{\theta}}(\mathrm{x})}{\phi(\mathrm{x})}\right)\phi(\mathrm{x})\mathrm{d}\mathrm{x}, \tag{6}$$

where $f:\mathbb{R}_{>0}\to\mathbb{R}$ is a continuously differentiable convex function with $f(1)=0$. For example, $f(t)=t\log t$, $f(t)=-\log t$, $f(t)=1-\sqrt{t}$ and $f(t)=t^2-1$ correspond to the Kullback-Leibler (KL) divergence, the reverse KL divergence, the Hellinger distance, and χ^2 divergence, respectively, which are a member of the family of α-divergences associated with $f(t)=\frac{1}{\alpha(\alpha-1)}(t^\alpha-1)$ for some $\alpha\in\mathbb{R}$ [41], [42]. Here, while the relations of the Hellinger distance and χ^2 divergence with the α-divergence are obvious, the α-divergence is known to simplify back to the (resp. reverse) KL divergence as $\alpha\to 1$ (resp. $\alpha\to 0$) [42]. Also note that, due to Jensen's inequality, an f-divergence of two probability distributions with the same support is always nonnegative and it is zero only if two distributions are identical [41], [42].

[4] A similar geometric condition is required to resolve the degeneracy of power diagrams, additively weighted generalized Voronoi diagrams whose generating objects are Euclidean balls [33].

[5] To handle unnormalized distribution functions, the f-divergence between distribution functions $\varphi:X\to\mathbb{R}_{>0}$ and $\phi:X\to\mathbb{R}_{>0}$ over a shared domain X can be defined as [42]

$$D_{f,X}(\varphi,\phi) := \int_X\left(f\left(\frac{\varphi(x)}{\phi(x)}\right)\phi(x)+\big(f'(1)-f(1)\big)\phi(x)-f'(1)\varphi(x)\right)\mathrm{d}x$$

where $f:\mathbb{R}_{>0}\to\mathbb{R}$ is a convex function. Here, the additional terms act as a regularizer that measures how much the mass distributions of φ and ϕ are concentrated over the domain X.

Since the event distribution function ϕ is normalized over $\mathcal{W}$, for the sake of simplifying the notational presentation and to be able to highlight the connection with the existing literature on Voronoi-based coverage control [9], we simply ignore these additional regularization terms in this paper, but our results can be simply extended to handle unnormalized distributions.

D. Some Important Bounds on f-Divergences

Since the overall event detection probability $\varphi_{\boldsymbol{\theta}}$ is defined as a convex combination of the individual event detection probabilities q_{s_i}, using Jensen's inequality, the f-divergence $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ can be bounded from below and above as:

Proposition 2 *For any convex function f over $(0,\infty)$, the f-divergence $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ in (6) satisfies*

$$f\left(\sum_{i=1}^{n}\pi_i\int_{\mathcal{W}}q_{\mathrm{s}_i}(\mathrm{x})\mathrm{dx}\right)\leq D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)\leq\sum_{i=1}^{n}\pi_i D_{f,\mathcal{W}}(q_{\mathrm{s}_i},\phi). \quad (7)$$

Proof. The lower bound simply follows from the continuous form of Jensen's inequality for probability distributions, i.e.,

$$D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)=\int_{\mathcal{W}}f\left(\frac{\varphi_{\boldsymbol{\theta}}(\mathrm{x})}{\phi(\mathrm{x})}\right)\phi(\mathrm{x})\mathrm{dx}\geq f\left(\int_{\mathcal{W}}\varphi_{\boldsymbol{\theta}}(\mathrm{x})\mathrm{dx}\right). \quad (8)$$

Similarly, since f is convex, the finite form of Jensen's inequality states that

$$f\left(\frac{\varphi_{\boldsymbol{\theta}}(\mathrm{x})}{\phi(\mathrm{x})}\right)=f\left(\sum_{i=1}^{n}\pi_i\frac{q_{\mathrm{s}_i}(\mathrm{x})}{\phi(\mathrm{x})}\right)\leq\sum_{i=1}^{n}\pi_i f\left(\frac{q_{\mathrm{s}_i}(\mathrm{x})}{\phi(\mathrm{x})}\right). \quad (9)$$

Thus, the upper bound follows from the monotonicity of the integral operator, which completes the proof. ■

Note that, for any monotone function f, both of the bounds in Proposition 2 can be considered as a coverage objective with no coordination constraint. More precisely, the lower bound measures how concentrated the individual event detection probabilities of sensors are over the environment $\mathcal{W}$ in average, whereas the upper bound quantifies how well the sensors' individual event detection probabilities approximate the event distribution ϕ over $\mathcal{W}$, in the sense of the f-divergence, in average. Hence, both of these coverage objectives can be independently optimized at the individual sensor level.

More interestingly, if f is monotonically decreasing (e.g., $f(t)=-\log t$) or increasing (e.g., $f(t)=t^2-1$) over $(0,\infty)$, then one can find upper and lower bounds on the f-divergence $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ in terms of the individual coverage performance of sensors over their respective Voronoi cells.

Proposition 3 *For any monotone decreasing convex function f over $(0,\infty)$ and any soft sensor network configuration $\boldsymbol{\theta}=(\mathbf{s},\boldsymbol{\pi})\in(\mathcal{W}\times\mathbb{R}_{>0})^n\times\Omega(n)$, with distinct sensor states, i.e., $\mathrm{s}_i\neq\mathrm{s}_j$ for all $i\neq j$, one has*

$$\sum_{i=1}^{n}D_{f,V_i}(q_{\mathrm{s}_i},\phi)\leq D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)\leq\sum_{i=1}^{n}D_{f,V_i}(\pi_i q_{\mathrm{s}_i},\phi), \quad (10)$$

and the inequalities in the reverse direction hold for an increasing convex function f. Here, $\mathcal{V}(\mathbf{s})=\{V_1,V_2,\ldots,V_n\}$ is the Voronoi diagram of $\mathcal{W}$ constructed as defined in (3).

Proof. By definitions (3) and (4), we have $q_{\mathrm{s}_i}(\mathrm{x})\geq q_{\mathrm{s}_j}(\mathrm{x})$ for all $\mathrm{x}\in V_i$ and $i\neq j$; and so $\pi_i q_{\mathrm{s}_i}(\mathrm{x})\leq\varphi_{\boldsymbol{\theta}}(\mathrm{x})\leq q_{\mathrm{s}_i}(\mathrm{x})$ for any $\mathrm{x}\in V_i$, since $q_{\mathrm{s}_i}(\mathrm{x})=q_{\mathrm{s}_i}(\mathrm{x})\sum_{j=1}^{n}\pi_j\geq\pi_i q_{\mathrm{s}_i}(\mathrm{x})+\sum_{j\neq i}\pi_j q_{\mathrm{s}_j}(\mathrm{x})$. Hence, for a monotone decreasing f, we have from the monotonicity of the integral operator that

$$D_{f,V_i}(q_{\mathrm{s}_i},\phi)\leq D_{f,V_i}(\varphi_{\boldsymbol{\theta}},\phi)\leq D_{f,V_i}(\pi_i q_{\mathrm{s}_i},\phi). \quad (11)$$

Thus, since a Voronoi diagram defines a partition of $\mathcal{W}$ for distinct sensor states, we have $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)=\sum_{i=1}^{n}D_{f,V_i}(\varphi_{\boldsymbol{\theta}},\phi)$ and so the inequalities in (10) directly follows.

Similarly, one can verify that the directions of inequalities are reversed for an increasing convex function f. ■

Once again, the bounds in Proposition 3 can be interpreted as a coverage objective with a Voronoi-based coordination constraint, and they measure how well the (scaled) individual event detection probabilities of sensors over the corresponding Voronoi cells approximate the event distribution ϕ in total. Although it does not hold in general, the bounds in Proposition 3 are usually tighter than the bounds in Proposition 2, especially around a locally optimal coverage configuration, because the individual event detection probabilities of the sensors are concentrated within their associated Voronoi cells at locally optimal coverage configurations (see Corollary 2).

To have a better intuitive understanding of the bounds on the f-divergence $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ in Proposition 3, one can relax the constraint of being a valid probability distribution on $\varphi_{\boldsymbol{\theta}}$ and redefine it as the maximum of the individual event detection probabilities of the sensors as

$$\widehat{\varphi}_{\mathbf{s}}(\mathrm{x}):=\max_{i\in\{1,2,\ldots,n\}}q_{\mathrm{s}_i}(\mathrm{x}), \quad (12)$$

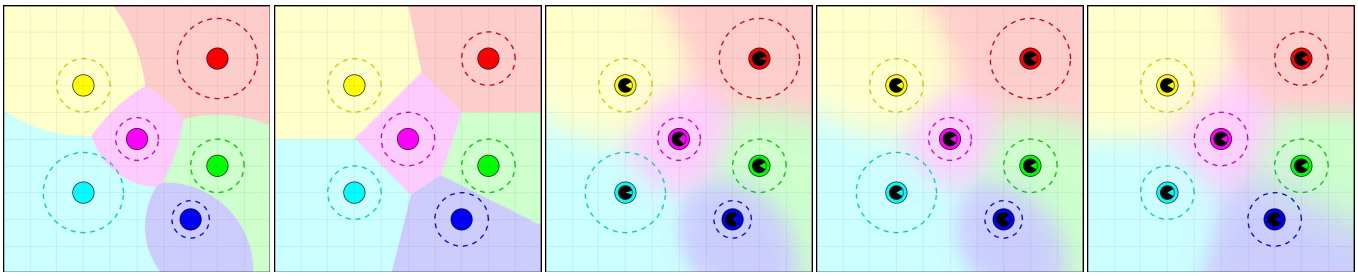
which is not necessarily a valid probability distribution, but the f-divergence in (6) is still well defined for comparison of any two positive functions over their shared support.[5] Note that $\varphi_{\boldsymbol{\theta}}$ in (4) is parametrized by a soft sensor network configuration $\boldsymbol{\theta}=(\mathbf{s},\boldsymbol{\pi})$, consisting of network state $\mathbf{s}$ and mixture (task) weights $\boldsymbol{\pi}$, whereas $\widehat{\varphi}_{\mathbf{s}}$ in (12) is parametrized by sensor network state $\mathbf{s}$. More importantly, observe from Proposition 3 that the f-divergences $D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ defines a bound on $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$, revealing an interesting connection between the f-divergence and Voronoi-based optimal sensor allocation of Section II-B, because

$$D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)=\sum\nolimits_{i=1}^{n}D_{f,V_i}(q_{\mathrm{s}_i},\phi), \quad (13)$$

where $\mathcal{V}(\mathbf{s})=\{V_1,V_2,\ldots,V_n\}$ is the additively and multiplicatively weighted Voronoi diagram of the environment $\mathcal{W}$, generated based on sensor configuration $\mathbf{s}$, as defined in (3).

Another significant connection with Voronoi-based coverage control approaches is that the objective function for the standard Voronoi-based coverage control of homogeneous (i.e., $r_i=r_j$ for all $i\neq j$) isotropic mobile sensors is the total expected event sensing cost, measured in terms of the squared Euclidean distance, of sensors within the respective Voronoi cells [9], which corresponds in our framework to setting the statistical distance to be the reverse Kullback-Leibler (KL) divergence (i.e., $f(t)=-\log t$), denoted by $D_{kl,\mathcal{W}}$, while assuming Voronoi-based sensor allocation.

Remark 1 The reverse KL divergence $D_{kl,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ measures the coverage quality of a sensor network configuration $\mathbf{s}\in(\mathcal{W}\times\mathbb{R}_{>0})^n$ in terms of an additively and multiplicatively weighted, expected squared Euclidean distance of events to sensors over their respective Voronoi cells as



(a) (b) (c) (d) (e)

Fig. 2. Hard and soft optimal sensor allocation. (a) Voronoi-based allocation of heterogeneous sensor networks with distinct sensing ranges (dashed circles), (b) Voronoi-based allocation of homogeneous sensor networks with identical sensing ranges, (c) Soft allocation of heterogeneous sensor networks with distinct sensing ranges and distinct mixture weights (pie-charts), (d) Soft allocation of heterogeneous sensor networks with distinct sensing ranges and identical mixture weights, (e) Soft allocation of homogeneous sensor networks with identical sensing ranges and identical mixture weights.

$$D_{kl,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)=\sum_{i=1}^{n}\int_{V_i} c_{\mathrm{s}_i}(\mathrm{x})\phi(\mathrm{x})\mathrm{d}\mathrm{x}-H_{\mathcal{W}}(\phi), \tag{14}$$

where $c_{\mathrm{s}_i}(\mathrm{x})$ is the spatial sensing cost of ith sensor as defined in (2) and $H_{\mathcal{W}}(\phi):=-\int_{\mathcal{W}}\phi(\mathrm{x})\log\phi(\mathrm{x})\mathrm{d}\mathrm{x}$ denotes the differential entropy of the event distribution ϕ over $\mathcal{W}$ and is constant. Moreover, for homogeneous sensors with a shared fixed sensing range parameter $r>0$ (i.e., $r_i=r$ for all i), the KL divergence $D_{kl,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ is proportional with the total expected squared Euclidean distance of events to sensors over the associated Voronoi cells, i.e.,

$$D_{kl,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)=\frac{1}{2r^2}\sum_{i=1}^{n}\int_{V_i}\|\mathrm{x}-\mathrm{p}_i\|^2\phi(\mathrm{x})\mathrm{d}\mathrm{x}+\frac{d}{2}\log(2\pi r)-H_{\mathcal{W}}(\phi).$$

E. Locally Optimal Coverage Configurations

We now provide a geometric characterization of locally optimal coverage configurations of the f-divergences $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ and $D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$. As we shall discuss below, at a locally optimal coverage configuration of $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$, each event location is assigned to every sensor based on its relative event detection probability, which is typically referred to as a *soft assignment*; whereas $D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$, by construction, relies on the Voronoi-based assignment of each event location to a particular sensor with the maximal event detection probability, which is typically referred to as a *hard assignment*, see Fig. 2.

1) Optimal Coverage Configurations for Soft Assignment: Soft assignment of event locations to sensors offers an intuitive formula for the gradient of the f-divergence $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ in terms of the associated workspace statistics:

Theorem 1 *The gradients of the f-divergence $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$, between the sensors' average event detection probability $\varphi_{\boldsymbol{\theta}}$ in (4) and the event distribution ϕ, with respect to sensor position p_i and squared sensing range r_i^2 are given by*

$$\nabla_{\mathrm{p}_i}D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)=\frac{m_{\mathcal{W}_i}}{r_i^2}(\mu_{\mathcal{W}_i}-\mathrm{p}_i), \tag{15a}$$

$$\nabla_{r_i^2}D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)=d\frac{m_{\mathcal{W}_i}}{2r_i^4}\left(\frac{\sigma^2_{\mathcal{W}_i}+\|\mu_{\mathcal{W}_i}-\mathrm{p}_i\|^2}{d}-r_i^2\right), \tag{15b}$$

where the mass $m_{\mathcal{W}_i}$, the centroid $\mu_{\mathcal{W}_i}$ and the variance $\sigma^2_{\mathcal{W}_i}$ of the environment $\mathcal{W}$ for the ith sensor are defined to be

$$m_{\mathcal{W}_i}:=\int_{\mathcal{W}}f'\left(\frac{\varphi_{\boldsymbol{\theta}}(\mathrm{x})}{\phi(\mathrm{x})}\right)\pi_i q_{\mathrm{s}_i}(\mathrm{x})\mathrm{d}\mathrm{x}, \tag{16a}$$

$$\mu_{\mathcal{W}_i}:=\frac{1}{m_{\mathcal{W}_i}}\int_{\mathcal{W}}\mathrm{x}\,f'\left(\frac{\varphi_{\boldsymbol{\theta}}(\mathrm{x})}{\phi(\mathrm{x})}\right)\pi_i q_{\mathrm{s}_i}(\mathrm{x})\mathrm{d}\mathrm{x}, \tag{16b}$$

$$\sigma^2_{\mathcal{W}_i}:=\frac{1}{m_{\mathcal{W}_i}}\int_{\mathcal{W}}\|\mathrm{x}-\mu_{\mathcal{W}_i}\|^2 f'\left(\frac{\varphi_{\boldsymbol{\theta}}(\mathrm{x})}{\phi(\mathrm{x})}\right)\pi_i q_{\mathrm{s}_i}(\mathrm{x})\mathrm{d}\mathrm{x}. \tag{16c}$$

Here, f' denotes the derivative of f.

Proof. Since the boundary of $\mathcal{W}$ is fixed, using the generalized Leibniz rule for differentiation under the integral sign [43], the partial derivative of $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ with respect to a sensor control parameter $\beta_i\in\{\mathrm{p}_i,r_i^2\}$ can be obtained as

$$\frac{\partial D_{f,\mathcal{W}}(\varphi_{\mathbf{s}},\phi)}{\partial\beta_i}=\int_{\mathcal{W}}\frac{\partial}{\partial\beta_i}f\left(\frac{\varphi_{\boldsymbol{\theta}}(\mathrm{x})}{\phi(\mathrm{x})}\right)\phi(\mathrm{x})\mathrm{d}\mathrm{x}, \tag{17a}$$

$$=\int_{\mathcal{W}}f'\left(\frac{\varphi_{\boldsymbol{\theta}}(\mathrm{x})}{\phi(\mathrm{x})}\right)\pi_i\frac{\partial q_{\mathrm{s}_i}(\mathrm{x})}{\partial\beta_i}\mathrm{d}\mathrm{x}, \tag{17b}$$

where

$$\frac{\partial q_{\mathrm{s}_i}(\mathrm{x})}{\partial\mathrm{p}_i}=q_{\mathrm{s}_i}(\mathrm{x})\frac{(\mathrm{x}-\mathrm{p}_i)^{\mathrm{T}}}{r_i^2}, \tag{18a}$$

$$\frac{\partial q_{\mathrm{s}_i}(\mathrm{x})}{\partial r_i^2}=q_{\mathrm{s}_i}(\mathrm{x})\frac{1}{2r_i^2}\left(\frac{\|\mathrm{x}-\mathrm{p}_i\|^2}{r_i^2}-d\right). \tag{18b}$$

Here, we can alternatively rewrite $\|\mathrm{x}-\mathrm{p}_i\|^2$ as

$$\|\mathrm{x}-\mathrm{p}_i\|^2=\|\mathrm{x}-\mu_{\mathcal{W}_i}\|^2+2(\mathrm{x}-\mu_{\mathcal{W}_i})^{\mathrm{T}}(\mu_{\mathcal{W}_i}-\mathrm{p}_i)+\|\mu_{\mathcal{W}_i}-\mathrm{p}_i\|^2, \tag{19}$$

Thus, after some arithmetic manipulation, one can verify the theorem. ■

Remark 2 While defining workspace statistics in (16), each event location $\mathrm{x}\in\mathcal{W}$ is *softly* assigned to every sensor with normalized assignment weight $\frac{\pi_i q_{\mathrm{s}_i}(\mathrm{x})}{\sum_{j=1}^{n}\pi_j q_{\mathrm{s}_j}(\mathrm{x})}$ that is proportional with the sensor's weighted event detection probability $\pi_i q_{\mathrm{s}_i}(\mathrm{x})$, as illustrated in Fig. 2.

Note that if sensors equally participate in the coverage task with identical mixture weights (i.e., $\pi_i=\frac{1}{n}$ for all i), as seen in Fig. 2(a)-(d) and Fig. 2(b)-(e), hard and soft sensor allocations are significantly consistent with each other, and soft sensor allocation yields a smooth transition in event-to-sensor assignment by offering overlapping coverage regions. Also observe that the mass $m_{\mathcal{W}_i}$ is negative for decreasing f (e.g., $f(t)=-\log t$ and $f(t)=1-\sqrt{t}$ corresponding to the reverse KL divergence and the Hellinger distance, respectively); however, the associated mean $\mu_{\mathcal{W}_i}$ is always contained in $\mathcal{W}$ and the variance $\sigma^2_{\mathcal{W}_i}$ is always positive for any

monotone function f. The sign of $m_{\mathcal{W}_i}$ will play a critical role in the design of coverage control in Section III and determine the sign of (negative or positive) feedback. Finally, we find it useful to emphasize that the global workspace statistics in (16) can be computed efficiently using a distributed consensus algorithm that combines local statistics obtained using local observations [44], [45].

A corollary of Theorem 1 is that a locally optimal soft coverage configuration satisfies:

Corollary 1 *At a locally optimal soft coverage configuration $\boldsymbol{\theta}=(\mathbf{p},\boldsymbol{r},\boldsymbol{\pi})\in(\mathcal{W}\times\mathbb{R}_{>0})^n\times\Omega(n)$ that locally minimizes $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$, all sensors are located at the associated workspace centroids with the centroidal sensing ranges, i.e.,*

$$\mathrm{p}_i=\mu_{\mathcal{W}_i},\quad \textit{and}\quad r_i^2=\tfrac{1}{d}\sigma^2_{\mathcal{W}_i},\quad \forall i. \tag{20}$$

Proof. By definition, a locally optimal coverage configuration is a critical point[6] of $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$, and so the result follows from Theorem 1. ■

Theorem 2 *For any monotone convex function f over $(0,\infty)$, at a locally optimal soft coverage configuration $\boldsymbol{\theta}=(\mathbf{p},\boldsymbol{r},\boldsymbol{\pi})\in(\mathcal{W}\times\mathbb{R}_{>0})^n\times\Omega(n)$ that locally minimizes $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$, the mixture weights π_i are proportional with the associated workspace masses $m_{\mathcal{W}_i}$ in (16a), i.e.,*

$$\pi_i=\frac{m_{\mathcal{W}_i}}{\sum_{j=1}^n m_{\mathcal{W}_j}},\quad \forall i. \tag{21}$$

Proof. Recall that mixture weights should be nonnegative and sum up to unity. The monotonicity of f ensures that the mixture weights in (21) are all positive because the workspace masses $m_{\mathcal{W}_i}$ in (16a) share the same sign for any monotone f and are all either positive or negative. Further, using the method of Lagrange multipliers for determining critical points of equality constrained optimization problems [46], a locally optimal choice of mixture weights can be determined as follows. Optimal mixture weights are a critical point of the following Lagrangian,

$$\mathcal{L}:=D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)+\lambda\Big(1-\sum\nolimits_{i=1}^n\pi_i\Big), \tag{22}$$

where $\lambda\in\mathbb{R}$ is a scalar Lagrange multiplier. The gradients of the Lagrangian $\mathcal{L}$, with respect to mixture weights and the Lagrange multiplier, are given by

$$\nabla_{\pi_i}\mathcal{L}=\int_{\mathcal{W}}f'\bigg(\frac{\varphi_{\boldsymbol{\theta}}(\mathrm{x})}{\phi(\mathrm{x})}\bigg)q_{\mathrm{s}_i}(\mathrm{x})\mathrm{dx}-\lambda=\frac{m_{\mathcal{W}_i}}{\pi_i}-\lambda, \tag{23a}$$

$$\nabla_{\lambda}\mathcal{L}=1-\sum\nolimits_{i=1}^n\pi_i. \tag{23b}$$

At a critical point of $\mathcal{L}$, $\nabla_\lambda\mathcal{L}=0$ guarantees $\sum_{i=1}^n\pi_i=1$, and, using $\nabla_{\pi_i}\mathcal{L}=0$, one can obtain that the optimal value of the Lagrange multiplier is given by $\lambda=\sum_{i=1}^n m_{\mathcal{W}_i}$, because

$$0=\sum\nolimits_{i=1}^n\pi_i\nabla_{\pi_i}\mathcal{L}=\Big(\sum\nolimits_{i=1}^n m_{\mathcal{W}_i}\Big)-\lambda. \tag{24}$$

Thus, we have from $\nabla_{\pi_i}\mathcal{L}=0$ that the optimal mixture weights should satisfy (21), which completes the proof. ■

[6]It is usually very hard to determine the type of critical points of $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$, but in practice gradient descent algorithms are usually observed to converge a local minimum of the $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$.

Finally, we find it important to note that, soft and hard assignments of homogeneous isotropic mobile sensors with a shared sensing range $r\in\mathbb{R}_{>0}$ (i.e., $r_i=r$ for all i) coincide as the shared sensing range r approaches to zero [39], which reveals another interesting connection between our statistical coverage control method and the standard Voronoi-based coverage control approaches [9].

Remark 3 For homogeneous isotropic sensor networks with a shared sensing range $r\in\mathbb{R}_{>0}$ and distinct positions (i.e., $r_i=r$ for all i and $\mathrm{p}_i\neq\mathrm{p}_j$ for all $i\neq j$), the soft assignment converges to the Voronoi-based hard assignment (almost everywhere over $\mathcal{W}$) as $r\to 0$, i.e.,

$$\lim_{r\to 0}\frac{\pi_i q_{\mathrm{s}_i}(\mathrm{x})}{\sum_{j=1}^n\pi_j q_{\mathrm{s}_j}(\mathrm{x})}=\begin{cases}1, & \mathrm{x}\in\mathring{V}_i,\\ \frac{\pi_i}{\sum_{j=1}^n\pi_j\mathbb{1}(V_i\cap V_j\neq\varnothing)}, & \mathrm{x}\in\partial V_i,\\ 0, & \text{otherwise},\end{cases} \tag{25}$$

where $\mathrm{s}_i=(\mathrm{p}_i,r)\in\mathcal{W}\times\mathbb{R}_{>0}$ and $\pi_i\in(0,1)$ denote the ith sensor's state and mixture weight; $\mathcal{V}(\mathbf{s})=\{V_1,V_2,\ldots,V_n\}$ is the Voronoi diagram of $\mathcal{W}$ defined in (3); $\mathring{V}_i$ and ∂V_i denote the interior and the boundary of Voronoi cell V_i, respectively; and $\mathbb{1}(.)$ is the indicator function that returns one if its argument is true, and zero otherwise.

Proof. For any $\mathrm{x}\in\mathcal{W}$, let $i^*:=\arg\max_{i\in\{1,2,\ldots,n\}}q_{\mathrm{s}_i}(\mathrm{x})$, and observe that $\lim_{r\to 0}\frac{q_{\mathrm{s}_i}(\mathrm{x})}{q_{\mathrm{s}_{i^*}}(\mathrm{x})}=1$ if $\mathrm{x}\in V_i$, and zero otherwise. Hence, since a Voronoi diagram defines a partition of $\mathcal{W}$ for distinct sensor positions, one can verify (25) by simply dividing the numerator and the denominator of the limit term in (25) by $q_{\mathrm{s}_{i^*}}(\mathrm{x})$. ■

2) Optimal Coverage Configurations for Hard Assignment: Similar to Theorem 1, the gradient of the f-divergence $D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ can be computed using Voronoi statistics as:

Theorem 3 *The gradients of the f-divergence $D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ in (13), between the maximum event detection probability $\widehat{\varphi}_{\mathbf{s}}$ of sensors in (12) and the event distribution ϕ, with respect to sensor position p_i and squared sensing range r_i^2 are given by*

$$\nabla_{\mathrm{p}_i}D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)=\frac{m_{V_i}}{r_i^2}(\mu_{V_i}-\mathrm{p}_i), \tag{26a}$$

$$\nabla_{r_i^2}D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)=d\frac{m_{V_i}}{2r_i^4}\bigg(\frac{\sigma^2_{V_i}+\|\mu_{\mathcal{W}_i}-\mathrm{p}_i\|^2}{d}-r_i^2\bigg), \tag{26b}$$

where the mass m_{V_i}, the centroid μ_{V_i} and the variance $\sigma^2_{V_i}$ of Voronoi cell V_i are defined to be

$$m_{V_i}:=\int_{V_i}f'\bigg(\frac{q_{\mathrm{s}_i}(\mathrm{x})}{\phi(\mathrm{x})}\bigg)q_{\mathrm{s}_i}(\mathrm{x})\mathrm{dx}, \tag{27a}$$

$$\mu_{V_i}:=\frac{1}{m_{V_i}}\int_{V_i}\mathrm{x}\,f'\bigg(\frac{q_{\mathrm{s}_i}(\mathrm{x})}{\phi(\mathrm{x})}\bigg)q_{\mathrm{s}_i}(\mathrm{x})\mathrm{dx}, \tag{27b}$$

$$\sigma^2_{V_i}:=\frac{1}{m_{V_i}}\int_{V_i}\|\mathrm{x}-\mu_{V_i}\|^2f'\bigg(\frac{q_{\mathrm{s}_i}(\mathrm{x})}{\phi(\mathrm{x})}\bigg)q_{\mathrm{s}_i}(\mathrm{x})\mathrm{dx}. \tag{27c}$$

Proof. Recall from (6) and (13) that

$$D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)=\sum_{i=1}^n D_{f,\mathcal{W}}(q_{\mathrm{s}_i},\phi)=\sum_{i=1}^n\int_{V_i}f\bigg(\frac{q_{\mathrm{s}_i}(\mathrm{x})}{\phi(\mathrm{x})}\bigg)\phi(\mathrm{x})\mathrm{dx}. \tag{28}$$

Hence, using the generalized Leibniz rule for differentiation under the integral sign [43] (also refer to Lemma 6.1 in [38]), the partial derivative of $D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ with respect to a sensor control parameter $\beta_i \in \{\mathrm{p}_i, r_i^2\}$ can be obtained as

$$\begin{aligned}\frac{\partial D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)}{\partial \beta_i} &= \int_{V_i} f'\left(\frac{q_{\mathrm{s}_i}(\mathrm{x})}{\phi(\mathrm{x})}\right)\frac{\partial q_{\mathrm{s}_i}(\mathrm{x})}{\partial \beta_i}\mathrm{dx} \\ &+ \int_{\partial V_i} f\left(\frac{q_{\mathrm{s}_i}(\mathrm{x})}{\phi(\mathrm{x})}\right)\phi(\mathrm{x})\,\mathrm{n}_{i_{\mathrm{x}}}\cdot\frac{\partial \mathrm{x}}{\partial \beta_i}\mathrm{dx} \\ &+ \sum_{j\neq i}\int_{\partial V_j} f\left(\frac{q_{\mathrm{s}_j}(\mathrm{x})}{\phi(\mathrm{x})}\right)\phi(\mathrm{x})\,\mathrm{n}_{j_{\mathrm{x}}}\cdot\frac{\partial \mathrm{x}}{\partial \beta_i}\mathrm{dx}, \quad (29)\end{aligned}$$

where $\mathrm{n}_{i_{\mathrm{x}}}$ denotes the outward surface normal of the Voronoi boundary ∂V_i at $\mathrm{x}\in\partial V_i$. By definition (3), $q_{\mathrm{s}_i}(\mathrm{x}) = q_{\mathrm{s}_j}(\mathrm{x})$ and $\mathrm{n}_{i_{\mathrm{x}}} = -\mathrm{n}_{j_{\mathrm{x}}}$ for any point $\mathrm{x}\in\partial V_i\cap\partial V_j$ on the shared boundary of adjacent Voronoi cells V_i and V_j (where $i\neq j$). Moreover, $\frac{\partial \mathrm{x}}{\partial \beta_i} = 0$ for any point $\mathrm{x}\in\partial\mathcal{W}$ on the workspace boundary $\partial\mathcal{W}$. Hence, the last two terms in (29) sum to zero, which yields

$$\frac{\partial D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)}{\partial \beta_i} = \int_{V_i} f'\left(\frac{q_{\mathrm{s}_i}(\mathrm{x})}{\phi(\mathrm{x})}\right)\frac{\partial q_{\mathrm{s}_i}(\mathrm{x})}{\partial \beta_i}\mathrm{dx}. \quad (30)$$

Thus, using (18) and (19), one can obtain the result. ■

Recall from Proposition 1 that a Voronoi cell can be empty. Hence, to resolve the indeterminacy due to an empty Voronoi cell, we set $m_{V_i}=0$, $\mu_{V_i}=\mathrm{p}_i$ and $\sigma^2_{V_i}=0$ if $V_i=\varnothing$. Moreover, observe that although the Voronoi mass m_{V_i} is negative for decreasing f, the associated mean μ_{V_i} is always in $\mathcal{W}$ (not necessarily in V_i, see Proposition 1) and the variance $\sigma^2_{V_i}$ is positive for any monotone f. The sign of m_{V_i} will become critical later in determining the sign of (negative or positive) feedback coverage control in Section III. As a final remark, it is useful to note that the Voronoi assignment of sensors and the associated Voronoi statistics in (27) can be efficiently computed using a distributed algorithm [9].

A direct consequence of Theorem 3 is that:

Corollary 2 *A locally optimal coverage configuration $\mathbf{s}=(\mathbf{p},\boldsymbol{r})\in(\mathcal{W}\times\mathbb{R}_{>0})^n$ is a critical point of $D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ where all sensors are located at the associated Voronoi centroids with Voronoi-centroidal sensing range, i.e.,*

$$\mathrm{p}_i = \mu_{V_i}, \quad \textit{and} \quad r_i^2 = \tfrac{1}{d}\sigma^2_{V_i}, \qquad \forall i. \quad (31)$$

III. Coverage Control via Statistical Distances

In this section, we first consider first-order continuous-time sensor dynamics and present a family of greedy coverage control algorithms for both soft and hard sensor allocation; then we present its adaptation to discrete-time sensor dynamics and describe its practical extensions for discrete event distributions, collision avoidance and differential drive sensor dynamics.

A. Continuous-Time Move-to-Centroidal-Configuration Laws

Assuming first-order (completely-actuated single-integrator) sensor dynamics,

$$\dot{\mathrm{p}}_i = u_{\mathrm{p}_i}, \quad \dot{r}_i^2 = u_{r_i^2}, \quad \dot{\pi}_i = u_{\pi_i}, \quad (32)$$

where u_{p_i}, $u_{r_i^2}$ and u_{π_i}, respectively, denote the control inputs for sensor position p_i, squared sensing range r_i^2 and mixture weight π_i, we now propose a new family of reactive coverage control policies, named the "move-to-centroidal-configuration" laws, that asymptotically bring the sensors to a locally optimal coverage configuration for both soft and hard sensor allocation.

1) The Move-to-Soft-Centroidal-Configuration Law: Using the workspace (mass $m_{\mathcal{W}_i}$, mean $\mu_{\mathcal{W}_i}$ and variance $\sigma^2_{\mathcal{W}_i}$) statistics in (16) associated with $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$, we define the "move-to-soft-centroidal-configuration" law at a soft coverage configuration $\boldsymbol{\theta}=(\mathbf{p},\boldsymbol{r},\boldsymbol{\pi})\in(\mathcal{W}\times\mathbb{R}_{>0})^n\times\Omega(n)$ to be

$$u_{\mathrm{p}_i} = \kappa_{\mathrm{p}_i}\,\mathrm{sgn}(m_{\mathcal{W}_i})(\mathrm{p}_i-\mu_{\mathcal{W}_i}), \quad (33a)$$

$$u_{r_i^2} = \kappa_{r_i^2}\,\mathrm{sgn}(m_{\mathcal{W}_i})\left(r_i^2 - \frac{\sigma^2_{\mathcal{W}_i}+\|\mathrm{p}_i-\mu_{\mathcal{W}_i}\|^2}{d}\right), \quad (33b)$$

$$u_{\pi_i} = \kappa_\pi \mathrm{sgn}(m_{\mathcal{W}_i})\left(\pi_i - \frac{m_{\mathcal{W}_i}}{\sum_{j=1}^n m_{\mathcal{W}_j}}\right), \quad (33c)$$

where $\kappa_{\mathrm{p}_i}>0$, $\kappa_{r_i^2}>0$ and $\kappa_\pi>0$ are fixed positive gains for position, sensing range and mixture weight control, respectively, and sgn returns the sign of its argument. Note that κ_π is a shared control gain for all mixture weights to ensure the positive invariance of $\Omega(n)$ in (5).

Theorem 4 *For any monotone decreasing[7] convex function f over $(0,\infty)$, the continuously differentiable "move-to-soft-centroidal-configuration" law leaves $(\mathcal{W}\times\mathbb{R}_{>0})^n\times\Omega(n)$ positively invariant, and its unique flow asymptotically brings sensors to a locally optimal coverage configuration of $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$, where $\mathrm{p}_i=\mu_{\mathcal{W}_i}$, $r_i^2=\frac{\sigma^2_{\mathcal{W}_i}}{d}$ and $\pi_i=\frac{m_{\mathcal{W}_i}}{\sum_{j=1}^n m_{\mathcal{W}_j}}$ for all i, while decreasing $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ along the way.*

Proof. The sign of the workspace mass $m_{\mathcal{W}_i}$ in (16a) is fixed for any monotone convex function f, because $m_{\mathcal{W}_i}$ is always positive (resp. negative) for increasing (resp. decreasing) f. Hence, the continuous differentiability of the "move-to-soft-centroidal-configuration" law follows from the integral form of the definition of the workspace statistics in (16). Further, as shown below, the level sets of $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ are positively invariant under the "move-to-soft-centroidal-configuration" law and so its continuous differentiability also guarantees the existence and uniqueness of the resulting flow over the positively invariant level sets of $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ [47], because a continuously differentiable function is also locally Lipschitz [48], and a locally Lipschitz function is globally Lipschitz over a compact domain [47].

The positive invariance of $(\mathcal{W}\times\mathbb{R}_{>0})^n\times\Omega(n)$ directly follows from the convexity of $\mathcal{W}$, $\mathbb{R}_{>0}$ and $\Omega(n)$, and the fact that, by definition (16), $\mu_{\mathcal{W}_i}\in\mathcal{W}$, $\sigma^2_{\mathcal{W}_i}\in\mathbb{R}_{>0}$, and $\frac{1}{\sum_{i=1}^n m_{\mathcal{W}_i}}(m_{\mathcal{W}_1},m_{\mathcal{W}_2},\ldots,m_{\mathcal{W}_n})\in\Omega(n)$, because for any monotone decreasing f, the "move-to-soft-centroidal-configuration" law continuously generates linear motion towards a moving goal point in the convex set $(\mathcal{W}\times\mathbb{R}_{>0})^n\times\Omega(n)$.

[7]Many f-divergences used in practice are monotone decreasing; for example, the reverse KL divergence associated with $f(t)=-\log t$, and the Hellinger distance associated with $f(t)=1-\sqrt{t}$. For any monotone increasing f, e.g., $f(t)=t^2-1$ corresponding to the χ^2 divergence, the soft workspace mass $m_{\mathcal{W}_i}$ in (16a) is positive, and so the positive feedback used in the coverage control in (33) cannot ensure the positive invariance of $(\mathcal{W}\times\mathbb{R}_{>0})^n\times\Omega(n)$. In general, the soft workspace mass $m_{\mathcal{W}_i}$ can be zero for any arbitrary choice of a convex function f over $(0,\infty)$, and one therefore needs to pay a special attention to resolve the associated indeterminacy.

To prove the asymptotic stability of the "move-to-soft-centroidal-configuration" law, one can consider $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ as a Lyapunov candidate, and verify from Theorem 1 that

$$\begin{aligned}\dot{D}_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi) = &-\sum_{i=1}^{n}\kappa_{\mathrm{p}_i}\pi_i\frac{|m_{\mathcal{W}_i}|}{r_i^2}\underbrace{\|\mathrm{p}_i-\mu_{\mathcal{W}_i}\|^2}_{\geq 0}\\ &-\sum_{i=1}^{n}\kappa_{r_i^2}\pi_i\frac{d|m_{\mathcal{W}_i}|}{2r_i^4}\underbrace{\left(r_i^2-\frac{r_{\mathcal{W}_i}^2+\|\mathrm{p}_i-\mu_{\mathcal{W}_i}\|^2}{d}\right)^2}_{\geq 0}\\ &-\frac{\kappa_\pi}{\sum_{i=1}^{n}|m_{\mathcal{W}_i}|}\underbrace{\left(\sum_{i=1}^{n}\frac{|m_{\mathcal{W}_i}|^2}{\pi_i}-\left(\sum_{i=1}^{n}|m_{\mathcal{W}_i}|\right)^2\right)}_{\geq 0}, \quad (34)\end{aligned}$$

$$\dot{D}_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)\leq 0, \quad (35)$$

which follows from $\frac{m_{\mathcal{W}_i}}{\sum_{j=1}^{n}m_{\mathcal{W}_j}}=\frac{|m_{\mathcal{W}_i}|}{\sum_{j=1}^{n}|m_{\mathcal{W}_j}|}$ for any monotone f and that we have from Jensen's inequality that

$$\sum_{i=1}^{n}\frac{|m_{\mathcal{W}_i}|^2}{\pi_i}=\sum_{i=1}^{n}\left(\frac{|m_{\mathcal{W}_i}|}{\pi_i}\right)^2\pi_i\geq\left(\sum_{i=1}^{n}|m_{\mathcal{W}_i}|\right)^2. \quad (36)$$

Hence, observe that $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ strictly decreases away from a locally optimal coverage configuration. Thus, it follows from LaSalle's Invariance Principle [47] that the sensors asymptotically converge to a critical point of $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$. ■

Note that the control inputs in (33) can be used separately, for example, for coverage optimization of sensor networks with static sensor positions or fixed sensing ranges or predefined constant mixture weights, while retaining stability properties.

2) The Move-to-Voronoi-Centroidal-Configuration Law: In the same spirit, using the Voronoi (mass m_{V_i}, mean μ_{V_i} and variance $\sigma_{V_i}^2$) statistics in (27) associated with $D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$, we define the "move-to-Voronoi-centroidal-configuration" law at any coverage configuration $\mathbf{s}=(\mathbf{p},\boldsymbol{r})\in(\mathcal{W}\times\mathbb{R}_{>0})^n$ to be

$$u_{\mathrm{p}_i}=\kappa_{\mathrm{p}_i}\operatorname{sgn}(m_{V_i})(\mathrm{p}_i-\mu_{V_i}), \quad (37a)$$

$$u_{r_i^2}=\kappa_{r_i^2}\operatorname{sgn}(m_{V_i})\left(r_i^2-\frac{\sigma_{V_i}^2+\|\mathrm{p}_i-\mu_{V_i}\|^2}{d}\right), \quad (37b)$$

where $\kappa_{\mathrm{p}_i}>0$ and $\kappa_{r_i^2}>0$ are fixed positive control gains.

Theorem 5 *For any monotone decreasing convex f over $(0,\infty)$, the configuration space $(\mathcal{W}\times\mathbb{R}_{>0})^n$ of sensors is positively invariant under the continuously differentiable "move-to-Voronoi-centroidal-configuration" law in (37) whose unique flow, starting at any configuration in $(\mathcal{W}\times\mathbb{R}_{>0})^n$, asymptotically decreases $D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ and reaches a locally optimal coverage configuration of $D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$.*

Proof. The monotonicity of f ensures that the Voronoi mass m_{V_i} has a fixed sign; more precisely, m_{V_i} is always positive for monotone increasing f and negative for monotone decreasing f. Hence, since the Voronoi statistics (mass, mean and variance) are continuously differentiable due to their integral form [49], the "move-to-Voronoi-centroidal-configuration" law is continuously differentiable and so is locally Lipschitz [48]. As we show below, the level sets of $D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ are positively invariant under the "move-to-Voronoi-centroidal-configuration" law. Therefore, since a locally Lipschitz function on a compact set is also globally Lipschitz on that set [47], the uniqueness and existence of its flow follows from its global Lipschitz continuity over the positively invariant level sets of $D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$.

Moreover, the convexity of $\mathcal{W}$ and $\mathbb{R}_{>0}$ implies the convexity of $(\mathcal{W}\times\mathbb{R}_{>0})^n$, and we also have, by definition (27), that $\mu_{V_i}\in\mathcal{W}$ and $\sigma_{V_i}^2\in\mathbb{R}_{>0}$. Thus, since the "move-to-Voronoi-centroidal-configuration" law continuously reconfigures the sensors towards a configuration in its convex domain $(\mathcal{W}\times\mathbb{R}_{>0})^n$, it leaves $(\mathcal{W}\times\mathbb{R}_{>0})^n$ positively invariant.

Finally, if $D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ is chosen as a Lyapunov function candidate, then we have from Theorem 3 that

$$\begin{aligned}\dot{D}_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi) = &-\sum_{i}^{n}\kappa_{\mathrm{p}_i}\frac{|m_{V_i}|}{r_i^2}\underbrace{\|\mathrm{p}_i-\mu_{V_i}\|^2}_{\geq 0}\\ &-\sum_{i}^{n}\kappa_{r_i^2}\frac{d|m_{V_i}|}{2r_i^4}\underbrace{\left(r_i^2-\frac{\sigma_{V_i}^2+\|\mathrm{p}_i-\mu_{V_i}\|^2}{d}\right)^2}_{\geq 0}\leq 0, \quad (38)\end{aligned}$$

which is strictly negative away from a critical point of $D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$. Hence, we have from LaSalle's Invariance Principle [47] that the sensors asymptotically reach a locally optimal coverage configuration of $D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$. ■

Once again, sensor positions and sensing ranges can be controlled separately as described in (37) for coverage optimization of sensor networks with static positions or fixed sensing ranges, while maintaining the stability properties summarized in Theorem 5. It is also useful to observe that:

Remark 4 For homogeneous mobile sensors with a shared fixed sensing range $r\in\mathbb{R}_{>0}$ (i.e., $r_i=r$ for all i), if the coverage quality is measured using the reverse KL divergence (i.e., $f(t)=-\log t$), then the "move-to-Voronoi-centroidal-configuration" law simplifies to the standard "move-to-Voronoi-centroid" law in [9] as the coverage objective $D_{kl,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ simplifies to the total expected squared Euclidean distance over a Voronoi diagram (see Remark 1).

B. Discrete-Time Move-to-Centroidal-Configuration Laws

A practical challenge of the continuous-time "move-to-centroidal-configuration" laws is the need for a proper adaptive step size selection [50] (e.g., based on line search [46]) in order to follow the associated coverage dynamics. Although it is difficult to resolve this issue in general, in this section we present the discrete-time version of "move-to-centroidal-configuration" laws constructed based on the forward Euler method and prove the asymptotic convergence of the resulting discrete-time coverage dynamics to a locally optimal coverage configuration, for both the reverse KL divergence and the α-divergence, using the MM (majorize-minimization or minorize-maximization) principle [21]. Hence, it is useful to briefly recall the idea of the MM principle.

Definition 1 ([21]) A scalar valued function $h:X\times X\to\mathbb{R}$ is said to *majorize* another scalar valued function $g:X\to\mathbb{R}$ at point $x'\in X$ if and only if $x\mapsto h(x,x')$ lies above $g(x)$ and is tangent to $g(x)$ at x', i.e.,

$$g(x)\leq h(x,x')\quad\forall x\in X,\quad\text{and}\quad g(x')=h(x',x'). \quad (39)$$

Similarly, $h(x,x')$ is said to *minorize* $g(x)$ iff $-h(x,x')$ majorizes $-h(x)$.

Therefore, instead of directly minimizing (resp. maximizing) a given objective function, the MM principle aims to construct an iterative optimization method that minimizes a majorizing (resp. maximizes a minorizing) surrogate function of the objective function, which is usually easier to optimize. Because, given a majorizing function $h(x,x')$ of $g(x)$ at $x' \in X$, if one can find $x^* \in X$ with $h(x^*,x') < h(x',x')$, then it also guarantees a strict decrease in g, i.e.,

$$g(x^*) \leq h(x^*,x') < h(x',x') = g(x'). \tag{40}$$

1) General Form of Discrete-Time Coverage Control: In order to iteratively decrease the f-divergence $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ associated with a differentiable convex function f over $(0,\infty)$, by simply applying the forward Euler method [50], one can construct a discrete-time version of the "move-to-soft-centroidal configuration" law in (33) that updates a soft sensor coverage configuration $\boldsymbol{\theta}[k] = (\mathbf{p}[k], \boldsymbol{r}[k], \boldsymbol{\pi}[k]) \in (\mathcal{W}\times\mathbb{R}_{>0})^n \times \Omega(n)$ at iteration $k \in \mathbb{N}$ as follows:

$$\mathrm{p}_i[k+1] = \mathrm{p}_i[k] + \delta_{\mathrm{p}_i}\mathrm{sgn}(m_{\mathcal{W}_i}[k])\big(\mathrm{p}_i[k] - \mu_{\mathcal{W}_i}[k]\big), \tag{41a}$$

$$r_i^2[k+1] = r_i^2[k] + \delta_{r_i^2}\mathrm{sgn}(m_{\mathcal{W}_i}[k])\bigg(r_i^2[k] - \frac{\sigma^2_{\mathcal{W}_i}[k] + \big\|\mu_{\mathcal{W}_i}[k] - \mathrm{p}_i[k]\big\|^2}{d}\bigg), \tag{41b}$$

$$\pi_i[k+1] = \pi_i[k] + \delta_{\pi}\mathrm{sgn}(m_{\mathcal{W}_i}[k])\bigg(\pi_i[k] - \frac{m_{\mathcal{W}_i}[k]}{\sum_{j=1}^n m_{\mathcal{W}_j}[k]}\bigg), \tag{41c}$$

where $\delta_{\mathrm{p}_i}, \delta_{r_i^2}, \delta_{\pi} \in (0,1]$ are fixed normalized step size parameters, sgn denotes the sign function, and the workspace statistics (mass $m_{\mathcal{W}_i}[k]$, mean $\mu_{\mathcal{W}_i}[k]$ and variance $\sigma^2_{\mathcal{W}_i}[k]$) of $\mathcal{W}$ associated with $D_{f,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k]},\phi\big)$ are defined as in (16).

Proposition 4 *For any decreasing convex f over $(0,\infty)$, the discrete-time "move-to-soft-centroid-configuration" law in (41) generates valid iterations over $(\mathcal{W}\times\mathbb{R}_{>0})^n \times \Omega(n)$, i.e., $\boldsymbol{\theta}[k] \in (\mathcal{W}\times\mathbb{R}_{>0})^n \times \Omega(n)$ implies $\boldsymbol{\theta}[k+1] \in (\mathcal{W}\times\mathbb{R}_{>0})^n \times \Omega(n)$.*

Proof. For any monotone decreasing function f, the workspace statistics in (16) satisfy $m_{\mathcal{W}_i}[k] < 0$, $\mu_{\mathcal{W}_i}[k] \in \mathcal{W}$ and $\frac{1}{d}(\sigma^2_{\mathcal{W}_i}[k] + \big\|\mu_{\mathcal{W}_i}[k] - \mathrm{p}_i[k]\big\|^2) \in \mathbb{R}_{>0}$. Hence, the result holds since the iterations in (41) is based on a convex combination of points in the convex set $(\mathcal{W}\times\mathbb{R}_{>0})^n \times \Omega(n)$. ■

Similarly, in the case of Voronoi-based sensor allocation, a discrete-time version of the "move-to-Voronoi-centroidal-configuration" law in (37) can be constructed to iteratively decrease the f-divergence $D_{f,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ at any coverage configuration $\mathbf{s}[k] = (\mathbf{p}[k], \boldsymbol{r}[k]) \in (\mathcal{W}\times\mathbb{R}_{>0})^n$ as

$$\mathrm{p}_i[k+1] = \mathrm{p}_i[k] + \delta_{\mathrm{p}_i}\mathrm{sgn}(m_{V_i}[k])\big(\mathrm{p}_i[k] - \mu_{V_i}[k]\big), \tag{42a}$$

$$r_i^2[k+1] = r_i^2[k] + \delta_{r_i^2}\mathrm{sgn}(m_{V_i}[k])\bigg(r_i^2[k] - \frac{\sigma^2_{V_i}[k] + \big\|\mu_{V_i}[k] - \mathrm{p}_i[k]\big\|^2}{d}\bigg), \tag{42b}$$

where $\delta_{\mathrm{p}_i}, \delta_{r_i^2} \in (0,1]$ are fixed normalized step sizes for sensor position and sensing range updates, and, the Voronoi statistics (mass $m_{V_i}[k]$, mean $\mu_{V_i}[k]$ and variance $\sigma^2_{V_i}[k]$), associated with $D_{f,\mathcal{W}}\big(\widehat{\varphi}_{\mathbf{s}[k]},\phi\big)$ and the Voronoi diagram $\mathcal{V}(\mathbf{s}[k]) = \{V_1[k], V_2[k], \ldots, V_n[k]\}$, are defined as in (27).

Proposition 5 *For any decreasing convex f over $(0,\infty)$, the discrete-time "move-to-Voronoi-centroid-configuration" law in (42) yields valid iterations over $(\mathcal{W}\times\mathbb{R}_{>0})^n$, that is to say, $\mathbf{s}[k] \in (\mathcal{W}\times\mathbb{R}_{>0})^n$ implies $\mathbf{s}[k+1] \in (\mathcal{W}\times\mathbb{R}_{>0})^n$.*

Proof. The proof follows similar lines as in the proof of Proposition 4 and hence is omitted. ■

Even though the update step sizes are limited to the unit interval above, this, in general, does not guarantee that the f-divergences $D_{f,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k]},\phi\big)$ and $D_{f,\mathcal{W}}\big(\widehat{\varphi}_{\mathbf{s}[k]},\phi\big)$ decrease at each iteration of the discrete-time "move-to-centroidal-configurations" laws in (41) and (42), respectively, away from a locally optimal coverage configuration. We shall show below that the restriction of step sizes to the unit interval guarantees a decrement in the coverage objective for the reverse KL divergence and the α-divergence, where $\alpha \in (0,1)$. Finally, it is important to observe that:

Remark 5 For a step size of unity (i.e., $\delta_{\mathrm{p}_i} = \delta_{r_i^2} = \delta_{\pi} = 1$) and the reverse KL divergence (i.e., $f(t) = -\log t$), the iterations of the discrete-time "move-to-centroidal-configuration" laws in (41) and (42) are identical to the iterations of the standard EM algorithm for learning (spherical) Gaussian mixture models and k-means clustering, respectively [39].

2) Discrete-Time Coverage Control via the KL Divergence: In this part, we focus our attention to the reverse KL divergence $D_{kl,\mathcal{W}}$, associated with $f(t) = -\log t$, and prove that the discrete-time "move-to-centroidal-configuration" laws in (41) and (42) asymptotically converge to a locally optimal coverage configuration for both soft and hard sensor allocation.

Our convergence analysis of the discrete-time "move-to-soft-centroidal configuration" law is strongly based on the idea of majorizing the reverse KL divergence $D_{kl,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$.

Lemma 1 *A majorizing surrogate function $M(\boldsymbol{\theta},\boldsymbol{\theta}[k])$ of the reverse KL divergence $D_{kl,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ at a soft coverage configuration $\boldsymbol{\theta}[k] \in (\mathcal{W}\times\mathbb{R}_{>0})^n \times \Omega(n)$ can be constructed as*

$$M(\boldsymbol{\theta},\boldsymbol{\theta}[k]) := \overline{M}(\boldsymbol{\theta},\boldsymbol{\theta}[k]) + \Delta M(\boldsymbol{\theta}[k]), \tag{43}$$

where

$$\overline{M}(\boldsymbol{\theta},\boldsymbol{\theta}[k]) := \sum_{i=1}^n -m_{\mathcal{W}_i}[k]\bigg(\frac{\sigma^2_{\mathcal{W}_i}[k] + \|\mathrm{p}_i - \mu_{\mathcal{W}_i}[k]\|^2}{2r_i^2} + \frac{d}{2}\log r_i^2 - \log\pi_i\bigg), \tag{44a}$$

$$\Delta M(\boldsymbol{\theta}[k]) := D_{f,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k]},\phi\big) - \overline{M}(\boldsymbol{\theta}[k],\boldsymbol{\theta}[k]), \tag{44b}$$

and the workspace statistics (mass, mean, variance) associated with $D_{kl,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k]},\phi\big)$ are defined as in (16) and given by

$$m_{\mathcal{W}_i}[k] := -\int_{\mathcal{W}} \frac{\phi(\mathrm{x})}{\varphi_{\boldsymbol{\theta}[k]}(\mathrm{x})}\pi_i[k]q_{\mathrm{s}_i[k]}(\mathrm{x})\mathrm{dx}, \tag{45a}$$

$$\mu_{\mathcal{W}_i}[k] := \frac{-1}{m_{\mathcal{W}_i}[k]}\int_{\mathcal{W}} \mathrm{x}\,\frac{\phi(\mathrm{x})}{\varphi_{\boldsymbol{\theta}[k]}(\mathrm{x})}\pi_i[k]q_{\mathrm{s}_i[k]}(\mathrm{x})\mathrm{dx}, \tag{45b}$$

$$\sigma^2_{\mathcal{W}_i}[k] := \frac{-1}{m_{\mathcal{W}_i}[k]}\int_{\mathcal{W}} \|\mathrm{x} - \mu_{\mathcal{W}_i}[k]\|^2\frac{\phi(\mathrm{x})}{\varphi_{\boldsymbol{\theta}[k]}(\mathrm{x})}\pi_i[k]q_{\mathrm{s}_i[k]}(\mathrm{x})\mathrm{dx}. \tag{45c}$$

Proof. We will provide a proof by construction. It is convenient to rewrite the reverse KL divergence as

$$D_{kl,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi) = \int_{\mathcal{W}} -\log\bigg(\frac{\varphi_{\boldsymbol{\theta}}(\mathrm{x})}{\phi(\mathrm{x})}\bigg)\phi(\mathrm{x})\mathrm{dx}, \tag{46}$$

$$= \int_{\mathcal{W}} -\log(\varphi_{\boldsymbol{\theta}}(\mathrm{x}))\phi(\mathrm{x})\mathrm{dx} - H_{\mathcal{W}}(\phi), \tag{47}$$

where $H_{\mathcal{W}}(\phi) = -\int_{\mathcal{W}} \phi(\mathrm{x})\log\phi(\mathrm{x})\mathrm{dx}$ is the differential entropy of the event distribution ϕ over $\mathcal{W}$ and is constant.

By applying Lemma 4 in the appendix, we can obtain a majorizing function of $-\log(\varphi_{\boldsymbol{\theta}}(\mathrm{x}))$ at around $\boldsymbol{\theta}[k]$ as

$$Q(\mathrm{x},\boldsymbol{\theta},\boldsymbol{\theta}[k]) := \sum_{i=1}^{n} -\log\left(\frac{\pi_i q_{\mathrm{s}_i}(\mathrm{x})}{\widehat{q}_{\mathrm{s}_i[k]}(\mathrm{x})}\right)\widehat{q}_{\mathrm{s}_i[k]}(\mathrm{x}), \quad (48)$$

$$\geq -\log(\varphi_{\boldsymbol{\theta}}(\mathrm{x})), \quad (49)$$

$$Q(\mathrm{x},\boldsymbol{\theta}[k],\boldsymbol{\theta}[k]) = -\log\big(\varphi_{\boldsymbol{\theta}[k]}(\mathrm{x})\big) \quad (50)$$

where

$$\widehat{q}_{\mathrm{s}_i[k]}(\mathrm{x}) := \frac{\pi_i[k] q_{\mathrm{s}_i[k]}(\mathrm{x})}{\varphi_{\boldsymbol{\theta}[k]}(\mathrm{x})}. \quad (51)$$

Hence, due to the monotonicity of the integral operator, this also provides a majorizing surrogate function

$$M(\boldsymbol{\theta},\boldsymbol{\theta}[k]) := \int_{\mathcal{W}} Q(\mathrm{x},\boldsymbol{\theta},\boldsymbol{\theta}[k])\phi(\mathrm{x})\mathrm{d}\mathrm{x} - H_{\mathcal{W}}(\phi), \quad (52)$$

of the reverse KL divergence $D_{kl,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ at $\boldsymbol{\theta}[k]$, i.e.,

$$D_{kl,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi) \leq M(\boldsymbol{\theta},\boldsymbol{\theta}[k]), \quad (53)$$

$$D_{kl,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k]},\phi\big) = M(\boldsymbol{\theta}[k],\boldsymbol{\theta}[k]). \quad (54)$$

To have a better intuitive understanding, one can alternatively rewrite $M(\boldsymbol{\theta},\boldsymbol{\theta}[k])$ using the workspace statistics in (45) as

$$M(\boldsymbol{\theta},\boldsymbol{\theta}[k]) = \sum_{i=1}^{n} -m_{\mathcal{W}_i}[k]\left(\frac{\sigma^2_{\mathcal{W}_i}[k] + \|\mathrm{p}_i - \mu_{\mathcal{W}_i}[k]\|^2}{2r_i^2} + \frac{d}{2}\log r_i^2 - \log\pi_i\right)$$
$$- \sum_{i=1}^{n} -m_{\mathcal{W}_i}[k]\left(\frac{\sigma^2_{\mathcal{W}_i}[k] + \|\mathrm{p}_i[k] - \mu_{\mathcal{W}_i}[k]\|^2}{2r_i^2[k]} + \frac{d}{2}\log r_i^2[k] - \log\pi_i[k]\right)$$
$$+ D_{kl,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k]},\phi\big), \quad (55)$$

which is identical to (43). Thus, the result follows. ■

Theorem 6 *Away from a locally optimal coverage configuration, the reverse KL divergence $D_{kl,\mathcal{W}}(\varphi_{\boldsymbol{\theta}[k]},\phi)$ decreases at each iteration $k \in \mathbb{N}$ of the discrete-time "move-to-soft-centroidal-configuration" law in (41) and so its iterations asymptotically converge to a locally optimal coverage configuration of $D_{kl,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ that satisfies*

$$\mathrm{p}_i[k] = \mu_{\mathcal{W}_i}[k], \quad r_i^2[k] = \tfrac{1}{d}\sigma^2_{\mathcal{W}_i}[k], \quad \pi_i[k] = \frac{m_{\mathcal{W}_i}[k]}{\sum_{j=1}^{n} m_{\mathcal{W}_j}[k]}. \quad (56)$$

Proof. As in the proof of Theorem 2, using the method of Lagrange multiplier, one can verify that the optimal Lagrange multiplier for the Lagrangian

$$\mathcal{L} = M(\boldsymbol{\theta},\boldsymbol{\theta}[k]) + \lambda\Big(1 - \sum\nolimits_{i=1}^{n}\pi_i\Big), \quad (57)$$

is $\lambda^* = \sum_{i=1}^{n} m_{\mathcal{W}_i}[k]$ and its unique global minimum with respect to each individual sensor control parameter is achieved, respectively, at

$$\mathrm{p}_i^* = \mu_{\mathcal{W}_i}[k], \quad (58a)$$

$$r_i^{*2} = \frac{1}{d}\Big(\sigma^2_{\mathcal{W}_i}[k] + \|\mathrm{p}_i - \mu_{\mathcal{W}_i}\|^2\Big), \quad (58b)$$

$$\pi_i^* = m_{\mathcal{W}_i}[k] \Big/ \sum\nolimits_{j=1}^{n} m_{\mathcal{W}_j}[k]. \quad (58c)$$

which are also, by definition, the optimal control parameters that uniquely and globally minimize $\boldsymbol{\theta} \to M(\boldsymbol{\theta},\boldsymbol{\theta}[k])$. Here the global optimality follows from the second-order sufficiency condition for optimality [46], because the Hessian of the Lagrangian $\mathcal{L}$ with respect to each control parameter is positive. Hence, at the globally optimal soft coverage configuration of $\boldsymbol{\theta} \mapsto M(\boldsymbol{\theta},\boldsymbol{\theta}[k])$, sensors satisfy

$$\mathrm{p}_i = \mu_{\mathcal{W}_i}[k], \quad {r_i}^2 = \frac{1}{d}\sigma^2_{\mathcal{W}_i}[k], \quad \pi_i = \frac{m_{\mathcal{W}_i}[k]}{\sum_{j=1}^{n} m_{\mathcal{W}_j}[k]}. \quad (59)$$

Now let us denote by $\boldsymbol{\theta}^{'}[k] = (\mathbf{p}[k], \boldsymbol{r}[k+1], \boldsymbol{\pi}[k])$ and $\boldsymbol{\theta}^{''}[k] = (\mathbf{p}[k+1], \boldsymbol{r}[k+1], \boldsymbol{\pi}[k])$ the intermediate coverage configurations between $\boldsymbol{\theta}[k]$ and $\boldsymbol{\theta}[k+1]$ after the sensing range and sensor position updates in (41b) and (41a), respectively. Note that, by definition, any convex function $g : X \to \mathbb{R}$ over its convex domain X with the global optimum $x^* \in X$ satisfies for any $x \in X$ and $\delta \in [0,1]$ that [51]

$$g(x^*) \leq g(\delta x^* + (1-\delta)x) \leq \delta g(x^*) + (1-\delta)g(x) \leq g(x). \quad (60)$$

Similarly, any function $h : I \to \mathbb{R}$ of a single variable over an interval $I \subset \mathbb{R}$ with the unique minimum $y^* \in I$ satisfies for any $y \in I$ and $\delta \in [0,1]$ that

$$h(y^*) \leq h(\delta y^* + (1-\delta)y) \leq h(y). \quad (61)$$

Hence, by applying (60) for $\mathrm{p}_i \to M(\boldsymbol{\theta},\boldsymbol{\theta}[k])$ over the convex domain $\mathcal{W}$ and for $\boldsymbol{\pi} \to M(\boldsymbol{\theta},\boldsymbol{\theta}[k])$ over the convex domain $\Omega(n)$, and by applying (61) for $r_i^2 \to M(\boldsymbol{\theta},\boldsymbol{\theta}[k])$ over the interval $\mathbb{R}_{>0}$, we obtain that the majorizing function $M(\boldsymbol{\theta},\boldsymbol{\theta}[k])$ of $D_{kl,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ decreases at each intermediate update in (41)(b-a-c), i.e.,

$$D_{kl,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k]},\phi\big) = M(\boldsymbol{\theta}[k],\boldsymbol{\theta}[k]) \geq M(\boldsymbol{\theta}^{'}[k],\boldsymbol{\theta}[k]) \geq M(\boldsymbol{\theta}^{''}[k],\boldsymbol{\theta}[k]),$$
$$\geq M(\boldsymbol{\theta}[k+1],\boldsymbol{\theta}[k]) \geq D_{kl,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k+1]},\phi\big). \quad (62)$$

Hence, the overall decrement on $D_{kl,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k]},\phi\big)$ at each iteration is bounded below by

$$D_{kl,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k]},\phi\big) - D_{kl,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k+1]},\phi\big) \geq M(\boldsymbol{\theta}[k],\boldsymbol{\theta}[k]) - M(\boldsymbol{\theta}[k+1],\boldsymbol{\theta}[k]),$$
$$\geq 0, \quad (63)$$

which is strictly positive away from a critical point of $D_{kl,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$. Thus, the iterations of the discrete-time "move-to-soft-centroidal-configuration" law asymptotically converge a locally optimal coverage configuration of $D_{kl,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$. ■

Similarly, for the reverse KL divergence, the discrete-time "move-to-Voronoi-centroidal configuration" law satisfies:

Theorem 7 *The discrete-time "move-to-Voronoi-centroidal configuration" law in (42) strictly decreases the reverse KL divergence $D_{kl,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}[k]},\phi)$ at each iteration $k \in \mathbb{N}$ away from a critical point of $D_{kl,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}[k]},\phi)$ and so asymptotically converges to a locally optimal configuration where*

$$\mathrm{p}_i[k] = \mu_{V_i}[k], \quad \textit{and} \quad r_i^2[k] = \tfrac{1}{d}\sigma^2_{V_i}[k], \qquad \forall i. \quad (64)$$

Proof. By definition (3), Voronoi diagrams determine optimal sensor allocation that maximizes the event detection probabilities of sensors. Hence, by fixing sensor allocation with the Voronoi diagram $\mathcal{V}(\mathbf{s}[k])$, a majorizing surrogate function of $D_{kl,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ in (13) at $\mathbf{s}[k]$ can be constructed as

$$Q(\mathbf{s},\mathbf{s}[k]) := \sum_{i=1}^{n} D_{kl,V_i[k]}(q_{\mathrm{s}_i},\phi), \tag{65}$$

$$= \sum_{i=1}^{n} -m_{V_i[k]}\left(\frac{\sigma_{V_i}^2[k]+\|\mathrm{p}_i-\mu_{V_i}[k]\|^2}{2r_i^2}+\frac{d}{2}\log(2\pi r_i^2)\right) - H_{\mathcal{W}}(\phi), \tag{66}$$

where $H_{\mathcal{W}}(\phi)$ denotes the differential entropy of ϕ over $\mathcal{W}$. Note that, since $m_{V_i[k]} < 0$ for the reverse KL divergence, $Q(\mathbf{s},\mathbf{s}[k])$ is a convex quadratic function of sensor position p_i over the convex set $\mathcal{W}$ whose unique global minimum is achieved at $\mathrm{p}_i = \mu_{V_i}[k]$. Similarly, observe that $Q(\mathbf{s},\mathbf{s}[k])$ is uniquely and globally minimized with respect to squared sensing range r_i^2 over $\mathbb{R}_{>0}$ at $r_i^2 = \frac{1}{d}\left(\sigma_{V_i}^2[k]+\|\mathrm{p}_i-\mu_{V_i}[k]\|^2\right)$. Hence, the unique global optimal of $\mathbf{s} \to Q(\mathbf{s},\mathbf{s}[k])$ over $(\mathcal{W}\times\mathbb{R}_{>0})^n$ is realized at

$$\mathrm{p}_i = \mu_{V_i}[k], \quad \text{and} \quad r_i^2 = \tfrac{1}{d}\sigma_{V_i}^2[k], \qquad \forall i. \tag{67}$$

Accordingly, if the intermediate state after the sensing range update in (42b) is denoted by $\mathbf{s}'[k] = (\mathbf{p}[k], \boldsymbol{r}[k+1])$, since $Q(\mathbf{s},\mathbf{s}[k])$ majorizes $D_{kl,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ at $\mathbf{s}[k]$, as discussed in the proof of Theorem 6, one can verify that $D_{kl,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ decreases at each iteration of (42), i.e.,

$$\begin{aligned} D_{kl,\mathcal{W}}\big(\widehat{\varphi}_{\mathbf{s}[k]},\phi\big) &= Q(\mathbf{s}[k],\mathbf{s}[k]) \geq Q(\mathbf{s}'[k],\mathbf{s}[k]), \\ &\geq Q(\mathbf{s}[k+1],\mathbf{s}[k]) \geq D_{kl,\mathcal{W}}\big(\widehat{\varphi}_{\mathbf{s}[k+1]},\phi\big), \end{aligned} \tag{68}$$

where the inequalities are strict away from a critical point of $D_{kl,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$. Thus, for the reverse KL divergence, the iterations of the discrete-time "move-to-Voronoi-centroidal-configuration" law asymptotically converge a locally optimal configuration of $D_{kl,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ that satisfies (64). ■

3) Discrete-Time Coverage Control via the α-Divergence: To extend the applicability of our results beyond the KL divergence, we now consider the α-divergence, associated with $f(t)=\frac{1}{\alpha(\alpha-1)}(t^\alpha-1)$ for some fixed $\alpha\in\mathbb{R}$, i.e.,

$$D_{\alpha,\mathcal{W}}(\varphi,\phi) := \frac{1}{\alpha(\alpha-1)}\left(\int_{\mathcal{W}} \varphi^\alpha(\mathrm{x})\phi^{1-\alpha}(\mathrm{x})\mathrm{dx}-1\right), \tag{69}$$

where the integral term is known to be the Hellinger integrals [41], denoted by

$$\mathcal{H}_{\alpha,\mathcal{W}}(\varphi,\phi) := \int_{\mathcal{W}} \varphi^\alpha(\mathrm{x})\phi^{1-\alpha}(\mathrm{x})\mathrm{dx}. \tag{70}$$

The α-divergence satisfies $D_{\alpha,\mathcal{W}}(\varphi,\phi) = D_{1-\alpha,\mathcal{W}}(\phi,\varphi)$ and is known to converge the (resp. reverse) KL divergence as $\alpha\to 1$ (resp. $\alpha\to 0$) [42].

Another very interesting, but not so obvious connection between the reverse KL divergence and the α-divergence is:

Lemma 2 *For any $\alpha\in\mathbb{R}$, the logarithm of the Hellinger integral $\mathcal{H}_{\alpha,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ can be minorized at around a soft coverage configuration $\boldsymbol{\theta}[k]\in(\mathcal{W}\times\mathbb{R}_{>0})^n\times\Omega(n)$ in terms of the reverse KL divergence by*

$$M_\alpha(\boldsymbol{\theta},\boldsymbol{\theta}[k]) := \overline{M}_\alpha(\boldsymbol{\theta},\boldsymbol{\theta}[k]) + \Delta M_\alpha(\boldsymbol{\theta}[k]), \tag{71}$$

where

$$\overline{M}_\alpha(\boldsymbol{\theta},\boldsymbol{\theta}[k]) := -\frac{\alpha}{\mathcal{H}_{\alpha,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k]},\phi\big)} D_{kl,\mathcal{W}}\Big(\varphi_{\boldsymbol{\theta}},\varphi_{\boldsymbol{\theta}[k]}^\alpha\cdot\phi^{1-\alpha}\Big), \tag{72a}$$

$$\Delta M_\alpha(\boldsymbol{\theta}[k]) := \log\mathcal{H}_{\alpha,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k]},\phi\big) - \overline{M}_\alpha(\boldsymbol{\theta}[k],\boldsymbol{\theta}[k]). \tag{72b}$$

Proof. By applying Lemma 3 in the appendix, a minorizing function of

$$\log\mathcal{H}_{\alpha,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi) = \log\left(\int_{\mathcal{W}} \varphi_{\boldsymbol{\theta}}^\alpha(\mathrm{x})\phi^{1-\alpha}(\mathrm{x})\mathrm{dx}\right). \tag{73}$$

can be constructed as

$$\begin{aligned} M_\alpha(\boldsymbol{\theta},\boldsymbol{\theta}[k]) = \int_{\mathcal{W}} \log\left(\frac{\varphi_{\boldsymbol{\theta}}^\alpha(\mathrm{x})}{\varphi_{\boldsymbol{\theta}[k]}^\alpha(\mathrm{x})}\right)\frac{\varphi_{\boldsymbol{\theta}[k]}^\alpha(\mathrm{x})\phi^{1-\alpha}(\mathrm{x})}{\mathcal{H}_{\alpha,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k]},\phi\big)}\mathrm{dx} \\ + \log\mathcal{H}_{\alpha,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k]},\phi\big), \end{aligned} \tag{74}$$

which can be further simplified and rearranged as in (71). Thus, the result follows. ■

Since $D_{\alpha,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ is inversely proportional with $\mathcal{H}_{\alpha,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ for any $\alpha\in(0,1)$, using the minorizing surrogate function in Lemma 2, one can show that the discrete-time "move-to-soft-centroidal-configuration" law iteratively maximizes $\mathcal{H}_{\alpha,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ and so minimizes $D_{\alpha,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$.

Theorem 8 *Away from any locally optimal coverage configuration, the α-divergence $D_{\alpha,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k]},\phi\big)$, for[8] $0<\alpha<1$, strictly decreases at each iteration, $k\in\mathbb{N}$, of the discrete-time "move-to-soft-centroidal-configuration" law in (41).*

Proof. By definition (69), for $\alpha\in(0,1)$, the α-divergence $D_{\alpha,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ is inversely proportional with the Hellinger integral $\mathcal{H}_{\alpha,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$. Since the logarithm of the Hellinger integral $\mathcal{H}_{\alpha,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ can be minorized in terms of the negated KL divergence $D_{kl,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\varphi_{\boldsymbol{\theta}[k]}^\alpha\cdot\phi^{1-\alpha})$ (Lemma 2), $\mathcal{H}_{\alpha,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ can be iteratively increased by iteratively decreasing the reverse KL divergence $D_{kl,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\varphi_{\boldsymbol{\theta}[k]}^\alpha\cdot\phi^{1-\alpha})$ as described in (41) in Theorem 6 where the event distribution at each iteration is specified by $\varphi_{\boldsymbol{\theta}[k]}^\alpha(\mathrm{x})\phi^{1-\alpha}(\mathrm{x})$. Thus, the result is simply inherited from Theorem 6 for $\alpha\in(0,1)$, because the workspace statistics, mean $\mu_{\mathcal{W}_i}[k]$ and variance $\sigma_{\mathcal{W}_i}^2[k]$, are the same for $D_{\alpha,\mathcal{W}}\big(\varphi_{\boldsymbol{\theta}[k]},\phi\big)$ and $D_{kl,\mathcal{W}}(\varphi_{\boldsymbol{\theta}[k]},\varphi_{\boldsymbol{\theta}[k]}^\alpha\cdot\phi^{1-\alpha})$ while the corresponding workspace masses, $m_{\mathcal{W}_i[k]}$, have the same sign (negative). ■

Likewise, one has for the discrete-time "move-to-Voronoi-centroidal-configuration" law that:

Theorem 9 *For any choice of $\alpha\in(0,1)$, the α-divergence $D_{\alpha,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}[k]},\phi)$, away from its critical points, strictly decreases at each iteration $k\in\mathbb{N}$ of the discrete-time "move-to-Voronoi-centroidal-configuration" law in (42).*

Proof. The α-divergence $D_{\alpha,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ can be written in terms of the Hellinger integrals as

$$D_{\alpha,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi) = \sum_{i=1}^{n} D_{\alpha,V_i}(q_{\mathbf{s}_i},\phi) = \sum_{i=1}^{n}\tfrac{1}{\alpha(\alpha-1)}(\mathcal{H}_{\alpha,V_i}(q_{\mathrm{s}_i},\phi)-1). \tag{75}$$

Hence, for $\alpha\in(0,1)$, the α-divergence $D_{\alpha,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ is inversely proportional with the Hellinger integrals $\mathcal{H}_{\alpha,V_i}(q_{\mathrm{s}_i},\phi)$.

[8]This restriction on α being in $(0,1)$ can be explained as follows: The α-divergence is "zero-avoiding" for $\alpha>1$ and yields an overestimate of the support of the event distribution (i.e., sensors might leave the region of interest and try to cover a larger domain); whereas it is "zero-forcing" for $\alpha<0$ and yields an underestimate of the support of the event distribution (i.e., sensors' sensing ranges asymptotically converge to zero so that their event sensing probabilities outside the environment approach zero.) [39], [42]. Such structural requirements are relaxed and mixed for $0<\alpha<1$.

Similar to Lemma 2, one can obtain a minorizing surrogate function $M_{\alpha,V_i}(\mathrm{s}_i,\mathrm{s}_i[k])$ of $\log\mathcal{H}_{\alpha,V_i}(q_{\mathrm{s}_i},\phi)$ at $\mathrm{s}_i[k]$ in terms of the reverse KL divergence as

$$\begin{aligned}M_{\alpha,V_i}(\mathrm{s}_i,\mathrm{s}_i[k]):=&-\frac{\alpha}{\mathcal{H}_{\alpha,V_i}\big(q_{\mathrm{s}_i[k]},\phi\big)}D_{kl,V_i}\Big(q_{\mathrm{s}_i},q^{\alpha}_{\mathrm{s}_i[k]}\cdot\phi^{1-\alpha}\Big)\\&+\frac{\alpha}{\mathcal{H}_{\alpha,V_i}\big(q_{\mathrm{s}_i[k]},\phi\big)}D_{kl,V_i}\Big(q_{\mathrm{s}_i[k]},q^{\alpha}_{\mathrm{s}_i[k]}\cdot\phi^{1-\alpha}\Big)\\&+\log\mathcal{H}_{\alpha,V_i}\big(q_{\mathrm{s}_i[k]},\phi\big).\end{aligned}\tag{76}$$

Thus, the Hellinger integrals $\mathcal{H}_{\alpha,V_i}\big(q_{\mathrm{s}_i[k]},\phi\big)$ can be iteratively increased away from a critical point of $D_{\alpha,\mathcal{W}}(\widehat{\varphi}_{\mathbf{s}},\phi)$ by iteratively decreasing the reverse KL divergence $D_{kl,V_i}(q_{\mathrm{s}_i},q^{\alpha}_{\mathrm{s}_i[k]}\cdot\phi^{1-\alpha})$ as described in Theorem 7 where the Voronoi statistics are computed assuming event distribution $q^{\alpha}_{\mathrm{s}_i[k]}(\mathrm{x})\phi^{1-\alpha}(\mathrm{x})$. Hence, the result follows for $\alpha\in(0,1)$ since the Voronoi statistics, mean and variance, are the same for $D_{\alpha,V_i}(q_{\mathrm{s}_i[k]},\phi)$ and $D_{kl,V_i}(q_{\mathrm{s}_i[k]},q^{\alpha}_{\mathrm{s}_i[k]}\cdot\phi^{1-\alpha})$ whereas the associated Voronoi masses are positively scaled. ■

C. Extensions of Move-to-Centroidal-Configuration Laws

In this part, to increase the practicability of our results, we briefly present three extensions of the "move-to-centroidal-configurations" laws for discrete event distributions, collision avoidance, and differential drive sensor dynamics.

1) Coverage of Discrete Event Distributions: A practical limitation of hard sensor allocation over soft allocation is that hard assignments usually yield coverage algorithms that suffer from discontinuities for discrete event distributions, which is often mitigated by assuming the availability of approximate continuous event distributions estimated based on a finite collection of events detected over an environment [22]–[25]. In this part, we assume that we are given a discrete event distribution Φ over the environment $\mathcal{W}$ that is represented as a finite collection $\Phi:=\{(\mathrm{x}_1,\rho_1),(\mathrm{x}_2,\rho_2),\ldots,(\mathrm{x}_K,\rho_K)\}$ of $K\in\mathbb{N}$ discrete event locations $\mathrm{x}_j\in\mathcal{W}$ with event probabilities $\rho_j\in\mathbb{R}_{\geq 0}$ (that satisfy $\sum_{j=1}^K\rho_j=1$). Accordingly, in order to asymptotically minimize the discrete f-divergence $D_f(\varphi_{\boldsymbol{\theta}},\Phi)$, associated with a differentiable convex function f over $(0,\infty)$ and a soft coverage configuration $\boldsymbol{\theta}\in(\mathcal{W}\times\mathbb{R}_{>0})^n\times\Omega(n)$,

$$D_f(\varphi_{\boldsymbol{\theta}},\Phi):=\sum_{j=1}^K f\left(\frac{\varphi_{\boldsymbol{\theta}}(\mathrm{x}_j)}{\rho_j}\right)\rho_j,\tag{77}$$

we propose the continuous-time "move-to-soft-sample-centroidal-configuration" law to be

$$u_{\mathrm{p}_i}=\kappa_{\mathrm{p}_i}\mathrm{sgn}(\overline{m}_{\mathcal{W}_i})\big(\mathrm{p}_i-\overline{\mu}_{\mathcal{W}_i}\big),\tag{78a}$$

$$u_{r_i^2}=\kappa_{r_i^2}\mathrm{sgn}(\overline{m}_{\mathcal{W}_i})\left(r_i^2-\frac{\overline{\sigma}^2_{\mathcal{W}_i}+\|\mathrm{p}_i-\overline{\mu}_{\mathcal{W}_i}\|^2}{d}\right),\tag{78b}$$

$$u_{\pi_i}=\kappa_{\pi}\mathrm{sgn}(\overline{m}_{\mathcal{W}_i})\left(\pi_i-\frac{\overline{m}_{\mathcal{W}_i}}{\sum_{j=1}^n\overline{m}_{\mathcal{W}_j}}\right),\tag{78c}$$

where $\kappa_{\mathrm{p}_i},\kappa_{r_i},\kappa_{\pi}\in\mathbb{R}_{>0}$ are fixed positive control gains, and the sample workspace statistics (mass, mean and variance) associated with $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\Phi)$ are defined to be

$$\overline{m}_{\mathcal{W}_i}:=\sum_{j=1}^K f'\left(\frac{\varphi_{\boldsymbol{\theta}}(\mathrm{x}_j)}{\rho_j}\right)\pi_i q_{\mathrm{s}_i}(\mathrm{x}_j),\tag{79a}$$

$$\overline{\mu}_{\mathcal{W}_i}:=\frac{1}{m_{\mathcal{W}_i}}\sum_{j=1}^K\mathrm{x}_j f'\left(\frac{\varphi_{\boldsymbol{\theta}}(\mathrm{x}_j)}{\rho_j}\right)\pi_i q_{\mathrm{s}_i}(\mathrm{x}_j),\tag{79b}$$

$$\overline{\sigma}^2_{\mathcal{W}_i}:=\frac{1}{m_{\mathcal{W}_i}}\sum_{j=1}^K\|\mathrm{x}_j-\mu_{\mathcal{W}_i}\|^2 f'\left(\frac{\varphi_{\boldsymbol{\theta}}(\mathrm{x}_j)}{\rho_j}\right)\pi_i q_{\mathrm{s}_i}(\mathrm{x}_j).\tag{79c}$$

The "move-to-soft-sample-centroidal-configuration" law inherits all the stability properties of the original construction in (33) summarized in Theorem 4.

Theorem 10 *For any strictly monotone convex function f over $(0,\infty)$, the continuously differentiable "move-to-soft-sample-centroidal-configuration" law in (78) asymptotically brings any initial soft coverage configuration $\boldsymbol{\theta}$ in its positively invariant domain $(\mathcal{W}\times\mathbb{R}_{>0})^n\times\Omega(n)$ to a locally optimal coverage configuration of the discrete f-divergence $D_f(\varphi_{\boldsymbol{\theta}},\Phi)$ in (77) that satisfies*

$$\mathrm{p}_i=\overline{\mu}_{\mathcal{W}_i},\ r_i^2=\tfrac{1}{d}\overline{\sigma}^2_{\mathcal{W}_i},\ \textit{and}\ \pi_i=\overline{m}_{\mathcal{W}_i}\Big/\sum\nolimits_{j=1}^n\overline{m}_{\mathcal{W}_j},\tag{80}$$

while decreasing $D_f(\varphi_{\boldsymbol{\theta}},\Phi)$ along the way.

Proof. The proof follows a similar pattern as the proof of Theorem 4 where the integrations over the environment is simply replaced with summations over discrete event locations, and so is omitted for the sake of brevity. ■

Similarly, the discrete-time "move-to-soft-centroidal-configuration" laws for the KL divergence in Section III-B2 and the α-divergence in Section III-B3 can be extended to handle discrete event distributions using sample workspace statistics while retaining the convergence properties. The "move-to-Voronoi-centroidal-configuration" laws can also be adapted to use with sample Voronoi statistics, but this yields discontinuities due to the hard event-to-sensor assignment.

2) Safe Coverage Control: In practice, sensors usually have finite body sizes and some limits on their adjustable sensing ranges. The "move-to-centroidal-configuration" laws of Section III-A and Section III-B, unfortunately, do not ensure collision-free coverage control for finite-size[9] mobile sensors and not respect any limits on their adjustable sensing ranges. In this part, assuming disk-shaped sensor bodies and some upper and lower bounds on adjustable sensing range, we propose another extension of our "move-to-centroidal-configuration" laws for safe coverage control of mobile sensor networks.

a) Encoding Collisions via Body Diagrams: Consider a group of disk-shaped mobile sensors in $\mathcal{W}$ located at $\mathbf{p}=(\mathrm{p}_1,\mathrm{p}_2,\ldots,\mathrm{p}_n)\in\mathcal{W}^n$ with a list of fixed body radii $\boldsymbol{\beta}:=(\beta_1,\beta_2,\ldots,\beta_n)\in(\mathbb{R}_{>0})^n$, and denote the configuration space of nonoverlapping disks of radii $\boldsymbol{\beta}$ in $\mathcal{W}$ by

$$\begin{aligned}\mathrm{Conf}(\mathcal{W},\boldsymbol{\beta}):=\Big\{\mathbf{p}\in\mathcal{W}^n\Big|&\|\mathrm{p}_i-\mathrm{p}_j\|\geq\beta_i+\beta_j,\ \forall i\neq j,\\&S(\mathrm{p}_i,\beta_i)\in\mathcal{W},\ \forall i\Big\},\end{aligned}\tag{81}$$

where $S(\mathrm{p},\beta):=\big\{\mathrm{x}\in\mathbb{R}^d\big|\|\mathrm{x}-\mathrm{p}\|<\beta\big\}$ is the d-dimensional open Euclidean sphere centered at $\mathrm{p}\in\mathbb{R}^d$ with radius $\beta>0$.

[9]The standard Voronoi-based coverage control of mobile sensors is only known to be collision-free for point-size homogeneous sensors [9], and its application to heteregenous sensor networks, especially with finite body sizes, usually requires an additional collision avoidance strategy [10], [33].

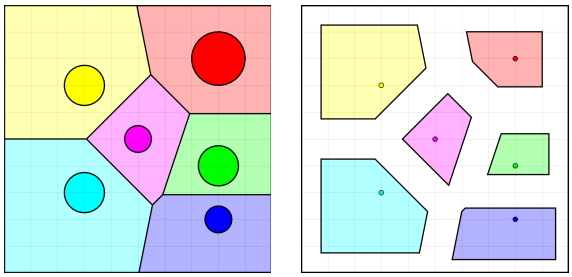


Fig. 3. Encoding collisions via body diagrams. (left) Body diagram and (right) free subdiagram of spherical sensor bodies that identify convex collision free neighborhoods of sensors for safe motion.

In [33], the *body diagram* $\mathcal{B}(\mathbf{p},\boldsymbol{\beta}) := \{B_1, B_2, \dots, B_n\}$ of the environment $\mathcal{W}$ associated with the disk-shaped sensor bodies is defined as the *power diagram* [52] of $\mathcal{W}$ generated by sensor body disks located at $\mathbf{p}$ and with body radii $\boldsymbol{\beta}$ as

$$B_i := \Big\{ \mathrm{x} \in \mathcal{W} \Big| \|\mathrm{x}-\mathrm{p}_i\|^2 - \beta_i^2 < \|\mathrm{x}-\mathrm{p}_j\|^2 - \beta_j^2,\ \forall i \neq j \Big\}, \quad (82)$$

which is an open convex polytope [52], illustrated in Fig. 3(left). Accordingly, in [33], the *free subdiagram* $\mathcal{F}(\mathbf{p},\boldsymbol{\beta}) := \{F_1, F_2, \dots, F_n\}$ of the body diagram $\mathcal{B}(\mathbf{p},\boldsymbol{\beta})$ is constructed by the erosion of each body cell with the associated sensor body shape[10], i.e.,

$$F_i := B_i \ominus S(\mathbf{0}, \beta_i) = \Big\{ \mathrm{x} \in B_i \Big| S(\mathrm{x}, \beta_i) \subseteq B_i \Big\}, \quad (83)$$

which is a closed convex polytope [33], see Fig. 3(right). Here, $\ominus$ denotes the Minkowski difference, and $\mathbf{0}$ is a vector of all zeros with appropriate length.

A critical use of body diagrams is for encoding collisions.

Proposition 6 ([33]) *A sensor placement $\mathbf{p} \in \mathcal{W}^n$ is a collision-free configuration in $\mathrm{Conf}(\mathcal{W},\boldsymbol{\beta})$ if and only if each sensor body (resp. position) is contained in the associated body cell (resp. free subcell), i.e.,*

$$\mathbf{p} \in \mathrm{Conf}(\mathcal{W},\boldsymbol{\beta}) \Leftrightarrow S(\mathrm{p}_i, \beta_i) \subseteq B_i,\ \forall i \Leftrightarrow \mathrm{p}_i \in F_i,\ \forall i. \quad (84)$$

b) Safe Coverage Control Laws : Accordingly, to ensure collision avoidance and respect any given limits on adjustable sensing ranges (i.e., $r_i \in [\underline{r}_i, \overline{r}_i] \subset \mathbb{R}_{>0}$), we propose a safe version of the "move-to-soft-centroidal configuration" law in (33), that is defined at a soft coverage configuration $\boldsymbol{\theta} = (\mathbf{p}, \boldsymbol{r}, \boldsymbol{\pi}) \in \mathrm{Conf}(\mathcal{W},\boldsymbol{\beta}) \times \prod_{i=1}^n [\underline{r}_i, \overline{r}_i] \times \Omega(n)$ as

$$u_{\mathrm{p}_i} = \kappa_{\mathrm{p}_i} \mathrm{sgn}(m_{\mathcal{W}_i})(\mathrm{p}_i - \Pi_{F_i}(\mu_{\mathcal{W}_i})), \quad (85a)$$

$$u_{r_i^2} = \kappa_{r_i^2} \mathrm{sgn}(m_{\mathcal{W}_i}) \left(r_i^2 - \Pi_{[\underline{r}_i^2, \overline{r}_i^2]} \left(\frac{\sigma_{\mathcal{W}_i}^2 + \|\mathrm{p}_i - \mu_{\mathcal{W}_i}\|^2}{d} \right) \right), \quad (85b)$$

$$u_{\pi_i} = \kappa_\pi \mathrm{sgn}(m_{\mathcal{W}_i}) \left(\pi_i - \frac{m_{\mathcal{W}_i}}{\sum_{j=1}^n m_{\mathcal{W}_j}} \right), \quad (85c)$$

where $\kappa_{\mathrm{p}_i} > 0$, $\kappa_{r_i^2} > 0$ and $\kappa_\pi > 0$ are fixed control gains; $m_{\mathcal{W}_i}$, $\mu_{\mathcal{W}_i}$ and $\sigma_{\mathcal{W}_i}^2$ are soft workspace statistics (mass, mean and variance, respectively) defined as in (16); $\mathcal{F}(\mathbf{p},\boldsymbol{\beta}) = \{F_1, F_2, \dots, F_n\}$ is the free subdiagram of $\mathcal{W}$ generated based on sensor positions $\mathbf{p}$ and sensor body radii $\boldsymbol{\beta}$ as in (83); and the adjustable sensing variance r_i^2 is assumed to be limited to some fixed interval $[\underline{r}_i^2, \overline{r}_i^2] \subset \mathbb{R}_{>0}$. Here, $\Pi_C(\mathrm{x}) := \arg\min_{\mathrm{c} \in C} \|\mathrm{c}-\mathrm{x}\|$ denotes the metric projection of a point $\mathrm{x} \in \mathbb{R}^N$ onto a convex set $C \subset \mathbb{R}^N$, which is piecewise continuously differentiable [57]–[59].

Some important properties of the "move-to-projected-soft-centroidal-configuration" law in (85) can be summarized as:

Theorem 11 *For any monotone decreasing convex function f over $(0,\infty)$, the piecewise continuously differentiable "move-to-projected-soft-centroidal-configuration" law in (85) leaves $\mathrm{Conf}(\mathcal{W},\boldsymbol{\beta}) \times \prod_{i=1}^n [\underline{r}_i, \overline{r}_i] \times \Omega(n)$ positively invariant, and its unique flow asymptotically decreases the f-divergence $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}}, \phi)$ in (6) and brings the sensors to a locally optimal coverage configuration of $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}}, \phi)$ over $\mathrm{Conf}(\mathcal{W},\boldsymbol{\beta}) \times \prod_{i=1}^n [\underline{r}_i, \overline{r}_i] \times \Omega(n)$ that satisfies for all i*

$$\mathrm{p}_i = \Pi_{F_i}(\mu_{\mathcal{W}_i}), \quad (86a)$$

$$r_i^2 = \Pi_{[\underline{r}_i^2, \overline{r}_i^2]} \left(\frac{\sigma_{\mathcal{W}_i}^2 + \|\mu_{\mathcal{W}_i} - \Pi_{F_i}(\mu_{\mathcal{W}_i})\|^2}{d} \right), \quad (86b)$$

$$\pi_i = m_{\mathcal{W}_i} \Big/ \sum\nolimits_{j=1}^n m_{\mathcal{W}_j}. \quad (86c)$$

Proof. By construction, the use of metric projections ensures the positive invariance of $\mathrm{Conf}(\mathcal{W},\boldsymbol{\beta}) \times \prod_{i=1}^n [\underline{r}_i, \overline{r}_i] \times \Omega(n)$ under the "move-to-projected-soft-centroidal-configuration" law because for any monotone decreasing f it generates a continuous linear motion towards a moving point in convex set $\prod_{i=1}^n F_i \times \prod_{i=1}^n [\underline{r}_i, \overline{r}_i] \times \Omega(n)$, and $\prod_{i=1}^n F_i$ is a collision-free convex subset of $\mathrm{Conf}(\mathcal{W},\boldsymbol{\beta})$ (see Proposition 6).

The monotonicity of f also guarantees that $m_{\mathcal{W}} \neq 0$ and $\mathrm{sgn}(m_{\mathcal{W}})$ is constant. Hence, since the soft workspace statistics are continuously differentiable, and the metric projection onto a convex set is piecewise continuous differentiable [57]–[59], and piecewise continuous differentiability is preserved under function composition [48], we have that the "move-to-projected-soft-centroidal-configuration" law is piecewise continuous differentiable and so is locally Lipschitz [48]. Since $\mathrm{Conf}(\mathcal{W},\boldsymbol{\beta}) \times \prod_{i=1}^n [\underline{r}_i, \overline{r}_i] \times \Omega(n)$ is compact, the "move-to-projected-soft-centroidal-configuration" law is globally Lipschitz over $\mathrm{Conf}(\mathcal{W},\boldsymbol{\beta}) \times \prod_{i=1}^n [\underline{r}_i, \overline{r}_i] \times \Omega(n)$, which implies the existence and uniqueness of its flow [47].

Finally, similar to the proof of Theorem 4, using Theorem 1 and (36), its asymptotic stability can be shown using $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}}, \phi)$ as a Lyapunov candidate function as follows:

$$\begin{aligned} \dot{D}_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}}, \phi) = & -\sum_{i=1}^n \kappa_{\mathrm{p}_i} \pi_i \frac{|m_{\mathcal{W}_i}|}{r_i^2} (\mathrm{p}_i - \mu_{\mathcal{W}_i})^{\mathrm{T}} (\mathrm{p}_i - \Pi_{F_i}(\mu_{\mathcal{W}_i})) \\ & - \sum_{i=1}^n \kappa_{r_i^2} \pi_i \frac{d |m_{\mathcal{W}_i}|}{2\sigma_{\mathcal{W}_i}^4} \left(r_i^2 - \frac{\sigma_{\mathcal{W}_i}^2 + \|\mathrm{p}_i - \mu_{\mathcal{W}_i}\|^2}{d} \right) \left(r_i^2 - \Pi_{[\underline{r}_i^2, \overline{r}_i^2]} \left(\frac{\sigma_{\mathcal{W}_i}^2 + \|\mathrm{p}_i - \mu_{\mathcal{W}_i}\|^2}{d} \right) \right), \\ & - \frac{\kappa_\pi}{\sum_{i=1}^n |m_{\mathcal{W}_i}|} \underbrace{\left(\sum_{i=1}^n \frac{|m_{\mathcal{W}_i}|^2}{\pi_i} - \left(\sum_{i=1}^n |m_{\mathcal{W}_i}| \right)^2 \right)}_{\geq 0, \text{ due to (36)}}, \end{aligned} \quad (87)$$

$$\begin{aligned} \dot{D}_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}}, \phi) \leq & -\sum_{i=1}^n \kappa_{\mathrm{p}_i} \pi_i \frac{|m_{\mathcal{W}_i}|}{2r_i^2} \|\mathrm{p}_i - \Pi_{F_i}(\mu_{\mathcal{W}_i})\|^2 \\ & - \sum_{i=1}^n \kappa_{r_i^2} \pi_i \frac{d |m_{\mathcal{W}_i}|}{4r_i^4} \left(r_i^2 - \Pi_{[\underline{r}_i^2, \overline{r}_i^2]} \left(\frac{\sigma_{\mathcal{W}_i}^2 + \|\mathrm{p}_i - \mu_{\mathcal{W}_i}\|^2}{d} \right) \right)^2, \end{aligned} \quad (88)$$

$$\dot{D}_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}}, \phi) \leq 0, \quad (89)$$

[10]Similar constructions are also utilized for sensor-based safe navigation of a single robot in cluttered environments [53]–[56], which also demonstrates how local sensory measurements, e.g., from LIDAR range scanners and depth sensors, can be used to construct body diagrams.

where the inequality follows from the fact that the metric projection $\Pi_C(x)$ of a point x onto a convex set C satisfies[11]

$$(c-x)^{\mathrm{T}}(c-\Pi_C(x)) \geq \frac{1}{2}\|c-\Pi_C(x)\|^2, \quad \forall c \in C. \quad (90)$$

Note that $\dot{D}_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ is strictly negative away from a locally optimal configuration satisfying (86). Thus, we have from LaSalle's Invariance Principle [47] that the sensors asymptotically reach a locally optimal configuration while decreasing $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ along the way. ■

It is important to note that, similar to (85), a safe version of the "move-to-Voronoi-centroidal-configuration" law in (37) can be constructed by using Voronoi statistics in (27) instead of workspace statistics in (16), while maintaining the stability and collision avoidance properties presented in Theorem 11. Moreover, the discrete-time "move-to-centroidal-configuration" laws presented in Section III-B can also be made collision-free by using projected statistics.

3) Coverage Control with Differential Drive Constraints: We now present a further extension of the safe "move-to-projected-soft-centroidal-configuration" law in (85) for the following kinematic differential drive sensor dynamics,

$$\dot{\mathrm{p}}_i = \begin{bmatrix}\cos\gamma_i \\ \sin\gamma_i\end{bmatrix} u_{\upsilon_i}, \quad \dot{\gamma}_i = u_{\omega_i}, \quad (91)$$

where $u_{\upsilon_i} \in \mathbb{R}$ and $u_{\omega_i} \in \mathbb{R}$ are the linear and angular velocity inputs, respectively, for controlling sensor position $\mathrm{p}_i \in \mathcal{W}$ and sensor orientation $\gamma_i \in [0,2\pi)$. Note that the differential-drive dynamics are underactuated due to the nonholonomic constraint $\begin{bmatrix}-\sin\gamma_i \\ \cos\gamma_i\end{bmatrix}^{\mathrm{T}} \mathrm{p}_i = 0$, and so we denote the straight-line motion range of each sensor by

$$H_i := \left\{ \mathrm{x} \in \mathbb{R}^2 \,\middle|\, \begin{bmatrix}-\sin\gamma_i \\ \cos\gamma_i\end{bmatrix}^{\mathrm{T}} (\mathrm{x}-\mathrm{p}_i) = 0 \right\}. \quad (92)$$

Following the lines of [33], using a standard differential drive controller [60], we define the "move-to-projected-soft-centroidal-configuration" law for differential drive sensors at an extended coverage configuration $\boldsymbol{\theta} = (\mathbf{p},\boldsymbol{\gamma},\boldsymbol{r},\boldsymbol{\pi})$ in $\mathrm{Conf}(\mathcal{W},\boldsymbol{\beta}) \times [0,2\pi)^n \times \prod_{i=1}^n [\underline{r}_i,\overline{r}_i] \times \Omega(n)$ to be

$$u_{\upsilon_i} = \kappa_{\upsilon_i}\mathrm{sgn}(m_{\mathcal{W}_i}) \begin{bmatrix}\cos\gamma_i \\ \sin\gamma_i\end{bmatrix}^{\mathrm{T}} (\mathrm{p}_i - \Pi_{F_i\cap H_i}(\mu_{\mathcal{W}_i})), \quad (93a)$$

$$u_{\omega_i} = \kappa_{\omega_i}\mathrm{atan}\left(\frac{\begin{bmatrix}-\sin\gamma_i \\ \cos\gamma_i\end{bmatrix}^{\mathrm{T}} \left(\mathrm{p}_i - \Pi_{F_{\varepsilon_i}}(\mu_{\mathcal{W}_i})\right)}{\begin{bmatrix}\cos\gamma_i \\ \sin\gamma_i\end{bmatrix}^{\mathrm{T}} \left(\mathrm{p}_i - \Pi_{F_{\varepsilon_i}}(\mu_{\mathcal{W}_i})\right)} \right), \quad (93b)$$

$$u_{r_i^2} = \kappa_{r_i^2}\mathrm{sgn}(m_{\mathcal{W}_i}) \left(r_i^2 - \Pi_{[\underline{r}_i^2,\overline{r}_i^2]} \left(\frac{\sigma^2_{\mathcal{W}_i} + \|\mathrm{p}_i - \mu_{\mathcal{W}_i}\|^2}{d} \right) \right), \quad (93c)$$

$$u_{\pi_i} = \kappa_{\pi}\mathrm{sgn}(m_{\mathcal{W}_i}) \left(\pi_i - \frac{m_{\mathcal{W}_i}}{\sum_{j=1}^n m_{\mathcal{W}_j}} \right), \quad (93d)$$

where $\kappa_{\upsilon_i}, \kappa_{\omega_i}, \kappa_{r_i^2}, \kappa_\pi \in \mathbb{R}_{>0}$ are fixed positive control gains; $m_{\mathcal{W}_i}$, $\mu_{\mathcal{W}_i}$ and $\sigma^2_{\mathcal{W}_i}$ are soft workspace statistics defined as in (16); H_i is the linear motion range in (92). Here, in addition to the free subdiagram $\mathcal{F}(\mathbf{p},\boldsymbol{\beta}) = \{F_1,F_2,\ldots,F_n\}$ in (83) of $\mathcal{W}$ associated with sensor positions $\mathbf{p}$ and body radii $\boldsymbol{\beta}$, we use its subdiagram $\mathcal{F}(\mathbf{p},\boldsymbol{\beta}+\boldsymbol{\varepsilon}) = \{F_{\varepsilon_1},F_{\varepsilon_2},\ldots,F_{\varepsilon_n}\}$ associated with safety margins $\boldsymbol{\varepsilon} := (\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n) \in (\mathbb{R}_{>0})^d$ to ensure that projected workspace mean $\Pi_{F_{\varepsilon_i}}(\mu_{\mathcal{W}_i})$ is strictly in the interior of $\mathcal{W}$. Note that while F_i is always nonempty and contains p_i for all $\mathbf{p} \in \mathrm{Conf}(\mathcal{W},\boldsymbol{\beta})$, the set F_{ε_i} can be empty. Hence, we set $\Pi_{F_{\varepsilon_i}}(\mu_{\mathcal{W}_i}) = \mathrm{p}_i$ whenever $F_{\varepsilon_i} = \varnothing$; and to resolve indeterminacy, we also set $u_{\omega_i} = 0$ when $\Pi_{F_{\varepsilon_i}}(\mu_{\mathcal{W}_i}) = \mathrm{p}_i$.

Theorem 12 *The "move-to-projected-soft-centroidal-configuration" law of differential drive sensors in (93) asymptotically brings any safe soft coverage configurations in its positively invariant domain* $\mathrm{Conf}(\mathcal{W},\boldsymbol{\beta}) \times [0,2\pi)^n \times \prod_{i=1}^n [\underline{r}_i,\overline{r}_i] \times \Omega(n)$ *towards a locally optimal coverage configuration that satisfies*

$$\mathrm{p}_i = \Pi_{F_i\cap H_i}(\mu_{\mathcal{W}_i}) = \Pi_{F_{\varepsilon_i}}(\mu_{\mathcal{W}_i}), \quad (94a)$$

$$r_i^2 = \Pi_{[\underline{r}_i^2,\overline{r}_i^2]} \left(\frac{\sigma^2_{\mathcal{W}_i} + \|\mu_{\mathcal{W}_i} - \Pi_{F_i}(\mu_{\mathcal{W}_i})\|^2}{d} \right), \quad (94b)$$

$$\pi_i = m_{\mathcal{W}_i} \Big/ \sum\nolimits_{j=1}^n m_{\mathcal{W}_j}. \quad (94c)$$

Proof. As in the proof of Theorem 11, the continuous motion towards the projected statistics in (93), by construction, ensures the positive invariance of $\mathrm{Conf}(\mathcal{W},\boldsymbol{\beta}) \times [0,2\pi)^n \times \prod_{i=1}^n [\underline{r}_i,\overline{r}_i] \times \Omega(n)$ for differential drive sensor dynamics. Also, the existence and uniqueness of its flow can be established using the flow properties of the 'move-to-projected-soft-centroidal-configuration" law in Theorem 11 and the flow properties of the differential drive controller in [60].

Using a similar argument as in the proof of Theorem 11, one can consider $D_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ as a Lyapunov function candidate and show that $\dot{D}_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi)$ is nonincreasing, i.e.,

$$\begin{aligned}\dot{D}_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi) \leq & -\sum_{i=1}^n \kappa_{\mathrm{p}_i}\pi_i \frac{|m_{\mathcal{W}_i}|}{2r_i^2}\|\mathrm{p}_i - \Pi_{F_i\cap H_i}(\mu_{\mathcal{W}_i})\|^2 \\ & - \sum_{i=1}^n \kappa_{r_i^2}\pi_i \frac{d|m_{\mathcal{W}_i}|}{4r_i^4} \left(r_i^2 - \Pi_{[\underline{r}_i^2,\overline{r}_i^2]} \left(\frac{\sigma^2_{\mathcal{W}_i} + \|\mathrm{p}_i - \mu_{\mathcal{W}_i}\|^2}{d} \right) \right)^2 \\ & - \frac{\kappa_\pi}{\sum_{i=1}^n |m_{\mathcal{W}_i}|} \underbrace{\left(\sum\nolimits_{i=1}^n \frac{|m_{\mathcal{W}_i}|^2}{\pi_i} - \left(\sum\nolimits_{i=1}^n |m_{\mathcal{W}_i}| \right)^2 \right)}_{\geq 0, \text{ due to (36)}}, \end{aligned} \quad (95)$$

$$\dot{D}_{f,\mathcal{W}}(\varphi_{\boldsymbol{\theta}},\phi) \leq 0 \quad (96)$$

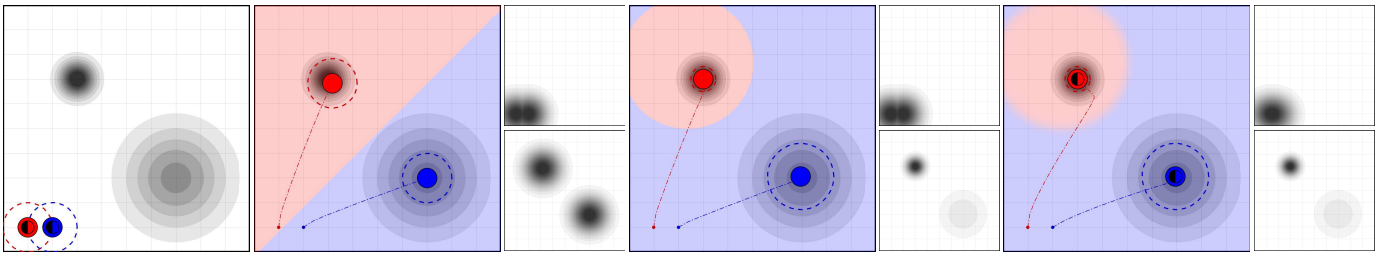
Moreover, since $F_{\varepsilon_i} \subset F_i$ and the standard different drive controller asymptotically aligns each sensor towards $\Pi_{F_{\varepsilon_i}}(\mu_{\mathcal{W}_i})$, i.e., $\begin{bmatrix}-\sin\gamma_i \\ \cos\gamma_i\end{bmatrix}^{\mathrm{T}} \left(\mathrm{p}_i - \Pi_{F_{\varepsilon_i}}(\mu_{\mathcal{W}_i})\right) = 0$ [60], the projected statistics $\Pi_{F_{\varepsilon_i}}(\mu_{\mathcal{W}_i})$ and $\Pi_{F_i\cap H_i}(\mu_{\mathcal{W}_i})$ asymptotically coincide. Hence, it follows from LaSalle's Invariance Principle [47] that the sensors asymptotically reach a locally optimal coverage configuration specified by (94). ■

[11]The distance between a point $x \in \mathbb{R}^N$ and a convex set $C \subset \mathbb{R}^N$ can be refactored via a point $c \in C$ as

$$\|x-\Pi_C(x)\|^2 = \|x-c\|^2 + 2(x-c)^{\mathrm{T}}(c-\Pi_C(x)) + \|c-\Pi_C(x)\|^2,$$

which can be further rearranged to observe

$$\begin{aligned} 2(c-x)^{\mathrm{T}}(c-\Pi_C(x)) &= \|c-\Pi_C(x)\|^2 + \underbrace{\|x-c\|^2 - \|x-\Pi_C(x)\|^2}_{\geq 0}, \\ &\geq \|c-\Pi_C(x)\|^2 \geq 0. \end{aligned}$$



(a) (b) (c) (d) (e) (f) (g)

Fig. 4. Statistical coverage control of mobile sensor networks with adjustable sensing range adaptively reconfigures sensor positions and sensing ranges to match both the first and the second moments (spatial mean and variance) of assigned event distributions for increased coverage performance. (a) Level sets of the event distribution in (97) of two well-separated activities with different spatial spreads and initial sensor coverage configuration. (b,d,f) Locally optimal Voronoi (b,d) and soft (f) centroidal coverage configurations, where sensors share identical fixed sensing range in (b) and have the same mixture weights in (f). (c,e,g) The maximum (c,e) and the average (g) of individual event detection probabilities of sensors at the initial (top) and the final (bottom) coverage configurations. Here, the reverse Kullback-Leibler (KL) divergence is used as a statistical coverage quality measure.[12] [13]

Using a similar construction as in (93), a safe version of the "move-to-Voronoi-centroidal-configuration" law in (37) can be designed for differential-drive sensors using the Voronoi statistics in (27). Moreover, the discrete-time "move-to-centroidal-configuration" laws in Section III-B can be extended for discrete-time differential drive dynamics by using the forward Euler method (with a step size no more than unity), while maintaining the convergence and safety properties.

IV. Numerical Simulations

To highlight the significance of adjustable sensing range for increased coverage performance at varying resolution, we first consider a 10×10 square environment with two well-separated activities with different spatial spreads, specified by the following event distribution function $\phi : [0, 10]^2 \to \mathbb{R}_{>0}$

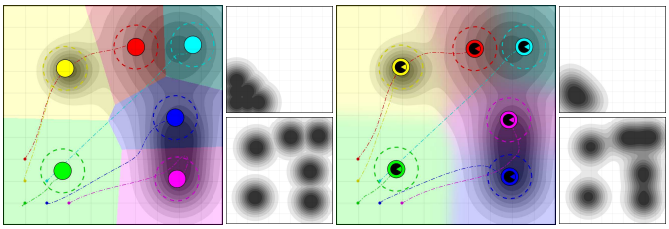
$$\phi(\mathrm{x}) \propto 2\exp\Big(-2\Big\|\mathrm{x}-\begin{bmatrix}3\\7\end{bmatrix}\Big\|^2\Big) + \exp\Big(-\frac{1}{4}\Big\|\mathrm{x}-\begin{bmatrix}7\\3\end{bmatrix}\Big\|^2\Big), \tag{97}$$

where events are assumed to be monitored by two mobile sensors that are initialized at around the lower left corner of the environment with equal sensing ranges and sensing weights, as illustrated in Fig. 4 (a). If the sensors have a fixed shared sensing range, as expected, they asymptotically move towards the central activity locations in order to match the first moment (i.e., mean) of assigned activity distributions for both soft and hard sensor allocation, which yields optimal coverage regions with almost the same volume irrespective of the spatial variances of assigned events, shown in Fig. 4 (b). On the other hand, if the sensors have adjustable sensing range, they still asymptotically move towards the central activity locations but also adjust their sensing ranges appropriately in order to match both the first and the second moments (i.e., mean and variance) of assigned activity distributions, which yields optimal coverage regions whose volumes are proportional with the spatial variance of assigned events, shown in Fig. 4 (d,f). Thus, instead of assuming a fixed proper sensing resolution selection, statistical coverage control of mobile sensors with adjustable sensing range offers increased coverage performance by adaptively adjusting sensors' sensing ranges in response to the spatial spread of assigned events in order to appropriately focus sensors based on event concentration.

To demonstrate coverage motion and locally optimal coverage configurations for the various settings discussed in Section III, we now consider a 10×10 square environment with the following more complicated event distribution function $\phi : [0, 10]^2 \to \mathbb{R}_{>0}$, which is also used in [11], [33],

$$\phi(\mathrm{x}) \propto 1+10\Big(\exp\Big(-\frac{1}{9}\Big\|\mathrm{x}-\begin{bmatrix}8\\8\end{bmatrix}\Big\|^2\Big)+\exp\Big(-\frac{1}{2}\Big\|\mathrm{x}-\begin{bmatrix}8\\2\end{bmatrix}\Big\|^2\Big) \exp\Big(-\frac{1}{2}\Big\|\mathrm{x}-\begin{bmatrix}8\\4\end{bmatrix}\Big\|^2\Big)+\exp\Big(-\Big\|\mathrm{x}-\begin{bmatrix}3\\7\end{bmatrix}\Big\|^2\Big)\Big). \tag{98}$$

and a group of six mobile sensors starting at around the lower left corner of the environment from a triangle formation with equal sensing ranges and equal mixture weights, as illustrated in Fig. 1(a). Since continuous- and discrete-time "move-to-centroidal-configuration" laws generate very similar coverage motion for small update step sizes[13], to save space, we here only present the simulation results for continuous-time coverage dynamics. Recall that the global and local approximation errors of the forward Euler method are, respectively, linear and quadratic with the update step size [50].

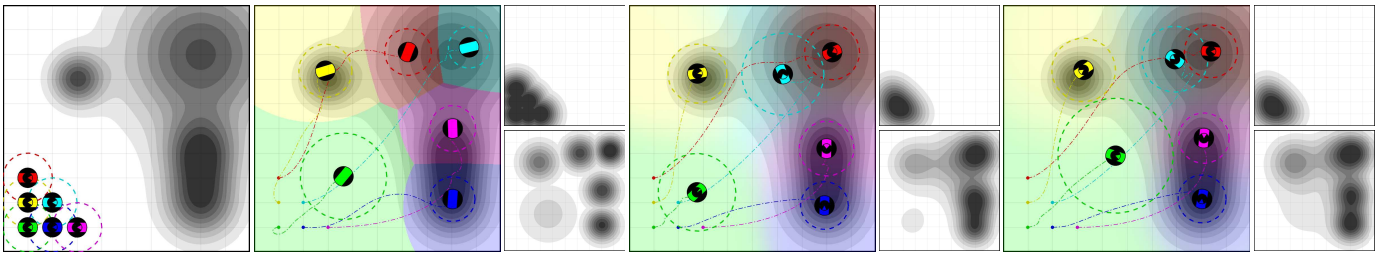


(a) (b) (c) (d)

Fig. 5. Statistical coverage control of homogeneous mobile isotropic sensors with fixed identical sensing ranges and mixture weights. Locally optimal (a) Voronoi-centroidal and (c) soft-centroidal coverage configurations. (b) The maximum and (d) the average of individual event detection probabilities of sensors at the initial (top) and the final (bottom) coverage configurations.[12] Here, the KL divergence is used as a measure of statistical coverage quality,[13], and the level sets of the event distribution in (98) and the initial sensor coverage configuration are illustrated in Fig. 1.

[12]Please refer to the accompanying video for an animated demonstration.

[13]For all simulations, inspired by k-means based initialization of the EM algorithm for Gaussian mixture learning [39], we set the control gains for continuous-time coverage control as $\kappa_{\mathrm{p}_i} = \kappa_{\upsilon_i} = \kappa_{\omega_i} = 1.0$ and $\kappa_{r_i^2} = \kappa_\pi = 0.1$ for all i, and the update step sizes for discrete-time coverage control as $\delta_{\mathrm{p}_i} = 0.1$ and $\delta_{r_i^2} = \delta_\pi = 0.01$ for all i; and the simulation results for continuous-time coverage control are obtained through numerical integration using the `ode45` function of MATLAB. Since the continuous- and discrete-time coverage control dynamics yield very similar coverage motion for this selection of control parameters, in order to save space, we only include the simulation results for continuous-time coverage control. Moreover, for safe coverage control, we assume the sensor body radius to be $\beta_i = 0.4$ and the safety margin to be $\varepsilon_i = 0.05$ for all i.



(a) (b) (c) (d) (e) (f) (g)

Fig. 6. Statistical coverage control of mobile differential-drive sensor networks. (a) Level sets of the event distribution in (98) and initial sensor coverage configuration. Locally optimal (b) Voronoi and (d,f) soft centroidal coverage configurations, where sensors share fixed equal mixture weights in (f). (c) The maximum and (e,g) the weighted average of individual event detection probabilities of sensors at the initial (top) and the final (bottom) coverage configuration.[12] Here, the statistical coverage quality is measured in terms of the KL divergence.[13]

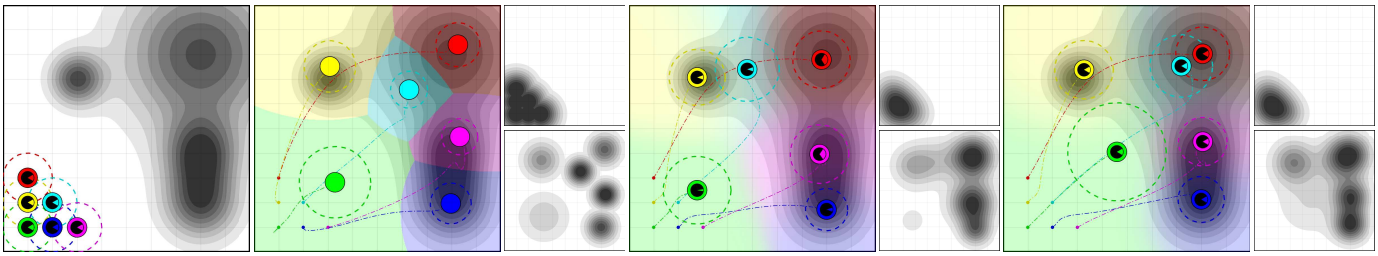
In Fig. 1, we select the KL divergence as our statistical coverage quality measure and illustrate the resulting coverage motion and the final locally optimal coverage configurations of the sensors for both the "move-to-soft-centroidal-configuration" law in (33) and the "move-to-Voronoi-centroidal-configuration" law in (37). Since soft sensor allocation provides a flexible choice of mixture (task) weights for determining how much each sensor participates in the coverage task, as expected, we observe that at a locally optimal coverage configuration, the weighted average of individual event detection probabilities of sensors approximates the event distribution more accurately than the maximum of individual event detection probabilities of sensors does.

To demonstrate the role of adjustable sensing range in coverage, we consider in Fig. 5 statistical coverage control of homogeneous sensors with fixed identical sensing ranges and mixture weights. As seen in Fig. 5, the sensors converge to very similar locally optimal coverage configurations (up to relabelling of sensors) for the "move-to-soft-centroidal-configuration" law and the "move-to-Voronoi-centroidal-configuration" law, because soft and hard assignments are very consistent with each other for sensors that participate equally in the coverage task (i.e., sensors with identical mixture weights). Not surprisingly, (homogeneous) sensors with fixed sensing ranges (in Fig. 5) usually perform less well in approximating the event distribution compared to (heterogeneous) sensors with variable sensing ranges (in Fig. 1).

In Fig. 6, we present the simulation results for safe statistical coverage control of differential drive mobile sensors. Although they start at the same initial positions with the fully actuated sensors in Fig. 1, we see that, due to the safety and nonholonomic motion constraints, differential drive sensors converge to slightly different locally optimal coverage configurations, especially for soft sensor allocation, via significantly different paths.

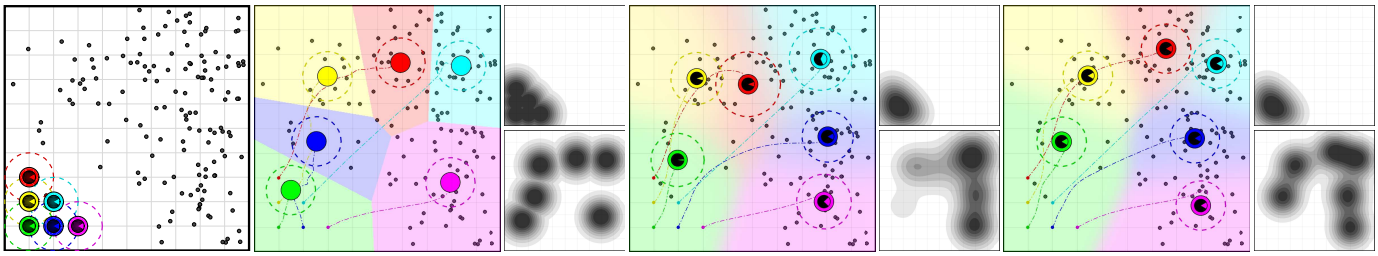
To illustrate the effect of selection of a statistical distance on the coverage performance, we consider in Fig. 7 statistical coverage control of mobile sensor network using the Hellinger distance. In our simulation experiments, we observe that locally optimal coverage configurations for the KL divergence and the Hellinger distance (even for the α-divergence for $\alpha \in (0,1)$) are structurally very similar, in general. We believe that this can be explained by the strong ("equivalence") relation of the α-divergence and the KL divergence observed in Lemma 2. While their locally optimal coverage configurations are significantly similar, the Hellinger distance yields smoother coverage motion than the KL divergence since the overall event detection probability of sensors acts as a smoothing filter on the event distribution; for example, for $\alpha \in (0,1)$, the workspace means and variances for $D_{\alpha,\mathcal{W}}(\varphi_{\boldsymbol{\theta}}, \phi)$ and $D_{kl,\mathcal{W}}(\varphi_{\boldsymbol{\theta}}, \varphi_{\boldsymbol{\theta}}^{\alpha} \cdot \phi^{1-\alpha})$ are the same.

Finally, we consider statistical coverage control of discrete event distributions and randomly sample a set of equally weighted discrete event locations from the continuous distribution function ϕ in (98) [39]. In Fig. 8 we present the resulting coverage motion and locally optimal coverage configurations for both hard and soft sensor allocation. Here, to resolve some potential technical issues due to the degeneracy of hard assignment (i.e., some sensors might be assigned to no or very few event locations), we assume that sensors share a fixed sensing range of unity for hard assignment. We observe that soft sensor allocation results in significantly better coverage configurations than hard sensor allocation does, because soft assignment leverages all event locations in estimating workspace statistics used in coverage control, whereas hard assignment can suffer from the lack of enough



(a) (b) (c) (d) (e) (f) (g)

Fig. 7. Statistical coverage control of mobile sensor networks using the Hellinger distance. (a) Level sets of the event distribution in (98) and initial sensor coverage configuration. Locally optimal Voronoi (b) and soft (d,f) coverage configurations. The maximum (c) and the weighted average (e,g) of individual event detection probabilities of the sensors at the initial (top) and the final (bottom) coverage configurations.[12] [13]



(a) (b) (c) (d) (e) (f) (g)

Fig. 8. Statistical coverage control of mobile sensor networks for a discrete event distribution. (a) Equally likely discrete event locations (randomly sampled from the event distribution function in (98)) and initial sensor coverage configuration. (b,d,f) Locally optimal Voronoi (b) and soft (d,f) centroidal coverage configurations, where sensors share identical fixed sensing range and mixture weights in (b) and (f). (c,e,g) The maximum (c) and the weighted average (e,g) of individual event detection probabilities of sensors at the initial (top) and the final (bottom) coverage configurations.[12] [13]

sample event assignment to sensors and so the (low) quality of estimated Voronoi statistics. Moreover, hard assignment causes discontinuous control inputs for discrete event distributions while soft assignment always generates smooth control inputs, which might be critical for certain application settings.

V. Conclusion

In this paper, we introduce a new statistical coverage quality measure for mobile sensor networks, with adjustable sensing range, that quantifies the total collective event coverage performance of sensors over an environment in terms of an f-divergence between the overall event detection probability of sensors and the event distribution over the environment. We consider soft and hard (Voronoi-based) sensor allocation in such mobile sensor networks and provide an explicit characterization of locally optimal coverage configurations for soft and hard sensor assignments, using soft workspace statistics and hard Voronoi statistics, respectively. Accordingly, for both continuous- and discrete-time sensor dynamics, we design a new family of simple, intuitive, reactive statistical coverage control policies, named the "move-to-centroidal-configuration" laws, that adaptively adjust sensor positions and sensing ranges in order to increase the coverage performance of sensors, and prove their asymptotic convergence to a locally optimal coverage configuration. We further present three practical extensions of our coverage control framework for discrete event distributions, collision avoidance, and differential drive sensor dynamics. In numerical simulations, we demonstrate the effectiveness of the proposed statistical coverage control strategies and the effect of various design parameters on the coverage performance.

Work now in progress targets statistical coverage control of limited-range anisotropic sensor networks [61]. We are also exploring other practical extensions of our framework for nonconvex environments [62] and camera networks [17]. Another interesting future extension is the use of mixture weights for balancing event and area coverage in sensor networks [63].

Appendix
Majorization/Minorization via Jensen's Inequality

In this part, we briefly present a generalization of some standard mathematical tricks for constructing a majorizing or minorizing surrogate function of a certain class of objective functions by exploiting convexity via Jensen's inequality [21].

Lemma 3 *Let $g : X \times \Theta \to \mathbb{R}_{>0}$ be a positive function over $X \times \Theta$, where $X \subseteq \mathbb{R}^d$, and define for any $\theta \in \Theta$ that*

$$G(\theta) := \int_X g(x,\theta)\mathrm{d}x, \quad \text{and} \quad \widehat{g}(x,\theta) := \frac{g(x,\theta)}{G(\theta)}. \tag{99}$$

Using a concave function $h : \mathbb{R}_{>0} \to \mathbb{R}$, one can construct a minorizing surrogate function of $h(G(\theta))$ at $\theta' \in \Theta$ as

$$H(\theta,\theta') := \int_X h\left(\frac{g(x,\theta)}{\widehat{g}(x,\theta')}\right)\widehat{g}(x,\theta')\mathrm{d}x \tag{100}$$

that satisfies $H(\theta,\theta') \le h(G(\theta))$ and $H(\theta',\theta') = h(G(\theta'))$.

Moreover, for any convex function $h : \mathbb{R}_{>0} \to \mathbb{R}$, $H(\theta,\theta')$ is a majorizing surrogate function of $h(G(\theta))$ at $\theta' \in \Theta$.

Proof. Since $x \mapsto \widehat{g}(x,\theta)$ defines a positive probability distribution over X, we directly have $H(\theta',\theta') = h(G(\theta'))$. Moreover, using Jensen's inequality, one can also obtain that

$$h(G(\theta)) = h\left(\int_X \frac{g(x,\theta)}{\widehat{g}(x,\theta')}\widehat{g}(x,\theta')\right)\mathrm{d}x, \tag{101}$$

$$\ge \int_X h\left(\frac{g(x,\theta)}{\widehat{g}(x,\theta')}\right)\widehat{g}(x,\theta')\mathrm{d}x = H(\theta,\theta'). \tag{102}$$

Similarly, the inequality in the reverse direction holds for a convex function h. Thus, the result follows. ■

Corollary 3 *For any monotone increasing concave function h over $(0,\infty)$, one has for any $\theta,\theta' \in \Theta$ that*

$$H(\theta,\theta') \ge H(\theta',\theta') \Longrightarrow G(\theta) \ge G(\theta') \tag{103}$$

where the inequalities change direction for a decreasing convex function h over $(0,\infty)$. Here, G and H are defined as in (99) and (100), respectively.

Proof. For a concave function h, we have from Lemma 3 that $H(\theta,\theta')$ is a minorizing function of $h(G(\theta))$ at θ', and so $H(\theta,\theta') \ge H(\theta',\theta')$ implies that

$$h(G(\theta)) \ge H(\theta,\theta') \ge H(\theta',\theta') = h(G(\theta')). \tag{104}$$

Hence, we obtain (103) because $G(\theta) \ge G(\theta') \iff h(G(\theta)) \ge h(G(\theta'))$ for any monotone increasing h. In a similar way, one can verify that the implication with the reversed inequalities holds for a monotone decreasing h. ■

Correspondingly, discrete counterparts of Lemma 3 and Corollary 3 can be stated as follows.

Lemma 4 *Let $g_i : X \times \Theta \to \mathbb{R}_{>0}$ be a set of positive functions, where $i \in \{1, 2 \ldots, m\}$, and define for any $x \in X$ and $\boldsymbol{\theta} := (\theta_1, \theta_2, \ldots, \theta_m) \in \Theta^m$ that*

$$G(x, \boldsymbol{\theta}) := \sum_{i=1}^{m} g_i(x, \theta_i), \quad \text{and} \quad \widehat{g}_i(x, \boldsymbol{\theta}) := \frac{g_i(x, \theta_i)}{G(x, \boldsymbol{\theta})}. \quad (105)$$

Then, for any concave function $h : \mathbb{R}_{>0} \to \mathbb{R}$,

$$H(x, \boldsymbol{\theta}, \boldsymbol{\theta'}) := \sum_{i=1}^{m} h\left(\frac{g_i(x, \theta_i)}{\widehat{g}_i(x, \boldsymbol{\theta'})}\right) \widehat{g}_i(x, \boldsymbol{\theta'}) \quad (106)$$

minorizes $\boldsymbol{\theta} \mapsto h(G(x, \boldsymbol{\theta}))$ at $\boldsymbol{\theta'} \in \Theta^m$, that is to say, $H(x, \boldsymbol{\theta}, \boldsymbol{\theta'}) \leq h(G(x, \boldsymbol{\theta}))$ and $H(x, \boldsymbol{\theta'}, \boldsymbol{\theta'}) = h(G(x, \boldsymbol{\theta'}))$; whereas, for any convex function $h : \mathbb{R}_{>0} \to \mathbb{R}$, $H(x, \boldsymbol{\theta}, \boldsymbol{\theta'})$ is a majorizing function of $\boldsymbol{\theta} \mapsto h(G(x, \boldsymbol{\theta}))$ at $\boldsymbol{\theta'} \in \Theta^m$.

Proof. By definition, since $\sum_{i=1}^{m} \widehat{g}_i(x, \boldsymbol{\theta}) = 1$, we have $H(\boldsymbol{\theta'}, \boldsymbol{\theta'}) = h(G(x, \boldsymbol{\theta'}))$. Further, using Jensen's inequality, we obtain the result for a concave function h as

$$h(G(x, \boldsymbol{\theta})) = h\left(\sum_{i=1}^{m} \frac{g_i(x, \theta_i)}{\widehat{g}_i(x, \boldsymbol{\theta'})} \widehat{g}_i(x, \boldsymbol{\theta'})\right), \quad (107)$$

$$\geq \sum_{i=1}^{m} h\left(\frac{g_i(x, \theta_i)}{\widehat{g}_i(x, \boldsymbol{\theta'})}\right) \widehat{g}_i(x, \boldsymbol{\theta'}) = H(x, \boldsymbol{\theta}, \boldsymbol{\theta'}), \quad (108)$$

where the inequality in the opposite direction holds for a convex function h. ■

Corollary 4 *For any increasing concave h over $(0, \infty)$,*

$$H(x, \boldsymbol{\theta}, \boldsymbol{\theta'}) \geq H(x, \boldsymbol{\theta'}, \boldsymbol{\theta'}) \Longrightarrow G(x, \boldsymbol{\theta}) \geq G(x, \boldsymbol{\theta'}), \quad (109)$$

whereas the implication with the reversed inequalities holds for any decreasing convex h over $(0, \infty)$, where G and H are, respectively, defined as in (105) and (106).

REFERENCES

[1] M. Schwager, B. J. Julian, M. Angermann, and D. Rus, "Eyes in the sky: Decentralized control for the deployment of robotic camera networks," *Proceedings of the IEEE*, vol. 99, no. 9, pp. 1541–1561, 2011.

[2] I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater acoustic sensor networks: Research challenges," *Ad Hoc Networks*, vol. 3, no. 3, pp. 257 – 279, 2005.

[3] N. Megiddo and K. J. Supowit, "On the complexity of some common geometric location problems," *SIAM Journal on Computing*, vol. 13, no. 1, pp. 182–196, 1984.

[4] P. S. Bradley and U. M. Fayyad, "Refining initial points for k-means clustering," in *Int Conf on Machine Learning*, 1998, pp. 91–99.

[5] A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," *Pattern Recognition*, vol. 36, no. 2, pp. 451–461, 2003.

[6] T. Lange, V. Roth, M. L. Braun, and J. M. Buhmann, "Stability-based validation of clustering solutions," *Neural Computation*, vol. 16, no. 6, pp. 1299–1323, 2004.

[7] U. Von Luxburg, "Clustering stability: An overview," *Foundations and Trends in Machine Learning*, vol. 2, no. 3, pp. 235–274, 2010.

[8] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial tessellations: concepts and applications of Voronoi diagrams*, 2nd ed., ser. Wiley Series in Probability and Statistics. John Wiley & Sons, 2000.

[9] J. Cortés, S. Martınez, T. Karatas, and F. Bullo, "Coverage control for mobile sensing networks," *IEEE Transactions on Robotics and Automation*, vol. 20, no. 2, pp. 243–255, 2004.

[10] L. Pimenta, V. Kumar, R. Mesquita, and G. Pereira, "Sensing and coverage for a network of heterogeneous robots," in *IEEE Conference on Decision and Control*, 2008, pp. 3947–3952.

[11] A. Kwok and S. Martınez, "Deployment algorithms for a power-constrained mobile sensor network," *International Journal of Robust and Nonlinear Control*, vol. 20, no. 7, pp. 745–763, 2010.

[12] W. Li and C. Cassandras, "Distributed cooperative coverage control of sensor networks," in *IEEE Conference on Decision and Control and European Control Conference*, 2005, pp. 2542–2547.

[13] B. Hexsel, N. Chakraborty, and K. Sycara, "Coverage control for mobile anisotropic sensor networks," in *IEEE International Conference on Robotics and Automation*, 2011, pp. 2878–2885.

[14] A. Gusrialdi, T. Hatanaka, and M. Fujita, "Coverage control for mobile networks with limited-range anisotropic sensors," in *IEEE Conference on Decision and Control*, 2008, pp. 4263–4268.

[15] A. Howard, M. Mataric, M., and G. S. Sukhatme, "Mobile sensor network deployment using potential fields: A distributed, scalable solution to the area coverage problem," in *Distributed Autonomous Robotic Systems 5*. Springer, 2002, pp. 299–308.

[16] M. Schwager, D. Rus, and J.-J. Slotine, "Unifying geometric, probabilistic, and potential field approaches to multi-robot deployment," *Int. J. Robotics Res.*, vol. 30, no. 3, pp. 371–383, 2011.

[17] O. Arslan, H. Min, and D. E. Koditschek, "Voronoi-based coverage control of pan/tilt/zoom camera networks," in *IEEE International Conference on Robotics and Automation*, 2018.

[18] C. Piciarelli, C. Micheloni, and G. L. Foresti, "PTZ camera network reconfiguration," in *ACM/IEEE International Conference on Distributed Smart Cameras*, 2009, pp. 1–7.

[19] ——, "Automatic reconfiguration of video sensor networks for optimal 3d coverage," in *ACM/IEEE International Conference on Distributed Smart Cameras*, 2011, pp. 1–6.

[20] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society. Series B*, vol. 39, no. 1, pp. 1–38, 1977.

[21] D. R. Hunter and K. Lange, "A tutorial on mm algorithms," *The American Statistician*, vol. 58, no. 1, pp. 30–37, 2004.

[22] M. Hasanbeig and L. Pavel, "Distributed coverage control by robot networks in unknown environments using a modified EM algorithm," *WASET International Journal of Computer and Information Engineering*, vol. 11, no. 7, pp. 805–813, 2017.

[23] M. Schwager, D. Rus, and J.-J. Slotine, "Decentralized, adaptive coverage control for networked robots," *Int. J. Robotics Res.*, vol. 28, no. 3, pp. 357–375, 2009.

[24] S. Rahili and W. Ren, "Game theory control solution for sensor coverage problem in unknown environment," in *IEEE Conference on Decision and Control*, 2014, pp. 1173–1178.

[25] Z. Guo, M. Zhou, and G. Jiang, "Adaptive sensor placement and boundary estimation for monitoring mass objects," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 38, no. 1, pp. 222–232, 2008.

[26] S. G. Lee, Y. Diaz-Mercado, and M. Egerstedt, "Multirobot control using time-varying density functions," *IEEE Transactions on Robotics*, vol. 31, no. 2, pp. 489–493, 2015.

[27] J. Wu and S. Yang, "Coverage issue in sensor networks with adjustable ranges," in *Workshops on Mobile and Wireless Networking/High Performance Scientific, Engineering Computing/Network Design and Architecture/Optical Networks Control and Management/Ad Hoc and Sensor Networks/Compil*, 2004, pp. 61–68.

[28] J. Wang and S. Medidi, "Energy efficient coverage with variable sensing radii in wireless sensor networks," in *IEEE Int Conf on Wireless and Mobile Computing, Networking and Communications*, 2007, pp. 61–61.

[29] M. Cardei, J. Wu, and M. Lu, "Improving network lifetime using sensors with adjustable sensing ranges," *International Journal of Sensor Networks*, vol. 1, no. 1-2, pp. 41–49, 2006.

[30] F. Aurenhammer, F. Hoffmann, and B. Aronov, "Minkowski-type theorems and least-squares clustering," *Algorithmica*, vol. 20, no. 1, pp. 61–76, 1998.

[31] J. Cortes, "Coverage optimization and spatial load balancing by robotic sensor networks," *IEEE Transactions on Automatic Control*, vol. 55, no. 3, pp. 749–754, 2010.

[32] M. Pavone, A. Arsie, E. Frazzoli, and F. Bullo, "Distributed algorithms for environment partitioning in mobile robotic networks," *IEEE Transactions on Automatic Control*, vol. 56, no. 8, pp. 1834–1848, 2011.

[33] O. Arslan and D. E. Koditschek, "Voronoi-based coverage control of heterogeneous disk-shaped robots," in *IEEE International Conference on Robotics and Automation*, 2016, pp. 4259–4266.

[34] Z. Li, L. P. Ligthart, P. Huang, W. Lu, and W. F. van der Zwan, "Trade-off between sensitivity and dynamic range in designing digital radar receivers," in *International Conference on Microwave and Millimeter Wave Technology*, vol. 3, 2008, pp. 1368–1371.

[35] M. M. Deza and E. Deza, *Encyclopedia of Distances*. Springer, 2014.

[36] H. W. Hamacher and Z. Drezner, *Facility Location: Applications and Theory*. Springer Science & Business Media, 2002.

[37] S. Lloyd, "Least squares quantization in PCM," *Information Theory, IEEE Transactions on*, vol. 28, no. 2, pp. 129–137, 1982.
[38] Q. Du, V. Faber, and M. Gunzburger, "Centroidal Voronoi tessellations: Applications and algorithms," *SIAM Review*, vol. 41, no. 4, pp. 637–676, 1999.
[39] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2007.
[40] S. M. Ali and S. D. Silvey, "A general class of coefficients of divergence of one distribution from another," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 28, no. 1, pp. 131–142, 1966.
[41] F. Liese and I. Vajda, "On divergences and informations in statistics and information theory," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4394–4412, 2006.
[42] T. Minka, "Divergence measures and message passing," Microsoft Research, Tech. Rep., 2005.
[43] H. Flanders, "Differentiation under the integral sign," *The American Mathematical Monthly*, vol. 80, no. 6, pp. 615–627, 1973.
[44] R. D. Nowak, "Distributed EM algorithms for density estimation and clustering in sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2245–2253, 2003.
[45] D. Gu, "Distributed EM algorithm for gaussian mixtures in sensor networks," *IEEE Transactions on Neural Networks*, vol. 19, no. 7, pp. 1154–1166, 2008.
[46] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.
[47] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice Hall, 2001.
[48] R. W. Chaney, "Piecewise C^k functions in nonsmooth analysis," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 15, no. 7, pp. 649–660, 1990.
[49] F. Bullo, J. Cortés, and S. Martinez, *Distributed Control of Robotic Networks: A Mathematical Approach to Motion Coordination Algorithms: A Mathematical Approach to Motion Coordination Algorithms*. Princeton University Press, 2009.
[50] J. C. Butcher, *Numerical Methods for Ordinary Differential Equations*. John Wiley & Sons, 2008.
[51] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
[52] F. Aurenhammer, "Power diagrams: Properties, algorithms and applications," *SIAM Journal on Computing*, vol. 16, no. 1, pp. 78–96, 1987.
[53] O. Arslan and D. E. Koditschek, "Exact robot navigation using power diagrams," in *IEEE Int Conf Robot Autom*, 2016, pp. 1–8.
[54] ——, "Sensor-based reactive navigation in unknown convex sphere worlds," in *Int Workshop on Algoritmic Foundations of Robotics*, 2016.
[55] O. Arslan, V. Pacelli, and D. E. Koditschek, "Sensory steering for sampling-based motion planning," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2017, pp. 3708–3715.
[56] O. Arslan and D. E. Koditschek, "Sensor-based reactive navigation in unknown convex sphere worlds," *The International Journal of Robotics Research*, vol. 38, no. 2-3, pp. 196–223, 2019.
[57] L. Kuntz and S. Scholtes, "Structural analysis of nonsmooth mappings, inverse functions, and metric projections," *Journal of Mathematical Analysis and Applications*, vol. 188, no. 2, pp. 346 – 386, 1994.
[58] A. Shapiro, "Sensitivity analysis of nonlinear programs and differentiability properties of metric projections," *SIAM Journal on Control and Optimization*, vol. 26, no. 3, pp. 628–645, 1988.
[59] J. Liu, "Sensitivity analysis in nonlinear programs and variational inequalities via continuous selections," *SIAM Journal on Control and Optimization*, vol. 33, no. 4, pp. 1040–1060, 1995.
[60] A. Astolfi, "Exponential stabilization of a wheeled mobile robot via discontinuous control," *Journal of Dynamic Systems, Measurement, and Control*, vol. 121, no. 1, pp. 121–126, 1999.
[61] A. Gusrialdi, S. Hirche, D. Asikin, T. Hatanaka, and M. Fujita, "Voronoi-based coverage control with anisotropic sensors and experimental case study," *Intelligent Service Robotics*, vol. 2, no. 4, pp. 195–204, 2009.
[62] S. Bhattacharya, R. Ghrist, and V. Kumar, "Multi-robot coverage and exploration on Riemannian manifolds with boundaries," *The International Journal of Robotics Research*, vol. 33, no. 1, pp. 113–137, 2014.
[63] H. Min, "On balancing event and area coverage in mobile sensor networks," Master's thesis, University of Pennsylvania, 2018.

Ömür Arslan (S'09–M'17) received the Ph.D. degree in electrical and systems engineering from the University of Pennsylvania, Philadelphia, PA, USA, in 2016 and the B.Sc. and M.Sc. degrees in electrical and electronics engineering from the Middle East Technical University, Ankara, Turkey, in 2007 and from Bilkent University, Ankara, in 2009, respectively. He is now a postdoctoral researcher in the Autonomous Motion Department at the Max Planck Institute for Intelligent Systems, Tübingen, Germany. His current research interests include robotics, motion planning, sensor networks, control theory, dynamical systems, optimization, machine learning, and machine perception.