Acoustic Sensor Network Design for Position Estimation

Volkan Cevher
Rice University
volkan@rice.edu
Lance M. Kaplan
U.S. Army Research Laboratory
lkaplan@ieee.org

In this paper, we develop tractable mathematical models and approximate solution algorithms for a class of integer optimization problems with probabilistic and deterministic constraints, with applications to the design of distributed sensor networks that have limited connectivity. For a given deployment region size, we calculate the Pareto frontier of the sensor network utility at the desired probabilities for d-connectivity and k-coverage. As a result of our analysis, we determine (i) the number of sensors of different types to deploy from a sensor pool, which offers a cost vs. performance trade-off for each type of sensor, (ii) the minimum required radio transmission ranges of the sensors to ensure connectivity, and (iii) the lifetime of the sensor network. For generality, we consider randomly deployed sensor networks and formulate constrained optimization techniques to obtain the localization performance. The approach is guided and validated using an unattended acoustic sensor network design. Finally, approximations of the complete statistical characterization of the acoustic sensor networks are given, which enable average network performance predictions of any combination of acoustic sensors.

Categories and Subject Descriptors: C.2.1 [**Computer-Communication Networks**]: Distributed networks; G.1.6 [**Numerical Analysis**]: Optimization—*Constrained optimization, convex programming, integer programming, nonlinear programming*; G.3 [**Probability and Statistics**]: Experimental design

General Terms: Algorithm, Design, Performance.
Additional Key Words and Phrases: Bayesian experimental design, dynamic programming, sensor networks.

1. INTRODUCTION

The design and operation of sensor networks are rich with integer and combinatorial optimization problems. Some of these problems concern (i) the dynamic assignment

Corresponding author's address: V. Cevher, Rice University, 6100 Main St., MS-380, Houston, TX 77005
Prepared through collaborative participation in the Advanced Sensors Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-02-0008.

of a subset of already deployed sensors for target tracking to minimize sensor network power consumption, [Liu et al. 2003; Subhlok et al. 1999; Sinha and Chandrakasan 2001], (ii) the automatic placement of a given set of sensors to guarantee coverage and connectivity, [Cowan and Kovesi 1988; Isler et al. 2004], and (iii) the optimal movement strategies for mobile sensors on a graph in the context of probabilistic pursuit evasion games, [Vidal et al. 2002; Isler et al. 2004]. Despite the advances that have been made, many difficult issues remain to be solved. The common theme of these sensor network problems is that the available resources are already spent on a set of sensors. Consequently, it is desired to obtain the best knowledge about a state-of-nature with minimum effort given the available sensors.

In this paper, we provide statistical models and mathematical programming solutions to determine the optimal sensor choices for deployment from a given sensor pool under a limited budget. Therefore, we focus the available resources on a set of complimentary sensors for the deployment purpose before the sensors are deployed. Optimality is obtained by simultaneously maximizing three criteria under a desired minimum connectivity level of the sensors. These criteria are (1) a desired utility of the sensor network deployment, (2) a desired minimum lifetime of the sensor network, and (3) a desired sensor coverage on the deployment area. For generality, we assume random sensor deployment, where each sensor's position comes from a homogenous Poisson point process over an area. We refer to the multi-criteria optimization as the network design strategy (NDS) problem.

In [Cevher and Kaplan 2008], we develop a theory to predict the localization performance of randomly distributed sensor networks consisting of various sensor modalities when only a constant active subset of sensors that minimize localization error is used for estimation. The characteristics of the modalities include measurement type (bearing or range) and error, sensor reliability, FOV, sensing range, and mobility. We show that the localization performance of a sensor network is a function of a weighted sum of the total number of each sensor modality. We also show that the optimization of this weighted sum is independent of how the sensor management strategy chooses the active sensors.

In this paper, we define a sensor network's localization utility in a Bayesian experimental design framework [Chaloner and Verdinelli 1995]. We then employ our results in determining the localization performance of heterogenous sensor networks in [Cevher and Kaplan 2008] to justify the form of the localization utility. We then discuss and summarize how the sensor network's lifetime depends on the connectivity of the sensors and what coverage implies for collective parameter estimation. Different from [Cevher and Kaplan 2008], we also provide continuous relaxation solutions that alleviate the computation and provide insight to the final form of the NDS. While the focus of [Cevher and Kaplan 2008] is the derivation of the localization accuracy of sensor networks, this paper focuses on the design aspects of sensor networks and how they can be efficiently obtained.

Our approach is guided and validated by the design of acoustic sensor networks connected with wireless links. Acoustic sensor networks use single microphone sensors and tethered acoustic arrays to detect, track, and classify acoustically loud targets and events such as ground vehicles, helicopters, and sniper fire. Target detection and parameter estimation performance of the acoustic sensors increase as

the number of microphones in the sensor is increased [Johnson and Dudgeon 1993; Cevher 2005]. Hence, it is possible to create a sensor pool using acoustic sensors with different number of microphones.

Our specific contributions are as follows:

(1) We provide approximations to the statistical characterization of acoustic sensors. These approximations enable average performance predictions for sensor networks that consist of any combination of acoustic sensors. In contrast, in [Cevher and Kaplan 2008], we use generic sensing models.
(2) We provide analytical solutions based on branch-and-bound relaxations that can upper and lower bound the integer programming solutions. These approximations enable quick calculations of optimal sensor network designs.
(3) We provide theorems that underlie the elements of the optimal sensor network design: concentration of resources and dominating sensor pairs. These conditions reduce the search space for the optimization problems and enable fast calculations. Specifically, we prove that given a total of N_{u} objectives, the optimal sensor network design chooses at most N_{u} different sensor types in the solution. We also prove that under conditions described in Section 3.3, there exists sensor pairs which automatically eliminate other sensors in the Pareto frontier.
(4) We extend the dynamical programming formulation in [Cevher and Kaplan 2008] to account for multiple utilities.

The organization of the paper is as follows. In Sect. 2, we describe the NDS problem in the context of Bayesian experimental design, d-connectivity, and k-coverage problems. In Sect. 3, we describe optimization solutions to the NDS problem. In Sect. 4, we derive the performance metrics for acoustic sensor networks. In Sect. 5, we demonstrate the NDS solution for acoustic sensor networks for position estimation where we determine the Pareto frontier for utility and lifetime of the sensor network.

2. THE SENSOR NETWORK DESIGN OBJECTIVES

The sensor network design simultaneously optimizes multiple objectives that summarize various performance criteria related to wireless sensor networks. These objectives are formulated as a function of the characteristics and quantities of each feasible sensor type and how the sensor network is managed. Typical characteristics of the sensor types include measurement modality (bearing or range) and error, sensor reliability, field-of-view, sensing range, and mobility. The design output is a vector $\boldsymbol{n} = [\, n_1 \;\; n_2 \;\; \dots \;\; n_T \,]'$, whose elements consist of the number of each sensor type from a catalogue of T-sensors, denoted as $\boldsymbol{s} = \{s_t | t = 1, \dots, T\}$, for deployment. Each sensor type has a different monetary cost and only a finite budget is available to purchase and operate these sensors. The sensor costs are expressed as a T-dimensional vector $\mathbf{c}$, whose t-th element c_t corresponds to the cost of s_t. Hence, a valid sensor design must satisfy the budget constraint $\mathbf{c}'\boldsymbol{n} \leq \$$.

The sensor network design relies on the ability to use the characteristics of the sensor types to predict the localization accuracy, lifetime, coverage, and reliability of the sensor network for any given combination of sensor types and deployments

satisfying the budget constraint. This mapping involves three entangled aspects of sensor networks: deployment plans, operational choices, and sensor choices. In order to make the problem tractable, we assume that the sensors are deployed randomly over a coverage area $\mathcal{A}$ of size A via a uniform distribution. By ignoring boundary effects, we assume that the node positions are realizations of a homogenous Poisson point process (PPP) [Ross 2007]. For the operation of the sensor network, a sensor management strategy based on selecting a total of q-sensors to minimize the localization error is assumed. Furthermore, the t-th sensor type has an effective sensing range of R_t^*, which is important in obtaining expressions for k-coverage. Finally, the transmission of data is the dominant user of battery resources, and the transmission range r_{tran} is determined based upon the density of the nodes to ensure a desired d-connectivity of the network. It should be noted that the objectives (or utilities) are defined to be non-negative real numbers where larger numbers represent higher performance.

2.1 Localization Accuracy

The purpose of the a sensor network is to exploit the disparate data collected from the various nodes in order to better infer the state of the objects (or targets) $\boldsymbol{\theta}$ comprising a scene, e.g., position, velocity, appearance, physical structure, pitch frequency, loudness, etc. Before the actual deployment of the sensor network, targets are assumed to be or to appear within the area $\mathcal{A}$, and their states have a prior distribution $p(\boldsymbol{\theta})$, e.g., targets have certain speed distributions, time-frequency and loudness characteristics. After deployment, the sensor network provides noisy observations of $\boldsymbol{\theta}$ via the sensory outputs, making it possible to judge the amount of information provided by the sensor network over the prior knowledge of $\boldsymbol{\theta}$. In this paper, our primary focus is the localization accuracy of the sensor network on $\mathcal{A}$.

After deployment, the location of the sensor nodes can be represented as the set $\boldsymbol{\zeta} = \{\boldsymbol{\zeta}_t | t = 1, \ldots, T\}$, where $\boldsymbol{\zeta}_t = [\, \boldsymbol{\zeta}_{t,1} \;\; \boldsymbol{\zeta}_{t,2} \;\; \cdots \;\; \boldsymbol{\zeta}_{t,n_t} \,]$ is the matrix of the locations of all n_t sensors of type-t and $\boldsymbol{\zeta}$ is the 2-D vector of actual sensor locations. Each sensor provides data $\boldsymbol{y} = \{y_{k,t} | k = 1, \ldots, n_t;\ t = 1, \ldots, T\}$, e.g., acoustic microphone recordings within some dynamic range. The characteristics of the sensors determine a probability density function (PDF) of the measurements as a function of the sensor/target geometry, i.e., $p(\boldsymbol{y}|\boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{n})$. Then, the measurements can be used to provide an estimate of the target state $\widehat{\boldsymbol{\theta}}$. Note that $\widehat{\boldsymbol{\theta}}$ can be viewed a function of the measurements $\boldsymbol{y}$, where the shape of the function is dictated by the measurement PDF $p(\boldsymbol{y}|\boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{n})$. As the data is not observed before $\boldsymbol{n}$ is chosen and an arbitrarily large number of real-world realizations is possible, it is meaningful to choose $\boldsymbol{n}$ to minimize the *average* estimation error of the target state over all random realizations

$$\varepsilon(\boldsymbol{n}) = \int\int\int p(\widehat{\boldsymbol{\theta}}(\boldsymbol{y})|\boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{n}) p(\boldsymbol{\theta}) p(\boldsymbol{\zeta}) \epsilon(\boldsymbol{y}, \boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{n}) d\boldsymbol{y} d\boldsymbol{\theta} d\boldsymbol{\zeta}, \tag{1}$$

where $\epsilon(\boldsymbol{y}, \boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{n})$ represents the accuracy of the state inference:

$$\epsilon(\boldsymbol{y}, \boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{n}) = \left(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}(\boldsymbol{y})\right)^T \boldsymbol{\Lambda} \left(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}(\boldsymbol{y})\right), \tag{2}$$

and $\boldsymbol{\Lambda}$ is a positive definite weighting matrix. In this work, we consider the mean squared error (mse) so that $\boldsymbol{\Lambda} = \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix.

When the relationship between $\boldsymbol{y}$, $\boldsymbol{\theta}$, $\boldsymbol{n}$, and $\boldsymbol{\zeta}$ is nonlinear, approximations based on Taylor series expansions are typically used to avoid the complicated integral in (1) [Berger 1993; Chaloner and Verdinelli 1995; Tierney and Kadane 1986]. These approximations use the expected Fisher information matrix or the matrix of the second derivatives of the log likelihood function and are extremely accurate even for small data sample sizes. Hence, assuming (i) a network design that guarantees $\upsilon \geq 3$ sensors contributing to the estimation at any spatial location (k-coverage), (ii) a network with high connectivity (d-connectivity), and (iii) that each sensor has sufficiently large number of data samples, we can use the following normal approximation for the data likelihood in (1) ($\mathcal{N}(\text{mean}, \text{variance})$ is the Gaussian density):

$$p(\widehat{\boldsymbol{\theta}}(\boldsymbol{y})|\boldsymbol{\theta}, \boldsymbol{\zeta}, \boldsymbol{n}) = \mathcal{N}\left(\overline{\boldsymbol{\theta}}, \left[\boldsymbol{H} + \mathfrak{F}(\overline{\boldsymbol{\theta}}, \boldsymbol{\zeta}, \boldsymbol{n})\right]^{-1}\right), \tag{3}$$

where $\overline{\boldsymbol{\theta}}$ is the mode of the posterior distribution, $\boldsymbol{H}$ is the Hessian of the logarithm of the prior density $p(\boldsymbol{\theta})$ (precision matrix of the prior), and $\mathfrak{F}(\boldsymbol{\theta}, \boldsymbol{\zeta}, \boldsymbol{n})$ is the expected Fisher information matrix (FIM) [Chaloner and Verdinelli 1995]. When the prior for the target state is uninformative, e.g., uniform, $\boldsymbol{H} = 0$. While some domain knowledge such as a road may lead to the use of an informative prior, this paper assumes that no such knowledge is available: $\boldsymbol{H} = 0$ and $p(\boldsymbol{\theta}) = 1/A$. With the normal approximation (3), the expected utility (1) can be expressed as follows [Chaloner and Verdinelli 1995]:

$$\varepsilon(\boldsymbol{n}) = \frac{1}{A} \iint \operatorname{tr}\left\{\mathfrak{F}^{-1}(\boldsymbol{\theta}, \boldsymbol{\zeta}, \boldsymbol{n})\right\} p(\boldsymbol{\zeta}) d\boldsymbol{\theta} d\boldsymbol{\zeta}. \tag{4}$$

We define the localization utility as the reciprocal of the expected MSE as defined in (3). Typically, the MSE expression is too complex to compute in closed form. Because the targets and sensor nodes are distributed uniformly over $\mathcal{A}$, we model the sensor/target geometry using 2-D PPP. As a result, the actual MSE scales with the density of the network normalized to the quality of the sensor measurements. As shown in [Cevher and Kaplan 2008; Kaplan and Cevher 2007], when the sensors are stationary, the localization utility (within a linear scale factor) is given by:

$$U(\boldsymbol{n}) = \left(\sum_{t=1}^{T} f_t n_t\right)^{\gamma}, \tag{5}$$

where f_t depends on the inherent reference sensing accuracy σ_t, field-of-view $\alpha_t \in (0, 1]$, and reliability probability $\beta_t \in (0, 1]$ of the sensor type-t:

$$f_t = \frac{\alpha_t \beta_t}{\sigma_t^{2/\gamma}}. \tag{6}$$

The derivations of (5) and (6) assume that only q-active sensors are used for estimation at each snapshot. The scale factor of (5) depends on the number of active sensors q and what objective the sensor management optimizes in active sensor selection. In [Cevher and Kaplan 2008], utility expressions including mobile sensors are also provided. The parameter $\gamma > 0$ depends on the signal propagation loss factor of the medium. For the case of no sensor management where all nodes within the sensing range to the target participate in the localization given that the obtain at detections, the MSE is approximately inversely related to (5). Monte

Carlo experiments in Section 5 demonstrate that the reciprocal of (5) is indeed a good approximation to the MSE. Therefore, we use (5) as the localization accuracy utility. The specific weights for acoustic sensors are obtained computationally in Section 5.

2.2 Effects of Connectivity and Lifetime

Connectivity (C) is an important property of multi-hop wireless networks, where each sensor cooperates in routing each other's packets. In distributed sensor network operation, it is desirable that the sensors do not communicate raw data and the communication among the sensors has a constant bandwidth to maximize the sensor lifetime and to minimize interference among the receivers sharing the same wireless channel when the sensors are communicating or when the network is queried for information [Akyildiz et al. 2002]. In addition, the network estimates are based on local sensor observations, which can determine the parameter of interest $\boldsymbol{\theta}$ after fusion. For example, when bearing and range sensors are used, determining the target position relies on collective observability, which requires connectivity.

In graph theory terms, the network is d-connected (with connectivity degree $C = d$) if, for any given pair of sensor nodes, there exists at least d mutually independent paths connecting them [Bollobás 1998]. Each sensor usually has a fixed transmission range r_{tran}, which is significantly smaller than the dimensions of $\mathcal{A}$. A communications link can be created between two sensors only if their physical distance is less than their transmission range (Fig. 1(a)). Hence, given transmission range r_{tran}, the probability of creating links between sensors increase as the number of sensors in the field increases.

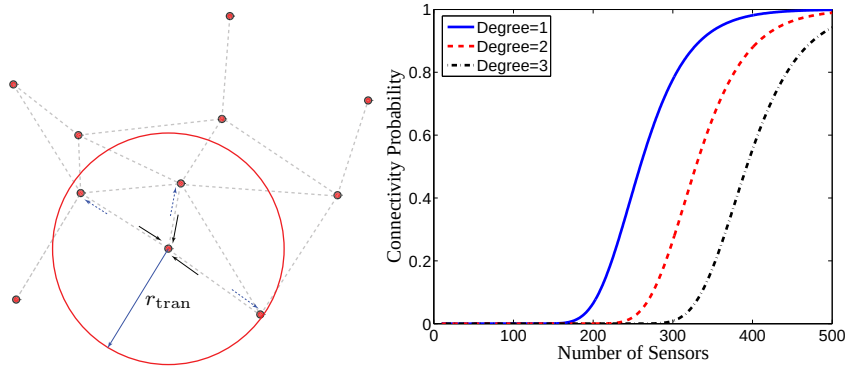


Fig. 1. (a) A sensor node (vertices) can connect to its neighbors by creating links (edges) only if other nodes are within its transmission range r_{tran}. Then, to connect to any other node in the network, the sensor can multi-hop its information using its neighboring nodes. (b) Approximate connectivity probability vs. the number of sensors for a sensor network (7) with the following parameters: $\mathcal{A} = 1000\text{m} \times 1000\text{m}$ and $r_{\text{tran}} = 100\text{m}$.

Under the random deployment assumption, it is possible to show that the pdf of the nearest neighbor sensor distance ρ is given by $p(\rho) = 2\pi\lambda\rho \exp\left(-\pi\lambda\rho^2\right)$,

where $\lambda = \frac{N}{A}$ is the spatial node density, and $N = \sum_{t=1}^{T} n_t$ (total number of sensors) [Bettstetter 2002]. Hence, the probability that each sensor has at least d-neighbors within the transmission range r_{tran} (node degree $D = d$) on a toroidal surface is given by the following [Bettstetter 2002]:

$$P(D \geq d|\boldsymbol{n}, \boldsymbol{\zeta}) = \exp\left\{ N \log\left(1 - \sum_{k=0}^{d-1} \frac{\left(\pi\lambda r_{\text{tran}}^2\right)^k}{k!} \exp\left(-\pi\lambda r_{\text{tran}}^2\right) \right) \right\}. \tag{7}$$

Figure 1(b) gives example plots of the node degree. It is shown in [Penrose 1999] that in a graph, as the number of vertices increase, the node degree D converges to the connectivity degree C with probability one. In this paper, we approximate the connectivity of the sensor network by (7). In general, the connectivity probability should be simulated for the specific geometry of $\mathcal{A}$ (e.g., see [Bettstetter 2002] for simulations on square $\mathcal{A}$). It is important to note that the connectivity probability monotonically increases as a function of N. Hence, to maximize the connectivity probability at any given degree, the total number of sensors N in the sensor network must be maximized.

Impacts of Connectivity on Power: Note that the connectivity can always be increased by increasing the sensor transmission range r_{tran}. However, the choice of r_{tran} also affects the lifetime (L) of the sensor network.[1] A sensor consumes power for ($\boldsymbol{\mathcal{A}}$) transmitting data, ($\boldsymbol{\mathcal{B}}$) receiving data, ($\boldsymbol{\mathcal{C}}$) sensing, ($\boldsymbol{\mathcal{D}}$) aggregating/fusing information, ($\boldsymbol{\mathcal{E}}$) idling, and ($\boldsymbol{\mathcal{F}}$) coping with radio interference/communications overhead/etc. [Heinzelman et al. 2002; Bhardwaj et al. 2001; Zhang and Hou 2004] [Bhardwaj and Chandrakasan 2002; Feeney and Nilsson 2001; Lindsey et al. 2002]. The sensor transmission range r_{tran} directly affects $\boldsymbol{\mathcal{A}}$ and also indirectly affects $\boldsymbol{\mathcal{F}}$.

In [Bhardwaj et al. 2001], $\boldsymbol{\mathcal{A}}$-$\boldsymbol{\mathcal{C}}$ are assumed to derive geometry-independent tight-bounds for the lifetime of a sensor network, where $L \propto N$. The reason for this proportionality is that the total energy of the sensor network scales with the total number of sensors. If the communication protocols are smart enough to minimize losses, then the actual lifetime should follow the same trend. The bound is not tight when the number of sensors is small as the characteristic distance, needed to make the cost of transmitting a bit linear with distance, cannot be realized. In [Bhardwaj and Chandrakasan 2002], $\boldsymbol{\mathcal{D}}$ is also assumed, and an improved bound is given which is also geometry dependent. In [Zhang and Hou 2004], $\boldsymbol{\mathcal{A}}$-$\boldsymbol{\mathcal{E}}$ is assumed for randomly deployed sensor networks and it is shown that the lifetime is also almost linear with the total number of sensors when algorithms that can adaptively turn on/off sensors for energy conservation are used. In [Giridhar and Kumar 2005], it is shown that message passing among the nearest neighbors yields nearly optimal lifetime.

The explicit dependence of the sensor network lifetime on r_{tran} and N is beyond the scope of this paper. This is because the analysis requires an underlying network model, which can quantify the MAC collision and communication delay issues. Without focusing on the specifics of the network models, we argue that the typical MAC collision and communication delay issues have negligible effect on the sensor network lifetime in our case because we only use $q \ll N$ active sensors per target

[1] The lifetime can be defined in various ways such as the time to first node failure due to battery depletion or the time to appearance of the first connectivity brake-down.

for localization at any given time. Moreover, we assume that the sensor network queried at random sensor locations, that is, there is no fixed base station; hence, the traffic load and the energy consumption per sensor are uniformly distributed across the sensor network. Then, based on the results from the literature, we can expect have a monotonic behavior for L as a function of N.

In the acoustic sensor network design, when $p_C = 1 - \epsilon$ ($\epsilon \ll 1$), we assume that the maximum network lifetime has approximately the following form:

$$L \propto N^{\delta}, \tag{8}$$

where $\delta > 0$ for generality and the proportionality is independent of r_{tran}. This form is motivated by the idealized case where the sensor batteries are depleted only by the r^p-propagation loss ($p \geq 4$): if p_C is high, then $L \propto C^{-1}(N) \times r^{-p}$, where $C(N)$ is the average hop count. For a square $\mathcal{A}$, the average hop count increases approximately proportional to $N^{-1/2}$ with N, whereas $r^{-p} \propto N^{p/2}$ (see footnote 2). Hence, $L \propto N^{(p-1)/2}$. Other forms for L can be assumed, which result in tractable solutions, e.g., $L \propto \exp(\delta N)$. In general, we emphasize again that the lifetime has a monotonic (quasiconvex) form for any given p_C for efficient solutions.

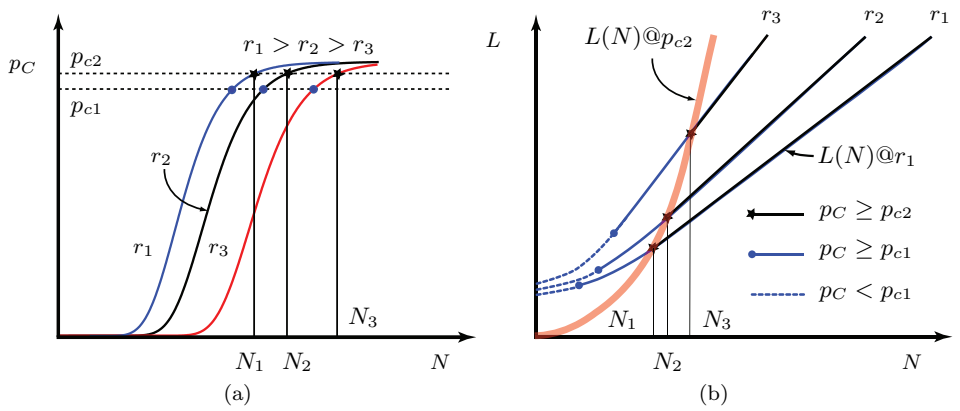


Fig. 2. (a) Connectivity probability p_C as a function of N and r_{tran}. (b) Sensor network lifetime L as a function of N and r_{tran}.

Figures 2(a) and (b) roughly illustrate the interplay between the sensor network lifetime and connectivity. As the total number of sensors is increased, the desired connectivity probability p_C can be achieved with smaller sensor transmission ranges r_{tran}. For a given transmission range that allows high connectivity, the lifetime of the sensor network linearly increases with the total number of sensors in the network as the total energy of the network is linearly increased (assuming efficient communication protocols that can minimize losses). Since the connectivity can be preserved with smaller transmission ranges as the number of sensors is increased,[2]

[2]As an example, if $p_C = 1 - \epsilon$ ($\epsilon \ll 1$), then from (7) with $d = 1$, we have $r_{\text{tran}} = \sqrt{-\frac{\log\left(1-p_C^{1/N}\right)|\mathcal{A}|}{\pi N}} \approx \sqrt{-|\mathcal{A}|\frac{\log\epsilon - \log N}{\pi N}} \propto N^{-1/2}$ for $|\log\epsilon| \gg |\log N|$, which is satisfied for typical N.

the actual lifetime of the sensor network grows faster than the growth obtained by increasing N while keeping r_{tran} fixed.

2.3 Effects of Coverage

The coverage (V) problem in sensor networks aims to quantify how well the area of interest $\mathcal{A}$ is monitored. The coverage problem has been extensively studied in the literature, see Sect. 2 in [Lazos and Poovendran 2006] for a survey. Given a planar $\mathcal{A}$ and random sensor network deployment, one can determine the number of sensors so that any given point in $\mathcal{A}$ is sensed by at least k-sensors with a given probability (k-coverage). For the sensor network design, the k-coverage probability is quite complicated as each sensor can have a heterogeneous sensing range.

In [Altman and Miorandi 2005], a Poisson approximation is given for the k-coverage probability for randomly deployed sensor networks with heterogeneous sensing ranges:

$$P(V \geq k|\boldsymbol{n}, \boldsymbol{\zeta}) = \sum_{l \geq k} \frac{\upsilon^l e^{-\upsilon}}{l!}, \text{ where } \upsilon = \sum_{t=1}^{T} \kappa_t n_t, \ \kappa_t = \frac{\pi R_t^{*2}}{A}. \tag{9}$$

Note that the coverage probability is a monotonically increasing function of υ. To increase k-coverage probability to a desired level p_V, υ in (9) must be maximized.

3. OPTIMIZATION PROBLEMS

The goal of NDS is to maximize three competing objectives: 1) localization accuracy, 2) lifetime, and 3) coverage. The exact form of these three objectives were discussed in the previous section. NDS is basically a multi criteria optimization problem over the simplex $\boldsymbol{\mathcal{N}}$ defining the feasible set of network designs, i.e., the vectors $\boldsymbol{n}$ whose elements are non-negative integers and also satisfy the budget constraint $\mathbf{c}'\boldsymbol{n} \leq \$$. Because of the integer nature of the network design vectors $\boldsymbol{n}$, the number of elements in the simplex $\boldsymbol{\mathcal{N}}$ is finite, albeit a very large number. In general, the objectives (or utilities) are expressed as an N_{u} dimensional vector $\mathbf{u}$ whose elements represent the N_{u} different utility values. For NDS, $N_{\text{u}} = 3$. Furthermore, u_1, u_2, and u_3 represent the localization accuracy, lifetime, and coverage utilities, respectively.

The following subsections describe general multiple-objective optimization, exact integer programming methods to find the "optimal" designs, and approximate non-integer optimization methods to quickly determine a "nearly optimal" design.

3.1 Multi-Objective Optimization

Multi-objective problems are typically addressed using Pareto optimality. In other words, an *optimal* network design is one such that one cannot find another sensor design in $\boldsymbol{\mathcal{N}}$ where one or more of the objectives increase but none decrease. The locus of all possible optimal network designs forms the Pareto frontier, i.e, the trade-off surface for the utilities. In other words, the Pareto frontier is the collection of feasible utilities that cannot be uniformly dominated by any other feasible utility. In general, enumeration of all possible parameter choices that satisfy the constraints is required to explore the Pareto frontier. However, in the end, some preference must be made to choose an operating point.

Many points on the Pareto frontier can be determined via scalarization (or weighting), where one simply maximizes a weighted sum of the utilities $\boldsymbol{\lambda}^T \mathbf{u}$ for any $\boldsymbol{\lambda}$ whose elements are non-negative [Boyd and Vandenberghe 2004]. The scalarization concepts also hold for the maximization of a product of utilities

$$\prod_{i=1}^{N_{\mathrm{u}}} u_i^{\lambda_i}, \tag{10}$$

because the log operator is monotonically increasing. Figure 3 illustrates an example Pareto surface for two objectives. If the Pareto frontier happens to be concave, then all the points can be determined via scalarization.

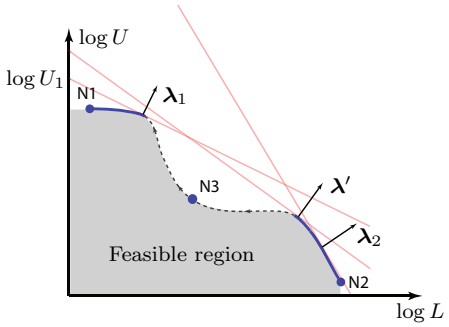


Fig. 3. An example efficient frontier between $\log U$ and $\log L$ is shown. Shaded region is the feasible region for the problem that can be achieved by increasing r_{tran}. The dashed part of the Pareto frontier cannot be determined using scalarization as there is no supporting plane to the set of feasible U and L values. Among the supporting planes, $\boldsymbol{\lambda}'$ maximizes $U \times L$.

We consider four general optimization problems relevant to the NDS:

—**Problem 1:** Determine the Pareto frontier for the N_{u} objectives,

—**Problem 2:** Maximize one objective over $\boldsymbol{\mathcal{N}}$ without regard to the values of the other $N_{\mathrm{u}} - 1$ objectives,

—**Problem 3:** Maximize the product scalarization of the utilities for a given $\boldsymbol{\lambda}$, and

—**Problem 4:** Maximize one objective while constraining the other objectives to exceed a threshold.

Once **Problem 1** is solved, it is straightforward to solve the other three single objective problems. For **Problems 2-3**, one simply determines the point on the Pareto frontier that maximizes the scalar quantity of interest. Finally, **Problem 4** searches for the maximal point on a subset of the Pareto frontier where the constraints are satisfied.

In general the relationship between the utilities and the network design is a vector of nonlinear functions, i.e., $\mathbf{u} = \mathbf{h}(n_1, \ldots, n_M)$. As shown in Section 2, all three NDS objectives can be expressed as a monotonically increasing function of a

weighted sum of sensor type populations for a given network design, i.e.,

$$u_i = h_i\left(\sum_{t=1}^{T}\omega_{i,t}n_t\right), \quad \text{for } i = 1,\ldots,3. \tag{11}$$

Note that the weights $\omega_{i,t}$ are all non-negative.

3.2 Integer-Programming

In this section, we discuss a dynamic programming solution to enumerate the parameter values comprising the Pareto frontier, thereby solving **Problem 1**. Note that the Pareto frontier is a subset of the surface of the feasible region (see Figure 3) where the feasible region is the locus of utility values corresponding to all sensor designs in the simplex $\boldsymbol{\mathcal{N}}$. It is useful to tabulate the associated network design and cost to a given utility in the feasible region. Therefore, we define the feasible triple as the set of triplets describing the utility, NDS vector, and cost for each point in the feasible region as follows

$$\mathcal{F} = \{(\text{utility: } \mathbf{h}(\boldsymbol{n}), \text{design: } \boldsymbol{n}, \text{cost: } \mathbf{c}'\boldsymbol{n}) : \boldsymbol{n} \in \boldsymbol{\mathcal{N}}\}. \tag{12}$$

Clearly, the feasible triple as given by (12) can be generated by considering each network design in $\boldsymbol{\mathcal{N}}$, and the feasible region is simply $\mathcal{F}_{\text{u}} = \mathcal{F}.\text{utility}$, i.e.,

$$\mathcal{F}_{\text{u}} = \{\mathfrak{f}.\text{utility} : \mathfrak{f} \in \mathcal{F}\}. \tag{13}$$

The finite cardinality of the simplex $\boldsymbol{\mathcal{N}}$ leads to a finite cardinality for $\mathcal{F}_{\text{u}}$ and $\mathcal{F}$.

Fortunately, the relationship between the network designs and the utilities is given by (11), and one need not resort to an exhaustive search over the simplex $\boldsymbol{\mathcal{N}}$. Many of the triples in the feasible region are redundant in the sense that they can be removed in $\mathcal{F}$ without affecting the feasible region $\mathcal{F}_{\text{u}}$ given by (13). Formally, a triple $\mathfrak{f}_1 \in \mathcal{F}_{\text{u}}$ is redundant if there exist another triple $\mathfrak{f}_2$ such that $\mathfrak{f}_1.\text{utility} = \mathfrak{f}_2.\text{utility}$ but $\mathfrak{f}_1.\text{cost} > \mathfrak{f}_2.\text{cost}$. The observation that the design $\boldsymbol{n}_1$ leads to a redundant triplet $(\mathbf{h}(\boldsymbol{n}_1), \boldsymbol{n}_1, \mathbf{c}'\boldsymbol{n}_1)$, i.e., there exist another design $\boldsymbol{n}_2$ such that $\mathbf{h}(\boldsymbol{n}_1) = \mathbf{h}(\boldsymbol{n}_2)$ but $\mathbf{c}'\boldsymbol{n}_2 < \mathbf{c}'\boldsymbol{n}_1$, instantly distinguishes many other network designs as also redundant. For instance, the network design $\boldsymbol{n}_1 + \boldsymbol{\Delta}$, where the elements of $\boldsymbol{\Delta}$ are non-negative integers, leads to a triplet that is redundant in light of the triplet associated to the design $\boldsymbol{n}_2 + \boldsymbol{\Delta}$. Algorithm 1 is a dynamic integer program to determine the feasible region that exploits the redundancy property to prune the search space.

In Algorithm 1, $\mathbf{0}$ is a column vector of zeros and $\mathbf{e}_t$ is the unit vector whose t-th element is one and whose other elements are zero. Parameter $N_{f,\max}$ is the feasible design with the most number of sensors, and n represents the total number of sensors associated to the network designs being analyzed. The sets $\mathcal{C}_0$ and $\mathcal{C}_1$ represents the set of triples associated to network designs of $n-1$ and n sensors, respectively, that are currently non-redundant, and the set $\mathcal{C}$ accumulates the non-redundant triples. For each non-redundant design consisting of n sensors, the algorithm sequentially adds one sensor of type t to the design in line 6 and tests if such a design leads to a redundant utility vector via lines 7 and 8. If the design is currently non-redundant and meets the cost constraint, its triple is added to $\mathcal{C}_1$ and $\mathcal{C}$. Furthermore, the redundant triple is removed from $\mathcal{C}$. Note that the redundancy

test is performed before applying the nonlinearity in (11). Therefore, $\mathcal{C}$ accumulates the minimal set of network designs to generate the feasible set, and line 17 transforms this set into the minimal feasible triple by incorporating the nonlinear function $\mathbf{h}(\cdot)$.

ALGORITHM 1.

(1) `Determine` $N_{f,\max} = \left\lfloor \frac{\$}{c^*} \right\rfloor$ `where` $c^* = \min_t c_t$.
(2) `Initialize` $\mathcal{C}_0 = \{(\mathbf{0}_{N_u \times 1}, \mathbf{0}_{T\times 1}, 0)\}, \mathcal{C} = \emptyset, \mathcal{C}_1 = \emptyset$.
(3) `for` $n = 1, \ldots, N_{f,\max}$,
(4) `for` $t = 1, \ldots, T$,
(5) `for all` $\mathfrak{f}^* \in \mathcal{C}_0$,
(6) $\mathfrak{f}^+ = (\mathfrak{f}^*.\texttt{utility} + \tilde{\mathbf{\Omega}}\mathbf{e}_t, \mathfrak{f}^*.\texttt{design} + \mathbf{e}_t, \mathfrak{f}^*.\texttt{cost} + c_t)$;
(7) `if` $\exists \mathfrak{f} \in \mathcal{C}$ `such that` $\mathfrak{f}^+.\texttt{utility} = \mathfrak{f}.\texttt{utility}$,
(8) `if` $\mathfrak{f}^+.\texttt{cost} < \mathfrak{f}.\texttt{cost}$,
(9) `remove` $\mathfrak{f}$ `from` $\mathcal{C}$ `and insert` $\mathfrak{f}^+$ `into` $\mathcal{C}$ `and` $\mathcal{C}_1$;
(10) `elseif` $\mathfrak{f}^+.\texttt{cost} \leq \$$,
(11) `insert novel` $\mathfrak{f}^+$ `into` $\mathcal{C}$ `and` $\mathcal{C}_1$;
(12) `end`;
(13) `end`;
(14) $\mathcal{C}_0 \leftarrow \mathcal{C}_1$;
(15) $\mathcal{C}_1 = \emptyset$;
(16) `end`;
(17) $\mathcal{F} = \{(\mathbf{h}(\mathfrak{f}.\texttt{utility}), \mathfrak{f}.\texttt{design}, \mathfrak{f}.\texttt{cost}) : \mathfrak{f} \in \mathcal{C}\}$ `forms the minimal triple.`

Algorithm 1 requires that the relationship between the utilities and the network design is given by (11), but it does not assume any more structure regarding the utilities. For the NDS, the second element of the utility vector is the lifetime objective, which is equivalent to the total number of sensors. Thus, the second element of the utility vectors associated to $\mathcal{C}_0$ and $\mathcal{C}_1$ are always $n-1$ and n respectively. The test in line 7 can only be positive if $\mathfrak{f} \in \mathcal{C}_1$.

Before the application of Algorithm 1, the weighting elements $\omega_{i,t}$ should be transformed into integers by an appropriate scaling within a desired level of accuracy. Integer $\omega_{i,t}$'s in turn lead to integer values for $\mathfrak{f}$.utility. Furthermore, these values can be bounded by $N_b = N_{f,\max} \max \omega_{i,t}$. In other words, the localization accuracy and coverage utilities form N_b^2 unique values and the existence and non-existence of these N_b^2 unique utility values in $\mathcal{C}_0$ and $\mathcal{C}_1$ can be stored in $N_b \times N_b$ matrices. This replaces the implicit search in line 7 with a table look-up. As a result, the overall complexity of Algorithm 1 is $O(N_{f,\max} T N_b^2)$.

Once the non-redundant feasible triple is determined by Algorithm 1, it is straightforward to determine the Pareto frontier. The Pareto frontier $\mathcal{P}$ is simply the maximal subset of $\mathcal{F}_u$ such for any $\mathbf{u} \in \mathcal{P}$.utility and any $\mathbf{u}^\dagger \in \mathcal{F}$.utility with $\mathbf{u} \neq \mathbf{u}^\dagger$,

then it is not the case that $u_i \leq u_i^\dagger$ for all $i \in \{1, \ldots, M\}$, i.e., $\mathbf{u}$ is not pointwise less than $\mathbf{u}^\dagger$. A simple method to compute $\mathcal{P}$ is to exhaustively remove any point in $\mathcal{F}_\mathrm{u}$ that is pointwise less than another. Let $\mathcal{P}_i$ represent the surface formed by maximizing the i-th objective over points in $\mathcal{F}_\mathrm{u}$ whose other objectives are equal. In other words, if $\mathbf{u} \in \mathcal{P}_i$.utility, $\mathbf{u}^\dagger \in \mathcal{F}_\mathrm{u}$, and $u_j = u_j^\dagger$ for $j \neq i$, then $u_i \geq u_i^\dagger$. Clearly, if one arbitrarily selects a utility $\mathbf{u}$ from $\mathcal{F}$.utility there exist an utility $\mathbf{u}^\dagger$ from $\mathcal{P}_i$.utility that is pointwise greater than or equal to $\mathbf{u}$. Thus, $\mathcal{P}_i$ is a superset of $\mathcal{P}$. As demonstrated in Section 5, this feasible surface is not necessarily equivalent to the Pareto frontier. As mentioned at the end of Section 3.1, once the collection of points forming the Pareto frontier is known (**Problem 1**), then the solutions to **Problems 2-4** are straightforward.

3.3 Non-integer Optimization

The optimal NDS solution to **Problem 1** determines the Pareto frontier where feasible designs occupy a simplex $\boldsymbol{\mathcal{N}}$ in the T-dimensional integer lattice $\mathbb{Z}^T$. By relaxing the simplex $\boldsymbol{\mathcal{N}}$ in the T-dimensional real space $\mathbb{R}^T$, optimization algorithms can become more computationally efficient. The solutions obtained by continuous relaxations are usually not optimal. However, when the total budget \$ is rather large, optimal solutions select the key sensor types in great numbers. Then, the difference between the solutions obtained by continuous relaxations and the optimal integer solutions become negligible. Continuous relaxations also allow us to gain further insights into the other NDS problems, as we will see this section.

For the ease of notation, we define the following transformations: $\tilde{n}_m = c_m n_m/\$$, $\tilde{\omega}_{i,t} = \omega_{i,t}/c_m$, and $[\tilde{\mathbf{\Omega}}]_{it} = \tilde{\omega}_{i,t}$ so that the NDS problem is equivalent to

$$\text{Find the Pareto frontier of } \tilde{\mathbf{\Omega}}\tilde{\boldsymbol{n}} \text{ subject to } \mathbf{1}'\tilde{\boldsymbol{n}} \leq 1 \text{ and } \tilde{\boldsymbol{n}} \succeq \mathbf{0}, \tag{14}$$

where $\mathbf{1}$ is a T dimensional column vector of ones. The following three theorems alleviate the search for the optimal points on the Pareto frontier.

THEOREM 1. *(Concentration of Resources) The optimal NDS design $\tilde{\boldsymbol{n}}^*$ associated with the Pareto frontier of the multi-objective optimization problem in (14) expense the available resources in at most N_u different sensor types.*

PROOF. The Pareto frontier is a subset of feasible utilities in the positive orthant of the N_u-dimensional space. Within the feasible region, any point can be viewed as a unit vector times a scale factor. A point in the feasible region where the scale factor is maximized for a given direction is located on the Pareto frontier of (14). Let us consider the an arbitrary unit vector in the positive orthant, i.e., $\mathbf{e}$ such that $\mathbf{e} \succeq 0$ and $\mathbf{e}'\mathbf{e} = 1$. Then, consider the following the linear program

$$\min \mathbf{1}'\tilde{\mathbf{n}} \;\text{ subject to }\; \tilde{\mathbf{\Omega}}\tilde{\boldsymbol{n}} = \mathbf{e} \;\text{ and }\; \tilde{\boldsymbol{n}} \succeq \mathbf{0}. \tag{15}$$

This linear program has N_u equality constraints; hence, the optimal solution $\tilde{\boldsymbol{n}}^*$ contains at most N_u nonzero elements (Theorem 20.1) [Moon and Sterling 2000]. Let $\tilde{\$}^* = \mathbf{1}'\tilde{\boldsymbol{n}}^*$ be the minimum objective value of the linear program in (15) at the optimal parameter value $\tilde{\boldsymbol{n}}^*$. Then, a network design $s\tilde{\boldsymbol{n}}^*$ where $s = 1/\tilde{\$}^*$ meets the cost constraint of (14) and leads to a feasible point $s\mathbf{e}$ that lies on the boundary. Otherwise, if there was another design $\tilde{\boldsymbol{n}}_1$ that leads to a feasible point in the same

direction but larger scale factor, then $\tilde{\boldsymbol{n}}^*$ could not solve the linear program in (15). Namely, if $\tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{n}}_1 = s_1\mathbf{e}$ where $s_1 > s$ and $\mathbf{1}'\tilde{\boldsymbol{n}}_1 \leq 1$, the NDS design $\tilde{\boldsymbol{n}}_2 = 1/s_1\tilde{\boldsymbol{n}}_1$ would solve the constraint in (15) at a cost of

$$\begin{aligned}\mathbf{1}'\tilde{\boldsymbol{n}}_2 &= 1/s_1\mathbf{1}'\tilde{\boldsymbol{n}}_1\\ &\leq 1/s_1,\\ &< 1/s = \tilde{\$}^* = \mathbf{1}'\tilde{\boldsymbol{n}}^*,\end{aligned}$$

which creates a contradiction because $\tilde{\boldsymbol{n}}^*$ is the optimal solution of (15).

The proof is true for any direction in the positive orthant; hence, any network design associated with the Pareto frontier has at most N_{u} nonzero elements. This result reduces the NDS sensor search space from $2^T = \sum_{t=1}^{T}\binom{T}{t}$ to $N_{\mathrm{sc}} = \binom{T}{N_{\mathrm{u}}}$ sensor combinations. □

THEOREM 2. *(Dominating Sensor Pairs–1) Consider two indices $(j,k) \in [1,T]$ and $j < k$ for which the row vectors forming the matrix $\tilde{\boldsymbol{\Omega}}$ in (14) are locally strictly convex $\forall m \in (j,k)$, i.e., $\tilde{\omega}_{i,m} < \frac{k-m}{k-j}\tilde{\omega}_{i,j} + \frac{m-j}{k-j}\tilde{\omega}_{i,k}, \forall i$. Then, no sensor m is assigned any resources on the Pareto frontier. That is, sensors j and k dominate any sensor m in the solution.*

PROOF. Assume a feasible $\tilde{\boldsymbol{n}}^*$ for which $\tilde{n}^*_m > 0$ and define a new resource distribution where $\tilde{n}_j = \tilde{n}^*_j + \beta\tilde{n}^*_m$, $\tilde{n}_k = \tilde{n}^*_k + (1-\beta)\tilde{n}^*_m$, and $\tilde{n}_m = 0$ while keeping all other elements the same from $\tilde{\boldsymbol{n}}^*$, where $\beta = \frac{k-m}{k-j}$. The utilities associated to these two designs are $\mathbf{u}^* = \tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{n}}^*$ and $\mathbf{u} = \tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{n}}$, and $(u_i - u^*_i)/\tilde{n}^*_m = \beta\tilde{\omega}_{i,j} + (1-\beta)\tilde{\omega}_{i,k} - \tilde{\omega}_{i,m}$. Thus by the convexity, $u_i > u^*_i$ for all objectives i. Hence, on the Pareto frontier $u^*_m = 0$, otherwise it is always possible to increase all the objectives in the multi-criteria optimization problem.

If the dominating sensor pairs are on the boundary of the sensor book, then the final solution of the sensor network design has at most three configuration choices: only sensor type $t = 1$, only sensor type $t = T$, or a combination of sensor types $t = 1$ and $t = T$. □

THEOREM 3. *(Dominating Sensor Pairs–2) Consider two indices $(j,k) \in [1,T]$ and $j < k$ for which $N_u - 1$ row vectors forming the matrix $\tilde{\boldsymbol{\Omega}}$ in (14) are locally strictly convex $\forall m \in (j,k)$ with respect to the ordering of remaining row vector, i.e., $\tilde{\omega}_{i,m} < \frac{\tilde{\omega}_{i',k}-\tilde{\omega}_{i',m}}{\tilde{\omega}_{i',k}-\tilde{\omega}_{i',j}}\tilde{\omega}_{i,j} + \frac{\tilde{\omega}_{i',m}-\tilde{\omega}_{i',j}}{\tilde{\omega}_{i',k}-\tilde{\omega}_{i',j}}\tilde{\omega}_{i,k}, \forall i \neq i'$. Then, no sensor m is assigned any resources on the Pareto frontier. That is, sensors j and k dominate any sensor m in the solution.*

PROOF. The proof is similar to that of Theorem 2, and it is omitted. □

The difference between Theorem 2 and Theorem 3 is the ordering of the feasible sensors. Whereas the ordering of the sensors in Theorem 2 is arbitrary, Theorem 3 requires that they are sorted with respect to the order of one of the rows of $\tilde{\boldsymbol{\Omega}}$. Theorems 1-3 provide the following simplifications of the NDS problems in Section 3.1:

3.3.1 *Solution of **Problem 2**.* In light of Theorem 1, the solution of **Problem 2** simply assigns all the resources to a single sensor type because $N_{\mathrm{u}} = 1$, i.e., $\tilde{\boldsymbol{n}} = \mathbf{e}_t$

for some t. The sensor modality that maximizes the utility is given by

$$t^* = \arg\max_t \tilde{\omega}_{1,t}. \tag{16}$$

Hence, one simply chooses the modality that provides the *biggest bang for the buck*.

3.3.2 *Solution of **Problem 3**.* The solution to **Problem 3** also assigns resources to at most N_{u} sensor types due to Theorem 1. Theorems 2 and 3 can be used to further decrease N_{sc} by removing sensors that are dominated in the feasible sensor set. Subsequently, given each sensor combination, one must maximize the utility over the resource allocation into each of the N_{u} sensor types.

Let us further examine the solution of **Problem 3** to see how further simplifications can be achieved in the resource allocation of the sensor combinations using Karush-Kuhn-Tucker (KKT) conditions. Without loss of generality, we specifically consider the case the overall utility is the product of the localization accuracy (5) and lifetime (8) utilities ($N_{\mathrm{u}} = 2$). Then, an equivalent problem is solve the objective product where $\boldsymbol{\lambda} = [1 \;\; \rho \;\; 0]$, and $\rho \triangleq \alpha/\gamma > 0$. The overall utility is

$$K(\tilde{\boldsymbol{n}}) = \left(\sum_{t=1}^{T} \tilde{\omega}_{1,t}\tilde{n}_t\right)\left(\sum_{t=1}^{T} \tilde{\omega}_{2,t}\tilde{n}_t\right)^{\rho}. \tag{17}$$

We start by noting that the following is a necessary KKT condition for the optimal solution $\tilde{\boldsymbol{n}}^*$ to satisfy over its simplex [Bertsekas 2003; Epelman et al. 2005]:

$$\tilde{n}_m^* > 0, \;\Rightarrow\; \left.\frac{\partial K(\tilde{\boldsymbol{n}})}{\partial \tilde{n}_m}\right|_{\tilde{n}_m=\tilde{n}_m^*} \geq \left.\frac{\partial K(\tilde{\boldsymbol{n}})}{\partial \tilde{n}_l}\right|_{\tilde{n}_l=\tilde{n}_l^*} \quad \forall\; l \text{ and } m, \tag{18}$$

which implies that at the optimum point, it is not possible to move resources from any non-zero u_m to any other u_l without decreasing the utility $K(\tilde{\boldsymbol{n}})$. Using the derivative of (17), it is straightforward to see that the necessary condition (18) can be explicitly written as follows:

$$\tilde{n}_m^* > 0, \;\Rightarrow\; \sum_{t=1}^{T} G_{mt}\tilde{n}_t \geq \sum_{t=1}^{T} G_{lt}\tilde{n}_t \;\forall\; l \text{ and } m. \tag{19}$$

where $G_{jk} = \tilde{\omega}_{1,j}\tilde{\omega}_{2,k} + \rho\tilde{\omega}_{1,k}\tilde{\omega}_{2,j}$. Then, the optimal resource allocation solution is given below without a detailed derivation:

$$\tilde{n}_j^* = \begin{cases} 1 & G_{jj} \geq G_{kj}; \\ 0 & G_{kk} \geq G_{jk}; \\ \frac{(\tilde{\omega}_{1,j}-\tilde{\omega}_{1,k})\tilde{\omega}_{2,k}+\rho(\tilde{\omega}_{2,j}-\tilde{\omega}_{2,k})\tilde{\omega}_{1,k}}{(1+\rho)(\tilde{\omega}_{1,j}-\tilde{\omega}_{1,k})(\tilde{\omega}_{2,j}-\tilde{\omega}_{2,k})} & G_{jj} < G_{kj} \text{ and } G_{kk} < G_{jk} \end{cases} \tag{20}$$

where $\tilde{n}_k^* = 1 - \tilde{n}_j^*$, and $\tilde{n}_m^* = 0$ for $m \neq j, k$. Hence, when $N_{\mathrm{sc}} = \binom{T}{2}$ possible combinations of the sensors in $\boldsymbol{s}$ are considered, (20) can be used to determine the optimal resource allocations. In the general case, (18) can be exploited to facilitate the resource allocation computations.

3.3.3 *Solution of **Problem 4**.* Using similar arguments in Theorem 1, the solution to **Problem 4** assigns resources to at most N_{u} sensor types. For this problem, the optimization problem is a linear program where there are N_{u} linear constraints ($N_{\mathrm{u}} - 1$ constraints are from objectives and one cost constraint). Therefore, it is straight forward to solve **Problem 4**.

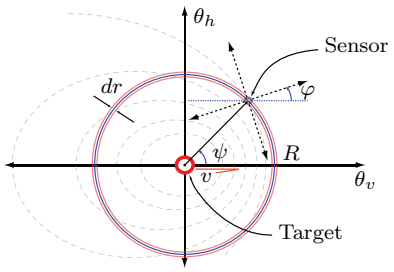


Fig. 4. This figure illustrates the geometry of the sensor-target configuration for a monopole acoustic source moving along the horizontal direction with a speed of v. The dashed lines represent the acoustic wave-fronts, which create the interaction between the target and the sensor.

Lastly, once the NDS $\tilde{\boldsymbol{n}}^*$ to **Problems 2, 3, or 4** is obtained via continuous relaxations, it is converted to an integer solution via

$$n_t^* = \text{floor}\left(\$\frac{\tilde{n}_t^*}{c_t}\right). \tag{21}$$

For large budgets, the NDS solution $\boldsymbol{n}^*$ can be slightly different from the solutions via integer programming as the support of the integer solution may be greater than N_{u}. In the integer solution, most of the resources are used in N_{u} sensor types, but a few other sensor types help achieve a network cost as close to the budget \$ as possible. In a sense, the rounding error in (21) allows for a few number of sensors outside of the dominant N_{u} to appear in the NDS solution.

4. ACOUSTIC SENSOR NETWORKS

In this section, we focus on the expressions that will lead to the localization utility of acoustic sensor networks for position estimation. We first provide data models, then give expressions for the Fisher information matrices (FIM) and the sensor receiver operating characteristics (ROC) curves for position estimation using acoustic sensors. We assume that the target and the sensors are on the same ground plane, measured in horizontal (h) and vertical (v) coordinates. The target position is denoted as $\boldsymbol{\theta} = \left[\ \theta_h,\ \theta_v\ \right]'$ and the position of the sensor is denoted as $\boldsymbol{\zeta} = \left[\ \zeta_h,\ \zeta_v\ \right]'$. The target bearing ϕ is measured counterclockwise with respect to the sensor's h-axis (sensor orientation: φ) and the target range $R = ||\boldsymbol{\theta} - \boldsymbol{\zeta}||$ is defined as the Euclidian distance between the target and the sensor (Fig. 4).

Our results use the analytical signal representations because time-shifts in real signals correspond to phase shifts in their analytical representations. We denote $s(t)$, $x(t)$, $n_i(t)$, and $y_i(t)$ as the complex envelopes of the source signal, the source signal at the array, the ith microphone additive noise, and the ith microphone output signal, respectively. To calculate target bearing or range, N snapshots of the observed acoustic data, taken at times $t_1, \ldots, t_N$, are used. We note that if the time samples are sufficiently apart, then successive samples of the source and the noise samples are uncorrelated [Davenport II and Root 1958; Weiss and Weinstein 1983]. We model the source signal samples as *i.i.d.*, zero mean, complex circularly symmetric Gaussian random variables with variance σ_s^2 and the

noise samples with *i.i.d.*, zero mean, complex circularly symmetric Gaussian random variables with covariance $\sigma^2 \boldsymbol{I}$.

4.1 Range Sensors

We first discuss the range estimation of a narrow-band source using an omni-directional microphone in an isotropic medium. We assume that there are no multipath effects. Assuming spherical propagation, we write the envelope of the microphone output signal at the target narrow band frequency f_w as follows, [Morse and Ingard 1968; Johnson and Dudgeon 1993]:

$$y(t) = x(t) + n(t) = \frac{s(t)}{\sqrt{\beta} R} e^{-j\frac{2\pi f_w R}{\beta c}} + n(t), \tag{22}$$

where $\beta(t) = 1 + \frac{v}{c}\cos\psi$ is called the Doppler shift factor and c is the speed of sound. Based on our signal and the noise signal assumptions, it is straightforward to prove that the microphone output signal $y(t)$ also has an *i.i.d.* zero mean circularly complex Gaussian distribution with variance $\sigma_y^2 = \sigma_x^2 + \sigma^2$, where $\sigma_x^2 = \frac{\sigma_s^2}{\beta R^2}$. Now, we denote the N-sample root-mean-squared microphone output as ε:

$$\varepsilon = \sqrt{\frac{1}{N}\sum_{i=1}^{N} |y(t_i)|^2} = \frac{\sigma_y}{\sqrt{2N}} \sqrt{\sum_{i=1}^{N} \left(\frac{y_{\text{real}}^2(t_i)}{\sigma_y^2/2} + \frac{y_{\text{imag}}^2(t_i)}{\sigma_y^2/2} \right)} = \frac{\sigma_y}{\sqrt{2N}} z, \tag{23}$$

where we define z as the second square-root summation term in (23). The variable z has a Chi distribution $p_Z(z)$ with $2N$ degrees of freedom [Evans et al. 2000].

Although we now have an analytical expression for ε, we apply the Laplacian approximation to further facilitate the derivations of the ROC curves. The Laplacian approximation uses the mode and the Hessian of the log likelihood at the mode to approximate it with a Gaussian. The resulting approximation is ($N \gg 1$)

$$\varepsilon \sim p(\varepsilon) = \mathcal{N}\left(\sqrt{\frac{2N-1}{2N}}\sigma_y, \frac{\sigma_y^2}{4N} \right) \approx \mathcal{N}\left(\sqrt{\sigma_x^2 + \sigma^2}, \frac{\sigma_x^2 + \sigma^2}{4N} \right). \tag{24}$$

The Fisher information matrix $\boldsymbol{F}$ is defined as follows [Lehmann and Casella 1998]:

$$\boldsymbol{F}(\boldsymbol{\theta}) = \int p(\varepsilon|\boldsymbol{\theta}) \left[\frac{\partial \log p(\varepsilon|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] \left[\frac{\partial \log p(\varepsilon|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right]^T d\varepsilon. \tag{25}$$

We use $p(\varepsilon)$ in (24) to determine an approximate FIM of the target position estimate given below without a detailed derivation:

$$\boldsymbol{F}_1(\boldsymbol{\theta}) = \frac{4N(\sigma_s^2/\sigma^2)^2}{R^2\left(\sigma_s^2/\sigma^2 + R^2\right)^2} \begin{bmatrix} \cos\psi \\ \sin\psi \end{bmatrix} \times \begin{bmatrix} \cos\psi & \sin\psi \end{bmatrix}, \tag{26}$$

where ψ is the sensor bearing with respect to the target (see Fig. 4) and it is assumed that $\sqrt{\beta} \approx 1$.

In the case of wideband sources, if the observation period of the N samples is much larger than the inverse bandwidth of the source signal $s(t)$, but short enough so that the target is approximately stationary, then the discrete Fourier transform (DFT) coefficients of the signal are statistically uncorrelated, [Davenport II and Root 1958; Weiss and Weinstein 1983; Chow and Schultheiss 1981]. Using the properties of the FIM, [Van Trees 1968], the wideband FIM is a summation of the individual narrow-band FIM's.

To derive the receiver operating characteristics (ROC) curves of the range sensors, we start with (24). Note that when $N \gg 1$, we can write $\varepsilon \approx \sigma_y + \frac{\sigma_y}{\sqrt{4N}}\mathcal{N}(0,1) \approx \sigma_y e^{\frac{\mathcal{N}(0,1)}{\sqrt{4N}}}$. Hence, a multiplicative noise model on the envelope is more appropriate than an additive model. Denoting $\mathcal{E} = \log \varepsilon$, we have

$$\mathcal{E} \sim \mathcal{N}\left(\log \sigma_y, \frac{1}{4N}\right) \tag{27}$$

with the following hypotheses:

$$\begin{aligned} &H_0\text{: } \sigma_y = \sigma_n, \\ &H_1\text{: } \sigma_y = \sqrt{\sigma^2 + \sigma_x^2} > \sigma_n. \end{aligned} \tag{28}$$

In this case, likelihood ratio test is equivalent to a simple linear threshold detector:

$$\mathcal{E} \underset{H_0}{\overset{H_1}{\gtrless}} \eta', \tag{29}$$

and this detector is also uniformly most powerful (UMP) [Van Trees 1968; Lehmann and Casella 1998]. It is easy to verify that the ROC curve is determined by

$$P_d = \Phi\left(\Phi^{-1}(P_f) - \sqrt{N}\left(\log(\sigma^2 + \sigma_x^2) - \log \sigma^2\right)\right), \tag{30}$$

where $\Phi(\cdot)$ is one minus the cumulative distribution function of the $\mathcal{N}(0,1)$-random variable, P_d and P_f are the detection and false alarm probabilities, respectively.

4.2 Bearing Sensors

Bearing sensors determine the direction-of-arrival (DOA) of a target by exploiting the time-delay information present in the elements of the acoustic array [Johnson and Dudgeon 1993]. When the number of acoustic data samples used to calculate the bearing are sufficiently high, the estimated bearing has the following Gaussian distribution, [Bell et al. 1996; Liu et al. ; Johnson and Dudgeon 1993]:

$$\phi = \tan^{-1}\left(\frac{\theta_v - \zeta_v}{\theta_h - \zeta_h}\right) + n_\phi, \tag{31}$$

where the bearing noise n_ϕ depends on the actual microphone noise on the array elements.

To determine the relationship between the bearing noise to the additive microphone noise, consider a narrow-band plane wave signal impinging on an m-microphone planar acoustic array. We first derive the Fisher information matrix for the bearing ϕ and approximate the variance of n_ϕ using the inverse of the FIM, [Johnson and Dudgeon 1993; Stoica and Nehorai 1989; Stoica et al. 2001; Bell et al. 1996]. The wavenumber vector $\boldsymbol{k}$ is given by

$$\boldsymbol{k} = \frac{2\pi f_w}{c}\boldsymbol{u}, \quad \boldsymbol{u} = \left[\begin{array}{cc} \cos\phi & \sin\phi \end{array}\right]', \tag{32}$$

where f_w is the narrow-band source temporal frequency, c is the speed of propagation, and $\boldsymbol{u}$ is the Cartesian bearing. The microphone locations in the acoustic array are given by the matrix $\boldsymbol{D}$ defined as follows:

$$\boldsymbol{D} = \left[\begin{array}{ccc} \boldsymbol{d}_1 & \ldots & \boldsymbol{d}_m \end{array}\right] = \left[\begin{array}{ccc} d_{h,1} & \ldots & d_{h,m} \\ d_{v,1} & \ldots & d_{v,m} \end{array}\right], \tag{33}$$

where $\boldsymbol{d}_i$ has the location of the ith microphone ($i = 1, \ldots, m$) in the Cartesian coordinate system. The time delay at the ith microphone relative to the origin of the array is given by $\tau_i = \frac{\boldsymbol{u}^T \boldsymbol{d}_i}{c}$. Then, the array output can be written as

$$\boldsymbol{y}(t) = \boldsymbol{a}(\boldsymbol{u})x(t) + \boldsymbol{n}(t), \tag{34}$$

where $\boldsymbol{y}(t) = \left[\, y_1(t) \;\; \ldots \;\; y_m(t) \,\right]'$, $\boldsymbol{n}(t) = \left[\, n_1(t) \;\; \ldots \;\; n_m(t) \,\right]'$, $\boldsymbol{a}(t) = \left[\, e^{-j2\pi f_w \tau_1} \;\; \ldots \;\; e^{-j2\pi f_w \tau_m} \,\right]'$. Under stationarity and Gaussian assumptions, it can be shown that the Fisher information matrix for a narrow-band source is given by [Bell et al. 1996]:

$$\boldsymbol{F}(\boldsymbol{u}) = \frac{2N}{m} \frac{\left(\frac{m\sigma_s^2}{R^2\sigma^2}\right)^2}{1 + \frac{m\sigma_s^2}{R^2\sigma^2}} \left(\frac{2\pi f_w}{c}\right)^2 \boldsymbol{D} \left(\boldsymbol{I} - \frac{1}{m}\boldsymbol{1}\boldsymbol{1}^T\right) \boldsymbol{D}^T, \tag{35}$$

where the matrix term involving the position matrix can be shown equal to $m(a^2/2)\boldsymbol{I}$ for uniform circular arrays where a is the array radius. Then, it is straightforward to determine the Fisher information matrix for the bearing ϕ via the following transformation, e.g., see [Van Trees 1968]:

$$F_\phi(\phi) = (\nabla_\phi \boldsymbol{u})^T \, \boldsymbol{F}(\boldsymbol{u}) \, (\nabla_\phi \boldsymbol{u}) \,. \tag{36}$$

Note that when the acoustic array has a uniform circular geometry, $F_\phi(\phi)$ does not depend on ϕ. Using (31) and (36), we present the array FIM's for the target position without a detailed derivation (see Fig. 4):

$$\boldsymbol{F}_m(\boldsymbol{\theta}) = \frac{F_\phi(\pi + \psi - \varphi)}{R^2} \begin{bmatrix} \sin\psi \\ \cos\psi \end{bmatrix} \times \left[\, \sin\psi \;\; \cos\psi \,\right]. \tag{37}$$

The wideband version for the bearing sensors is a summation of the corresponding narrowband FIM's similar to the range sensors.

To derive the ROC curves for bearing sensors, we use the following detector:

$$\max_m \mathcal{E}_m \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{38}$$

where $\mathcal{E}_m = \log \varepsilon_m$ for the mth sensor. Probability that the maximum of m statistically independent Gaussian random variables $\mathcal{E}_m$ with the mean $\log\sigma^2$ and variance $\frac{1}{4N}$ exceeds the threshold η is given by

$$P_f = 1 - \prod_{k=1}^{m} \left[1 - \Phi\left(\sqrt{4N}\left(\eta - 0.5\log\sigma^2\right)\right)\right]. \tag{39}$$

Similarly, the detection probability is given by

$$P_d = 1 - \left[1 - Q\left(Q^{-1}\left((1-P_f)^{\frac{1}{M}}\right) - \sqrt{N}\left(\log(\sigma^2 + \sigma_x^2) - \log\sigma^2\right)\right)\right]^m. \tag{40}$$

5. SIMULATIONS

5.1 Example Dynamic Programming and Approximate Branch-and-Bound Solutions

Using a synthetic example, we demonstrate the dynamic programming solution of the NDS problem and also compare them with the approximate solutions to establish their closeness. This comparison is quite important as the approximate solutions obtained by continuous relaxations sometimes fail to be useful in finding a branch-and-bound solution [Nemhauser and Wolsey 1988]. We use the following

synthetic problem: $T = 6$, $c_t = 1 + t$, $f_t = t^2$, $\delta = 1.2$, $\gamma = 2.2$, \$ = 500, and $R^* = [\,1,\ 2,\ 2,\ 2,\ 3,\ 3\,]'$. With these parameters, sensor type 6 has the maximum $\tilde{\omega}_{1,t}$, whereas sensor type 5 has the maximum $\tilde{\omega}_{3,t}$.

Figures 5(a) and (b) show the evolution of the dynamic programming solution for **Problem 1** and the Pareto frontier surface. In Fig. 5(a), u_1 is maximized at $N = 72$ where $\boldsymbol{n} = [\,0\ 1\ 0\ 0\ 0\ 71\,]'$ (**Problem 2**). When $N \leq 71$, the optimal $\boldsymbol{n}$ consists of only s_6 (see Fig. 5(c) (*bottom*)). Due to the integer nature of the problem, $N = 72$ gives slightly better result, because, when $N = 71$, the cost $c = 71 \times 7 = 497 \leq$ \$ results in $l = 71 \times 6^2 = 2556$, whereas, when $N = 71 + 1 = 72$, $c = 71 \times 7 + 1 \times 3 = 500 \leq$ \$ results in $l = 71 \times 6^2 + 1 \times 2^2 = 2560$. Note that this is less than $l = (71 + 3/7) \times 6^2 = 2571.4$ if it were possible to buy fractional sensors. However, most resources are only spent on the best performing sensors for their cost as expected. In Fig. 5(a), u_3 is maximized at $N = 84$ where $\boldsymbol{n} = [\,1\ 0\ 0\ 0\ 83\ 0\,]'$. As expected, most of the resources are spent on sensor type 5. Figure 5(c) demonstrates the distribution of the resources (*top*) at a constant coverage V while trading lifetime L for localization utility U starting from maximum utility, (*middle*) at a constant lifetime while trading coverage V for utility U starting from the maximum coverage, and (*bottom*) at the boundary of the $\mathcal{F}$ solutions where the lifetime L is traded for coverage V. In all cases, sensor type 3 rarely gets assigned any resources. This is because the rows of $\tilde{\boldsymbol{\Omega}}$ are locally convex around sensor type 3, hence sensor type 3 gets dominated by the combinations of sensor types 2 and 4. In fact, on the whole Pareto frontier, $n_3 \leq 1$, where it is only assigned resources due to the integer nature of the problem. Hence, for all practical purposes, sensor type 3 could have been eliminated before any solution is attempted.

For the solution of **Problem 3**, it is straightforward to see that the convexity conditions for the approximate solutions are satisfied. Analytical formula (20) for the solution of **Problem 3** results in $u_1 = 0.1321$ and $u_6 = 0.8679$ with $u_m = 0$ for $m = 2, \ldots, 5$. The continuous resource variables u_m's correspond to $\boldsymbol{n}$ estimates as $n_1 = 33.0317$, $n_6 = 61.9910$, and $n_m = 0$ for $m = 2, \ldots, 5$ via $n_m = u_m\$/c_m$. The dynamic programming solution in Algorithm 1 results in $\boldsymbol{n}^* = [\,33\ 0\ 0\ 0\ 0\ 62\,]'$, which closely exhibits the boundary effect phenomena as predicted by the branch-and-bound solution.

Figures 6(a) and (b) illustrate the evolution of the dynamic programming solution and Fig. 6(c) plots the Pareto frontier for the solution of **Problem 3**. In Fig. 6(b), the solution of **Problem 3** is shown, where $N = 95$, where we decrease the maximum utility by 1.305 times to increase the lifetime by 1.418. In Fig. 6(c), we show other operating points on the Pareto frontier. For example, to increase the lifetime 3.47 times, the utility needs to be decreased 10 times. Finally, we observed that all the points on the Pareto frontier exhibit the boundary behavior similar to what we expected for the solution of **Problem 3**.

5.2 Acoustic Sensor Network Design

In this section, we use our results to design an acoustic sensor network for position estimation. Typical target acoustic time-frequency signatures are shown in Fig. 7. Based on the typical vehicle time-frequency characteristics, we model the power of a vehicle as wideband with flat spectral characteristics between $0 - 300$Hz. The intersensor spacing of the uniform circular acoustic arrays is determined using the

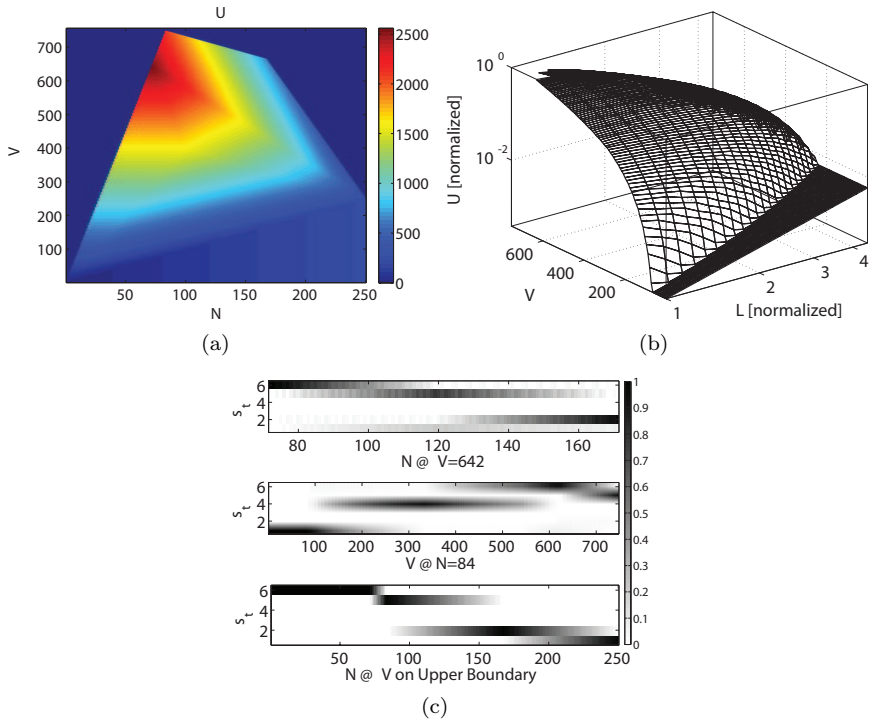


Fig. 5. (a) The dynamic programming solution via Algorithm 1 for $N_{\mathrm{u}} = 3$. (b) Pareto efficient frontier is shown with different operating points. The utility U and the lifetime L are plotted in logarithmic scale and are normalized. The maximum U is -1. Also, $L(N_{f,\min})$ is 1. (c) Resource distribution is shown for various paths on (a). Darker colors imply more resources. Note that sensor type 3 is rarely assigned resources.

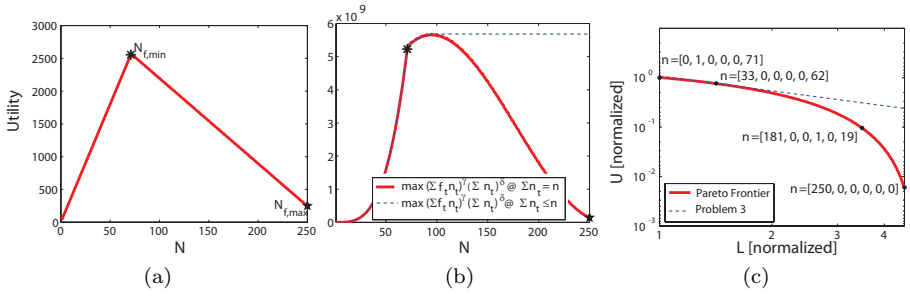
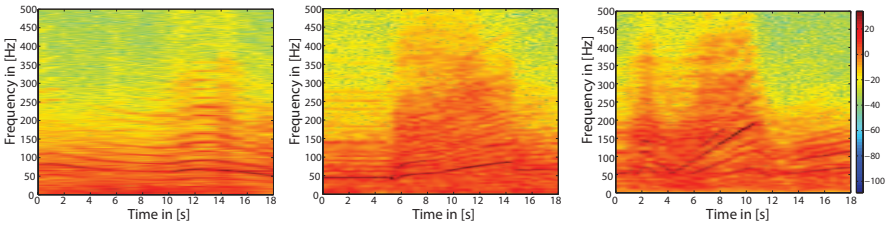


Fig. 6. (a) The dynamic programming solution via Algorithm 1 for $N_{\mathrm{u}} = 2$. (b) Corresponding solution of **Problem 3** is illustrated (c) Pareto efficient frontier is shown with different operating points. The dashed line shows the supporting line with the maximum slope going through the origin.

bandwidth of the source signals. We constrain the sensor sizes to be less than or equal to 1m for practical reasons.



(a) Compact (b) SUV (c) Truck

Fig. 7. Time frequency images in dB scale for (a) a compact vehicle (b) a sports utility vehicle (SUV), and (c) a truck, each moving at approximately 40km/h. The sampling rate for the microphone is $F_s = 44100$Hz. The microphone has flat spectral response up to 1000Hz. A notch filter is used to minimize the electrical interference at 60Hz. At moderate vehicle speeds, most of the vehicle signal energy lies within 0-300Hz region. Moreover, most of the energy is approximately uniformly distributed in addition to the concentration around engine cylinder firing frequencies.

To characterize the typical vehicle signal power levels, we use the published measurements of the U.S. Department of Transportation Federal Highway Administration [Department-of-Transportation]. For vehicles, a functional form of the vehicle noise power is given by

$$\sigma_s^2 = 10\log_{10}\left(10^{(C/10)} + v^{(A/10)}10^{(B/10)}\right) + 10\log_{10}\left(15^2\right), \tag{41}$$

where v is the vehicle speed in units of km-per-hour, $A = 41.74$, $B = 1.149$, and $C = 50.13$. In (41), the second logarithmic term accounts for the measurement distance, which is 15m. Hence, (41) illustrates the average vehicle signal power for automobiles, medium trucks, heavy trucks, buses, and motorcycles. The signal power has a flat level at $\sigma_s^2 \approx 73$dB when $v \leq 10$km/h due to the engine and exhaust, then has a linear increase when $v \geq 10$km/h ($\sigma_s^2 \approx 107$dB at $v = 100$km/h) attributed to the tire and pavement noise. For our example design in this section, we exclude the heavy trucks, e.g., 18-wheelers, and assume that the average signal power over the ambient noise is uniformly distributed between 50-80dB (i.e., $10\log_{10}\left(\sigma_s^2/\sigma^2\right) \sim \mathcal{U}(50, 80)$).

We use $c_1 = 5$ and $c_t = 25 + 5t$ to design an acoustic sensor network using a budget of 10K units to cover a surveillance region of $\mathcal{A} = 1500\text{m}\times 1500\text{m}$ and impose that the connectivity probability p_C of the network to be greater than 0.99. In cost schedule, acoustic arrays are more expensive than range only sensors because (i) they require special deployment mechanisms and (ii) they require additional orientation calibration. It is easy to see that the convexity assumption is satisfied for $\tilde{\omega}_{2,t}$.

To determine $\tilde{\omega}_{1,t}$, we first estimated a target's position situated at the center of a disk of radius $\max R_t^*$ over 1000 Monte Carlo realizations to determine estimation variances σ_h^2 and σ_v^2 for each sensor type by varying n_t. To estimate the position variances for range only sensors, we generated acoustic data using (22) with $v = 0$, used range estimates from the sensors with positive detections according to (29), then used a Newton-Raphson recursion with the known sensor positions and estimated a target position. We used $R_1^* = 50$m as the operational range of the range only sensors. To estimate the position variances for bearing only sensors, we used

the inverse of the wideband version of the FIM[3] in (36) to generate bearing estimates for each array, a Newton-Raphson recursion with the known sensor positions and estimated a target position ($R_t^* = 300$m). We then took the inverse of $\sigma_h^2 + \sigma_v^2$ for each sensor type to estimate the f_t and γ via (4).

Figures 8(a)-(f) show some system simulations using the single sensor and pairwise sensor combinations (solid curves). In the figures, the dashed lines represent the results of our approximation in (5), where we have estimated $\gamma = 2.1842$.[4] The corresponding f_t estimates are shown in Fig. 9(a) and (b). In Fig. 9(c), we plot the performance per cost curve $\tilde{\omega}_{1,t} = f_t/c_t$ for the assumed cost structure. Unfortunately, $\tilde{\omega}_{1,t}$ does not have a convex form, however, it can be closely upper and lower bounded by straight lines as marked with stars and squares in Fig. 9(c).

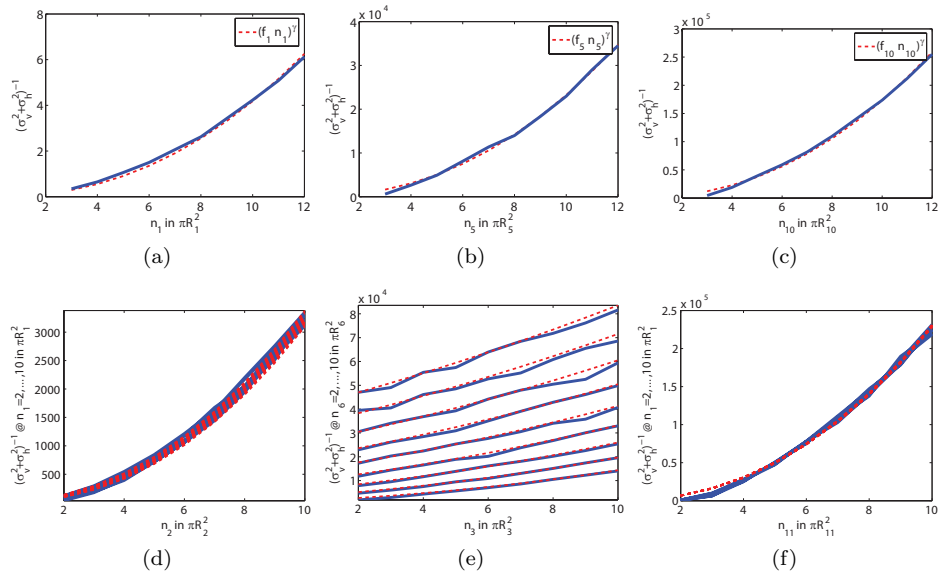


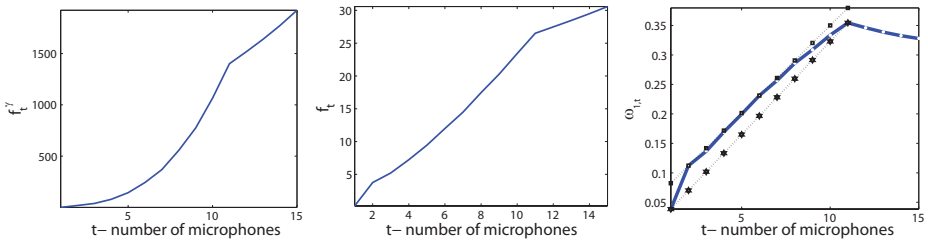
Fig. 8. System simulations (solid lines) and the solution of (5) (dashed lines). (a) s_1 alone. (b) s_5 alone. (c) s_{10} alone. (d) s_1 and s_2. (e) s_3 and s_6. (f) s_1 and s_{11}.

Figures 10(a) and (b) illustrate the evolution of the dynamic programming solution and Fig. 10(c) plots the Pareto frontier for the acoustic sensor network design along with the solution of **Problem 3** and the corresponding upper and lower bounds using (20). Figure 10(d) shows the resource allocation on the Pareto frontier. The dynamic programming solution of **Problem 3** results in $n_1 = 720$, $n_{11} = 80$ and zero for the rest of the sensor types. These numbers correspond

[3]The wideband version uses a flat spectrum assumption and is a summation of (36) at each FFT frequency determined by the sampling rate. This FIM is known to be tight for bearing errors when the number of data samples is high, which is the case here. For example, MUSIC algorithm achieves this bound, e.g., see [Stoica and Nehorai 1989].

[4]The value of the parameter γ depends on the target SNR range. When the SNR decreases, γ increases and vice versa.

to approximately 2.5-s_1 and 10.05-s_{11} sensors in areas of size $\pi {R_1^*}^2$ and $\pi {R_{11}^*}^2$, respectively. Hence, each point on $\mathcal{A}$ is well covered at least by 3 sensors. The minimum radio transmission range should be chosen as $r^*_{\text{tran}} = 100.51$m via (7) to guarantee a connectivity probability of $p_C = 0.99$ ($d = 1$). The lower bound results in $u_1 = 0.3549$ and $u_{11} = 0.6451$ via (19) corresponding to $n_1 = 709.7365$ and $n_{11} = 80.6415$. If the resources for 0.6415-s_{11} sensors are spent on sensors of s_1 type, the approximate design results in $n_1 = 720$ and $n_{11} = 80$, which incidentally coincides with the dynamic programming solution. The upper bound results in $u_1 = 0.4098$ and $u_{11} = 0.5902$ via (19) corresponding to $n_1 = 819.5500$ and $n_{11} = 73.7781$.



(a) (b) (c)

Fig. 9. (a) It is instructive to plot f^{γ}_m as it demonstrates the relative performance of the sensor networks that consist of only one sensor type s_m. There is approximately a cubic increase up to s_{11}, where the arrays hit the radius limit of 1m. After s_{11}, the increase is approximately $m^{1.4}$. The cubic increase can be explained by (i) the aperture gain due to the a^2-term in the FIMs, and (ii) SNR gain in the FIMs. At any single SNR, only the aperture gain is present. (b) Inherent performance f_m is shown. (c) Performance per cost $\tilde{\omega}_{1,t}$ for the given cost structure is shown. Convex upper and lower bounds are marked with squares and stars. The dynamic programming solution uses the 10 times the values of the dots, which are integer, that approximate κ_m.

6. CONCLUSIONS

In summary, we presented mathematical programming and approximation algorithms for the NDS problem for acoustic sensor networks using position estimation as an example. We provided approximation algorithms that can bound the results as well as methods to identify dominated sensors to alleviate computation. For a given deployment region size, our dynamic programming solution can calculate the exact integer Pareto frontier of the sensor network utility at the desired probabilities for d-connectivity and k-coverage. However, the continuous relaxations are shown to be quite useful in obtaining quick solutions to the problem.

To the Pareto problems in this paper, other linear matrix inequalities can be added without changing the solution strategies. As an example, the total weight of the sensor network might also be constrained in space exploration, where a mobile robot carries the sensors for deployment. This results in an affine constraint, which can be treated as convex or concave as required.

As future work, we will extend the formulation in this paper to hypothesis testing. Suppose that based on the state-of-nature, the binary classification problem is linear, where we choose H_0, when $\boldsymbol{b}^T\boldsymbol{\theta} > 0$, and H_1 otherwise [Duda et al. 2001].

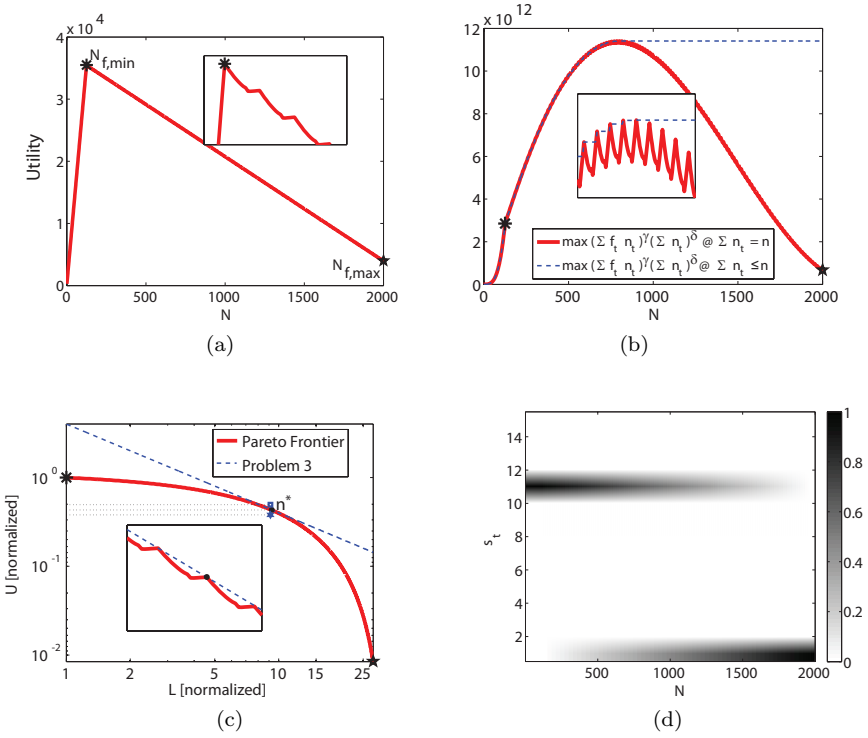


Fig. 10. (a) and (b) Dynamic programming solution of the NDS problem with the minimum cost are plotted. (c) Pareto efficient frontier is shown with different operating points. Due to the non-convexity of $\tilde{\omega}_{1,t}$, not all the point on the Pareto frontier can be determined by scalarization. The obtained upper and lower bounds (the square and the star, respectively) are quite close to the dynamic programming solution. (d) The resource distribution on the Pareto frontier is shown.

In this case, it can be shown that the Bayesian design criteria becomes the following [Toman 1996]: $U(\boldsymbol{n}) \propto -\mathrm{tr}\left\{\boldsymbol{b}\boldsymbol{b}^T \boldsymbol{\Lambda}\mathfrak{F}^{-1}(\boldsymbol{n})\right\}$. For sensor networks with acoustic and video modalities, the discriminant features can be the acoustic amplitude and the object size to classify vehicles, e.g., into compact vs. SUV categories [Cevher et al. 2007; Cevher et al. 2007]. We plan to investigate Pareto frontiers for joint parameter estimation and classification. The resulting NDS problems will then address to a broader class of sensor network problems and present further intellectual challenges for mathematical programming.

7. ACKNOWLEDGEMENTS

We would like to thank Prof. Arkadi Nemirovski for his useful comments and help on the integer programming solution. We also would like to thank Prof. Rama Chellappa for inspiring discussions on the statistical modeling of the problem.

REFERENCES

Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., and Cayirci, E. 2002. A survey on sensor networks. *IEEE Communications Magazine 40,* 8, 102–114.

Altman, E. and Miorandi, D. 2005. Coverage and connectivity of ad–hoc networks in presence of channel randomness. In *INFOCOM*. Vol. 1.

Bell, K. L., Ephraim, Y., and Van Trees, H. L. 1996. Explicit Ziv-Zakai lower bound for bearing estimation. *IEEE Transactions on Signal Processing 44,* 11, 2810–2824.

Berger, J. 1993. *Statistical Decision Theory and Bayesian Analysis*. Springer.

Bertsekas, D. P. 2003. *Nonlinear Programming*. Athena Scientific.

Bettstetter, C. June 9-11, 2002. On the minimum node degree and connectivity of a wireless multihop network. In *MOBIHOC 2002*. EPF Lausanne, Switzerland, 80–91.

Bhardwaj, M. and Chandrakasan, A. P. 2002. Bounding the lifetime of sensor networks via optimal role assignments. In *INFOCOM 2002*. Vol. 3.

Bhardwaj, M., Garnett, T., and Chandrakasan, A. P. 2001. Upper bounds on the lifetime of sensor networks. In *IEEE International Conference on Communications 2001*. Vol. 3.

Bollobás, B. 1998. *Modern Graph Theory*. Springer Verlag.

Boyd, S. P. and Vandenberghe, L. 2004. *Convex Optimization*. Cambridge University Press.

Cevher, V. 2005. A Bayesian framework for target tracking using acoustic and image measurements. Ph.D. thesis, Georgia Institute of Technology, Atlanta, GA.

Cevher, V., Chellappa, R., and McClellan, J. H. 15–20 April 2007. Joint acoustic-video fingerprinting of vehicles, part I. In *ICASSP 2007*. Honolulu, Hawaii.

Cevher, V., Guo, F., Sankaranarayanan, A. C., and Chellappa, R. 15–20 April 2007. Joint acoustic-video fingerprinting of vehicles, part II. In *ICASSP 2007*. Honolulu, Hawaii.

Cevher, V. and Kaplan, L. M. 2008. Pareto frontiers of sensor networks for localization. In *IPSN*.

Chaloner, K. and Verdinelli, I. 1995. Bayesian experimental design: A review. *Statistical Science 10,* 3, 273–304.

Chow, S. K. and Schultheiss, P. 1981. Delay estimation using narrow-band processes. *IEEE Transactions on Signal Processing 29,* 3, 478–484.

Cowan, C. K. and Kovesi, P. D. 1988. Automatic sensor placement from vision task requirements. *IEEE Transactions on Pattern Analysis and Machine Intelligence 10,* 3, 407–416.

Davenport II, W. B. and Root, W. L. 1958. *An Introduction to the Theory of Random Signals and Noise*. New York: McGraw-Hill. ch. 6, sec. 6-4, pp. 93–96.

Department-of-Transportation. United States Federal Highway Administration. http://www.fhwa.dot.gov/environment/noise/measure/chap5.htm.

Duda, R. O., Hart, P. E., and Stork, D. G. 2001. *Pattern Classification*. Wiley New York.

Epelman, M. A., Pollock, S., Netter, B., and Low, B. S. 2005. Anisogamy, expenditure of reproductive effort, and the optimality of having two sexes. *Operations Research 53,* 3, 560–567.

Evans, M., Hastings, N., and Peacock, B. 2000. *Statistical distributions. Ed*. Wiley & Sons. New York.

Feeney, L. and Nilsson, M. 2001. Investigating the energy consumption of a wireless network interface in an ad hoc networking environment. In *INFOCOM 2001*. Vol. 3.

Giridhar, A. and Kumar, P. R. 2005. Maximizing the functional lifetime of sensor networks. In *IPSN*. IEEE Press Piscataway, NJ, USA.

Heinzelman, W. B., Chandrakasan, A. P., and Balakrishnan, H. 2002. An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications 1,* 4, 660–670.

Isler, V., Kannan, S., and Daniilidis, K. 2004. Sampling based sensor-network deployment. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*.

Isler, V., Kannan, S., and Khanna, S. 2004. Randomized pursuit-evasion with limited visibility. In *The fifteenth annual ACM-SIAM Symposium on Discrete Algorithms*. Society for Industrial and Applied Mathematics Philadelphia, PA, USA, 1060–1069.

Johnson, D. H. and Dudgeon, D. E. 1993. *Array Signal Processing: Concepts and Techniques*. Prentice Hall.

Kaplan, L. M. and Cevher, V. 2007. Design considerations for a heterogeneous network of bearings-only sensors using sensor management. In *Proc. of the IEEE/AIAA Aerospace Conference*. Big Sky, MT.

Lazos, L. and Poovendran, R. 2006. Stochastic coverage in heterogeneous sensor networks. *ACM Transactions Sensor Networks 2,* 3, 325–358.

Lehmann, E. L. and Casella, G. 1998. *Theory of Point Estimation*. Springer.

Lindsey, S., Raghavendra, C., and Sivalingam, K. M. 2002. Data gathering algorithms in sensor networks using energy metrics. *IEEE Transactions on Parallel and Distributed Systems 13,* 9, 924–935.

Liu, J., Koutsoukos, X., Reich, J., and Zhao, F. Sensing field: coverage characterization in distributed sensor networks. In *ICASSP 2003*. Vol. 5. 173–176.

Liu, J., Reich, J., and Zhao, F. 2003. Collaborative in-network processing for target tracking. *EURASIP Journal on Applied Signal Processing 4*, 378–391.

Moon, T. K. and Sterling, W. C. 2000. *Mathematical Methods and Algorithms for Signal Processing*. Prentice Hall.

Morse, P. and Ingard, K. 1968. *Theoretical Acoustics*. McGraw-Hill.

Nemhauser, G. L. and Wolsey, L. A. 1988. *Integer and Combinatorial Optimization*. Wiley-Interscience New York, NY, USA.

Penrose, M. D. 1999. On k-connectivity for a geometric random graph. *Random Structures and Algorithms 15,* 2, 145–164.

Ross, S. M. 2007. *Introduction to Probability Models*. Academic Press.

Sinha, A. and Chandrakasan, A. 2001. Dynamic power management in wireless sensor networks. *IEEE Design and Test of Computers, 18,* 2, 62–74.

Stoica, P., Larsson, E. G., and Gershman, A. B. 2001. The stochastic CRB for array processing: a textbook derivation. *IEEE Signal Processing Letters 8,* 5, 148–150.

Stoica, P. and Nehorai, A. May 1989. Music, maximum likelihood, and Cramér-Rao bound. *IEEE Transactions on ASSP 37,* 5, 720–741.

Subhlok, J., Lieu, P., and Lowekamp, B. 1999. Automatic node selection for high performance applications on networks. In *the Seventh ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*. ACM Press New York, NY, USA, 163–172.

Tierney, L. and Kadane, J. B. 1986. Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association* 81, 82–86.

Toman, B. 1996. Bayesian experimental design for multiple hypothesis testing. *Journal of the American Statistical Association 91,* 433.

Van Trees, H. L. 1968. *Detection, Estimation, and Modulation Theory, Part I*. John Wiley & Sons, Inc.

Vidal, R., Shakernia, O., Kim, H. J., Shim, D. H., and Sastry, S. 2002. Probabilistic pursuit-evasion games: theory, implementation, and experimental evaluation. *IEEE Transactions on Robotics and Automation 18,* 5, 662–669.

Weiss, A. and Weinstein, E. 1983. Fundamental limitations in passive time delay estimation–Part I: Narrow-band systems. *IEEE Transactions on Signal Processing 31,* 2, 472–486.

Zhang, H. and Hou, J. 2004. On deriving the upper bound of α-lifetime for large sensor networks. In *Proceedings of the 5th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. ACM Press New York, NY, USA, 121–132.

Received 1 July 2007