Cost-Aware Bayesian Sequential Decision-Making for Search and Classification

Y. WANG, Member, IEEE
Clemson University

I. I. HUSSEIN, Member, IEEE
D. R. BROWN, III, Senior Member, IEEE
Worcester Polytechnic Institute

R. S. ERWIN, Senior Member, IEEE
Air Force Research Laboratory

A cost-aware Bayesian sequential decision-making strategy for domain search and object classification using a limited-range sensor is presented. On one hand, it is risky to allocate all available sensing resources at a single location while ignoring other regions. On the other hand, the sensor may miss-detect or miss-classify a critical object with insufficient observations. Therefore, we develop a decision-making strategy that balances the tolerable risks and the desired decision precision under limited resources.

I. INTRODUCTION

In many domain search and object classification problems, for example, aerial search and rescue/destroy, surveillance, space imaging systems, mine countermeasures, and wildfire control, the effective management of limited available sensing resources is key to mission success [1, 2]. In a search task, the objective is to find every unknown object in a mission domain and fix its position. In a classification task, the objective is to take enough measurements to determine the nature of the object.

On one hand, with limited available observations in the presence of sensor errors, a sensor may give a false alarm of object presence while there is actually none, miss detecting a critical object, or report incorrect classifications. However, taking exhaustive observations at one particular location of interest may result in losing the opportunity to find and classify possibly more critical objects at other locations within the domain. For example, in search and rescue, a sensor-equipped robot may come across a human victim and, at the cost of missing it, decide to continue to search and classify other objects. Conversely, locating and analyzing a nonhuman object may come at the cost of delaying the detection of or altogether missing a live human victim.

Therefore, it is crucial to choose whether to search and classify more unknown objects over a domain with probably lower accuracy, or to keep taking observations at a certain location until full certainty. This decision is especially critical when the mission domain is too large to be covered by statically distributed sensors with limited sensory ranges, or the number of unknown objects is far more than that of the available sensors [3–5]. To accomplish these competing tasks under limited sensory resources with minimum costs, in this paper, we develop a decision-making strategy that dynamically chooses the task to be performed based on an overall risk assessment associated with the decision.

We first review some related literature on probabilistic object search, classification, and tracking, which has been studied mainly for optimal path planning and state estimation. Note that classification of stationary objects and tracking of mobile objects both require observations associated with a single detected object. However, we focus on classification in this paper and future work will extend current research to mobile object tracking. Inspired by work on particle filtering, in [6] the authors develop a strategy to dynamically control the relative configuration of sensor teams in order to get optimal estimates for target tracking through multi-sensor fusion. In [7], [8] the authors use the Beta distribution to model the level of confidence of target existence for an unmanned aerial vehicle search task. The minimum number of observations required to achieve

a probability threshold is derived. In [9] the authors investigate search-and-tracking using recursive Bayesian filtering with foreknown targets' positions. The results are extended in [10] for dynamic search spaces, where a target might not be within a static search space at the next time step. In [11] the author proposes a Bayesian-based multisensor-multitarget sensor management scheme. The approximation strategy maximizes the expected number of targets. In [12] the target existence probability gain from searching a point is defined as a cost function used to determine the vehicle's optimal path.

It is worth noting that in the above literature there is no explicit decision-making strategy for search and classification/tracking. To remedy this, the authors developed a decision-making strategy for search versus tracking within a deterministic framework in [3] and extended the result to a probabilistic framework in [4] and [5]. In [3] the model describes how "aware" the vehicle fleet is of events over a given domain. The proposed strategy guarantees the detection of all objects of interest and the tracking of each object's state for a minimum guaranteed amount of time. In [4] and [5] a probabilistic Bayesian counterpart was developed for unknown object search versus classification treated as two competing tasks. A probabilistic framework is desirable as it takes into account sensor errors, as well as allows for future incorporation of other tasks such as object tracking, data association, data/decision fusion, sensor registration, and clutter resolution [4, 5, 13]. In the probabilistic setting, in this paper information entropy is utilized to model the uncertainty level of the search and classification tasks. Some related work can be found in the literature of statistics for optimal data selection. In [14] the variance of posterior density is used as a criterion for active data selection and a confidence measure for test point rejection. It is shown that high confidence level (i.e., small variance) is achieved when the training data are near the test data. The work in [15] provides a brief discussion of the optimum or effective sampling methods based on information entropy. The observation sample with maximum entropy is chosen. The proposed method therein may be used in either a search or a classification mission (treated as independent problems). However, note that the key aspect of the problem tackled in this paper is the treatment of search and tracking as competing tasks.

Sequential detection [16] allows the number of observations to vary in order to achieve an optimal decision. The Bayesian sequential detection method used in this paper is such that the Bayes risk (to be formally defined in Section IV) is minimized at each time step [17]. Two types of costs are taken into account in the risk calculation: 1) the cost of making a wrong decision, i.e., the probability of missed/false detection, or incorrect classification,

and 2) the cost of taking more observations for a possibly better decision. The observation cost is computed in real time based on the progress of the task. Due to the randomness of observations and the dynamic observation cost, a decision may be made with a few observation samples, whereas for other cases one would rather take more samples to minimize risk. In [18] a sequential Bayes classifier is utilized for the real-time classification of detected targets under a neural network based framework, however, without consideration of observation costs. Another sequential detection method is the sequential probability ratio test (SPRT) [16, 19] based on binary Neyman-Pearson formulation where no prior probability information is required. On average a smaller number of observations are needed to make a decision in SPRT compared with an equally reliable method with a predetermined fixed number of observations [20]. The change-point detection theory [21, 22] is a generalization and modification of SPRT. It detects a change in the probability distribution of a stochastic process or time series. Existing techniques include the Shyriaev-Roberts (SR) [21, 23] and the cumulative sum control chart (CUSUM, a.k.a. Page test) [24] tests. We focus on Bayes risk analysis with prior information in this paper and will investigate other solution approaches in future work.

In the literature, sequential decision-making via tradeoffs between exploration and exploitation has been investigated in a risk-neutral context. The work in [25], [26] and references therein provide an overview of techniques that trade off between expected information gain (or equivalently, rewards) and the cost incurred by applying a control action for partially observable Markov decision process (POMDP). The planning problem is addressed under no constraints of decision error, and is hence, risk neutral.

In this work we focus on a cost-aware Bayesian sequential decision-making strategy for search and classification in the presence of sensor errors. The contribution of this paper is the real-time decision-making for object existence and its classification with minimum Bayes risks under limited sensory resources. It is shown that the real-time data collection and optimal decision-making strategy is robust to sensor errors. A key novelty in this paper relative to previous work lies in the treatment of search/classification and taking observations as tasks competing for the same limited resources while seeking to minimize the combined risk of decision errors and observation costs.

Figure 1 shows the block diagram of the proposed strategy and the organization of the paper. At time t, the sensor takes an observation at a cell $\tilde{\mathbf{c}}_j$ in the search domain based on a ternary sensor model proposed in Section II. Next, the posterior probability of object existence or its classification at $\tilde{\mathbf{c}}_j$ gets
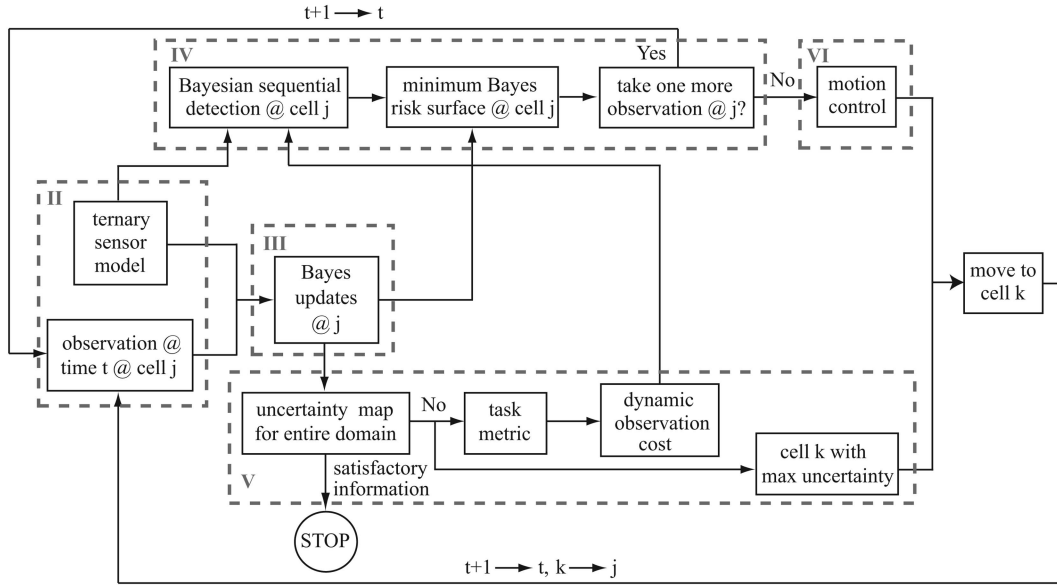
Fig. 1. Block diagram of cost-aware Bayesian sequential decision-making strategy.

updated via the Bayes update equations formulated in Section III. In Section IV we introduce the Bayesian sequential detection method for a single cell $\tilde{\mathbf{c}}_j$, which depends on the sensor model as well as the dynamic observation cost. Its output is the minimum Bayes risk surface at cell $\tilde{\mathbf{c}}_j$. Combined with the updated probabilities, the sensor makes a decision (whether or not to take one more observation at $\tilde{\mathbf{c}}_j$) that minimizes the Bayes risk at time t. In Section V an uncertainty map is constructed based on the updated probabilities of every cell within the domain. If the desired certainty level has not been achieved yet, a task metric is developed to formulate the dynamic observation cost. Finally, we combine the results from Sections IV and V: if the decision is to stop taking observation at the current cell $\tilde{\mathbf{c}}_j$, a sensor motion control scheme is provided in Section VI, which drives the sensor to the cell $\tilde{\mathbf{c}}_k$ that has the maximum uncertainty in the domain. This process is repeated over time until the uncertainty level over the entire domain is satisfactorily low. To illustrate the performance of the proposed strategy, we provide a simulation-based study in Section VII. The paper is concluded with a summary of current and future work in Section VIII.

II. SETUP AND SENSOR MODEL

We introduce the problem formulation in this section. For the sake of clarity, Table I summarizes a list of key variables used in this paper.

A. Problem Setup

Let $\mathcal{D} \subset \mathbb{R}^2$ be the domain in which objects to be found and classified are located. We discretize the domain into N_{tot} cells. Let $\tilde{\mathbf{c}}$ be an arbitrary cell in $\mathcal{D}$ (we omit the subscript j with the understanding that

TABLE I
List of Key Variables

$\mathcal{D}$	Mission domain
$\tilde{\mathbf{c}}$	A cell in $\mathcal{D}$
$\tilde{\mathbf{q}}$	Centroid of $\tilde{\mathbf{c}}$
P_p	Initial probability of object presence at $\tilde{\mathbf{c}}$
$X(\tilde{\mathbf{c}})$	Ternary state variable at $\tilde{\mathbf{c}}$
$Y_t(\tilde{\mathbf{c}})$	Ternary observation variable at $\tilde{\mathbf{c}}$ at t
β_{ij}	Probability of observing $Y(\tilde{\mathbf{c}}) = j$ given state $X(\tilde{\mathbf{c}}) = i$
L	Observation length
$Z_j(\tilde{\mathbf{c}})$	Number of times observing $Y(\tilde{\mathbf{c}}) = j$ during L observations
C_{ij}	Cost of deciding i given state $X(\tilde{\mathbf{c}}) = j$
c_j	$[C_{0j} \ C_{1j} \ C_{2j}]$
Δ	Deterministic decision rule
t	Time index
B	General conditional probability matrix
b_j	jth column of B
c_{obs}	Observation cost
$\tilde{R}_j$	Conditional Bayes risk without c_{obs}
R_j	Conditional Bayes risk with c_{obs}
$E_j[N(\phi)]$	Expected stopped time given state $X(\tilde{\mathbf{c}}) = j$
π_j	Prior probability of state being $X(\tilde{\mathbf{c}}) = j$
r	Bayes risk
$r^*_{\min}$	Minimum Bayes risk
H	Uncertainty function
γ	Observation cost weighting parameter
$\mathcal{Q}_d(t)$	Set of cells with highest H and shortest distance to the current cell
ϵ	Desired uncertainty level

$\tilde{\mathbf{c}}$ represents a single unique cell in $\mathcal{D}$) and $\tilde{\mathbf{q}}$ is the centroid of cell $\tilde{\mathbf{c}}$. Define $1 \leq N_{\text{o}} \leq N_{\text{tot}}$ as the total number of objects in $\mathcal{D}$ and the position of the static object $\mathcal{O}_j$, $j \in \{1, 2, \dots, N_{\text{o}}\}$, as $\mathbf{p}_j$. Both N_{o} and $\mathbf{p}_j$ are unknown beforehand. Assume that the objects

are independent and identically distributed (IID) over $\mathcal{D}$, and the partitioning of the domain is fine enough so that at most one object can exist in a cell. An object can be assigned as many property types as needed, but without loss of generality, we assume that an object can have one of two properties, either property "F" or property "G". Let N_o^F (N_o^G) be the (unknown) number of objects with property F (G), with $N_o^F + N_o^G = N_o$. Let $\mathcal{S}_F$ ($\mathcal{S}_G$) be the set containing those cells and P_F (P_G) be the corresponding initial probability mass function. Hence, N_o is a binomial random variable with parameters N_{tot} and $P_p = P_F + P_G$, where P_p gives the initial total probability of object present at a cell $\tilde{\mathbf{c}}$ and is identical for all $\tilde{\mathbf{c}} \in \mathcal{D}$ and independent. Therefore, the probability of k cells in the domain containing an object is given by

$$\Pr(N_o = k) = \binom{N_{tot}}{k} P_p^k (1 - P_p)^{N_{tot} - k}$$

where $k = 1, 2, \ldots, N_{tot}$ and the expectation of N_o is equal to the number of total cells multiplied by P_p, i.e.,

$$E[N_o] = N_{tot} P_p. \tag{1}$$

Similarly, the probability of k cells within the domain containing an object with property F (G) follows a binomial distribution with parameters N_{tot} and P_F (P_G).

Let $X(\tilde{\mathbf{c}})$ be a ternary state random variable at cell $\tilde{\mathbf{c}}$, where 0 corresponds to object absent, 1 corresponds to object having property F, and 2 corresponds to object having property G. The realization of $X(\tilde{\mathbf{c}})$ depends on the cell $\tilde{\mathbf{c}}$ being observed, i.e.,

$$X(\tilde{\mathbf{c}}) = \begin{cases} 1 & \text{if} \quad \tilde{\mathbf{c}} \in \mathcal{S}_F \\ 2 & \text{if} \quad \tilde{\mathbf{c}} \in \mathcal{S}_G \\ 0 & \text{otherwise} \end{cases}.$$

Since both $\mathcal{S}_F$ and $\mathcal{S}_G$ are unknown and random, $X(\tilde{\mathbf{c}})$ is a random variable with respect to every $\tilde{\mathbf{c}} \in \mathcal{D}$. Because we assume that the objects are immobile, $X(\tilde{\mathbf{c}})$ is invariant with respect to time.

B. Sensor Model

For the sake of illustrative clarity, we make the following assumptions for the sensor model.

1) A sensor is able to observe only one cell at a time. Extension to other sensor models that are capable of observing multiple cells at the same time (e.g., the sensor models with limited sensory range proposed in [3]–[5], [27]–[33]) is straightforward.

2) A sensor is able to move to any cell within the domain. Other motion schemes, such as gradient-based, awareness-based, and information-driven control laws ([3–5, 27–33]), can be adopted without difficulty.

3) We consider the extreme case in which the resources available are at a minimum (a single sensor as opposed to multiple cooperating ones).
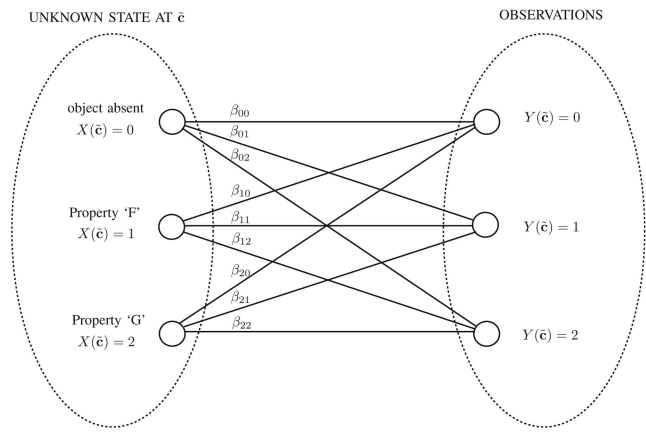

Fig. 2. Model showing how unknown state at particular cell $\tilde{\mathbf{c}}$ and observations obtained by sensor are related.

Let $Y(\tilde{\mathbf{c}})$ be a ternary observation random variable, where 0 corresponds to an observation indicating object absent at cell $\tilde{\mathbf{c}}$, and 1 (2) corresponds to an observation indicating that there exists an object having property F (G).

The sensor model follows a ternary discrete probability distribution. For a cell $\tilde{\mathbf{c}}$, given a state $X(\tilde{\mathbf{c}}) = i$, $i = 0, 1, 2$, the probability mass function f of the observation distribution is given by

$$f_Y(y \mid X(\tilde{\mathbf{c}}) = i) = \begin{cases} \beta_{i0} & \text{if} \quad y = 0 \\ \beta_{i1} & \text{if} \quad y = 1 \\ \beta_{i2} & \text{if} \quad y = 2 \end{cases} \tag{2}$$

where $\sum_{j=0}^{2} \beta_{ij} = 1$, Y corresponds to the ternary random variable and y is the dummy variable. Because the states $X(\tilde{\mathbf{c}})$ and the observations $Y(\tilde{\mathbf{c}})$ are spatially IID, the probability distribution for every $\tilde{\mathbf{c}} \in \mathcal{D}$ follows the same structure. Figure 2 shows the relationship between the unknown state $X(\tilde{\mathbf{c}})$ and an observation $Y(\tilde{\mathbf{c}})$. This is a simplified but reasonable sensor model because it abstracts away the complexities in sensor noise, image processing algorithm errors, and lays a foundation for decision fusion in multiple sensor systems [34, 35].

Conditioned on the true state $X(\tilde{\mathbf{c}})$, the observations $Y_t(\tilde{\mathbf{c}})$ taken along time are temporally IID Define an integer random variable $Z_j(\tilde{\mathbf{c}})$, $j = 0, 1, 2$ as the number of times that observation $Y(\tilde{\mathbf{c}}) = j$ appears during a window of L time steps. The quantity $Z_j(\tilde{\mathbf{c}})$ satisfies $\sum_{j=0}^{2} Z_j(\tilde{\mathbf{c}}) = L$, $Z_j(\tilde{\mathbf{c}}) \in [0, L]$. Therefore, given state $X(\tilde{\mathbf{c}}) = i$, $i = 0, 1, 2$, the probability of having observation (z_0, z_1, z_2) in a window of L time steps follows a multinomial distribution

$$\Pr(Z_0(\tilde{\mathbf{c}}) = z_0, Z_1(\tilde{\mathbf{c}}) = z_1, Z_2(\tilde{\mathbf{c}}) = z_2 \mid X(\tilde{\mathbf{c}}) = i)$$

$$= \frac{L!}{z_0! z_1! z_2!} \beta_{i0}^{z_0} \beta_{i1}^{z_1} \beta_{i2}^{z_2}, \qquad \sum_{j=0}^{2} z_j = L. \tag{3}$$

The sensor's probabilities of making a correct observation, i.e., the detection probabilities, are β_{00},

β_{11}, and β_{22}. Here we assume that the sensor is "good" and restrict these values to be $\beta_{00}, \beta_{11}, \beta_{22} > 0.5$. This implies that the sensor is able to better distinguish the true state from the other two states and returns a higher likely observation of the true state at that location. More general values within $[0, 1]$ can be considered, however, introducing extra analytical complexity does not contribute any new insights.

REMARK

1) If we have more than two possible classification properties, one can extend the current ternary sensor model to a more general M-ary model without difficulty.

2) A key feature of the proposed approach is that the sensor may have a limited range. Existing literature on cooperative coverage control for the sensor redeployment problem usually assumes infinite sensory range [36–38]. This assumption is not made here. This is very important in applications where $\mathcal{D}$ is large scale (i.e., too large to be optimally or suboptimally covered by a single set of static sensors). The application of limited-range sensors is consistent with previous work for dynamic coverage control of multi-sensor network with flocking and guaranteed collision avoidance [27–29, 31], awareness-based coverage control and decision-making for search versus tracking using multiple autonomous vehicles with intermittent communications [3, 33], Bayesian-based binary decision-making for search versus characterization/classfication [4, 5], and underwater effective coverage [30, 32].

III. BAYESIAN UPDATES FOR SEARCH AND CLASSIFICATION

Based on the ternary sensor model, in this section, we employ Bayes' rule to update the probability of object existence and its classification at a single cell $\tilde{\mathbf{c}}$. Under the IID assumption, the Bayesian updates equations developed in this section are deployed to multi-cell domain in Sections V, VI, and VII.

According to Bayes' rule, given a single observation $Y_t(\tilde{\mathbf{c}}) = j$ taken at cell $\tilde{\mathbf{c}}$ at time step t, we have

$$P(X(\tilde{\mathbf{c}}) = i \mid Y_t(\tilde{\mathbf{c}}) = j; t+1)$$
$$= \alpha_j P(Y_t(\tilde{\mathbf{c}}) = j \mid X(\tilde{\mathbf{c}}) = i) P(X(\tilde{\mathbf{c}}) = i; t),$$
$$i, j = 0, 1, 2 \qquad (4)$$

where $P(X(\tilde{\mathbf{c}}) = i \mid Y_t(\tilde{\mathbf{c}}) = j; t+1)$ is the posterior probability that the true state is $X(\tilde{\mathbf{c}}) = i$ at time step $t+1$, $P(Y_t(\tilde{\mathbf{c}}) = j \mid X(\tilde{\mathbf{c}}) = i)$ is the probability of the particular observation $Y_t(\tilde{\mathbf{c}}) = j$ being taken given $X(\tilde{\mathbf{c}}) = i$, which is determined by the ternary sensor model (2). $P(X(\tilde{\mathbf{c}}) = i; t)$ is the prior probability of being state type $X(\tilde{\mathbf{c}}) = i$ at t, and α_j serves as a normalizing function that ensures $\sum_{i=0}^{2} P(X(\tilde{\mathbf{c}}) = i \mid Y_t(\tilde{\mathbf{c}}) = j; t+1) = 1$. According to the law of total probability, α_j is given as follows,

$$\alpha_j = \frac{1}{P(Y_t(\tilde{\mathbf{c}}) = j)}$$
$$= \frac{1}{\beta_{0j} P(X(\tilde{\mathbf{c}}) = 0; t) + \beta_{1j} P(X(\tilde{\mathbf{c}}) = 1; t) + \beta_{2j} P(X(\tilde{\mathbf{c}}) = 2; t)}$$

and thus we can get the posterior probabilities by substituting the value of α_j into (4).

Remark about Extension to Multiple Sensors: When we have multiple sensors, each sensor will give its own observation for a certain cell $\tilde{\mathbf{c}}$. Hence, there are 3^m different combinations of unordered observations at each time step if we have m sensors in all. Under the IID assumption, a general update equation for the posterior probabilities can be obtained by using the products of the sensing capability functions β_{ij} of each sensor. Current research focuses on the development of decentralized decision criteria and sensor data/decision fusion for the search and tracking of multiple mobile objects using a team of mobile sensors.

IV. COST-AWARE BAYESIAN SEQUENTIAL DECISION-MAKING

In this section we use a ternary cost-aware Bayesian sequential decision-making strategy to determine the state at a cell $\tilde{\mathbf{c}}$ with minimum Bayes risk. It is extended from the standard binary Bayesian sequential detection method [16, 20, 39] in signal detection theory [16, 40–42]. Instead of deriving an optimal detector given a fixed number of observations as in classical Bayesian, minimax, or Neyman-Pearson hypothesis testing methods [16, 40–42], the Bayesian sequential detector takes observations until a decision can be made with lowest Bayes risk. This results in a random number of observations.

A. Conditional Bayes Risk without Observation Cost

Bayes risk is the expected value of the sum of two types of costs: decision cost and observation cost. Here, we first investigate the conditional Bayes risk without taken into account the observation cost, that is, the cost of making a wrong decision.

Assuming a uniform cost assignment (UCA) [16], we define the decision cost components as

$$C_{ij} = \begin{cases} 0 & \text{if } i = j \\ 1 & \text{if } i \neq j \end{cases}$$

where $i = 0, 1, 2$ represents the decision: "deciding that the state at $\tilde{\mathbf{c}}$ is $X(\tilde{\mathbf{c}}) = i$" and j corresponds to the true state $X(\tilde{\mathbf{c}}) = j$. Hence, C_{ij} is the cost of deciding i when the state is j. Note that the cost generally depends on the cell $\tilde{\mathbf{c}}$ being observed. For the sake of simplicity, we assume that the cost is IID over the domain and omit the argument $\tilde{\mathbf{c}}$ throughout the paper.

Let $\tilde{R}_j(\tilde{\mathbf{c}}, L, \Delta)$, $j = 0, 1, 2$, $L \geq 1$, be the conditional Bayes risk of deciding $X(\tilde{\mathbf{c}}) \neq j$ given that the true

state is $X(\tilde{\mathbf{c}}) = j$,

$$\tilde{R}_j(\tilde{\mathbf{c}}, L, \Delta) = c_j \Delta b_j \qquad (5)$$

where

1) $c_j = [C_{0j} \ C_{1j} \ C_{2j}]$.

2) $\Delta = [\Delta(i,n)]$ is the deterministic decision rule: Let N be the total number of possible observation combinations (z_0, z_1, z_2) that the sensor can take according to the multinomial distribution (3) over a window of L time steps. The matrix element $\Delta(i,n)$, $i = 0, 1, 2$, $n = 0, \ldots, N-1$ can be either 0 or 1, and $\sum_{i=0}^{1} \Delta(i,n) = 1$. When $\Delta(i,n) = 1$, it means that decision i is made given that the observation combination corresponds to the nth column of Δ. For $L \geq 1$, the dimension of Δ is $3 \times N$ because there are three possible realizations of the states. For $L = 0$, i.e., there are no observations taken, Δ could be "always decide there is no object," "always decide there is an object with property F" or "always decide there is an object with property G," regardless of the observations, and there will be no explicit matrix form.

3) b_j is the jth column of the general conditional probability matrix $B = [B_{ij}]$ for $L \geq 1$. The element B_{ij} gives the probability of having observation combination i given state j. According to the probability axiom, $\sum_{i=0}^{N-1} B_{ij} = 1$, $j = 0, 1, 2$. For $L \geq 1$, B is an $N \times 3$ matrix.

Therefore, under UCA, there is no cost if the decision is the true state, and the conditional Bayes risk $\tilde{R}_0, \tilde{R}_1, \tilde{R}_2$ can be interpreted as the error probability of making a wrong decision under a certain decision rule Δ over L observations at cell $\tilde{\mathbf{c}}$.

Remark about "Reasonable" Deterministic Decision Rules: As mentioned in Section II-B, in this paper, we assume that the sensor is a "good" one. That is to say, the detection probability is higher than the error probability of the sensor, i.e., $\beta_{ii} > \frac{1}{2}$. Under this assumption, there are only a small number of reasonable deterministic decision rules that need to be considered. Given L observations, the set of reasonable deterministic decision rules is the set of all rules of the type

$$\Delta(i,n) = \begin{cases} 1 & z_i = \max(z_0, z_1, z_2) \\ 0 & \text{otherwise} \end{cases}$$

where $i = 0, 1, 2$, $n = 0, \ldots, N-1$. This means, take $L = 2$ for instance, we only need to consider decision rule matrices such as

$$\Delta = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{but not}$$

$$\Delta = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

where the columns correspond to observation combinations $(z_0 = 2, z_1 = z_2 = 0)$, $(z_1 = 2, z_0 = z_2 = 0)$, $(z_2 = 2, z_0 = z_1 = 0)$, $(z_0 = z_1 = 1, z_2 = 0)$, $(z_0 = z_2 = 1, z_1 = 0)$, and $(z_0 = 0, z_1 = z_2 = 1)$ respectively. Note that reasonable decision rules grow with L and dominates any other type of decision rules with the same value of N.

B. Conditional Bayes Risk with Observation Cost

Now let us assign an observation cost c_{obs} each time the sensor makes a new observation. This cost could be based on energy, amount of observation time, etc. Here, we assume it is a constant. In Section V, a dynamic cost function $c_{\text{obs}}(t)$ is developed to relate the observation cost with the task metrics for real-time decision-making in multi-cell domains.

We first define $\phi = \{\phi_k\}_{k=0}^{\infty}$ as the stopping rule and $\delta = \{\delta_k\}_{k=0}^{\infty}$ as the intermediate decision rule. If $\phi_k = 0$, the sensor takes another measurement, if $\phi_k = 1$, the sensor stops taking further observations. At every time step k, δ_k can be either one of four intermediate decisions: 1) deciding object absent, 2) deciding object having property F, 3) deciding object having property G, or 4) taking one more observation and postpone making a decision to the following time step. Let the stopping time be the minimum amount of time it takes to make a final decision, i.e., $N(\phi) = \min\{k : \phi_k = 1\}$, which is a random variable due to the randomness of the observations. The expected stopping time under state $X(\tilde{\mathbf{c}}) = j$ is then given by $E_j[N(\phi)] = E[N(\phi) \mid X(\tilde{\mathbf{c}}) = j]$.

Since now we assign a cost c_{obs} for each observation, the conditional Bayes risk (5) under UCA over $L \geq 0$ observations can be modified as

$$R_j(\tilde{\mathbf{c}}, L, \Delta) = \Pr(\text{decide } X(\tilde{\mathbf{c}}) \neq j \mid X(\tilde{\mathbf{c}}) = j) + c_{\text{obs}} E_j[N(\phi)],$$
$$j = 0, 1, 2. \quad (6)$$

If $L \geq 1$, Δ has explicit matrix form and we can further rewrite the above equations as

$$R_j(\tilde{\mathbf{c}}, L, \Delta) = c_j \Delta b_j + c_{\text{obs}} E_j[N(\phi)], \qquad j = 0, 1, 2. \quad (7)$$

C. Bayes Risk

We now define the Bayes risk as the expected conditional Bayes risk under decision rule Δ over L observations at cell $\tilde{\mathbf{c}}$:

$$r(\tilde{\mathbf{c}}, L, \pi_1, \pi_2, \Delta) = (1 - \pi_1 - \pi_2) R_0(\tilde{\mathbf{c}}, L, \Delta) + \pi_1 R_1(\tilde{\mathbf{c}}, L, \Delta) + \pi_2 R_2(\tilde{\mathbf{c}}, L, \Delta),$$
$$L \geq 0 \quad (8)$$

where $\pi_j = P(X(\tilde{\mathbf{c}}) = j; t = t_\nu)$, $j = 0, 1, 2$ is the prior probability of state being $X(\tilde{\mathbf{c}}) = j$ at time instant t_ν when an observation is taken at cell $\tilde{\mathbf{c}}$. At each cell $\tilde{\mathbf{c}}$ at every time step t, given a fixed pair (π_1, π_2)
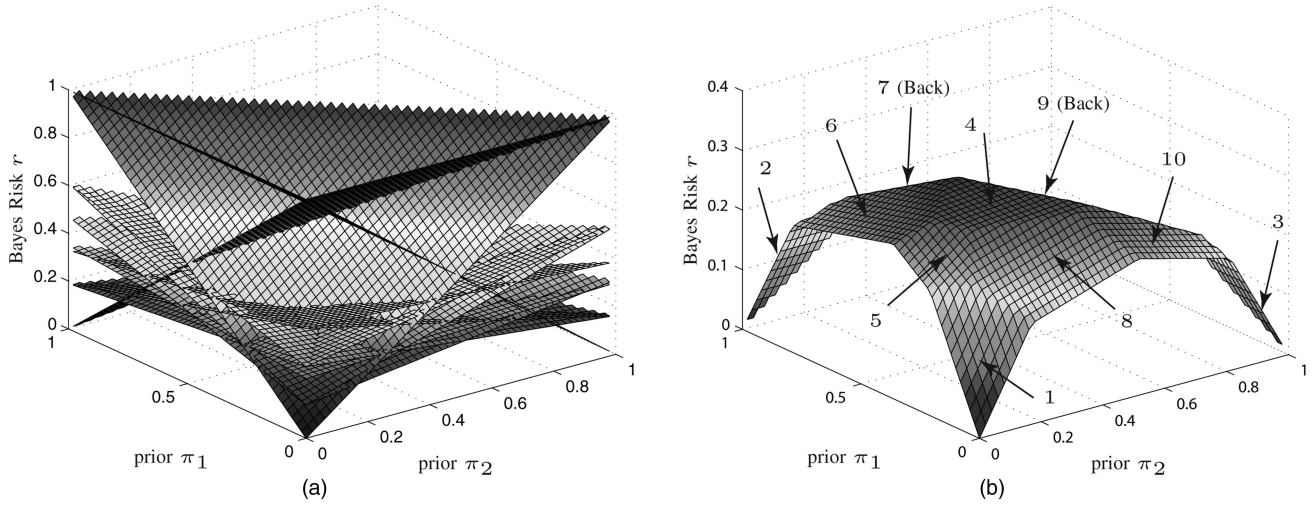
Fig. 3. (a) Bayes risk functions under 0, 1, 2 observations. (b) Minimum Bayes risk function $r^*_{\min}$. Overall minimum risk surface is composed of several enumerated risk planes, each of which is described in this section.

under the constraints $\pi_i \in [0,1]$ and $\sum_{i=1}^{3} \pi_i = 1$, the sensor chooses a combination of $(L \geq 0, \Delta)$ that yields the minimum value of the Bayes risk r. This same procedure is repeated until the cost of making a wrong decision based on the current observation is less than that of taking one more observation for a possibly better decision.

D. Bayesian Sequential Detection

We now elaborate on the decision-making procedure. If the sensor does not take any observations $(L = 0)$ and directly make a decision, according to (6) and (8), the Bayes risks of 3 different decision rules Δ are as follows

$$r(\tilde{\mathbf{c}}, L = 0, \pi_1, \pi_2, \Delta = \text{always decide object absent})$$
$$= \pi_1 + \pi_2$$

$$r(\tilde{\mathbf{c}}, L = 0, \pi_1, \pi_2, \Delta = \text{always decide object having}$$
$$\text{property F}) = 1 - \pi_1$$

$$r(\tilde{\mathbf{c}}, L = 0, \pi_1, \pi_2, \Delta = \text{always decide object having}$$
$$\text{property G}) = 1 - \pi_2.$$

If the sensor decides to take an observation $(L \geq 1)$, the minimum Bayes risk over all possible choices of Δ with L observations is

$$r_{\min}(\tilde{\mathbf{c}}, L \geq 1, \pi_1, \pi_2) = \min_{\Delta \in \mathcal{G}_L}(1 - \pi_1 - \pi_2)R_0(\tilde{\mathbf{c}}, L \geq 1, \Delta)$$
$$+ \pi_1 R_1(\tilde{\mathbf{c}}, L \geq 1, \Delta) + \pi_2 R_2(\tilde{\mathbf{c}}, L \geq 1, \Delta)$$
$$\geq L c_{\text{obs}}$$

where $\mathcal{G}_L$ is defined as the set of all deterministic decision rules that are based on exactly L observations.

Following similar procedure, we compute the minimum Bayes risk functions $r^*_{\min}$ under different observation numbers $(L \geq 0)$ and find the overall minimum Bayes risk over all possible combinations of (Δ, L),

$$r^*_{\min}(\tilde{\mathbf{c}}, \pi_1, \pi_2) = \min_{L=0,1,2,\dots} r_{\min}(\tilde{\mathbf{c}}, L, \pi_1, \pi_2).$$

The basic procedure of Bayesian sequential detection is summarized as follows. With initial priors $\pi_j = P(X(\tilde{\mathbf{c}}) = j; t = 0)$, check the corresponding $r^*_{\min}$ value. If $r^*_{\min}$ is given by the risk function with $L \geq 1$, the sensor takes an observation $Y_{t=0}(\tilde{\mathbf{c}})$. Compute the posteriors $P(X(\tilde{\mathbf{c}}) = j \mid Y_{t=0}(\tilde{\mathbf{c}}); t = 1)$ according to (4) and again check $r^*_{\min}$ to make decisions. The process is repeated using these posteriors as the new priors. The key is that an observation is taken if and only if $r_{\min}(\tilde{\mathbf{c}}, L \geq 1, \pi_1, \pi_2) < \min(\pi_1 + \pi_2, 1 - \pi_1, 1 - \pi_2)$. When $r^*_{\min} = r_{\min}(\tilde{\mathbf{c}}, L = 0, \pi_1, \pi_2)$, the sensor stops taking observations and a decision is made at $\tilde{\mathbf{c}}$.

E. Simulation for a Single Cell

Let us illustrate the proposed scheme via the following preliminary simulation for a single cell. We fix a cell $\tilde{\mathbf{c}}$ and assume that the sensor is located at the centroid of this cell. The sensing parameters are chosen as follows:

$$\beta_{00} = 0.8, \qquad \beta_{01} = 0.1, \qquad \beta_{02} = 0.1$$
$$\beta_{10} = 0.2, \qquad \beta_{11} = 0.7, \qquad \beta_{12} = 0.1 \qquad (9)$$
$$\beta_{20} = 0.1, \qquad \beta_{21} = 0.15, \qquad \beta_{22} = 0.75.$$

The observation cost is set as $c_{\text{obs}} = 0.05$. Figure 3(a) shows all the Bayes risk functions r under $L = 0, 1$ or 2 observations under the constraints $\pi_i \in [0,1]$, and $\sum_{i=1}^{2} \pi_i \leq 1$. Figure 3(b) shows the overall minimum Bayes risk surface $r^*_{\min}(\tilde{\mathbf{c}}, \pi_1, \pi_2)$, which is the minimum value of all $r(\tilde{\mathbf{c}}, L, \pi_1, \pi_2, \Delta)$, $L \geq 0$, under each fixed prior probability pair (π_1, π_2). Here, we

only provide a brief derivation of the risk planes that constitute $r_{\min}^*$. The Bayes risk functions under more than 3 observations ($L \geq 3$) have larger r values and do not contribute to $r_{\min}^*(\tilde{\mathbf{c}}, \pi_1, \pi_2)$ for the particular choice of β and c_{obs} here.

Let us interpret each of these risk planes in Fig. 3(b) annotated by the numerals 1–10.

Risk Plane 1. $r(\tilde{\mathbf{c}}, L = 0, \pi_1, \pi_2, \Delta = \text{always decide there is no object}) = \pi_1 + \pi_2$.

Risk Plane 2. $r(\tilde{\mathbf{c}}, L = 0, \pi_1, \pi_2, \Delta = \text{always decide object present with property F}) = 1 - \pi_1$.

Risk Plane 3. $r(\tilde{\mathbf{c}}, L = 0, \pi_1, \pi_2, \Delta = \text{always decide object present with property G}) = 1 - \pi_2$.

Risk Plane 4.   This plane corresponds to the decision rule after taking one observation. The general conditional probability matrix for $L = 1$ is given as

$$B(L = 1) = \begin{bmatrix} \beta_{00} & \beta_{10} & \beta_{20} \\ \beta_{01} & \beta_{11} & \beta_{21} \\ \beta_{02} & \beta_{12} & \beta_{22} \end{bmatrix}$$

where the rows correspond to the observations $(z_0 = 1, z_1 = 0, z_2 = 0)$, $(z_0 = 0, z_1 = 1, z_2 = 0)$, and $(z_0 = 0, z_1 = 0, z_2 = 1)$, respectively. Risk plane 4 corresponds to the following decision rule,

$$\Delta_{11} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

That is, decide the state according to the only one observation taken. This is the only reasonable decision rule for $L = 1$. Therefore, according to (7), we have $R_0(\tilde{\mathbf{c}}, L = 1, \Delta = \Delta_{11}) = \beta_{01} + \beta_{02} + c_{obs}$, $R_1(\tilde{\mathbf{c}}, L = 1, \Delta = \Delta_{11}) = \beta_{10} + \beta_{12} + c_{obs}$, and $R_2(\tilde{\mathbf{c}}, L = 1, \Delta = \Delta_{11}) = \beta_{20} + \beta_{21} + c_{obs}$. Hence, $r(\tilde{\mathbf{c}}, L = 1, \pi_1, \pi_2, \Delta = \Delta_{11})$ is given directly by (8).

Risk Plane 5–10.   These planes give the decision rules after two observations. The general conditional probability matrix for $L = 2$ is given as

$$B(L = 2) = \begin{bmatrix} \beta_{00}^2 & \beta_{10}^2 & \beta_{20}^2 \\ \beta_{01}^2 & \beta_{11}^2 & \beta_{21}^2 \\ \beta_{02}^2 & \beta_{12}^2 & \beta_{22}^2 \\ 2\beta_{00}\beta_{01} & 2\beta_{10}\beta_{11} & 2\beta_{20}\beta_{21} \\ 2\beta_{00}\beta_{02} & 2\beta_{10}\beta_{12} & 2\beta_{20}\beta_{22} \\ 2\beta_{01}\beta_{02} & 2\beta_{11}\beta_{12} & 2\beta_{21}\beta_{22} \end{bmatrix}$$

where the rows correspond to the observations $(z_0 = 2, z_1 = 0, z_2 = 0)$, $(z_0 = 0, z_1 = 2, z_2 = 0)$, $(z_0 = 0, z_1 = 0, z_2 = 2)$, $(z_0 = 1, z_1 = 1, z_2 = 0)$, $(z_0 = 1, z_1 = 0, z_2 = 1)$, and $(z_0 = 0, z_1 = 1, z_2 = 1)$, respectively. The

corresponding decision rules are,

$$\Delta_{21} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{(risk plane 5)}$$

$$\Delta_{22} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{(risk plane 6)}$$

$$\Delta_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \quad \text{(risk plane 7)}$$

$$\Delta_{24} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \quad \text{(risk plane 8)}$$

$$\Delta_{25} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \quad \text{(risk plane 9)}$$

$$\Delta_{26} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \quad \text{(risk plane 10)}.$$

The Bayes risks follow according to (7) and (8).

When $r_{\min}^*$ is given by risk plane 1, 2, or 3, the sensor stops taking observation and makes the final decision, otherwise, it always takes one more observation. Note that according to the remark about reasonable decision rules, there should be $2^3 = 8$ reasonable decision rules for $L = 2$, but there is one decision rule with higher Bayes risks parallel to risk plane 4 ($L = 1$), and another one overlapping with risk plane 10.

We now give a simple example of how to utilize the minimum Bayes risk surface $r_{\min}^*$ for decision-making. At a cell $\tilde{\mathbf{c}}$, assume we start with the priors $P(X(\tilde{\mathbf{c}}) = 0; t = 0) = 0.7$, $P(X(\tilde{\mathbf{c}}) = 1; t = 0) = 0.1$, and $P(X(\tilde{\mathbf{c}}) = 2; t = 0) = 0.2$. The corresponding minimum Bayes risk for the prior pair $(0.1, 0.2)$ is given by risk plane 8 ($L = 2$). So the sensor takes an observation, and if the observation is $Y_{t=0}(\tilde{\mathbf{c}}) = 0$, the posterior probabilities are updated according to (4). The posteriors are $P(X(\tilde{\mathbf{c}}) = 0; t = 1) = \frac{14}{15}$, $P(X(\tilde{\mathbf{c}}) = 1; t = 1) = \frac{1}{30}$, and $P(X(\tilde{\mathbf{c}}) = 2; t = 1) = \frac{1}{30}$. Now $r_{\min}^*$ is given by risk plane 1 ($L = 0$). Therefore, the sensor decides not to take any more observations and decides that there is no object at $\tilde{\mathbf{c}}$ with an associated Bayes risk of $r = r_{\min}^* = \frac{1}{15}$.

V.   THE UNCERTAINTY MAP AND TASK METRIC

We have discussed the mechanics of the Bayesian probability updates (Section III) and sequential decision-making (Section IV) for a single cell. In this section, we define the uncertainty map based on these posterior probabilities and the metric for the

search and classification tasks in general multi-cell domains. We also relate the task metric with a dynamic observation cost for the Bayesian sequential decision-making strategy in multi-cell domains.

A. The Uncertainty Map

We use an information-based approach to construct an uncertainty map for a multi-cell task domain. The uncertainty map is used to guide the sensor to regions of high uncertainty in the domain. The information entropy function of a probability distribution is used to evaluate uncertainty [43]. Let $P(\tilde{\mathbf{c}}, t)$ be the probability distribution for object absent and its classification at cell $\tilde{\mathbf{c}}$ at time t and is given by $P(\tilde{\mathbf{c}}, t) = \{P(X(\tilde{\mathbf{c}}) = 0; t), P(X(\tilde{\mathbf{c}}) = 1; t), P(X(\tilde{\mathbf{c}}) = 2; t)\}$. We define its information entropy as

$$H(P(\tilde{\mathbf{c}}, t)) = -\sum_{j=0}^{2} P(X(\tilde{\mathbf{c}}) = j; t) \ln P(X(\tilde{\mathbf{c}}) = j; t).$$
(10)

If $P(X(\tilde{\mathbf{c}}) = j; t) = 0$, we set the term $P(X(\tilde{\mathbf{c}}) = j; t) \cdot \ln P(X(\tilde{\mathbf{c}}) = j; t) = 0$ by convention because there is no uncertainty about object existence or its property. It also follows that $\lim_{P(X(\tilde{\mathbf{c}}) = j; t) \to 0} P(X(\tilde{\mathbf{c}}) = j; t) \ln P(X(\tilde{\mathbf{c}}) = j; t) = 0$. $H(P(\tilde{\mathbf{c}}, t)) \geq 0$ measures the uncertainty level of object existence and its classification at cell $\tilde{\mathbf{c}}$ at time t. The greater the value of H, the larger the uncertainty is. The desired uncertainty level is $H(P(\tilde{\mathbf{c}}, t)) = 0$. Defining $H_{\max}$ as the maximum uncertainty value attainable by $H(P(\tilde{\mathbf{c}}, t))$, we have $H_{\max} = \ln(3)$ when $P(X(\tilde{\mathbf{c}}) = j; t) = \frac{1}{3}$. The information entropy distribution at time step t over the domain forms an uncertainty map at that time instant.

The initial uncertainty distribution $\{1 - P_p, P_F, P_G\}$, $\forall \tilde{\mathbf{c}} \in \mathcal{D}$, reflects the certainty level at the outset of the mission and is consistent with the assumption of IID objects. Other initial uncertainty distributions may be used such that it reflects any previous knowledge of regions with higher likelihood of object presence. Such distributions affect the probabilities of object existence and its classification at other cells within the domain whenever a decision is made at a certain cell.

When the sensor is taking observations at a cell $\tilde{\mathbf{c}}$, the value of $H(P(\tilde{\mathbf{c}}, t))$ varies with $P(X(\tilde{\mathbf{c}}) = j; t)$, which is updated according to Bayes rule (4) in Section III. When the sensor makes a decision and leaves $\tilde{\mathbf{c}}$, $H(P(\tilde{\mathbf{c}}, t))$ remains constant until the sensor comes back when possible. This is repeated at each cell in the domain and is iterated until the uncertainty level of each cell is within a small neighborhood of zero. At that moment, the search and classification mission is said to be completed.

B. Task Metric

When the observation cost is low, the Bayes risk may be minimized by taking one more observation with possible reduction in decision errors. Under this scenario, the sensor will decide not to proceed searching for and classifying more objects but, instead, to stop and take an observation at the current cell. We define the associated cost of not carrying further search and classification as follows:

$$\mathcal{J}(t) = \frac{\sum_{\tilde{\mathbf{c}} \in \mathcal{D}} H(P(\tilde{\mathbf{c}}, t))}{H_{\max} A_{\mathcal{D}}}.$$
(11)

The cost $\mathcal{J}$ is proportional to the sum of uncertainty over $\mathcal{D}$. To normalize $\mathcal{J}(t)$, we divide this summation by the area of the domain $A_{\mathcal{D}}$ multiplied by $H_{\max}$. According to this definition, we have $0 \leq \mathcal{J}(t) \leq 1$. Initially, we have $H(P(\tilde{\mathbf{c}}, 0)) = -(1 - P_p) \ln(1 - P_p) - P_F \ln P_F - P_G \ln P_G$, $\forall \tilde{\mathbf{c}} \in \mathcal{D}$, this is the uncertainty level before any observation is taken, then $0 \leq \mathcal{J}(0) = H(P(\tilde{\mathbf{c}}, 0))/H_{\max} \leq 1$. On the other hand, if $H(P(\tilde{\mathbf{c}}, t_s)) = 0$ at some $t = t_s$ for all $\tilde{\mathbf{c}} \in \mathcal{D}$, then $\mathcal{J}(t_s) = 0$ and the entire domain has been satisfactorily covered and we know with 100% certainty that there are no more objects yet to be found and classified.

Now let us associate a dynamic observation cost $c_{\text{obs}}(t)$ with the cost function $\mathcal{J}(t)$,

$$c_{\text{obs}}(t) = \gamma \mathcal{J}(t)$$
(12)

where $\gamma > 0$ is some positive weighting parameter. At the outset of the mission, few regions in the domain have been covered, therefore, the cost, $\mathcal{J}$, of not searching and classifying anywhere else is high. Equivalently, taking an observation at the current cell is "expensive," i.e., $c_{\text{obs}}(t)$ is large. In this case, the cost-aware Bayesian sequential decision-making strategy tends to make a decision with a few observations, which may yield a large number of wrong decisions (however, it still gives the minimum Bayes risk over all decisions given the limited available observations), but increase the potential of rapidly detecting and classifying more critical objects in the domain. When the sensor stops taking observations, makes a decision, and leaves the current cell, it will move to another cell and again take an observation there. Because the uncertainty level associated with that cell changes (see (4), (10)), the values for $\mathcal{J}$ (see (11)) and c_{obs} (see (12)) over the entire domain differ accordingly. Additional information is gained by changing the cell to be observed. When the sensor has surveyed more regions in the domain, the uncertainty level at all the visited cells is reduced with respect to the initial uncertainty, and hence both $\mathcal{J}$ and c_{obs} decrease. The process will be repeated until $\mathcal{J}(t) \to 0$, i.e., all the unknown objects of interest within the domain have been found and classified with a desired uncertainty level in a small neighborhood of zero. Note that the observation cost is assigned according to the real-time progress of the search and classification task and facilitates
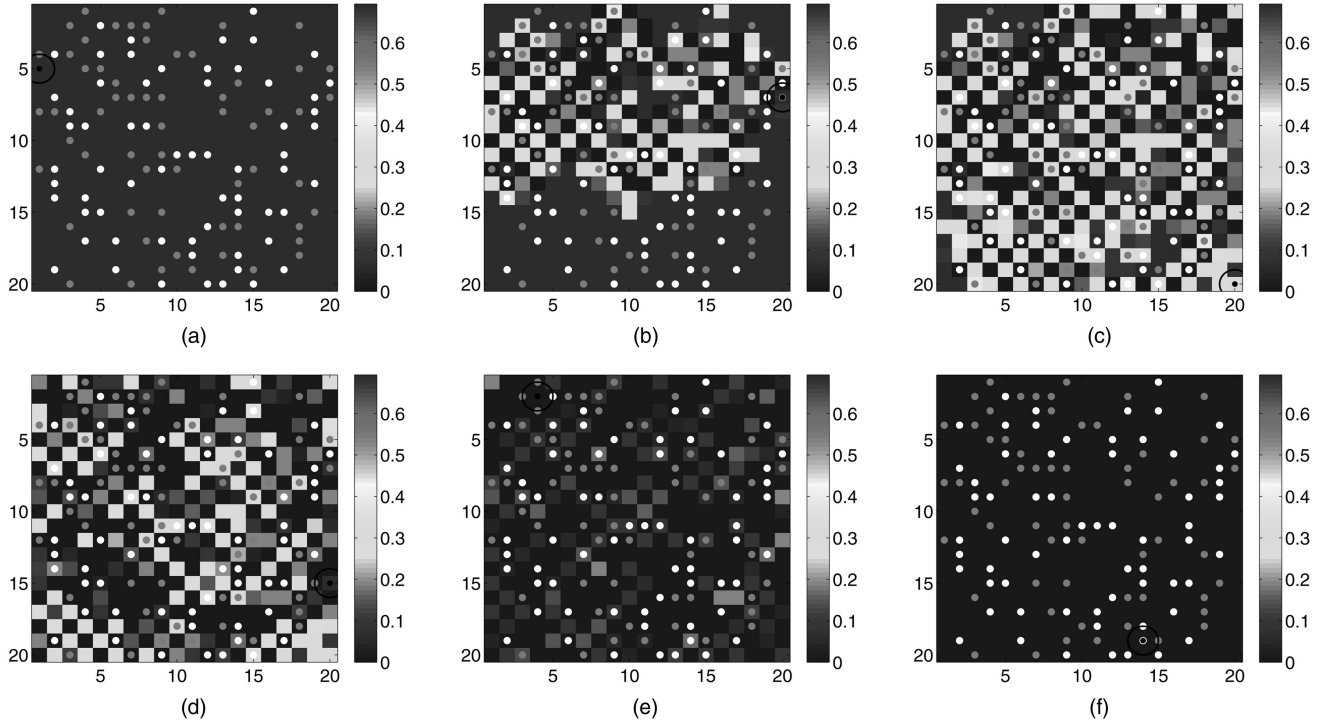
Fig. 4. Uncertainty map (dark red (a) for highest uncertainty and dark blue (f) for lowest uncertainty) at (a) $t = 1$, (b) $t = 167$, (c) $t = 334$, (d) $t = 501$, (e) $t = 668$, and (f) $t = 835$ (with initial uncertainty $H_0(P(\tilde{\mathbf{c}}, 0)) = -(1 - P_p)\ln(1 - P_p) - P_F \ln P_F - P_G \ln P_G$, where $P_p = 0.7$, $P_F = 0, 1$, $P_G = 0.1$.)

real-time decision-making based on the available observations.

Remark about the Weighting Rarameter γ: A small value of γ corresponds to the case where the sensor will stay in a cell until a high certainty about the object existence or its classification at this cell is achieved before moving on. A large value gives the opposite case, i.e., the sensor will not linger long in any cell until it has had a chance to survey more regions in the domain.

VI. SENSOR MOTION CONTROL

In this section we consider a control strategy for the sensor motion over the mission domain $\mathcal{D}$. Combining with the Bayesian sequential decision-making strategy, it seeks to find and classify all objects in $\mathcal{D}$ with a desired confidence level (i.e., achieve $\mathcal{J} \to 0$) under a dynamic observation cost and the minimum Bayes risk at every time step. As mentioned in Section II-B, we assume that there is no speed limit on the sensor, i.e., the sensor is able to move to any cell within $\mathcal{D}$ from its current location.

Consider the set

$$\mathcal{Q}_H(t) = \{\tilde{\mathbf{c}} \in \mathcal{D} : \arg\max_{\tilde{\mathbf{c}}} H(P(\tilde{\mathbf{c}}, t))\}$$

which is the set of cells with highest uncertainty level H within $\mathcal{D}$ at time t. Next, let $\tilde{\mathbf{q}}_c(t)$ be the centroid of the cell that the sensor is currently located at and define the subset $\mathcal{Q}_d(t) \subseteq \mathcal{Q}_H(t)$ as

$$\mathcal{Q}_d(t) = \{\tilde{\mathbf{c}} \in \mathcal{Q}_H(t) : \arg\min_{\tilde{\mathbf{c}}} \|\tilde{\mathbf{q}}_c(t) - \tilde{\mathbf{q}}\|\}$$

where $\tilde{\mathbf{q}}$ is the centroid of $\tilde{\mathbf{c}}$. The set $\mathcal{Q}_d(t)$ contains the cells which have both the shortest distance from the current cell and the highest uncertainty.

When the sensor finishes taking observations in a current cell via the Bayesian sequential decision-making strategy and decides to move to a new cell, it will choose the next cell to go to from $\mathcal{Q}_d(t)$. Note that $\mathcal{Q}_d(t)$ may have more than one cell. Let $N_{\mathrm{Hd}} = |\mathcal{Q}_d(t)|$, the sensor will randomly pick a cell from $\mathcal{Q}_d(t)$ with probability $1/N_{\mathrm{Hd}}$.

VII. FULL-SCALE DOMAIN SIMULATIONS

In this section we provide a detailed numerical simulation that illustrates the performance of the decision-making strategy, and a Monte-Carlo simulation comparison between the proposed strategy and the classical fixed-sample Bayesian hypothesis testings. All the simulations are implemented on a 2.80-GHz, i7-860m processor with 4.0 GB RAM, and Matlab-compiled codes.

A. Simulation Example

We consider a 20×20 square domain $\mathcal{D}$. For each $\tilde{\mathbf{c}} \in \mathcal{D}$, we assume an IID prior probability distribution: $P(X(\tilde{\mathbf{c}}) = 0; t = 0) = 0.7$, $P(X(\tilde{\mathbf{c}}) = 1; t = 0) = 0.1$, and $P(X(\tilde{\mathbf{c}}) = 2; t = 0) = 0.2$. The number, locations, and properties of the objects are randomly generated. The initial position of the sensor is also selected randomly. The diameter of the sensor (one grid length) is shown by the black circle in Fig. 4. The black dot represents
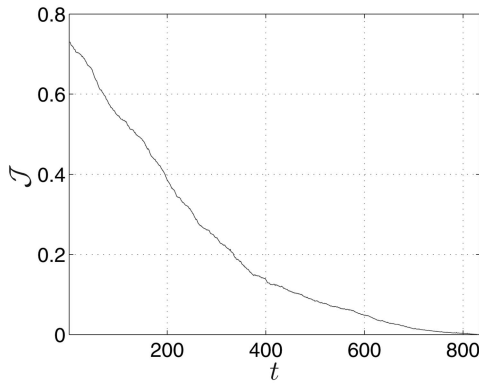
Fig. 5. Evolution of the search cost $\mathcal{J}(t)$.

the position of the sensor. The sensing parameters β_{ij} are the same as in (9). The observation cost weighting parameter γ in (12) is set as 0.05 and the desired uncertainty for every cell is $\epsilon = 0.02$.

The number of objects generated for this simulation turns out to be 125 (the expected number of objects is 120 according to (1)) with locations as indicated by the 64 white dots (objects with property F) and 61 magenta dots (objects with property G) in Fig. 4. Figure 4 shows the evolution of H. At $t = 835$, $H = 0 < \epsilon$ has been achieved everywhere within $\mathcal{D}$. Figure 5 shows the evolution of the search cost function $\mathcal{J}(t)$ and can be seen to converge to zero.

Figure 6(a) shows the number of missed and false detections versus time. The total number of missed detections during the entire mission is 10, and that of false detections is 4. Figure 6(b) shows that the number of times that property F has been decided given property G is 3, and property G has been decided given property F is 1. Note that the number of missed detections is much larger than that of false detections or incorrect classifications. This

is because the initial priors $P(X(\tilde{\mathbf{c}}) = i, t = 0)$, $i = 1, 2$ that we start with are much closer to zero than one. If the sensor makes an erroneous observation $Y(\tilde{\mathbf{c}}) = 0$ given there is actually an object, according to the Bayes probability updates (4), the posterior probability of object absent increases and the sensor tends to make a wrong decision of object absent. Another reason that the number of incorrect classifications is relatively smaller is because the number of objects is smaller compared with the total number of cells in the domain and only the wrong decisions between $X(\tilde{\mathbf{c}}) = 1$ and $X(\tilde{\mathbf{c}}) = 2$ are taken into account. In both figures the numbers of erroneous decisions are small relative to the total number of cells within the domain. This suggests that the cost-aware Bayesian sequential decision-making strategy is efficient in making good decisions given limited available observations. It is also shown in the figures that as time increases, the number of errors decreases.

Figure 7(a) shows that object 25 (located at (4, 17)) has property F with classification uncertainty $H = 0 < \epsilon$ at time step 835. Figure 7(b) shows that object 66 (located at (10, 12)) has property G with classification uncertainty $H = 0 < \epsilon$ at time step 835. The results are consistent with the simulation setup. The properties of other objects are also satisfactorily classified with the desired uncertainty level less than ϵ and can be shown like Figure 7(a) and 7(b) without difficulty.

Table II shows the mean percentage of missed detections, false detections, and incorrect classifications during time period 1–200, 201–400, 401–600, and 601–800, respectively. 100 runs are carried out in 800 time steps with the same parameter settings as above. From the table, most of the errors occur at the earlier stage of the mission and the number of errors decreases with time. This implies that we can balance between the number of errors
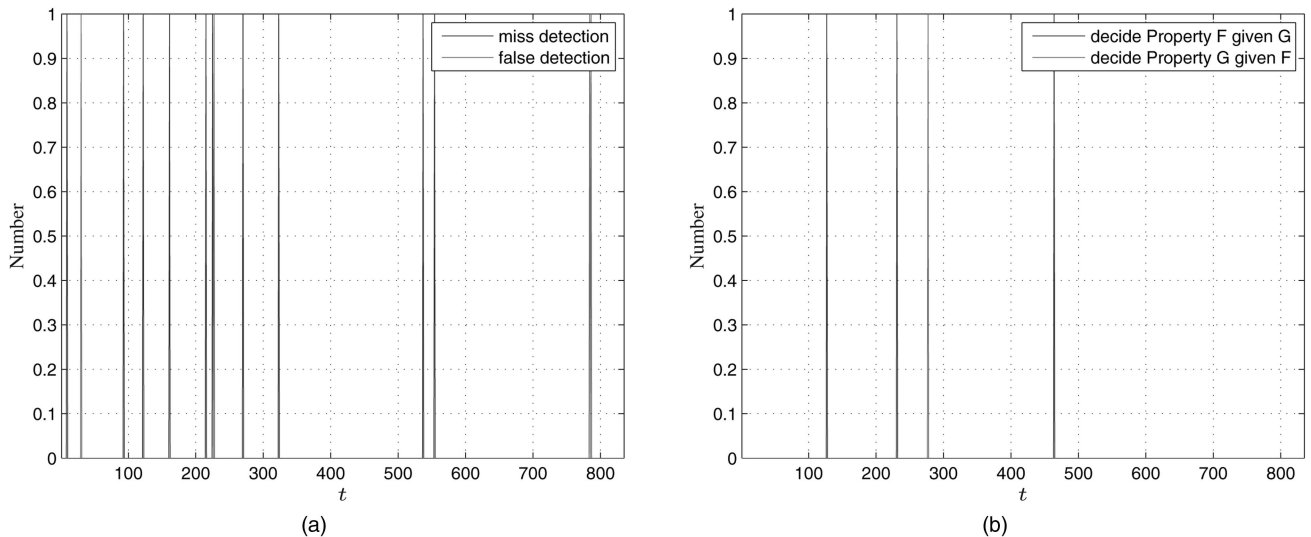

(a)

(b)

Fig. 6. (a) Number of missed detections and false detections. (b) Number of times that property F is decided given property G, and property G is decided given property F.
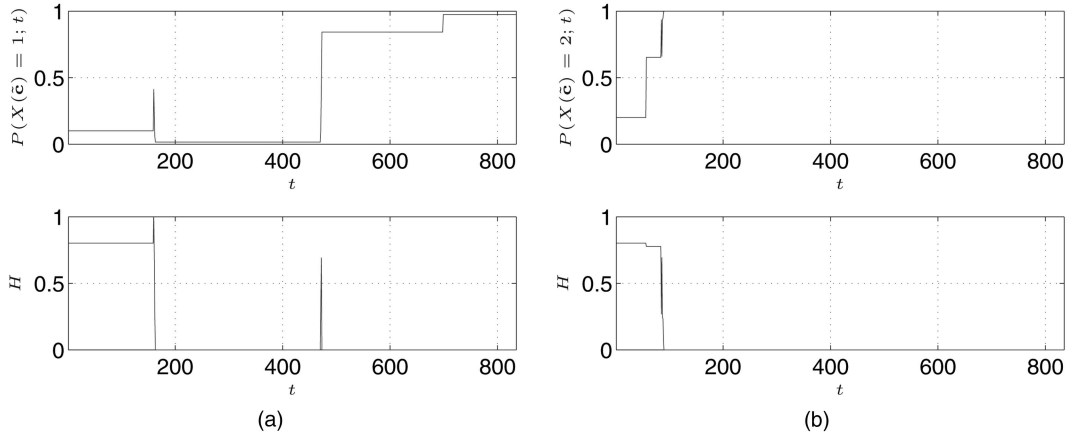
Fig. 7. (a) Probability of object 25 having property F and corresponding uncertainty function H. (b) Probability of object 66 having property G and corresponding uncertainty function H.

TABLE II
Mean Percentage of Missed Detections, False Detections, and Incorrect Classifications During Time Period $T_1 : 1$–200, $T_2 : 201$–400, $T_3 : 401$–600, and $T_4 : 601$–800, Respectively

Mean Percentage (%)	T_1	T_2	T_3	T_4
missed detection	40.43	34.51	17.35	7.71
false detection	43.49	25.43	21.64	9.44
incorrect classification	42.09	24.48	18.52	14.91

within the tolerance range and the limited time we have to decide when to stop.

B. Monte-Carlo Simulation Comparison

Now we perform a Monte-Carlo simulation to compare the performance of the proposed Bayesian sequential strategy and the classical fixed-sample Bayesian hypothesis testing [16, 42, 43]. Under UCA, the fixed-sample Bayesian hypothesis testing is the maximum a posterior (MAP) estimator. That is, the optimal decision corresponds to the state that gives the maximum posterior probability after L observations. Note that this is an off-line batch technique where a decision is made if and only if all the fixed L observations have been taken. Here we use it as a benchmark performance criterion.

From the simulation results, the expected number of observations taken at each cell under the Bayesian sequential method is 1.988. Therefore, it is reasonable that we compare the statistics of this method with 1–4 fixed-sample Bayesian hypothesis testing. Five metrics are considered: the final achieved maximum uncertainty $H_{\max,t_f}$, the final value for the cost function $\mathcal{J}(t_f)$, the total number of missed detections n_m, the total number of false detections n_f, and the total number of incorrect classifications n_i. For each case 100 runs are carried out. For the sake of comparison, we use same settings for object number, positions, properties, and initial position of the vehicle. All the other parameters are the same as in Section VII-A.

Figures 8(a)–(e) show the comparison of the metrics, respectively, and Table III summarizes the statistical results. In order to achieve a similar small amount of decision errors, the fixed-sample hypothesis testing method requires $L = 4$ observations at each cell. The cost-aware Bayesian sequential decision-making strategy outperforms the classical methods by 1) reducing decision errors, and 2) minimizing observation numbers. Therefore, according to (7) and (8), under UCA, the proposed strategy leads to minimum Bayes risk with a same performance level.

VIII. CONCLUSION

Built on a probabilistic framework, a cost-aware decision-making strategy was developed to seek the detection and satisfactory classification of all objects in a domain using Bayesian sequential risk analysis. The proposed strategy guarantees the detection and classification decisions made with minimum Bayes risk in the presence of sensor errors under limited sensory resources. Detailed numerical simulations are provided to demonstrate the performance of the strategy. Current work includes the investigation of a generalized choice of the observation cost weighting parameter γ and task metrics to guarantee zero uncertainty convergence. Future research will focus on the following aspects:

1) tracking and classification of mobile objects using multiple autonomous sensor vehicles,
2) exploration in unknown environment geometries (i.e., unknown $\mathcal{D}$),
3) non IID object distributions over the domain, where decision-making at one cell is affected by all the decisions made at other cells,
4) SPRT method for the cases where no prior information is available, and
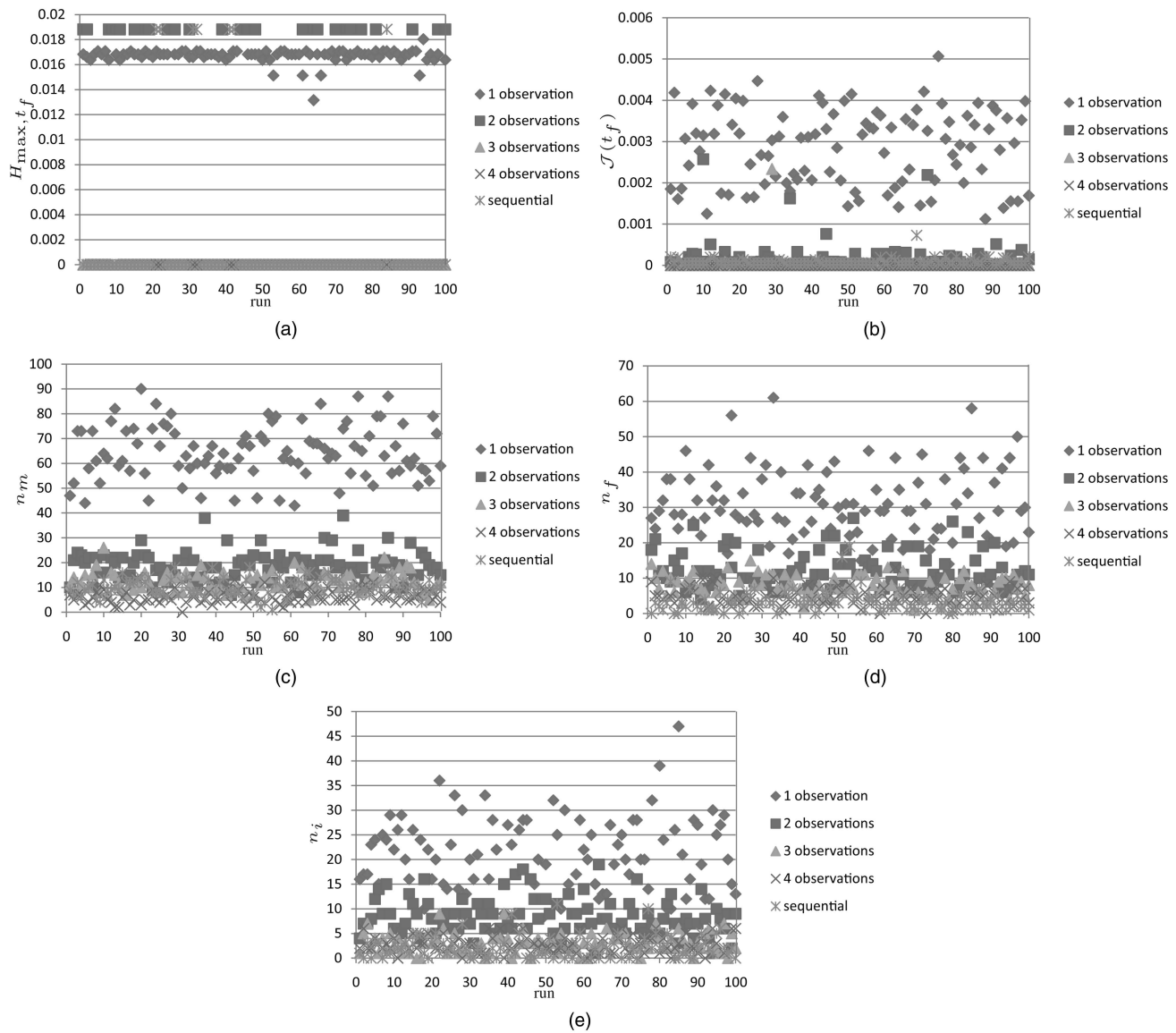5) SR and CUSUM tests for change-point detection.

Fig. 8.   Performance comparison in terms of (a) $H_{\max,t_f}$, (b) $\mathcal{J}(t_f)$, (c) n_m, (d) n_f, and (e) n_i between fixed-sample Bayesian hypothesis testings with 1, 2, 3, 4 observations and Bayesian sequential detection.

TABLE III
Comparison Between Fixed-Sample Hypothesis Testing and Sequential Detection under Bayes Framework

	$H_{\max,t_f}$	$\mathcal{J}(t_f)$	n_m	n_f	n_i
$L=1$	1.68E-2	2.85E-3	64.49	30.99	20.51
$L=2$	5.08E-3	1.7E-4	18.72	12.09	8.48
$L=3$	1.07E-3	2.55E-5	12.02	7.43	2.91
$L=4$	6.85E-4	4.39E-6	6.41	4.63	2.74
Sequential	1.18E-3	8.86E-5	9.07	3.98	2.15

Different from Bayesian sequential detection, the main focus of SPRT, SR, and CUSUM is to minimize the expected number of observation samples instead of the overall risk, and there is no explicit risk-based analysis available in the literature for these methods. For the purpose of risk analysis, the exact missed/false detection rates at the stopping times, which will probably overshoot the preset thresholds, need to be derived.

REFERENCES

[1]   Hernandez, M. L., Kirubarajan, T., and Bar-Shalom, Y.
      Multisensor resource deployment using posterior
      Cramér-Rao bounds.
      IEEE Transactions on Aerospace and Electronic Systems,
      40, 2 (Apr. 2004), 399–416.

[2]   Wintenby, J. and Krishnamurthy, V.
      Hierarchical resource management in adaptive airborne
      surveillance radars.
      IEEE Transactions on Aerospace and Electronic Systems,
      42, 2 (Apr. 2006), 401–420.

[3]   Wang, Y., Hussein, I. I., and Erwin, R. S.
      Awareness-based decision making for search and
      tracking.
      Proceedings of the American Control Conference, Seattle,
      WA, June 2008, pp. 3169–3175, invited paper.

[4]   Wang, Y. and Hussein, I. I.
      Bayesian-based decision-making for object search and
      classification.
      IEEE Transactions on Control Systems Technology, **19**, 6
      (Nov. 2011), 1639–1647.

[5]   Wang, Y., et al.
      Cost-aware sequential Bayesian tasking and
      decision-making for search and classification.
      Proceedings of the American Control Conference,
      Baltimore, MD, 2010, pp. 6423–6428.

[6]   Spletzer, J. R. and Taylor, C. J.
      Dynamic sensor planning and control for optimally
      tracking targets.
      The International Journal of Robotics Research, 1 (Jan.
      2003), 7–20.

[7]   Bertuccelli, L. F. and How, J. P.
      Robust UAV search for environments with imprecise
      probability maps.
      *Proceedings of the 44th IEEE Conference on Decision and
      Control, and the European Control Conference*, Dec. 2005.

[8]   Bertuccelli, L. F. and How, J. P.
      Bayesian forecasting in multi-vehicle search operations.
      Presented at the AIAA Guidance, Navigation, and Control
      Conference and Exhibit, Aug. 2006.

[9]   Furukawa, T., et al.
      Recursive Bayesian search-and-tracking using coordinated
      UAVs for lost targets.
      *Proceedings of the 2006 IEEE International Conference on
      Robotics and Automation*, May 2006.

[10] Lavis, B., Furukawa, T., and Durrant-Whyte, H. F.
      Dynamic space reconfiguration for Bayesian
      search-and-tracking with moving targets.
      Autonomous Robots, **24** (May 2008), 387–399.

[11] Mahler, R.
      Objective functions for Bayesian control-theoretic sensor
      management, I: Multitarget first-moment approximation.
      Proceedings of IEEE Aerospace Conference, vol. 4, 2003,
      pp. 4-1905–4-1923.

[12] Flint, M., Polycarpou, M., and Fernández-Gaucherand, E.
      Cooperative control for multiple autonomous UAVs
      searching for targets.
      *Proceedings of the 41st IEEE Conference on Decision and
      Control*, vol. 3, Dec. 2002, pp. 2823–2828.

[13] Kalandros, M. K., et al.
      Tutorial on multisensor management and fusion
      algorithms for target tracking.
      Proceedings of the American Control Conference, vol. 5,
      Boston, MA, 2004, pp. 4743–4748.

[14] Seo, S., et al.
      Gaussian process regression: Active data selection and
      test point rejection.
      *Proceedings of the IEEE International Joint Conference on
      Neural Networks*, July 2000, pp. 241–246.

[15] Shewry, M. C. and Wynn, H. P.
      Maximum entropy sampling.
      Journal of Applied Statistics, **14**, 2 (1987), 165–170.

[16] Poor, H. V.
      An Introduction to Signal Detection and Estimation (2nd
      ed.).
      New York: Springer-Verlag, 1994.

[17] Wald, A. and Wolfowitz, J.
      Bayes solutions of sequential decision problems.
      The Annals of Mathematical Statistics, **21**, 1 (1950),
      82–99.

[18] Yu, X. and Azimi-Sadjadi, M. R.
      Neural network directed Bayes decision rule for moving
      target classification.
      IEEE Transactions on Aerospace and Electronic Systems,
      36, 1 (Jan. 2000), 176–188.

[19] Wald, A.
      Sequential tests of statistical hypotheses.
      The Annals of Mathematical Statistics, **16**, 2 (June 1945),
      117–186.

[20] Wald, A.
      Sequential Analysis.
      New York: Dover Publications, 2004.

[21] Shiryaev, A. N.
      Optimal Stopping Rules.
      New York: Springer, Nov. 2007.

[22] Basseville, M. E. and Nikiforov, I. V.
      Detection of Abrupt Changes: Theory and Application
      (Information and System Sciences Series).
      Upper Saddle River, NJ: Prentice-Hall, Apr. 1993.

[23] Roberts, S. W.
      A comparison of some control chart procedures.
      Technometrics, **8**, 3 (Aug. 1966), 411–430.

[24] Page, E. S.
      Continuous inspection scheme.
      Biometrika, **41**, 1/2 (June 1954), 100–115.

[25] Thrun, S., Burgard, W., and Fox, D.
      Probabilistic Robotics (Intelligent Robotics and
      Autonomous Agents Series), R. C. Arkin (Ed.).
      Cambridge, MA: The MIT Press, Sept. 2005.

[26] Smith, T.
      Probabilistic planning for robotic exploration.
      Ph.D. dissertation, Carnegie Mellon University,
      Pittsburgh, PA, July 2007.

[27] Hussein, I. I. and Stipanović, D.
      Effective coverage control using dynamic sensor networks
      with flocking and guaranteed collision avoidance.
      Proceedings of the 2007 American Control Conference,
      July 2007, pp. 3420–3425.

[28] Hussein, I. I. and Stipanović, D.
      Effective coverage control using dynamic sensor
      networks.
      *Proceedings of the 45th IEEE Conference on Decision and
      Control*, Dec. 2006, pp. 2747–2752.

[29] Hussein, I. I. and Stipanović, D.
      Effective coverage control for mobile sensor networks
      with guaranteed collision avoidance.
      IEEE Transactions on Control Systems Technology, **15**, 4
      (July 2007), 642–657.

[30] Wang, Y. and Hussein, I. I.
      Cooperative vision-based multi-vehicle dynamic coverage
      control for underwater applications.
      *Proceedings of the IEEE Multiconference on Systems and
      Control*, 2007, invited paper.

[31] Hussein, I. I., Stipanović, D., and Wang, Y.
      Reliable coverage control using heterogeneous vehicles.
      *Proceedings of the 46th IEEE Conference on Decision and
      Control*, New Orleans, LA, Dec. 2007, pp. 6142–6147.

[32] Wang, Y. and Hussein, I. I.
      Underwater acoustic imaging using autonomous vehicles.
      *2008 IFAC Workshop on Navigation, Guidance and Control
      of Underwater Vehicles*, vol. 2, Apr. 2008.

[33] Wang, Y. and Hussein, I. I.
      Awareness coverage control over large scale domains
      with intermittent communications.
      IEEE Transactions on Automatic Control, **55**, 8 (Aug.
      2010), 1850–1859.

[34] Viswanathan, R. and Varshney, P.
      Distributed detection with multiple sensors: Part
      1—Fundamentals.
      Proceedings of the IEEE, **85**, 1 (Jan. 1997).

[35] Blum, R. S., Kassam, A., and Poor, H.
      Distributed detection with multiple sensors: Part 2—
      Advanced topics.
      Proceedings of the IEEE, **85**, 1 (Jan. 1997).

[36] Cortés, J., et al.
Coverage control for mobile sensing networks.
IEEE Transactions on Robotics and Automation, **20**, 2 (2004), 243–255.

[37] Ganguli, A., et al.
On collective motion in sensor networks: Sample problems and distributed algorithms.
Proceedings of 44th IEEE Conference on Decision and Control and 2005 European Control Conference, Dec. 2005, pp. 4239–4244.

[38] Grocholsky, B.
Information-theoretic control of multiple sensor platforms.
Ph.D. dissertation, The University of Sydney, Australia, 2002.

[39] Poor, H. V. and Hadjiliadis, O.
Quickest Detection (1st ed.).
New York: Cambridge, Dec. 2008.

[40] Van Trees, H. L.
Detection, Estimation, and Modulation Theory, vol. I.
Hoboken, NJ: Wiley, 2003.

Yue Wang (M'07) received a Ph.D. in mechanical engineering in 2011, an M.S. degree in mechanical engineering from Worcester Polytechnic Institute in 2008, and a B.S. degree in mechanical engineering and automation from Shanghai University, China in 2005.

She is an assistant professor in the Department of Mechanical Engineering at Clemson University. From 2011 to 2012, she was a postdoctoral research associate in the Electrical Engineering Department at the University of Notre Dame. Her research interests include multi-agent systems, cyber-physical systems, networked control systems. In particular, she has been focused on decision-making and sensor management using multiple autonomous vehicles, the tight integration of cyber-physical systems using passivity-based method, coverage control of sensor networks over large-scale domains, and space-augmented space situational awareness.

Islam I. Hussein (M'04) was awarded the Ph.D. degree in aerospace engineering in 2005 and the M.Sc. degrees in aerospace engineering and applied mathematics in 2002 from the University of Michigan. He received his B.Sc. degree in mechanical engineering in 2000 from the American University in Cairo.

He is an assistant professor in the Department of Mechanical Engineering at Worcester Polytechnic Institute. From 2005 to 2006, he held a postdoctoral research associate position at the Coordinated Science Laboratory at the University of Illinois at Urbana–Champaign. His research interests are in the areas of cooperative coverage control of multi-vehicle sensor network systems, dynamics and control of virtual tether satellite systems, and nonlinear constrained dynamics and control, in particular from the geometric point of view.

Professor Hussein is the recipient of the 2010 and 2009 ASEE/AFRL Summer Faculty Fellowship Awards. He was a finalist for the best student paper award at the 2003 and 2004 American Control Conferences. He received the Presidential Scholarship at the American University in Cairo in 1997 and 1999. He is a member of the AIAA and ASME.

D. Richard Brown, III (M'97—SM'09) received a Ph.D. in electrical engineering from Cornell University in 2000 and M.S. and B.S. degrees in electrical engineering from The University of Connecticut in 1996 and 1992, respectively.

He is an associate professor in the Department of Electrical and Computer Engineering at Worcester Polytechnic Institute.

Dr. Brown, III held an appointment as a visiting associate professor at Princeton University from August 2007 to June 2008.

R. Scott Erwin (M'94—SM'07) received a B.S. in aeronautical engineering from Rensselaer Polytechnic Institute in 1991, and the M.S. and Ph.D. degrees in aerospace engineering from the University of Michigan in 1993 and 1997, respectively.

He has been an employee of the Air Force Research Laboratory, Space Vehicles Directorate (AFRL/RV) located at Kirtland AFB, NM, from 1997 to the present. He is currently a principal research aerospace engineer and assistant to the chief scientist of Space Vehicles Directorate. His responsibilities include performing basic and applied research in space communications systems, spacecraft control, and satellite autonomy as well as providing advice to and coordinating efforts on behalf of the AFRL/RV Chief Scientist across all areas of the AFRL/RV portfolio. He oversees all University-level outreach and collaboration programs for Space Vehicles Directorate, including the Space Scholars Program, the University NanoSat Program, and the ASEE Summer Faculty Fellowship and National Research Council Faculty and Postdoctoral Research Fellowship Programs. His current research interests are autonomous spacecraft, satellite communications, and the interplay between communications, estimation, and control in networked space systems.

Dr. Erwin is an Associate Fellow of AIAA. He serves on the IEEE Control Systems Society Aerospace Controls Technical Committee and the AIAA Guidance, Navigation, and Control Technical Committee and is a Member of the IEEE Communications Society. Dr. Erwin was selected as the recipient of the 2001 United States Air Force Science and Engineering Award for Exploratory and Advanced Technology Development, the 2004 Young Engineer of the Year and 2008 Outstanding Engineer of the Year Awards from the Albuquerque Section of the IEEE, and the 2006 Air Force Material Command General Lester L. Lyles Award for mentoring, professional development and educational outreach. He is the author or co-author of over sixty technical publications in the areas of spacecraft dynamics and controls.