Non-Myopic Sensor Control for Target Search and Track Using a Sample-Based GOSPA Implementation

Marcel Hernandez, Ángel F. García-Fernández, and Simon Maskell

Department of Electrical Engineering, Electronics and Computer Science,
The University of Liverpool, Brownlow Hill, Liverpool, L69 3GJ

Abstract— **This paper is concerned with sensor management for target search and track using the generalised optimal subpattern assignment (GOSPA) metric. Utilising the GOSPA metric to predict future system performance is computationally challenging, because of the need to account for uncertainties within the scenario, notably the number of targets, the locations of targets, and the measurements generated by the targets subsequent to performing sensing actions. In this paper, efficient sample-based techniques are developed to calculate the predicted mean square GOSPA metric. These techniques allow for missed detections and false alarms, and thereby enable the metric to be exploited in scenarios more complex than those previously considered. Furthermore, the GOSPA methodology is extended to perform non-myopic (i.e. multi-step) sensor management via the development of a Bellman-type recursion that optimises a conditional GOSPA-based metric. Simulations for scenarios with missed detections, false alarms, and planning horizons of up to three time steps demonstrate the approach, in particular showing that optimal plans align with an intuitive understanding of how taking into account the opportunity to make future observations should influence the current action. It is concluded that the GOSPA-based, non-myopic search and track algorithm offers a powerful mechanism for sensor management.**

Keywords— **GOSPA metric, myopic planning, multi-step planning, optimal control, sensor management, search, target tracking, efficient sampling, Bellman recursion.**

I. INTRODUCTION

Recently, there has been great interest in sensor management [1] for search and track of multiple targets (e.g. see [2], [3] and references therein). In [2], [3], non-myopic (i.e. multi-step) planning was performed using a Poisson multi-Bernoulli mixture (PMBM) filter [4], [5] as the basis for search, and the "predicted ideal measurement set approach" [6][1] to predict the benefit of observing/updating a target track. The optimal control problem then used a cost function that was a weighted sum of the integrated intensity of the PMBM density (for search) and the sum total of the trace of each updated target covariance (for tracking). In other recent work (see [7], [8]), non-myopic planning approaches were developed that assumed the origin of measurements was known, allowing each detected target to be tracked independently. An occupancy grid filter was then used to represent the presence of undetected targets.

In [9], a sensor management approach was developed for scenarios with an unknown number of targets that allowed for noisy measurements, with measurement origin uncertainty, and potentially both missed detections and false alarms. The basis of the approach was to use the posterior Cramér-Rao bound (PCRB) [10] to predict multi-target tracking performance, accounting for measurement origin uncertainty via a matrix of "information reduction factors" (e.g. see [11]). In conducting search to detect new targets, "particles" (i.e. target hypotheses) were distributed uniformly along the perimeter of the surveillance region, and the probability of detecting a new target was estimated if the sensor observed a region containing one or more hypothesis. The bi-criterion optimisation problem then used a cost function that was a weighted total of the multi-target PCRB and the probability of detecting a new target. It is noted that the approach developed in [9] built on earlier sensor management research for target tracking that used the PCRB as the objective function[2] [12] (see also [13]).

The generalised optimal subpattern assignment (GOSPA) metric [14] provides a mechanism for combining together the costs corresponding to localisation errors for properly detected targets, and errors for missed and false targets. These three types of errors are of major interest in multiple target estimation. Importantly, the GOSPA metric is not prone to two erroneous effects that hinder its predecessor, the optimal subpattern assignment metric (OSPA) [15], [16], specifically: (i): the OSPA metric does not necessarily increase as the number of false targets increases [14]; and: (ii): the OSPA metric is prone to the "spooky effect" in optimal metric-based estimation [17].

In order to use the GOSPA metric as the basis for *predictive* sensor management, it is necessary to account for uncertainties within the scenario, namely the number of targets, the locations of targets, and the measurements generated by the targets as a result of performing sensing actions. To this end, the GOSPA metric is averaged over these uncertainties to provide the average minimum mean squared GOSPA (AMMS-GOSPA) error [18].

Calculating the AMMS-GOSPA is computationally challenging, because of the required averaging across the uncertainties, and the minimisation (within the GOSPA metric). Consequently, in [18], sensor management was demonstrated for relatively simple scenarios with just a single action (i.e. myopic planning), and a single target state hypothesis, (represented by a Dirac delta function) per target. Either one target or two well separated targets were considered[3]. The AMMS-GOSPA was calculated analytically, and the resulting sensor management strategies demonstrated the tradeoff between sensing costs and the probability of target existence, with (all else being equal) the target(s) more likely to be observed as the probability of ex-

[1]The ideal measurement set approach assumes that the measurements are error-free and there are no missed detections or false alarms.

[2]The PCRB was shown to offer an accurate mechanism for predicting tracking performance, allowing sensor management to be performed in time-critical applications.

[3]The separation of the targets allowed the GOSPA metric to be summed across the two targets.

istence increased or the sensing costs decreased. Also of note, in [19, Chapter 6], myopic GOSPA-based sensor management was performed using a Bernoulli-Gaussian approximation of the conditional squared GOSPA error. However, the algorithm developed in [19, Chapter 6] is for track-before-detect applications, and is therefore not suitable for the detection-based measurement model that is the focus of the current paper.

Building on this previous research, this article provides details of the following theoretical advancements:

1. Analytical calculation of the MS-GOSPA for different combinations of measurements and actions.

2. The development of efficient sampling techniques (i.e. of the measurements), to reduce the computational complexity of calculating the AMMS-GOSPA error.

3. The development of an optimal non-myopic planning (Bellman type, e.g. [20]) recursion that exploits the conditional AMMS-GOSPA error.

Efficient calculation of the AMMS-GOSPA allows the GOSPA metric to be exploited in scenarios significantly more complex than those considered in [17], [18]. To this end, the approach is demonstrated in performing sensor management with: (i): a high degree of uncertainty in each target location[4], with the prior distribution represented by a mixture of weighted Dirac delta functions, e.g. as in a particle filter estimate [21]; and (ii): time horizons with multiple time steps (i.e. non-myopic planning).

The remainder of this paper is as follows. In Section II, a review of the GOSPA metric is provided. In Section III, details are provided of the existing state-of-the-art regarding utilising the GOSPA metric for predictive sensor management via optimising the AMMS-GOSPA error. In Section IV, the AMMS-GOSPA is generalised to multiple time step scenarios, and analytical equations for the MMS-GOSPA error are calculated. In Section V, efficient sample-based approximations of the AMMS-GOSPA error are presented, along with baseline tests applying the approaches to a scenario from [18] in which the optimal solution was determined analytically. It is shown that an efficient sampling approach determines the optimal solution with a circa 250 times reduction in computational expensive compared to a sampling approach that does not take into account the possible measurement sequences. In Section VI, the optimal non-myopic planning approach is presented, along with a suboptimal approach and a baseline approach that minimises target localisation errors. In Section VII, simulation results demonstrate the multi-step planning approaches, and in particular show that the optimal actions align with an intuitive understanding of how taking into account the ability to make further observations should influence the current observation. In Section VIII, a Summary and Conclusions are provided. Finally, in an Appendix, it is proven that when the ideal measurement set approximation is used (but allowing $P_d < 1$), the optimal non-myopic planning approach generates identical solutions to a commonly implemented suboptimal approach.

II. GOSPA METRIC

This section reviews the GOSPA metric. We first present the notation and then its definition.

[4]E.g. representing a search track, or a target that has not yet been accurately geo-located.

A. Notation

Let c and p be two real numbers such that $c > 0$ and $1 \leq p < \infty$. Let $d(\cdot, \cdot)$ denote a metric on the single target space[5], which is typically $\mathbb{R}^{n_x}$.

Let $X = \{x_1, ..., x_{|X|}\}$ and $Y = \{y_1, ..., y_{|Y|}\}$ denote two finite sets of targets, with $|X| \leq |Y|$, and $|X|$ being the cardinality (number of elements) of the set X. In the context of target tracking, X typically represents the set of target ground-truth states, and Y the set of target state estimates.

Let γ be an assignment set between $\{1, ..., |X|\}$ and $\{1, ..., |Y|\}$, which satisfies $\gamma \subseteq \{1, ..., |X|\} \times \{1, ..., |Y|\}$, $(i,j), (i,j') \in \gamma \to j = j'$, and $(i,j), (i',j) \in \gamma \to i = i'$. The last two properties ensure that every i and j have at most one assignment. The set of all possible γ is denoted by Γ.

B. Metric

The GOSPA metric, with parameters p and c, between X and Y (for $\alpha = 2$) is given as follows (see [14, Proposition 1]):

$$d_p^{(c,2)}(X, Y) \triangleq \min_{\gamma\in\Gamma} \left(\sum_{(i,j)\in\gamma} [d(x_i, y_j)]^p + \frac{c^p}{2}(|X| - |\gamma| + |Y| - |\gamma|) \right)^{1/p} \quad (1)$$

The first term in (1) represents the localisation errors (to the p-th power) for assigned targets (properly detected ones), which meet $(i,j) \in \gamma$. The terms $\frac{c^p}{2}(|X| - |\gamma|)$ and $\frac{c^p}{2}(|Y| - |\gamma|)$ represent the costs (to the p-th power) for missed and false targets respectively.

Compared to the OSPA metric, the GOSPA metric has an additional parameter α that controls the cardinality mismatch penalty. Importantly, as shown in equation (1), only for $\alpha = 2$ can the GOSPA metric be written in terms of costs corresponding to localisation errors for properly detected targets, missed and false targets, which are usually the penalties of interest in multiple target estimation. Consequently, $\alpha = 2$ is used throughout this paper.

III. EXPLOITATION OF THE GOSPA METRIC FOR MYOPIC SENSOR MANAGEMENT

In order to use the GOSPA metric for predictive sensor management it is necessary to account for potential target state and measurement origin/accuracy uncertainties. To this end, the following errors are determined (see [18]).

A. Mean Squared GOSPA (MS-GOSPA) Error

Given an action $a \in \mathbb{A}$ (e.g. sensor mode or steer direction), and a resulting measurement z, a posterior multi-target state estimate $\hat{X}(z, a)$ can be determined. The resulting mean squared GOSPA (MS-GOSPA) error given z and a is then defined as follows:

$$\begin{aligned} &\text{MS-GOSPA}(\hat{X}; z, a) \\ &\triangleq \mathbb{E}_X \left[\left(d_2^{(c,2)}(X, \hat{X}(z,a)) \right)^2 \middle| z; a \right] \quad (2) \\ &= \int_X \left(d_2^{(c,2)}(X, \hat{X}(z,a)) \right)^2 p(X|z;a)\, dX \quad (3) \end{aligned}$$

[5]In this paper, the metric $d(\cdot, \cdot)$ will denote the geo-location distance (error) between the target (typically denoted by X) and the target state estimate (typically denoted by $\hat{X}$).

B. Minimum MS-GOSPA (MMS-GOSPA) Error

The minimum MS-GOSPA (MMS-GOSPA) error is achieved by selecting the estimate $\hat{X}(z,a)$ to minimise equation (3), i.e.:

$$\begin{aligned}&\text{MMS-GOSPA}(z,a) \qquad (4)\\ &= \min_{\hat{X}(z,a)} \int_X \left(d_2^{(c,2)}\left(X,\hat{X}(z,a)\right)\right)^2 p\left(X|z;a\right)dX\end{aligned}$$

C. Average MMS-GOSPA (AMMS-GOSPA) Error

The average MMS-GOSPA (AMMS-GOSPA) for action a, averaged over the measurement z is then given as follows:

$$\begin{aligned}&\text{AMMS-GOSPA}(a)\\ &\triangleq \mathbb{E}_z\left[\text{MMS-GOSPA}(z,a)\right] \qquad (5)\\ &= \int_z \text{MMS-GOSPA}(z,a)p(z;a)dz \qquad (6)\end{aligned}$$

For each pair (z,a), the AMMS-GOSPA selects the estimate $\hat{X}(z,a)$ that minimises the MS-GOSPA metric, and then averages over the value of z. The AMMS-GOSPA therefore gives the average minimum MS-GOSPA, taking into account both uncertainty in the target state X, and uncertainty in the measurement z for each potential action a.

D. Optimal Action

Selecting the action a that minimises the sum total of the AMMS-GOSPA error and sensing costs (if applicable) provides a mechanism for performing (myopic) sensor management, enabling the system to balance the objectives of (i): minimising localisation errors for properly detected targets, and: (ii): minimising cardinality errors resulting from missed and false targets, (iii): minimising sensing costs.

E. Computational Complexity

Clearly, calculating the AMMS-GOSPA is computationally challenging, because of the averaging performed (over X given (z,a), and then z given a), and the minimisations (within the GOSPA metric, and then over all potential posterior estimates). Consequently, e.g. in [18], sensor management was demonstrated for a relatively simple scenario with just a single action (i.e. myopic planning), and a single potential target location hypothesis.

In this paper, efficient techniques are developed to calculate the AMMS-GOSPA, enabling the metric to be used as a basis for performing sensor management in more complex scenarios (i.e. with targets whose states are highly uncertain, and multiple time steps).

IV. Multiple Time Step AMMS-GOSPA Error

In this section the AMMS-GOSPA error is determined for a time window containing multiple time steps.

A. Decomposition Over Time Steps

To extend the methodology to calculate the multiple time step AMMS-GOSPA error, some additional notation is required. Let $T \geq 1$ denote the number of time steps considered (e.g. $T=1$ denotes myopic planning). Let:

$$\hat{X}_{1:T} \triangleq \left(\hat{X}_1,\ldots,\hat{X}_T\right) \qquad (7)$$

where $\hat{X}_{1:T}$ is a sequence of sets of target state estimates at times 1 through T. Similarly, let $a_{1:T} \triangleq (a_1,\ldots,a_T)$ denote the sensor actions at times 1 through T; and let $z_{1:T} \triangleq (z_1,\ldots,z_T)$ denote the sensor measurements at times 1 through T.

The MS-GOSPA error at time t is then given as follows:

$$\begin{aligned}&\text{MS-GOSPA}(\hat{X}_t;z_{1:t},a_{1:t}) \qquad (8)\\ &= \int_{X_t}\left(d_2^{(c,2)}(X_t,\hat{X}_t(z_{1:t},a_{1:t}))\right)^2 p(X_t|z_{1:t},a_{1:t})dX_t\end{aligned}$$

The multiple step MS-GOSPA error is the discounted sum of MS-GOSPA errors at times 1 through T, i.e.:

$$\begin{aligned}&\text{MS-GOSPA}_{1:T}(\hat{X}_{1:T};z_{1:T},a_{1:T}) \qquad (9)\\ &= \sum_{t=1}^{T}\lambda^{t-1}\text{MS-GOSPA}(\hat{X}_t;z_{1:t},a_{1:t})\end{aligned}$$

where the parameter $\lambda \in [0,1]$ is the discount factor.

The multiple step MMS-GOSPA is then defined as follows:

$$\begin{aligned}&\text{MMS-GOSPA}_{1:T}(z_{1:T},a_{1:T})\\ &\triangleq \min_{\hat{X}_{1:T}}\Big[\sum_{t=1}^{T}\lambda^{t-1}\text{MS-GOSPA}(\hat{X}_t;z_{1:t},a_{1:t})\Big] \qquad (10)\\ &= \sum_{t=1}^{T}\lambda^{t-1}\min_{\hat{X}_t}\Big[\text{MS-GOSPA}(\hat{X}_t;z_{1:t},a_{1:t})\Big] \qquad (11)\\ &= \sum_{t=1}^{T}\lambda^{t-1}\text{MMS-GOSPA}(z_{1:t},a_{1:t}) \qquad (12)\end{aligned}$$

where:

$$\begin{aligned}&\text{MMS-GOSPA}(z_{1:t},a_{1:t}) \qquad (13)\\ &\triangleq \min_{\hat{X}_t}\Big[\text{MS-GOSPA}(\hat{X}_t;z_{1:t},a_{1:t})\Big]\end{aligned}$$

The simplification in (11) is because given $z_{1:t}$ and $a_{1:t}$, the MS-GOSPA at each time step t is dependent only on the target state estimate $\hat{X}_t$ at that time, and critically is independent of $\hat{X}_{1:t-1}$, and also independent of $\hat{X}_{t+1:T}$ (i.e. estimates that perform smoothing are not considered).

The AMMS-GOSPA error at times 1 through T is then given as follows:

$$\begin{aligned}&\text{AMMS-GOSPA}_{1:T}(a_{1:T})\\ &\triangleq \int_{z_{1:T}}\text{MMS-GOSPA}_{1:T}(z_{1:T},a_{1:T}) \qquad (14)\\ &\qquad \times p(z_{1:T};a_{1:T})dz_{1:T}\\ &= \sum_{t=1}^{T}\lambda^{t-1}\text{AMMS-GOSPA}(a_{1:t}) \qquad (15)\end{aligned}$$

where:

$$\begin{aligned}&\text{AMMS-GOSPA}(a_{1:t}) \qquad (16)\\ &\triangleq \int_{z_{1:t}}\text{MMS-GOSPA}(z_{1:t},a_{1:t})p(z_{1:t};a_{1:t})dz_{1:t}\end{aligned}$$

It is noted that equation (15) does not rely on any assumptions, e.g. regarding the number of targets or the number of sensors.

B. Target Hypotheses

For the remainder of this article, it is assumed that there is either zero or one target in the focal scenario[6]. Let the set of potential target hypotheses be denoted by:

$$\text{target hypotheses, } X = \{x_0, x_1, \dots, x_n\} \tag{17}$$

where "$x_0 = \phi$" denotes the hypothesis that a target is not present, and x_i, $i = 1, \dots, n$ denotes the hypothesis that a target exists and its state is given by x_i. The prior probability of X is given as follows:

$$f(X) = \begin{cases} r w_i \delta(X - x_i) & \text{for } i = 1, \dots, n \\ 1 - r & \text{otherwise} \end{cases} \tag{18}$$

where r is the probability of target existence, and w_i denotes the weight of each location hypothesis, with $\sum_{i=1}^{n} w_i = 1$.

Note that $f(X)$ is a Bernoulli density with probability of existence r and whose single target density is a mixture of weighted Dirac delta functions (e.g. as in a particle filter implementation [21]). The probability of target existence remains fixed throughout (i.e. it is assumed that the probability of survival is unity). It is also noted that x_i ($i \geq 1$) will generally depend on t (unless the target is stationary). For brevity, this dependency is omitted from the notation.

In applications in which a particle filter is not used to estimate the target state(s), the hypotheses represent samples drawn from the target probability density functions (in tracking applications), or samples drawn from a region under surveillance (in search applications).

C. Calculation of the MS-GOSPA Error at Each Time Step

The MS-GOSPA error at time t (given by equation (8)) can be written in terms of the target hypotheses as follows:

$$\begin{aligned} &\text{MS-GOSPA}(\hat{X}_t; z_{1:t}, a_{1:t}) \qquad (19) \\ &= \sum_{i=0}^{n} \left(d_2^{(c,2)}(x_i, \hat{X}_t(z_{1:t}, a_{1:t})) \right)^2 p(x_i | z_{1:t}, a_{1:t}) \end{aligned}$$

The probability $p(x_i | z_{1:t}, a_{1:t})$ can be written as follows (for $i \geq 0$):

$$\begin{aligned} p(x_i | z_{1:t}, a_{1:t}) &= \frac{p(z_{1:t} | x_i, a_{1:t}) p(x_i)}{\gamma_t} & (20) \\ &= \frac{p(x_i)}{\gamma_t} \prod_{j=1}^{t} p(z_j | x_i, a_j) & (21) \end{aligned}$$

where $p(x_i)$ denotes the prior probability of hypothesis i being true, i.e.:

$$p(x_i) = \begin{cases} r w_i & \text{for } i > 0 \\ (1 - r) & \text{otherwise} \end{cases} \tag{22}$$

and γ_t is a normalising constant that ensures that:

$$\sum_{i=0}^{n} p(x_i | z_{1:t}, a_{1:t}) = 1$$

[6]The generalisation to scenarios to multiple targets is straightforward, provided that the targets are well separated. Indeed, in [18], it was shown that the AMMS-GOSPA was additive across the Bernoulli components for well separated targets. In scenarios with targets that are in close proximity to each other, the analytical results presented later in Section IV-D no longer hold, and the AMMS-GOSPA error must also consider the optimal assignment between the hypothesised targets and the target state estimates.

D. Analytical Calculation of the MMS-GOSPA Error

Based on the measurements $z_{1:t}$ and action sequence $a_{1:t}$, the following two target state estimates $\hat{X}_t$ are considered:

$$\begin{aligned} \hat{X}_t &= \phi \qquad \text{(i.e. no target is present)} & (23) \\ \hat{X}_t &= \sum_{i=1}^{n} w_{1:t}^i x_i \triangleq \hat{X}_e & (24) \end{aligned}$$

where $w_{1:t}^i$ are the posterior hypothesis weights based on the measurements $z_{1:t}$ and action sequence $a_{1:t}$, given as follows (for $i \geq 1$):

$$\begin{aligned} w_{1:t}^i &\propto p(x_i | z_{1:t}, a_{1:t}) & (25) \\ &= \frac{p(z_{1:t} | x_i, a_{1:t}) p(x_i)}{\lambda_t} & (26) \\ &= \frac{p(x_i)}{\lambda_t} \prod_{j=1}^{t} p(z_j | x_i, a_j) & (27) \end{aligned}$$

and λ_t is a normalising constant that ensures that $\sum_{i=1}^{n} w_{1:t}^i = 1$. It then follows from equations (21) and (27) that:

$$p(x_i | z_{1:t}, a_{1:t}) = \lambda_t w_{1:t}^i / \gamma_t \tag{28}$$

Summing equation (28) over $i = 1, \dots, n$, it follows that λ_t / γ_t gives the posterior probability that a target is not present.

To proceed, it is straightforward to show the following:

$$d_2^{(c,2)}(x_0, \phi) = 0 \tag{29}$$

(i.e. no target is present or estimated to be present)

$$d_2^{(c,2)}(x_i, \phi) = c/\sqrt{2} \qquad \text{for } i > 0 \tag{30}$$

(i.e. the target is missed, cardinality error = 1)

$$d_2^{(c,2)}(x_0, \hat{X}_e) = c/\sqrt{2} \tag{31}$$

(i.e. a target is incorrectly estimated to be present, cardinality error = 1)

$$d_2^{(c,2)}(x_i, \hat{X}_e) = \min\left(d(x_i, \hat{X}_e), c\right) \tag{32}$$

where $d(x_i, \hat{X}_e)$ is the distance between the hypothesis x_i and the estimate $\hat{X}_e$, and c is the cut-off. It is noted that $d_2^{(c,2)}()$ is actually a distance between sets, so one could write $\{x_i\}$ instead of x_i, and $\{\hat{X}_e\}$ instead of $\hat{X}_e$ etc. The simplified notation is used for brevity.

The square of the truncated distance error in equation (32) can be written as follows:

$$\begin{aligned} \left[d_2^{(c,2)}(x_i, \hat{X}_e)\right]^2 &= i_d(x_i, \hat{X}_e) d(x_i, \hat{X}_e)^2 & (33) \\ &\quad + \left(1 - i_d(x_i, \hat{X}_e)\right) c^2 \end{aligned}$$

where $i_d(x_i, \hat{X}_e) = 1$ if $d(x_i, \hat{X}_e) \leq c$, and $i_d(x_i, \hat{X}_e) = 0$ otherwise.

It then follows from equations (19) and (29) – (30) that:

$$\begin{aligned} &\text{MS-GOSPA}(\hat{X}_t = \phi; z_{1:t}, a_{1:t}) \\ &= \frac{c^2}{2} \sum_{i=1}^{n} p(x_i | z_{1:t}, a_{1:t}) & (34) \\ &= \frac{c^2}{2} \left(1 - p(x_0 | z_{1:t}, a_{1:t})\right) & (35) \end{aligned}$$

Equation (35) gives the squared cost of a cardinality error, weighted by the posterior probability that a target is present, based on the measurements $z_{1:t}$ and action sequence $a_{1:t}$.

To continue, from equations (19) and (31) – (33) it follows that:

$$\begin{aligned} &\text{MS-GOSPA}(\hat{X}_t = \hat{X}_e; z_{1:t}, a_{1:t}) \\ &= \frac{c^2}{2} p(x_0|z_{1:t}, a_{1:t}) + \sum_{i=1}^{n} \Big[i_d(x_i, \hat{X}_e) d(x_i, \hat{X}_e)^2 \\ &\quad + \big(1 - i_d(x_i, \hat{X}_e) c^2\big) \Big] p(x_i|z_{1:t}, a_{1:t}) \end{aligned} \tag{36}$$

This can be written as follows:

$$\begin{aligned} &\text{MS-GOSPA}(\hat{X}_t = \hat{X}_e; z_{1:t}, a_{1:t}) \\ &= \frac{c^2}{2} p(x_0|z_{1:t}, a_{1:t}) \\ &+ \big(1 - p(x_0|z_{1:t}, a_{1:t})\big) \left(\sum_{l=1}^{2} \hat{\mathcal{V}}_l(x) + c^2 T_{w_{1:t}}(\hat{X}_e) \right) \end{aligned} \tag{37}$$

where:

$$\begin{aligned} \hat{\mathcal{V}}_l(x) &= \mathcal{E}_{\underline{w}_{1:t}}\left[x(l)^2\right] + \mathcal{E}_{w_{1:t}}\left[x(l)\right]^2 \\ &\quad -2\mathcal{E}_{\underline{w}_{1:t}}\left[x(l)\right] \mathcal{E}_{w_{1:t}}\left[x(l)\right] \end{aligned} \tag{38}$$

$$\underline{w}_{1:t}^i = w_{1:t}^i i_d(x_i, \hat{X}_e) \quad \text{for } i = 1, \ldots, n \tag{39}$$

$$\mathcal{E}_{\underline{w}_{1:t}}\left[x(l)^m\right] = \sum_{i=1}^{n} \underline{w}_{1:t}^i x_i(l)^m \quad \text{for } m = 1, 2 \tag{40}$$

$$T_{w_{1:t}}(\hat{X}_e) = \left(1 - \sum_{i=1}^{n} i_d(x_i, \hat{X}_e) w_{1:t}^i \right) \tag{41}$$

where $\hat{\mathcal{V}}_l(x)$ denotes the cross-covariance of the (posterior) weighted samples of the l-th dimension (denoted $x(l)$) of the target state (herein $x(1)$ and $x(2)$ denote the x- and y- coordinates of the target state respectively), based on the measurements $z_{1:t}$ and action sequence $a_{1:t}$. $T_{w_{1:t}}(\hat{X}_e)$ denotes the sum of weighted samples for which the estimation error exceeds c, given that a target exists.

The first term on the right-hand side of equation (37) is the squared cost of a cardinality error, multiplied by the posterior probability that a target is not present. The second term is the average estimation error squared (adjusted for a maximum estimation error squared of c^2) of the posterior estimate $\hat{X}_e$, multiplied by the posterior probability that a target is present.

The MMS-GOSPA$(z_{1:t}, a_{1:t})$ is then given by the minimum of (35) and (37).

E. Target State Estimation in the Presence of Clutter

In the presence of clutter, the expected likelihood particle filter (ELPF) approach is used [23] to determine the posterior hypothesis weights and target state estimates. These are used in the calculation of the performance metrics for the three (i.e. optimal, suboptimal, and baseline) control approaches. The ELPF approach is analogous to the probabilistic data association Kalman filter (e.g. see [24]).

If the target is within the field-of-view (FOV) of the sensor, it is detected with probability P_d. Conditional on target hypothesis $i > 0$, each target generated measurement is sampled from a Gaussian distribution with mean x_i and covariance Σ. It is assumed that the number of false alarms at each sampling time has a Poisson distribution with mean $\lambda_{FA} V$, where λ_{FA} is the false alarm rate per unit volume of the observation region, and V is the volume of the FOV of the sensor. False alarms are uniformly distributed within the FOV of the sensor.

Excusing an abuse of notation, let $Z = \{z_1, \ldots, z_N\}$ denote a vector of N measurements (i.e. a maximum of one target generated measurement plus false alarms) at any given sampling time. Using the ELPF approach, the measurement likelihood (used to calculate the posterior hypothesis weights, and the subsequent posterior target state estimate), for $N > 0$, is given as follows [23]:

$$p(Z|x_i, a) \propto \begin{cases} \lambda_{FA}(1 - P_d) + P_d \sum_{k=1}^{N} \mathcal{N}(z_k; x_i, \Sigma) \\ \quad \text{if } i > 0 \text{ and } x_i \in \text{FOV}(a) \\ \lambda_{FA} \\ \quad \text{otherwise} \end{cases} \tag{42}$$

where $\mathcal{N}(z_k; x_i, \Sigma)$ denotes the Gaussian probability density function evaluated at z_k for a distribution with mean x_i and error covariance Σ. For $N = 0$, $Pr(N = 0|x_i, a) \propto (1 - P_d)$ if $i > 0$ and $x_i \in$ FOV(a); and $Pr(N = 0|x_i, a) \propto 1$ otherwise. The posterior hypothesis weights at each sampling time t are then given by equation (27), from which the posterior target state estimate can be determined via equation (24).

V. Sample-Based Estimation of the AMMS-GOSPA Error

A. General Sampling Approach

The AMMS-GOSPA at time t can be written as follows:

$$\begin{aligned} &\text{AMMS-GOSPA}(a_{1:t}) \\ &= \int_{z_{1:t}} \text{MMS-GOSPA}(z_{1:t}, a_{1:t}) p(z_{1:t}; a_{1:t}) dz_{1:t} \qquad (43) \\ &= \int_X \int_{z_{1:t}} \text{MMS-GOSPA}(z_{1:t}, a_{1:t}) \qquad (44) \\ &\quad \times p(z_{1:t}|X, a_{1:t}) p(X) dz_{1:t} dX \\ &= \sum_{i=0}^{n} p(x_i) \int_{z_{1:t}} \text{MMS-GOSPA}(z_{1:t}, a_{1:t}) \qquad (45) \\ &\quad \times p(z_{1:t}|x_i, a_{1:t}) dz_{1:t} \end{aligned}$$

where $p(x_i)$ is given by equation (22).

A sample-based approximation is then given as follows:

$$\begin{aligned} &\text{AMMS-GOSPA}(a_{1:t}) \\ &\approx \frac{1}{m} \sum_{i=0}^{n} p(x_i) \sum_{l=1}^{m} \text{MMS-GOSPA}(z_{1:t}^l(x_i), a_{1:t}) \end{aligned} \tag{46}$$

where $z_{1:t}^l(x_i)$ are m measurement samples, conditional on the target hypothesis x_i. In the most general case there can be multiple measurements per time step, which can include both target generated measurements and false alarms.

B. Efficient Sampling Approach – Conditioning on the Measurement Sequence

In this section, the sample-based AMMS-GOSPA error is simplified to take into account the sequence of target detections and missed detections.

$$\text{AMMS-GOSPA}(a_{1:t}) \approx \frac{1}{n_h}\sum_{i=0}^{n} p(x_i)\sum_{m=1}^{2^t} Pr\big(s_{1:t}(m)|x_i,a_{1:t}\big)\sum_{j=1}^{n_h}\text{MMS-GOSPA}\big(z^{ij}_{1:t}(m),a_{1:t}\big) \tag{47}$$

The measurement vector at each time step can include a maximum of one target generated measurement plus false alarms. For notational brevity and ease of understanding, the derivations are presented for a single sensor scenario. However, the approach can be readily applied to scenarios in which there are multiple sensors.

Let $s_{1:t}(m) = (s_1(m), \ldots, s_t(m))$ denote a sequence of target detections and missed detections at times 1 through t, with $s_k(m) = 0$ or 1. The number of potential detection sequences is 2^t. $Pr\big(s_{1:t}(m)|x_i, a_{1:t}\big)$ is the probability of the sequence $s_{1:t}(m)$ occurring, conditional on the target hypothesis i and actions $a_{1:t}$. This is given as follows:

$$Pr\big(s_{1:t}(m)|x_i,a_{1:t}\big) = \prod_{k=1}^{t} Pr(s_k(m)|x_i,a_k) \tag{48}$$

where:

$$\begin{aligned}
&\text{if } i=0, \text{ or } i>0 \text{ and } x_i \notin \text{FOV}(a_k):\\
&\quad Pr(s_k(m)=0|x_i,a_k)=1\\
&\text{if } i=0, \text{ or } i>0 \text{ and } x_i \notin \text{FOV}(a_k):\\
&\quad Pr(s_k(m)=1|x_i,a_k)=0\\
&\text{if } i>0 \text{ and } x_i \in \text{FOV}(a_k):\\
&\quad Pr(s_k(m)|x_i,a_k)=P_d^{s_k(m)}(1-P_d)^{1-s_k(m)}\\
&\qquad \text{for } s_k(m)=\{0,1\}
\end{aligned} \tag{49}$$

and P_d is again the probability of detection on each sensor scan for a the target that is within the FOV of the sensor. Equation (45) can then be approximated as shown in equation (47), where n_h is the number of measurement samples per target hypothesis and measurement sequence pair.

The measurement sequences $z^{ij}_{1:t}(m) \triangleq \Big(z^{ij}_1(m), .., z^{ij}_t(m)\Big)$, $i = 0, \ldots, n$, $j = 1, \ldots, n_h$ and $m = 1, \ldots, 2^t$ can include both target generated measurements and false measurements, with a target generated measurement occurring at time k only if $s_k(m) = 1$.

In the approximation (47), the computation is simplified by the fact that the summations need only be performed for detection sequences for which $Pr\big(s_{1:t}(m)|x_i, a_{1:t}\big) > 0$. For example, for $i = 0$ the only possible sequence has no target detection on any time step (because there is no target under this hypothesis). Equally, for $i > 0$, sequences for which $s_k = 1$ when $x_i \notin \text{FOV}(a_k)$ can also be discounted (i.e. a target cannot be detected if it is outside of the sensor FOV).

Therefore, the total number of samples of $z_{1:t}$ used in calculating the AMMS-GOSPA via equation (47) is $(2^t n + 1)n_h$, and increases exponentially with the number of time steps t.

C. Computational Complexity

On first viewing, the approximation (47) would appear to be significantly more complex, and require a greater number of samples, than the approximation (46). However, the approximation (47) accounts for all potential sequences of target detections and missed detections, whereas the approximation (46) does not. Consequently, the number of samples m required in the approximation (46) is significantly greater than the number of samples n_h in the approximation (47). This is because the approximation (46) must contain enough samples to account for the target probability of detection.

For example, if the measurements are extremely accurate, it is feasible to allow $n_h = 1$ in (47), whereas $m > 100$ is required for an accurate estimate using (46), even in the simplest case with $t = 1$. If measurements are less accurate, a large number of samples n_h may be required irrespective of the approximation used. This significantly increases the computational expense of the optimisation algorithm, which is already computationally expensive for $t > 1$.

D. Demonstration

This scenario is identical to analysis I in [18]. There is one potential target location hypothesis x_1 (therefore, $n = 1$), and just a single time step (i.e. $t = 1$). If the sensor observes the potential location of the target, it generates a perfect measurement of the target state[7] with probability $P_d = 0.6$. There are no false alarms. The cost of a cardinality error is $c = 10$km. It is noted that the units of cardinality errors and sensing costs are both kilometres, in order to balance the units with the distance metric $d(x_i, \hat{X}_e)$, which is measured in kilometres.

There are two possibilities for the target state estimate, $\hat{X}(z, a)$:

$$\begin{aligned}
\hat{X}(z,a) &= \phi \quad \text{(no target is present)}\\
\hat{X}(z,a) &= x_1 \quad \text{(the potential target location)}
\end{aligned} \tag{50}$$

There are also two possible actions a, these being “do not attempt an observation/measurement” and “observe the potential target location”.

The analytical solution determined in [18] is compared to the solutions generated via the following three sample-based approaches:

1. **Approach 1:** The general sampling approach described in Section V-A, with the MMS-GOSPA for each sample calculated via equation (4).
2. **Approach 2:** The general sampling approach, exploiting the analytical calculations of the MMS-GOSPA provided in Section IV-D.
3. **Approach 3:** The efficient sampling approach described in Section V-B, exploiting the analytical calculations of the MMS-GOSPA and the known probability of detection.

The results are shown in Table I and Figure 1, with a range of values of the sensing cost $s \in (0, 20]$ km (which are added to the AMM-GOSPA error) and $r \in [0, 1]$ evaluated.

[7] In [18], the measurement probability density function was specified via the Dirac delta function. Herein, the measurement has a Gaussian distribution with mean x_1 and error covariance $\Sigma = \text{diag}(10^{-10}, 10^{-10})$.

TABLE I

RUN-TIME PER OPTIMISATION (AVERAGED ACROSS EACH COMBINATION OF r AND s THAT IS CONSIDERED) AS A FUNCTION OF THE NUMBER OF SAMPLES. **(A):** GENERAL SAMPLING APPROACH. **(B):** AS IN (A), BUT EXPLOITING THE ANALYTICAL RESULTS OF SECTION IV-D. **(C):** THE EFFICIENT SAMPLING APPROACH AGAIN EXPLOITING THE ANALYTICAL RESULTS. ALL COMPUTATIONS WERE PERFORMED USING AN INTEL® CORE™ I7-8750H (2.6GHZ) PROCESSOR.

(A)		**(B)**		**(C)**	
Samples (m)	**Run-time (ms)**	**Samples** (m)	**Run-time (ms)**	**Samples** (n_h)	**Run-time (ms)**
10	0.3	10	0.1	1	0.08
100	2	100	0.3		
1,000	20	1,000	3		

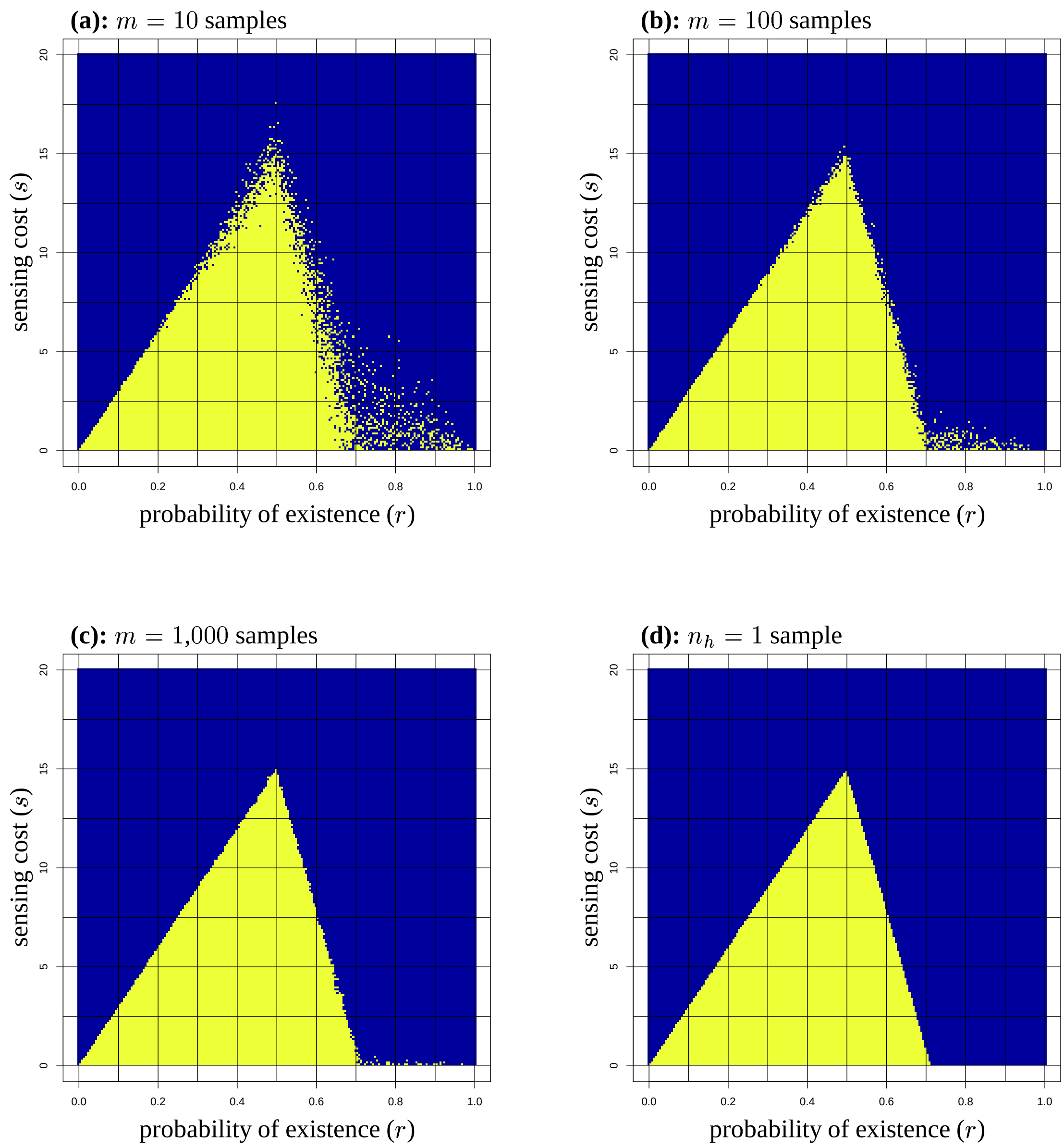

Fig. 1. Optimal action for each sensing cost and probability of target existence. Key: blue regions: do not make an observation/measurement, yellow regions: observe the potential target location. In **(a) – (c):** the general sampling approach of Section V-A are used. In **(d):** the efficient sampling approach of Section V-B is used.

The solutions generated using the efficient sampling approach shown in Figure 1(d) are identical to the analytical results shown in Figure 2 in [18][8]. Moreover, these solutions require only a single sample (i.e. $n_h = 1$).

However, the general sampling approaches 1–2 generate solutions that can be different at the decision boundary (see Figures 1(a)–(c)), particularly if a relatively small number of samples are used. Furthermore, the general sampling approach that calculates the MMS-GOSPA from first principles, via equation (4) (i.e. Approach 1) has a computational time circa 250 times greater than the efficient sampling approach, in order to determine comparable (but still occasionally sub-optimal) solutions (e.g. compare Table I(A): $m = 1000$ with Table I(C): $n_h = 1$).

VI. Non-Myopic (Multi-Step) Sensor Planning

A. *Suboptimal Control Approach*

A.1 Motivation

In this section, a suboptimal planning approach is introduced. The primary reason it is suboptimal is that due to the way in which averaging across measurements is performed, the action calculated at each sampling time depends only on the actions at the previous sampling times, and is independent of the measurements generated at those times. Expressing this another way, the first action does not anticipate how the measurements that are subsequently generated will impact on future actions that are dependent on the realisation of the measurements.

This formulation is shown later (in Section VI-B) to generate a solution with an overall cost that is an upper bound on the overall cost of a solution calculated via a second recursive formulation that conditions on previous measurements. The second formulation is referred to as the "optimal control approach".

A.2 Overall Cost Function

The cost function is the AMMS-GOSPA error, and it is assumed that there are no sensing costs. Let:

$$r_t(z_{1:t}, a_{1:t}) \triangleq \text{MMS-GOSPA}(z_{1:t}, a_{1:t}) \quad (51)$$

$$V_{t:T}(a_{1:T}) \triangleq \mathbb{E}_{z_{t:T}\,|\,z_{1:t-1}} \left[\sum_{l=t}^{T} \lambda^{l-1} r_l(z_{1:l}, a_{1:l}) \right] \quad (52)$$

where MMS-GOSPA$(z_{1:t}, a_{1:t})$ is given by the minimum of (35) and (37). The overall cost function of the actions $a_{1:T}$ is then given as follows:

$$C(a_{1:T}) = \text{AMMS-GOSPA}_{1:T}(a_{1:T}) \quad (53)$$

It then follows from equation (15) that $C(a_{1:T}) = V_{1:T}(a_{1:T})$.

A.3 Optimisation

An action $a_1^\star$ at each time step can be determined to minimise the overall cost function $C(a_{1:T})$, i.e.:

$$a_1^\star = \arg\min_{a_1} \left[\min_{a_{2:T}} \left[C(a_{1:T}) \right] \right] \quad (54)$$

The minimum value of the cost function $C(a_{1:T})$ can then be written as follows:

$$\min_{a_{1:T}} \left[C(a_{1:T}) \right] = \min_{a_{1:T}} \left[V_{1:T}(a_{1:T}) \right] \triangleq V_{1:T}^\star \quad (55)$$

This is a commonly used approach (e.g. see [2, equation (24)] and [3, equation (28)]).

It is noted that each action a_l, $l = 2, \ldots, T$ is dependent on the previous actions $a_1, \ldots, a_{l-1}$ but is independent of the potential measurements, due to the expectation over $z_{1:T}$ in the cost function $C(a_{1:T})$. It is also noted that actions a_l, $l = 2, \ldots, T$ may never be realised, because at the next time step, the situation awareness picture will be updated (it is dependent on the measurement generated following the action $a_1^\star$), and the optimisation (54) will be repeated for the new sliding window of T time steps. This is the basis of standard receding horizon planning (e.g. see [22]).

B. *Optimal Control Approach*

B.1 Formulation

$V_{t:T}^\star$ can be written as shown in equations (60) – (62). The inequalities (63) – (66) are then true, via interchanging $\min_{a_{t+1:T}}$ and $\mathbb{E}_{z_t\,|\,z_{1:t-1}}$ in the second term of equation (62)[9].

Inequalities (65) and (66) (of Bellman type, e.g. see [20]) follow by recursion of inequality (63). For the suboptimal control approach, $C_{1:T}^\star = V_{1:T}^\star$. Clearly this cost in an upper bound on the Bellman type equation (66) for all values of T.

Motivated by this finding, the "optimal control approach" determines each action $\hat{a}_1^\star$ via equation (68). It is important to note that each minimisation (i.e. minimum with regard to a_t) in equation (68) is conditional on both the previous actions $a_{1:t-1}$ and the previous measurements $z_{1:t-1}$.

As an illustrative example of the approach, for $T = 2$:

$$\hat{a}_1^\star = \arg\min_{a_1} \mathbb{E}_{z_1} \Big[r_1(z_1, a_1) + \lambda \min_{a_2} \mathbb{E}_{z_2\,|\,z_1} \left[r_2(z_{1:2}, a_{1:2}) \right] \Big] \quad (56)$$

where:

$$\hat{a}_2^\star(z_1, a_1) = \arg\min_{a_2} \mathbb{E}_{z_2\,|\,z_1} \left[r_2(z_{1:2}, a_{1:2}) \right] \quad (57)$$

$$\hat{V}_{2:2}^\star = \lambda \min_{a_2} \mathbb{E}_{z_2\,|\,z_1} \left[r_2(z_{1:2}, a_{1:2}) \right] \quad (58)$$

$$\hat{V}_{1:2}^\star = \min_{a_1} \mathbb{E}_{z_1} \left[r_1(z_1, a_1) + \hat{V}_{2:2}^\star \right] \quad (59)$$

When using the ideal measurement approach (but allowing $P_d < 1$), the optimal and suboptimal approaches generate identical actions, i.e. $a_1^\star = \hat{a}_1^\star$. This follows from the fact that in this case, the MMS-GOSPA error at each time t is only non-zero if measurements are not generated at times 1 through t (see Appendix for full details).

[8]The analytical results demonstrate that it is optimal to attempt a measurement if the probability of existence is not too low or too high, with the decision boundary dependent on the magnitude of the sensing cost s. For very high sensing costs, it is never optimal to attempt a measurement. Conversely. for $s = 0$ (not shown, as $s \geq 0.1$ in Figure 1), is is optimal to always attempt a measurement irrespective of the existence probability.

[9]These inequalities follow from the fact that $\min_{a_{t+1:T}} \mathbb{E}_{z_t\,|\,z_{1:t-1}} [\Phi] \geq \mathbb{E}_{z_t\,|\,z_{1:t-1}} \left[\min_{a_{t+1:T}} [\Phi] \right]$ with equality only if the function Φ is independent of the variable z_t. Putting it another way, the minimum of a sum is greater than the sum of the individual minimums.

$$V^\star_{t:T} = \min_{a_{t:T}} \mathbb{E}_{z_t\,|z_{1:t-1}} \left[\mathbb{E}_{z_{t+1:T}\,|z_{1:t}} \left[\sum_{l=t}^{T} \lambda^{l-1} r_l(z_{1:l}, a_{1:l}) \right]\right] \qquad \text{(using Bayes' rule)} \tag{60}$$

$$= \min_{a_{t:T}} \mathbb{E}_{z_t\,|z_{1:t-1}} \left[\lambda^{t-1} r_t(z_{1:t}, a_{1:t}) + \mathbb{E}_{z_{t+1:T}\,|z_{1:t}} \left[\sum_{l=t+1}^{T} \lambda^{l-1} r_l(z_{1:l}, a_{1:l}) \right]\right] \tag{61}$$

(using the fact that $r_t(z_{1:t}, a_{1:t})$ is conditional only on $z_{1:t}$ and $a_{1:t}$)

$$= \min_{a_t} \left\{ \mathbb{E}_{z_t\,|z_{1:t-1}} \left[\lambda^{t-1} r_t(z_{1:t}, a_{1:t})\right] + \min_{a_{t+1:T}} \left\{ \mathbb{E}_{z_t\,|z_{1:t-1}} \left[\mathbb{E}_{z_{t+1:T}\,|z_{1:t}} \left[\sum_{l=t+1}^{T} \lambda^{l-1} r_l(z_{1:l}, a_{1:l}) \right]\right] \right\} \right\} \tag{62}$$

(using the fact that $\mathbb{E}_{z_t\,|z_{1:t-1}} \left[r_t(z_{1:t}, a_{1:t})\right]$ is independent of $a_{t+1:T}$)

$$V^\star_{t:T} \geq \min_{a_t} \mathbb{E}_{z_t\,|z_{1:t-1}} \left[\lambda^{t-1} r_t(z_{1:t}, a_{1:t}) + \min_{a_{t+1:T}} \mathbb{E}_{z_{t+1:T}\,|z_{1:t}} \left[\sum_{l=t+1}^{T} \lambda^{l-1} r_l(z_{1:l}, a_{1:l}) \right]\right] \tag{63}$$

$$= \min_{a_t} \mathbb{E}_{z_t\,|z_{1:t-1}} \left[\lambda^{t-1} r_t(z_{1:t}, a_{1:t}) + V^\star_{t+1:T} \right] \tag{64}$$

$$\geq \min_{a_t} \mathbb{E}_{z_t\,|z_{1:t-1}} \left[\lambda^{t-1} r_t(z_{1:t}, a_{1:t}) + \min_{a_{t+1}} \mathbb{E}_{z_{t+1}\,|z_{1:t}} \left[\lambda^{t} r_{t+1}(z_{1:t+1}, a_{1:t+1}) + V^\star_{t+2:T} \right]\right] \tag{65}$$

$$\vdots$$

$$\geq \lambda^{t-1} \min_{a_t} \mathbb{E}_{z_t\,|z_{1:t-1}} \Bigg[r_t(z_{1:t}, a_{1:t}) + \lambda \min_{a_{t+1}} \mathbb{E}_{z_{t+1}\,|z_{1:t}} \Bigg[r_{t+1}(z_{1:t+1}, a_{1:t+1}) \tag{66}$$

$$+\lambda \min_{a_{t+2}} \mathbb{E}_{z_{t+2}\,|z_{1:t+1}} \Big[r_{t+2}(z_{1:t+2}, a_{1:t+2}) + \ldots + \lambda \min_{a_T} \mathbb{E}_{z_T\,|z_{1:T-1}} \left[r_T(z_{1:T}, a_{1:T}) \right]\Big]\Bigg]\Bigg]$$

$$\triangleq \hat{V}^\star_{t:T} \qquad (= \text{optimal overall cost}) \tag{67}$$

$$\hat{a}^\star_1 = \arg\min_{a_1} \mathbb{E}_{z_1} \Bigg[r_1(z_1, a_1) + \lambda \min_{a_2} \mathbb{E}_{z_2\,|z_1} \Bigg[r_2(z_{1:2}, a_{1:2}) \tag{68}$$

$$+\lambda \min_{a_3} \mathbb{E}_{z_3\,|z_{1:2}} \Big[r_3(z_{1:3}, a_{1:3}) + \ldots + \lambda \min_{a_T} \mathbb{E}_{z_T\,|z_{1:T-1}} \left[r_T(z_{1:T}, a_{1:T}) \right]\Big]\Bigg]\Bigg]$$

B.2 Calculating the Conditional AMMS-GOSPA Error

The optimal control approach requires calculation of the AMMS-GOSPA at each time $t+1$, conditional on the previous measurements $z_{1:t}$ and all actions $a_{1:t+1}$. Similar to the calculation of the (unconditional) AMMS-GOSPA in Section V, the conditional AMMS-GOSPA is given as follows:

$$\mathbb{E}_{z_{t+1}\,|z_{1:t}} \left[\text{MMS-GOSPA}(z_{1:t+1}, a_{1:t+1})\right]$$

$$= \int_{z_{t+1}} \text{MMS-GOSPA}(z_{1:t+1}, a_{1:t+1}) \times p(z_{t+1}|z_{1:t}, a_{1:t+1}) dz_{t+1} \tag{69}$$

$$= \sum_{i=0}^{n} \int_{z_{t+1}} \text{MMS-GOSPA}(z_{1:t+1}, a_{1:t+1}) \times p(z_{t+1}|x_i, z_{1:t}, a_{1:t+1}) p(x_i|z_{1:t}, a_{1:t}) dz_{t+1} \tag{70}$$

$$= \sum_{i=0}^{n} \hat{w}_i \int_{z_{t+1}} \text{MMS-GOSPA}(z_{1:t+1}, a_{1:t+1}) \times p(z_{t+1}|x_i, a_{t+1}) dz_{t+1} \tag{71}$$

where $\hat{w}_i = p(x_i|z_{1:t}, a_{1:t})$ is the posterior probability of each target hypothesis (including "no target present"), conditional on the previous measurements $z_{1:t}$ and previous actions $a_{1:t}$. Equation (71) is approximated as in equation (47) but only requires samples to be generated from $p(z_{t+1}|x_i, a_{t+1})$, This approximation of the conditional AMMS-GOSPA error is given as follows:

$$\mathbb{E}_{z_{t+1}\,|z_{1:t}} \left[\text{MMS-GOSPA}(z_{1:t+1}, a_{1:t+1})\right] \approx \frac{1}{n_h} \left\{ \sum_{i=0}^{n} \hat{w}_i \sum_{m=0}^{1} Pr\big(s_{t+1} = m | x_i, a_{t+1}\big) \times \sum_{j=1}^{n_h} \text{MMS-GOSPA}\big(z_{1:t}, z^{ij}_{t+1}(m), a_{1:t+1}\big) \right\} \tag{72}$$

where to remind the reader, $s_{t+1} = 1$ denotes a target detection at sampling time $t+1$. $Pr(s_{t+1} = m|x_i, a_{t+1})$ is the probability that $m = 0, 1$ target measurements are generated on the $(t+1)$-th time step, conditional on the target state being given by x_i and action a_{t+1}. $Pr\big(s_{t+1}(m)|x_i, a_{t+1}\big)$ is again calculated via equation (49). Each measurement sample $z^{ij}_{t+1}(m)$ can be a vector of multiple measurements, which includes a maximum of one target generated measurement (sampled from $\mathcal{N}(x_i, \Sigma)$ if $s_{t+1} = 1$) plus false alarms.

$$\hat{a}_1^\star = \arg\min_{a_1} \frac{1}{n_h}\sum_{j=1}^{n_h}\sum_{m=0}^{1}\sum_{i=0}^{n} p(x_i)Pr(s_1=m|x_i,a_1)\Big\{r_1(z_1^{ij}(m),a_1)+\lambda\min_{a_2}\mathbb{E}_{z_2|z_1^{ij}(m)}\Big[r_2(z_1^{ij}(m),z_2,a_{1:2}))\Big]\Big\} \quad (73)$$

$$= \arg\min_{a_1} \frac{1}{n_h}\sum_{j=1}^{n_h}\sum_{m=0}^{1}\sum_{i=0}^{n} p(x_i)Pr(s_1=m|x_i,a_1)\Big\{r_1(z_1^{ij}(m),a_1) \quad (74)$$

$$+\min_{a_2}\frac{\lambda}{n_h}\sum_{\underline{j}=1}^{n_h}\sum_{\underline{m}=0}^{1}\sum_{\underline{i}=0}^{n} p(x_i|z_1^{ij}(m))Pr(s_2=\underline{m}|x_{\underline{i}},a_2)r_2\big(z_1^{ij}(m),z_2^{\underline{ij}}(\underline{m}),a_{1:2}\big)\Big\}$$

$$\hat{V}_{1:2}^\star = \frac{1}{n_h}\sum_{j=1}^{n_h}\sum_{m=0}^{1}\sum_{i=0}^{n} p(x_i)Pr(s_1=m|x_i,a_1)\Big\{r_1(z_1^{ij}(m),a_1^\star)+\lambda\mathbb{E}_{z_2|z_1^{ij}(m)}\Big[r_2(z_1^{ij}(m),z_2,a_1^\star,a_2^\star\big(a_1^\star,z_1^{ij}(m)\big))\Big]\Big\} \quad (75)$$

$$\bar{\mathcal{MSE}}_{1:2}(a_{1:2}) = \frac{1}{n_h}\sum_{j=1}^{n_h}\sum_{m=0}^{1}\sum_{i=1}^{n} w_i Pr(s_1=m|x_i,a_1)\Big\{\mathcal{SE}_1(z_1^{ij}(m),a_1,i) \quad (76)$$

$$+\frac{\lambda}{n_h}\sum_{\underline{j}=1}^{n_h}\sum_{\underline{m}=0}^{1}\sum_{\underline{i}=1}^{n} p(x_i|z_1^{ij}(m),x_i\neq x_0)Pr(s_2=\underline{m}|x_{\underline{i}},a_2)\mathcal{SE}_2(z_1^{ij}(m),z_2^{\underline{ij}}(\underline{m}),a_{1:2},\underline{i})\Big\}$$

B.3 Implementation of the Optimal Control Approach

Using the efficient sampling approach, for two-step planning, the optimal action $a_1^\star$ (given by equation (68)) satisfies equations (73) – (74), where r_t denotes the cost at time t conditional on the actions $a_{1:t}$ and measurements $z_{1:t}$ generated at times 1 through t. $Pr(s_1=m|x_i,a_1)$ is the probability that $m=0$ or 1 target measurements are generated on the first time step, conditional on the target state being given by x_i and action a_1. This is calculated via equation (49). The measurement samples $z_1^{ij}(m)$ include a maximum of one target generated measurement (sampled from $\mathcal{N}(x_i,\Sigma)$) plus a Poisson distributed number of false alarms.

The corresponding optimal total cost incurred (the AMMS-GOSPA error, given by equations (66) – (67)) is then given in equation (75). Equations (73) – (75) generalise in an obvious manner for $T>2$.

B.4 Computational Complexity

The optimisation (73) is computationally expensive, because for each potential action a_1, the minimisation of each second time step action a_2 must be performed for each sampled first time step measurement $z_1^{ij}(m)$. Hence, the number of minimisations that must be performed at the second time step is $(2n+1)n_h$ for each potential action a_1 (i.e. $2n_h$ minimisations corresponding to hypotheses $i>0$; and n_h minimisations for $i=0$, because only $s_1=0$ is possible for $i=0$). Hence the total number of minimisations is $n_a(2n+1)n_h$, where n_a is the number of possible actions. This generally makes the approach computationally prohibitive for time horizons greater than two time steps[10], unless the algorithm is parallelised.

By way of comparison, the suboptimal approach performs just a single minimisation (given by equation (54)), with n_a^T potential combinations of actions for $T\geq 1$. Hence, for $T=2$ the computational expense of the suboptimal algorithm is typically much lower (assuming that $n_a \ll (2n+1)n_h$), though it should be noted that in the suboptimal approach, the cost function in each minimisation is computationally more expensive as it is the sum total of the costs incurred across all time steps.

Special case – If $\lambda_{FA}=0$ and $\Sigma\approx 0$ it can easily be shown that $Pr(s_1=1|x_i,a_1)r_1(z_1^{ij}(1),a)=0$ for all values of i. This is due to the fact that whenever a measurement is generated by a target hypothesis, $r_1(z_1^{ij}(1),a)=0$, as a target can be inferred to be present without geo-location or cardinality errors. As a result, in the optimisation (74) one can set $n_h=1, m=\underline{m}=0$ and $z_1=z_2=\phi$. Consequently, the minimisation on the second time step need only to be performed if no measurement is generated on the first time step, and the optimisation (74) reduces to equation (78) in the Appendix. If a measurement is generated, it is not necessary to attempt a second observation. Therefore, only one minimisation need be performed on the second time step for each potential action a_1. To remind the reader, it is shown in the Appendix that in this case, the optimal and suboptimal actions are the same on the first time step (i.e. $\hat{a}_1^\star=a_1^\star$), irrespective of the length of the planning horizon.

C. *Baseline Approach – Minimisation of Target Localisation Error*

As a baseline for comparison, the optimal action sequence is determined in order to minimise the target localisation error within the time window under consideration. Minimisation of the localisation error is a widely used approach in target tracking, with the PCRB often used to predict future performance (e.g. again see [13]). In the current paper, the average estimated target location root mean squared error (RMSE), conditional on a target existing, is used to quantify performance. This metric is similar to the PCRB, and takes into account the potential for missed detections, as well as the impact of false alarms and measurement errors.

Similar to the calculation of the AMMS-GOSPA error, but not considering cardinality errors, the average target estimated location mean squared error (MSE) for $T=2$ is calculated as shown in equation (76), where $\mathcal{SE}_t(z_{1:t},a_{1:t},i)$ denotes the squared distance between the particle filter based posterior state

[10] E.g. for three-step planning, the number of minimisations performed on the third step alone is $((2n+1)n_hn_a)^2$.

estimate $p(X|z_{1:t}, a_{1:t})$ and the hypothesis x_i. It is noted that unlike, e.g., (74), the average mean squared error (76) only considers hypotheses under which a target exists (i.e. giving $i > 0$ and $\bar{i} > 0$ in the summations in (76)). The first term on the right-hand side of equation (76) gives the average MSE on the first time step (denoted $\bar{\mathcal{MSE}}_1(a_1^\star)$), with the second term giving the average MSE on the second time step (denoted $\bar{\mathcal{MSE}}_2(a_{1:2}^\star)$). Equation (76) generalises for $T > 2$ in a straightforward manner via adding the average MSE at subsequent time steps.
The baseline control approach then determines the action a_1^b at each time step to minimise $\bar{\mathcal{MSE}}_{1:T}(a_{1:T})$, i.e.:

$$a_1^b = \arg\min_{a_1} \left[\min_{a_{2:T}} \left[\bar{\mathcal{MSE}}_{1:T}(a_{1:T}) \right] \right] \quad (77)$$

This is analogous to the suboptimal control approach (i.e. it does not use a full Bellman recursion).

VII. Simulations

A. Scenario Specification

Three target distributions are considered, unimodal, bimodal and trimodal, representing scenarios concerned with sensor control for wide-area search. Target hypotheses are sampled from these distributions as outlined in Table II. In each case, the hypotheses are given equal weights (i.e. $w_i = 1/n$) and are time-invariant (i.e. stationary).

A single sensor has a circular FOV (i.e. operates in a "spotlight" mode) of radius 10km centred on the action $a \in A$ (which specifies the coordinates of the centre of the spotlight). The action hypotheses consist of uniformly spaced spotlight centres that overlay each target mode, as shown in Figure 2. In each scenario there is also the action hypothesis of not attempting to make a target observation[11].

Conditional on target hypothesis x_i being true and action a being performed, a measurement of the Cartesian coordinates of the target is generated with probability P_d if the hypothesis is within the FOV of the sensor. Each measurement error has a zero-mean Gaussian distribution with covariance Σ. There are either no false alarms (i.e. $\lambda_{FA} = 0$) or false alarms are generated with rate $\lambda_{FA} = 0.01$ per unit volume of the sensor FOV. Parameter settings are summarised in Table III.

TABLE II
SAMPLING DISTRIBUTIONS OF THE TARGET HYPOTHESES IN THE THREE SCENARIOS CONSIDERED.

Target distribution	**Number of hypotheses (n)**	**Sampling distribution ($x_i \sim \mathcal{N}(\bar{x}, \Sigma_x)$)**	
Unimodal	100	$\bar{x} = (100\text{km } 100\text{km})'$	$\Sigma_x = \text{diag}(100\text{km}^2, 100\text{km}^2)$
Bimodal	50	$\bar{x} = (92\text{km } 100\text{km})'$	$\Sigma_x = \text{diag}(2.5^2\text{km}^2, 2.5^2\text{km}^2)$
	50	$\bar{x} = (108\text{km } 100\text{km})'$	
Trimodal	33	$\bar{x} = (92\text{km } 100\text{km})'$	$\Sigma_x = \text{diag}(2.5^2\text{km}^2, 2.5^2\text{km}^2)$
	33	$\bar{x} = (108\text{km } 100\text{km})'$	
	33	$\bar{x} = (100\text{km } (100 + \sqrt{192})\text{km})'$	

TABLE III
SUMMARY OF PARAMETER SETTINGS USED IN THE SIMULATIONS.

Parameter	**Value**
Number of time steps, T	1, 2 or 3
Number of target hypotheses, n	100 (unimodal, bimodal) or 99 (trimodal)
Number of possible actions, n_a	26 (unimodal), 20 (bimodal), 29 (trimodal)
Probability that a target exists, r	0.8
Probability of detection, P_d	0.6, 0.9 or 1.0
False alarm rate, λ_{FA} (m^{-2})	0.0 or 0.01
Measurement error standard deviation, σ $\left(\Sigma = \text{diag}(\sigma^2, \sigma^2)\right)$	10^{-5}km (when $\lambda = 0$) 10^{-2}km (when $\lambda = 0.01$)
Sensor FOV	circular, with radius 10km
Measurement samples n_h per target hypothesis	1 (when $\lambda_{FA} = 0$) 10 (when $\lambda_{FA} = 0.01$, baseline/suboptimal approaches) 1 (when $\lambda_{FA} = 0.01$, optimal approaches)
Cardinality error cost, c	10km
Discount factor for non-myopic planning, λ	1.0

[11] The default is to not attempt an observation unless doing so offers a performance improvement. As noted earlier, when $\lambda_{FA} = 0$ and $\Sigma \approx 0$, observations are only required up to the time instance at which the target is detected.

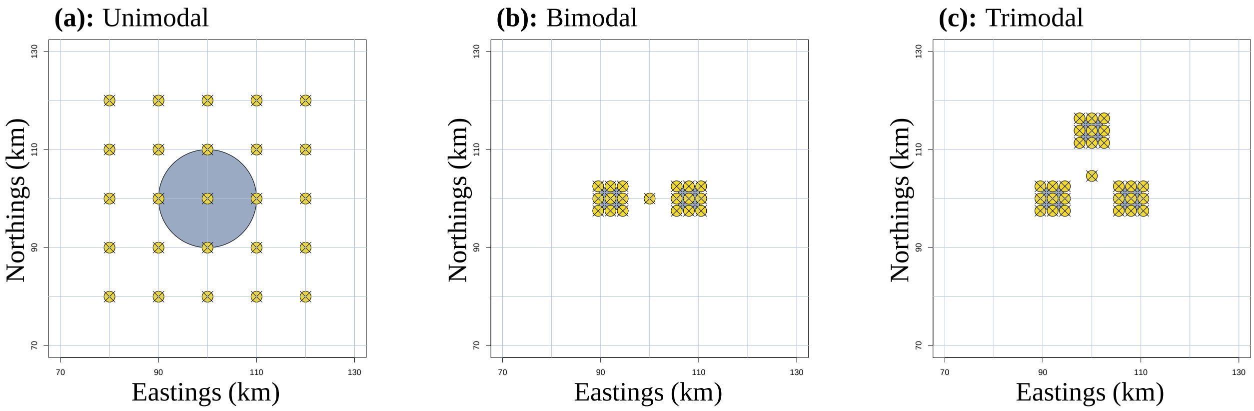

Fig. 2. Action hypotheses for the three scenarios. In each case, the yellow circles show the potential actions, with the cross showing the "spotlight" centre. The grey regions show the one standard deviation uncertainty regions for each target distribution. Note that the sensor FOV is circular, centred on the spotlight centre with a FOV of 10km.

B. *Simulation Results*

B.1 Baseline vs. GOSPA-Based Control Approaches

Exemplar optimal actions for $T = \{2, 3\}$ and $P_d \in \{0.6, 0.9, 1.0\}$ are shown in Figure 3, for each of the target distributions and for scenarios with no clutter (i.e. $\lambda_{FA} = 0$) and insignificant measurement errors (i.e. $\Sigma = \text{diag}(10^{-10}\text{km}^2, 10^{-10}\text{km}^2)$). The second and third time step actions (shown in green and blue respectively) are only necessary if a measurement is not generated on any previous time step. If a measurement is generated at any time, the target state can be estimated extremely accurately and without cardinality error (because there is no clutter, and so the presence of a measurement indicates that a target is present), resulting in an AMMS-GOSPA error of zero. The suboptimal approach generates identical solutions on the first time step (due to the ideal measurement set assumptions). However, the actions on subsequent time steps are independent of the previous measurements, as discussed in Section VI-A.1. As a result, under the suboptimal approach the second and third time step actions are made irrespective of whether previous measurements have been generated.

Overall location RMSE and AMMS-GOSPA values for the baseline and optimal approaches are shown in Table IV. Because there is no cost of sensing, the AMMS-GOSPA cost is identical for the suboptimal and optimal approaches. It can be seen that although, by design, the baseline approach generates slightly lower RMSE values, the corresponding AMMS-GOSPA values are significantly higher than for the GOSPA-based approaches. Hence, the baseline approach does a poor job in balancing the tradeoff between estimation accuracy and cardinality errors, with cardinality errors far greater than for the GOSPA-based approaches[12].

Exemplar suboptimal and optimal actions for $T = 2$ and $P_d \in \{0.6, 0.9, 1.0\}$ are shown in Figure 4, for each of the target distributions and for scenarios with clutter (i.e. $\lambda_{FA} = 0.01$) and measurement errors (i.e. $\Sigma = \text{diag}(10^{-4}\text{km}^2, 10^{-4}\text{km}^2)$). In these cases, because of the measurement origin uncertainty, the generation of measurements does not guarantee that a target is present, and the second step actions are always required. For the optimal approach, the second step action is dependent on the measurement(s) generated on the first time step, and the four most commonly selected second step actions are shown in green in Figures 4(g) – 4(l). Generally, in the scenarios considered, the first step action is the same for the optimal and suboptimal approaches. The one exception is shown in 4(j), in which the the optimal action offsets the first sensor look and subsequently does not always view the centre of the target distribution.

Overall location RMSE and AMMS-GOSPA values for the three control approaches, for $\lambda_{FA} = 0.01$, are shown in Table V. It is observed that the baseline approach again results in slightly lower estimation errors, but with significantly greater AMMS-GOSPA errors (and therefore significantly greater cardinality errors). The suboptimal approach generates AMMS-GOSPA values that are greater than the optimal approach (reaffirming the derivations (63) – (67)), with a maximum difference of around 7%. Future work will identify scenarios in which performance differences between the optimal and suboptimal approaches are more significant.

B.2 Myopic vs. Non-Myopic Planning

In each scenario, the optimal myopic strategy maximises the probability of detecting the target via a single measurement, thereby always viewing the centre/midpoint of the target hypotheses (i.e. shown by the red circle in Figures 3(a) – 3(c) and Figures 4(a) – 4(c)). However, myopic planning does not have the foresight to appreciate that further measurements can be generated. Clearly, observing the centre point can be suboptimal in the multi-modal scenarios, e.g. for the bimodal distribution, it may then be necessary to make (at least) two further observations (one for each mode) to guarantee that the target is detected.

[12]It was observed that the baseline approach often favours making multiple observations of the same region (particularly if $P_d < 1$ and there is clutter), thereby allowing the target to be accurately geo-located if (by chance) it is within that region and is subsequently detected at least once. However, the lack of exploration makes it difficult to infer whether a target is present if this tactic fails (i.e. the target is not detected).

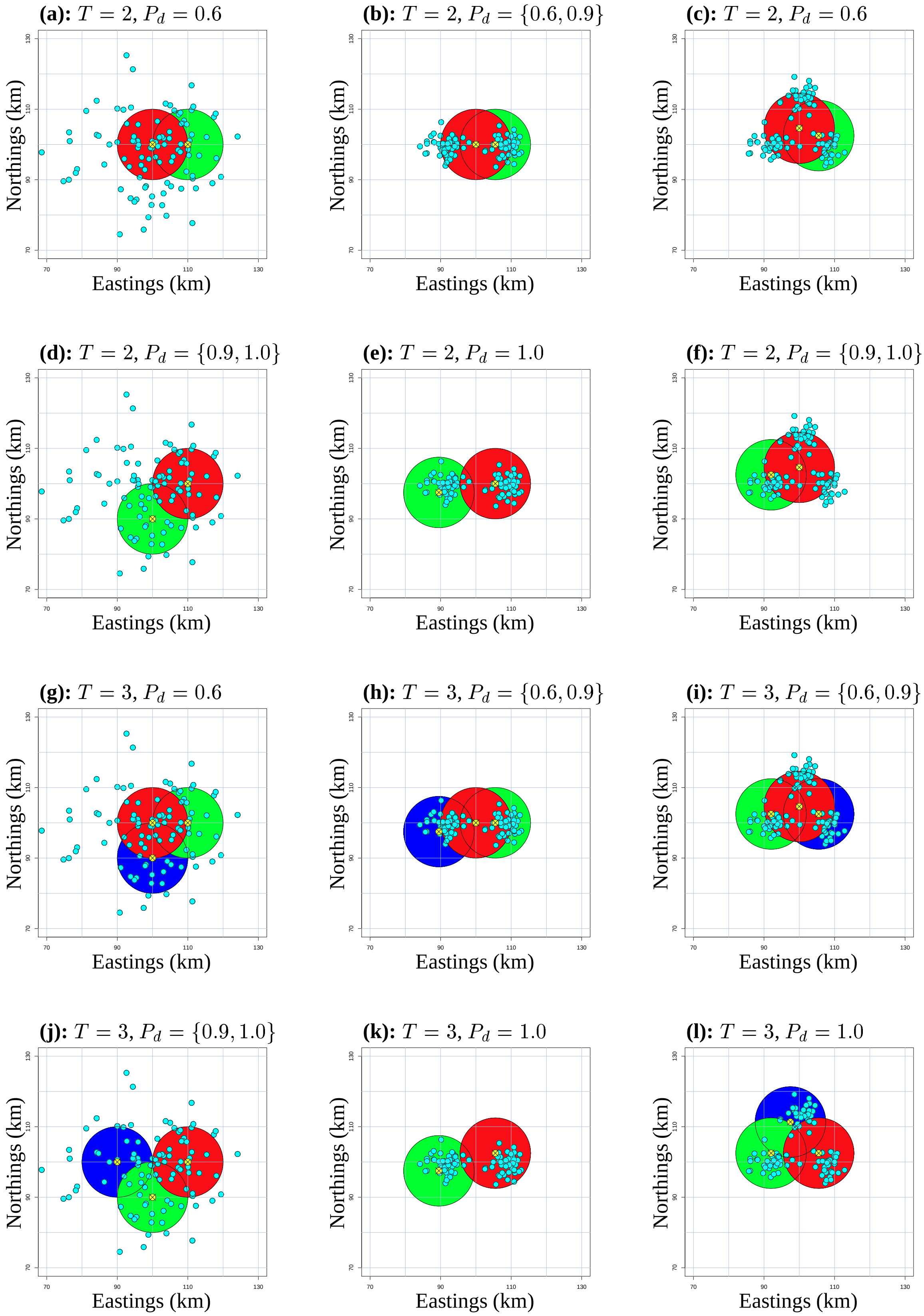

Fig. 3. Optimal actions for exemplar scenarios with no clutter (i.e. $\lambda_{FA} = 0.0$), $P_d \in \{0.6, 0.9, 1.0\}$, and non-myopic time horizons $T \in \{2, 3\}$. Left column: unimodal target prior distribution, middle column: bimodal distribution, right column: trimodal distribution. Red circles: FOV of 1st action, green circles: FOV of 2nd action, blue circles: FOV of 3rd action. For the optimal approach, the 2nd and 3rd actions are only necessary if a measurement is not generated on any previous timestep. For the suboptimal control approach, the 2nd and 3rd actions occur regardless of whether previous measurements have been generated. The optimal myopic action is always to observe the centre of the distribution (i.e. with a sensor spotlight centre as shown by the red circles in (a) – (c)).

TABLE IV

OVERALL LOCATION RMSE AND AMMM-GOSPA COSTS INCURRED (IN KM) BY THE BASELINE AND OPTIMAL CONTROL APPROACHES, FOR EACH OF THE SCENARIOS CONSIDERED WITH NO CLUTTER (I.E. $\lambda_{FA} = 0$) AND $P_d \in \{0.6, 0.9, 1.0\}$. RESULTS ARE AVERAGED OVER 20 RUNS, WITH THE MEAN VALUE $\pm$ ONE STANDARD DEVIATION SHOWN. TO REMIND THE READER, THE OVERALL AMMS-GOSPA ERROR IS IDENTICAL IN THE SUBOPTIMAL AND OPTIMAL APPROACHES AND GIVEN BY $V^{\star}_{1:T}$ (EQUATION (55)) OR $\hat{V}^{\star}_{1:T}$ (EQUATION (66)).

Target Prior	P_d	**Control Approach**	$T = 2$		$T = 3$	
			RMSE	**AMMS-GOSPA**	**RMSE**	**AMMS-GOSPA**
Unimodal	0.6	Baseline	25.58 ± 1.01	65.53 ± 3.99	37.21 ± 1.57	92.57 ± 5.94
		Optimal	26.72 ± 1.10	55.90 ± 1.85	38.99 ± 1.67	77.20 ± 2.65
	0.9	Baseline	23.77 ± 1.08	57.84 ± 5.85	33.61 ± 1.74	78.76 ± 7.70
		Optimal	25.53 ± 1.14	45.53 ± 2.42	36.43 ± 1.96	60.00 ± 3.03
	1.0	Baseline	23.05 ± 1.12	55.62 ± 6.24	32.09 ± 1.87	74.52 ± 7.66
		Optimal	25.17 ± 1.24	42.18 ± 2.57	35.41 ± 1.93	54.56 ± 3.09
Bimodal	0.6	Baseline	11.64 ± 0.30	49.43 ± 1.20	15.29 ± 0.40	69.55 ± 1.98
		Optimal	12.87 ± 0.46	36.36 ± 2.20	17.36 ± 0.66	45.14 ± 2.43
	0.9	Baseline	6.70 ± 0.20	39.52 ± 1.31	8.02 ± 0.25	41.59 ± 1.40
		Optimal	9.02 ± 0.66	19.29 ± 2.34	10.94 ± 0.81	20.56 ± 2.50
	1.0	Baseline	2.20 ± 0.12	14.03 ± 0.52	2.20 ± 0.12	14.18 ± 0.65
		Optimal	4.01 ± 1.98	13.03 ± 1.39	4.01 ± 1.98	13.03 ± 1.39
Trimodal	0.6	Baseline	15.68 ± 0.46	52.01 ± 3.21	21.85 ± 0.66	68.18 ± 4.88
		Optimal	16.40 ± 0.70	46.57 ± 2.57	22.79 ± 1.08	61.36 ± 3.45
	0.9	Baseline	11.62 ± 0.36	40.30 ± 1.44	14.69 ± 0.45	43.76 ± 1.42
		Optimal	14.13 ± 0.93	32.51 ± 3.38	16.77 ± 1.20	38.83 ± 3.44
	1.0	Baseline	8.36 ± 0.38	34.91 ± 1.86	8.36 ± 0.38	35.00 ± 1.84
		Optimal	13.28 ± 0.97	28.12 ± 0.97	11.47 ± 2.42	31.80 ± 3.06

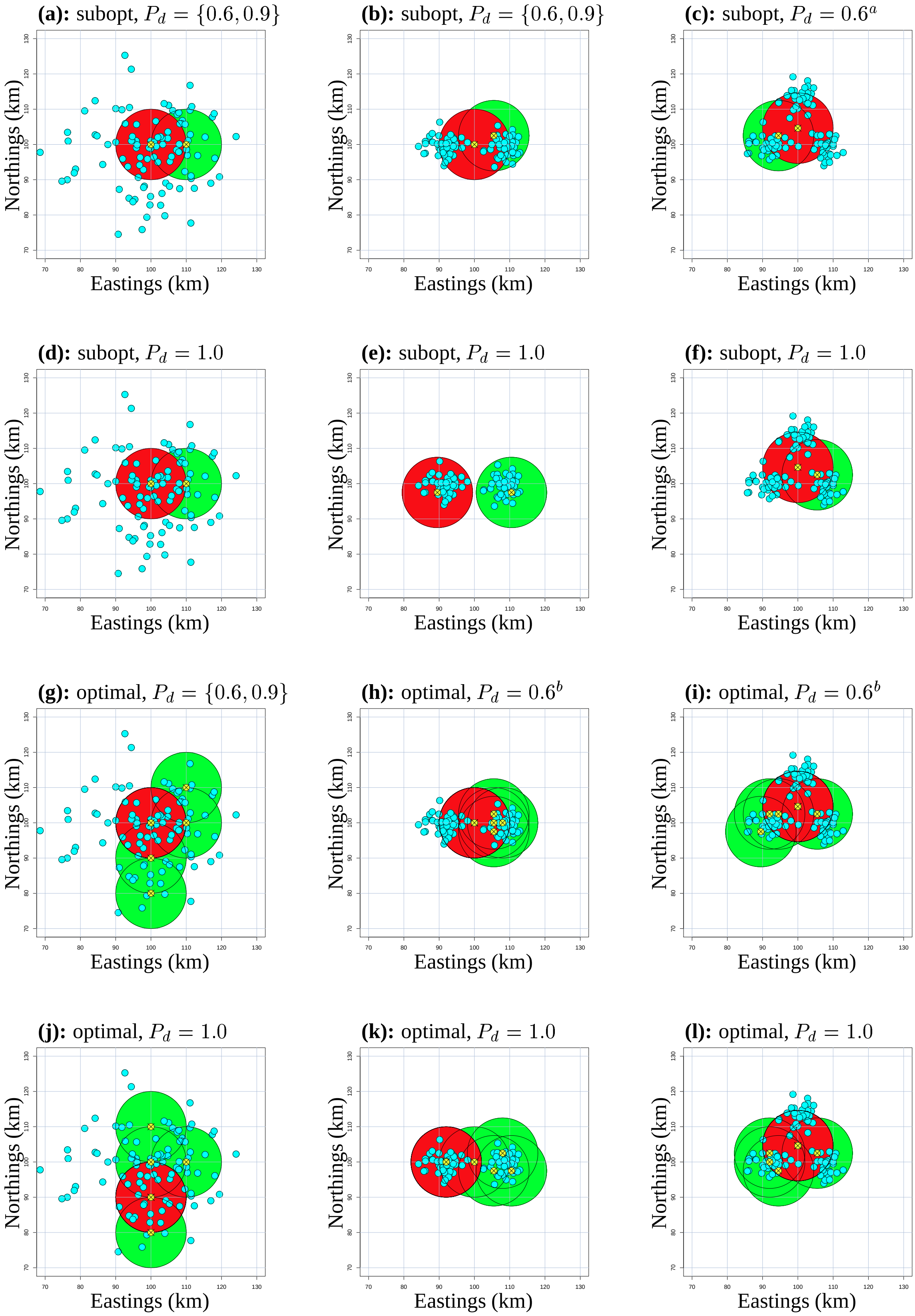

Fig. 4. Suboptimal and optimal actions for exemplar scenarios with clutter (i.e. $\lambda_{FA} = 0.01$), $P_d \in \{0.6, 0.9, 1.0\}$, and $T = 2$. Left column: unimodal target prior distribution, middle column: bimodal distribution, right column: trimodal distribution. Red circles: FOV of 1st action, green circles: FOV of 2nd action (for the optimal solution, the four most commonly occurring 2nd step actions are shown). Again, the optimal myopic action is always to observe the centre of the distribution (i.e. with a sensor spotlight centre as shown by the red circles in (a) – (c)).

^{a}The solution for $P_d = 0.9$ has a slightly different 2nd action which also prioritises attempting to detect a target in the left most mode.

^{b}The solution for $P_d = 0.9$ is virtually identical to that for $P_d = 0.6$, but with a very slight difference to the most commonly occuring 2nd step actions.

TABLE V

OVERALL ERRORS INCURRED (IN KM) BY THE BASELINE, SUBOPTIMAL AND OPTIMAL CONTROL APPROACHES. THERE IS CLUTTER (I.E. $\lambda_{FA} = 0.01$) AND $T = 2$. AGAIN, RESULTS ARE AVERAGED OVER 20 RUNS, WITH THE MEAN VALUE ± ONE STANDARD DEVIATION SHOWN.

Target prior	P_d	Baseline Approach		Suboptimal Control		Optimal Control	
		RMSE	AMMS-GOSPA	RMSE	AMMS-GOSPA	RMSE	AMMS-GOSPA
Unimodal	0.6	25.59 ± 1.01	74.65 ± 1.97	26.83 ± 1.11	69.08 ± 1.29	26.63 ± 1.09	67.74 ± 1.40
	0.9	23.79 ± 1.08	68.16 ± 3.92	25.88 ± 1.18	58.04 ± 2.17	25.63 ± 1.21	55.39 ± 2.34
	1.0	23.07 ± 1.12	66.39 ± 4.30	25.71 ± 1.42	53.55 ± 2.60	25.23 ± 1.24	50.28 ± 2.54
Bimodal	0.6	11.69 ± 0.30	64.49 ± 0.86	12.90 ± 0.51	50.22 ± 2.03	12.82 ± 0.49	49.03 ± 2.03
	0.9	6.79 ± 0.21	49.85 ± 3.39	9.08 ± 0.45	27.52 ± 2.73	8.92 ± 0.64	26.69 ± 2.67
	1.0	2.33 ± 0.20	21.40 ± 0.76	5.48 ± 1.73	18.48 ± 2.45	5.52 ± 1.75	18.22 ± 2.47
Trimodal	0.6	15.72 ± 0.46	66.48 ± 2.63	16.47 ± 0.67	61.82 ± 1.96	16.34 ± 0.78	60.45 ± 2.24
	0.9	11.67 ± 0.37	54.15 ± 1.74	14.17 ± 0.93	44.68 ± 3.36	14.05 ± 0.98	42.72 ± 3.60
	1.0	8.46 ± 0.36	48.32 ± 2.53	13.33 ± 1.01	37.88 ± 4.01	13.32 ± 0.98	35.37 ± 4.15

For the non-myopic approaches, it can be seen from Figures 3 – 4 that:

1. When $P_d = \{0.6, 0.9\}$, multi-step planning also almost always favours initially viewing the centre/midpoint of the target hypotheses (except in the scenarios shown in Figures 3(d) and 3(j)). Consequently, there is considerable overlap between the action spotlights on each time step. This is because when the probability of detection is less than unity, multiple sensor observations in regions of high target probability mass provides an improved opportunity to generate at least one detection of a target.

2. When $P_d = 1.0$, multi-step planning almost always avoids viewing the centre of the distribution of target hypotheses[13], and instead offsets each sensor spotlight to achieve optimal surveillance coverage over multiple time steps by observing each mode in turn, e.g. see Figures 3(e), 3(k) and 3(l)).

3. When $P_d = 1.0$ and $\lambda_{FA} = 0$, one observation is sufficient to detect the target in any given region. Therefore, in these cases, a "cookie-cutter" strategy with limited overlap between the sensor spotlights is optimal.

These observations are in agreement with an intuitive understanding of how taking into account the ability to make further observations should influence the first observation.

B.3 Concluding Remarks

It is noted that multi-step planning often generates the same first action as myopic planning when there is a low probability of detection. This is because there is no guarantee that the target will be detected across multiple time steps, and so regions of high probability mass offer the greatest opportunity achieve at least one target detection. By design, myopic planning always favours these high probability mass regions because its decision-making lacks the foresight to appreciate that further observations are possible.

Conversely, in scenarios with a high probability of detection, offsetting the first observation and using a cookie-cutter strategy often allows the multi-step approach to decrease the number of time steps required to provide complete surveillance of the region of interest. This approach should therefore be favoured if time constraints allow.

VIII. SUMMARY AND CONCLUSIONS

This paper has proposed a sample-based approach for myopic and non-myopic sensor management for a Bernoulli target using the GOSPA metric. We have provided the following contributions: analytical calculation of the MS-GOSPA for different measurements and actions, development of efficient sampling techniques to calculate the AMMS-GOSPA error, and the development of an optimal non-myopic (Bellman type [20]) planning recursion that exploits the conditional AMMS-GOSPA error. Simulations demonstrate the approach in scenarios with: (i): missed detections, (ii): false alarms, (iii): a high degree of uncertainty in the target location, with the prior distribution represented by large number of potential hypotheses, and: (iv): a planning horizon of up to three time steps.

Various behavioural patterns are identified, notably demonstrating the benefits of non-myopic planning, and in particular showing that optimal plans align with an intuitive understanding of how taking into account the opportunity to make further observations should influence the current action. It is concluded that the GOSPA-based, non-myopic search and track algorithm offers a powerful mechanism for sensor management in order to minimise estimation errors and errors due to missed and false targets in a unified way.

The current approach is directly applicable to multi-target scenarios with well-separated targets, due to the separability of the optimal actions when using the GOSPA metric [18]. Future work will extend the approach to multi-sensor, multi-target scenarios in which targets may move in close proximity. We will also implement Monte Carlo roll-out as a mechanism for efficiently estimating the long-term impact of actions. Furthermore, we will work to identify scenarios in which the suboptimal and optimal approaches generate markedly different solutions, thereby highlighting the importance of accounting for the potential sequences of future measurements at each decision epoch.

ACKNOWLEDGEMENTS

This research was funded by the UK Ministry of Defence (MOD) through the Fusion and Information Theory (FIT) project, DSTLX-1000143908.

[13] Notable exceptions are shown in Figures 3(f), 4(f), and 4(l), in which two-step planning cannot provide coverage of all three modes, and consequently prioritises attempting early detection by viewing the centre of the target distribution.

$$\hat{a}_1^\star = \arg\min_{a_1} \Big[p(z_1=\phi|a_1) \Big[r_1(z_1=\phi, a_1) + \lambda \min_{a_2} p(z_2=\phi|z_1=\phi, a_{1:2}) \big[r_2(z_{1:2}=\phi, a_{1:2}) \tag{78}$$

$$+ \ldots + \lambda \min_{a_T} p(z_T=\phi|z_{1:T-1}=\phi, a_{1:T}) \big[r_T(z_{1:T}=\phi, a_{1:T}) \big] \Big] \Big]$$

$$= \arg\min_{a_1} \Big[\mathbb{E}_{z_1} \left[r_1(z_1, a_1) \right] + \lambda \min_{a_2} \Big[\mathbb{E}_{z_{1:2}} \left[r_2(z_{1:2}, a_{1:2}) \right] + \ldots + \lambda \min_{a_T} \left[\mathbb{E}_{z_{1:T}} \left[r_T(z_{1:T}, a_{1:T}) \right] \right] \Big] \Big] \tag{79}$$

$$= \arg\min_{a_1} \left[\min_{a_{2:T}} \left[\mathbb{E}_{z_{1:T}} \left[\sum_{t=1}^{T} \lambda^{t-1} r_t(z_{1:t}, a_{1:t}) \right] \right] \right] \quad \text{(as the conditional expectations have been removed)} \tag{80}$$

$$= a_1^\star \quad \text{(given by equation (54))} \tag{81}$$

The authors would like to thank Professor Gustav Hendeby (Linköping University) and Professor Rahul Savani (University of Liverpool) for insightful discussions regarding sensor management.

APPENDIX

In this appendix, we prove the following proposition.

Proposition

Using the ideal measurement set assumptions: (i): $\Sigma = 0$ (i.e. target generated measurements are error-free), and: (ii): $\lambda_{FA} = 0$ (i.e. there are no false alarms), but allowing for missed detections (i.e. allowing $P_d < 1$), the suboptimal and optimal control approaches generate identical solutions (i.e. $\hat{a}_1^\star = a_1^\star$) irrespective of the length of the planning horizon.

Proof

In this case, MMS-GOSPA$(z_{1:t}, a_{1:t}) = 0$ unless $z_i = \phi$ for $i = 1, \ldots, t$. This is because, with extremely accurate measurements, and no clutter, the presence of just a single measurement signifies that a target is present, and the accurate measurement allows the target to be geo-located without error.

Hence, in determining the GOSPA-based cost function, it is necessary to only consider cases in which there are no previous measurements. The optimal action (68) can then be manipulated as shown in equations (78) – (81).

This completes the proof.

□

REFERENCES

[1] V. Krishnamurthy, *Partially Observed Markov Decision Processes: From Filtering to Controlled Sensing*, Cambridge University Press, 2016.

[2] P. Boström-Rost, D. Axehill, and G. Hendeby, "Sensor Management for Search and Track Using the Poisson Multi-Bernoulli Mixture Filter", *IEEE Transactions on Aerospace and Electronic Systems*, **57**(5), pp. 2771–2783, 2021.

[3] P. Boström-Rost, D. Axehill, and G. Hendeby, "PMBM Filter With Partially Grid-Based Birth Model With Applications in Sensor Management", *IEEE Transactions on Aerospace and Electronic Systems*, **58**(1), pp. 530–540, 2022.

[4] J. L. Williams, "Marginal Multi-Bernoulli Filters: RFS Derivation of MHT, JIPDA, and Association-Based MeMBer", *IEEE Transactions on Aerospace and Electronic Systems*, **51**(3), pp. 1664–1687, 2015.

[5] Á. F. García-Fernández, J. L. Williams, K. Granström, L. Svensson, "Poisson multi-Bernoulli mixture filter: direct derivation and implementation", *IEEE Transactions on Aerospace and Electronic Systems*, **54**(4), pp. 1883–1901, 2018.

[6] R. Mahler, "Multitarget Sensor Management of Dispersed Mobile Sensors", in: *Theory and Algorithms for Cooperative Systems*, (eds. D. Grundel, R. Murphey, and P. Pardalos), chapter 12, pp. 239–310, World Scientific Publishing Co, Singapore, 2004.

[7] B. Charrow, N. Michael, and V. Kumar, "Active Control Strategies for Discovering and Localizing Devices with Range-Only Sensors", *Algorithmic Foundations of Robotics XI* (eds. H. L. Akin, N. M. Amato, V. Isler, and A. F. van der Stappen), pp. 55–71, Springer, 2015.

[8] H. Van Nguyen, H. Rezatofighi, B.-N. Vo, and D. C. Ranasinghe, "Multiobjective Multi-Agent Planning for Jointly Discovering and Tracking Mobile Objects", *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, New York, USA, 2020.

[9] R. Tharmarasa, T. Kirubarajan, M. L. Hernandez, and A. Sinha, "PCRLB Based Multisensor Array Management for Multitarget Tracking", *IEEE Transactions on Aerospace and Electronic Systems*, **43**(2), pp. 539–555, 2007.

[10] H. Van Trees, K. L. Bell, and Z. Tian, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Filtering*, 2nd Edition, John Wiley and Sons, New York, USA, 2013.

[11] X. Zhang, and P. K. Willett, "Cramer-Rao Bounds for Discrete Time Linear Filtering with Measurement Origin Uncertainty", *Proceedings of the Workshop on Estimation, Tracking and Fusion: A Tribute to Yaakov Bar-Shalom*, Monterey, California, USA, 2001.

[12] M. L. Hernandez, T. Kirubarajan, and Y. Bar-Shalom, "Multisensor Resource Deployment Using Posterior Cramér-Rao Bounds", *IEEE Transactions on Aerospace and Electronic Systems*, **40**(2), 399–416, 2004.

[13] M. L. Hernandez, "Performance Bounds for Target Tracking: Computationally Efficient Formulations and Associated Applications", in: *Integrated Tracking, Classification, and Sensor Management: Theory and Applications* (eds. M. Mallick, V. Krishnamurthy, and B.-N. Vo), John Wiley and Sons Inc., 2013.

[14] A. S. Rahmathullah, Á. F. García-Fernández, and L. Svensson, "Generalized Optimal Sub-Pattern Assignment Metric", *Proceedings of the 20th International Conference on Information Fusion*, pp. 1–8, Xi'an, China, 2017.

[15] D. Schuhmacher, and A. Xia, "A New Metric Between Distributions of Point Processes", *Advances in Applied Probability*, **40**(3), pp. 651–672, 2008.

[16] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A Consistent Metric for Performance Evaluation of Multi-Object Filters", *IEEE Transactions on Signal Processing*, **56**(8), pp. 3447–3457, 2008.

[17] Á. F. García-Fernández, and L. Svensson, "Spooky Effect in Optimal OSPA Estimation and How GOSPA Solves It", *Proceedings of the 22nd International Conference on Information Fusion*, Ottawa, Canada, 2019.

[18] Á. F. García-Fernández, M. Hernandez, and S. Maskell, "An Analysis on Metric-Driven Multi-Target Sensor Management: GOSPA versus OSPA", *Proceedings of the 24th IEEE International Conference on Information Fusion*, Sun City, South Africa, 2021.

[19] L. Úbeda-Medina, "Robust Techniques for Multiple Target Tracking and Fully Adaptive Radar", Ph.D. Dissertation, Universidad Politecnica de Madrid, 2018. Available online at: http://oa.upm.es/53209/

[20] R. Bellman, "On the Theory of Dynamic Programming", *Proceedings of the National Academy of Sciences of the United States of America*, **38**(8), pp. 716–719, 1952.

[21] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking", *IEEE Transactions on Signal Processing*, **50**(2), pp. 174–188, 2002.

[22] J. M. Maciejowski, *Predictive Control: With Constraints*, Prentice Hall, Englewood Cliffs, New Jersey, USA, 2002.

[23] A. Marrs, S. Maskell, and Y. Bar-shalom, "Expected Likelihood for Tracking in Clutter with Particle Filters", *Proceedings of the SPIE Conference on Signal and Data Processing of Small Targets*, **4728**, pp. 230-239, Orlando, Florida, USA, 2002.

[24] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*, John Wiley and Sons Inc., New York, USA, 2001.