Receding Horizon Tracking of an Unknown Number of Mobile Targets using a Bearings-Only Sensor

James D. Turner, James McMahon, and Michael M. Zavlanos

***Abstract*— Planning the motion of bearings-only sensors is critical for enabling accurate tracking of the positions of moving targets. In this paper, we demonstrate planning the observer's motion over horizons greater than one step for estimating an unknown and varying number of indistinguishable, maneuvering targets of interest using a probability hypothesis density (PHD) filter, with a Rényi divergence reward for selecting actions. We describe approximations to make this approach computationally feasible, and we propose using Monte Carlo tree search (MCTS) to further reduce the cost. Finally, we present simulation results showing that longer planning horizons reduce the error in the estimates and that MCTS can reduce the cost of planning without sacrificing the quality of the estimates.**

I. Introduction

An important application of autonomous robotics is monitoring multiple maneuvering targets within a region of interest, such as vehicles, animals, athletes, or pedestrians; see [1] for a review. A particularly interesting case is when the observer carries a bearings-only sensor with limited field-of-view (FoV) in order to estimate an unknown, varying number of indistinguishable targets of interest. This situation arises, e.g., in acoustic tracking of ships [2] and in tracking the motion of pedestrians with monocular cameras [3]. With a bearings-only sensor, the motion of the observer significantly affects its ability to correctly estimate targets' states; therefore, planning its motion can result in significant performance gains.

In this paper, the states of the targets are estimated using a probability hypothesis density (PHD) particle filter [4], [5]. Then, to plan the motion of the observer, a receding horizon controller uses the Rényi divergence between the predicted and future posterior PHDs within the region of interest at the end of the planning horizon as a heuristic reward [6]. Two different types of planners are presented. The first exhaustively evaluates the reward for every possible sequence of actions. The second uses Monte Carlo tree search (MCTS) to reduce the cost of planning by evaluating fewer action sequences. We show that longer planning horizons reduce the error in the estimates and that MCTS can reduce the cost of planning without sacrificing the quality of the estimates.

James D. Turner is with the Physical Acoustics Branch, Code 7130, Naval Research Laboratory, Washington, D.C., 20375, USA, and with the Department of Mechanical Engineering & Materials Science, Duke University, Durham, NC, 27708, USA.

James McMahon is with the Physical Acoustics Branch, Code 7130, Naval Research Laboratory, Washington, D.C., 20375, USA.

Michael M. Zavlanos is with the Department of Mechanical Engineering & Materials Science, Duke University, Durham, NC, 27708, USA.

The work of J. McMahon and J. Turner is supported by the Office of Naval Research through the U.S. Naval Research Laboratory Base program.

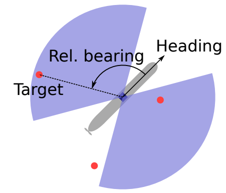
A. Related Work

To our knowledge, there are only a few works that consider the problem of planning the motion of a single observer in 2-D space carrying a bearings-only sensor for estimating the positions of an unknown number of indistinguishable, maneuvering targets. Particularly, Beard, et al., demonstrated planning with a horizon of a single action step and multiple observation steps using a GLMB filter with a Cauchy–Schwarz divergence reward [7], while Wolek, et al., demonstrated a basic "keep broadside" behavior [2]. However, these approaches have not been extended to planning horizons with multiple action steps.

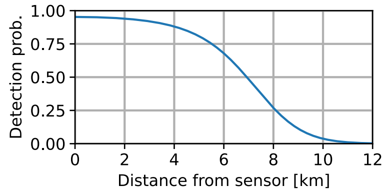
Related is also literature which studies variations of the estimation problem considered in this paper. For example, Dames and Kumar considered localization of stationary targets using multiple bearings-only sensors [8], while El-Fallah, et al., considered planning with a horizon of 1 for reorienting bearings-only sensors for tracking satellites [9]. Range-only sensors have limitations similar to bearings-only sensors; one-step planning methods using range-only sensors have been investigated in [6], [10]–[12]. Sensors providing position or range-and-bearing measurements, such as radar, active sonar, or downward-facing cameras on aircraft, are qualitatively different since they provide an estimate of the targets' positions without requiring movement of the sensor. Research that employs these sensors generally focuses on cases where the noise or probability of detection depends on the relative position of the target; works with these types of sensors include [13]–[30]. Finally, other literature has considered sensor planning for a known number of targets or distinguishable targets, e.g. [31]–[36].

B. Contributions

To our knowledge, this is the first approach designed to plan the motion of a single observer over horizons greater than one action step for bearings-only estimation of an unknown and varying number of indistinguishable, maneuvering targets of interest. An additional contribution of this work is the development of approximations to make this approach computationally feasible despite the particle representation of the filter's belief. We show that planning with a horizon greater than one step reduces



(a) Sensor FoV. Lengths are not to scale.



(b) Detection probability for targets within the FoV.

Fig. 1: Properties of the sensor.

the error in the target estimates, as described by the optimal subpattern assignment (OSPA) metric. Finally, we demonstrate that Monte Carlo tree search (MCTS) can reduce the cost of planning without sacrificing the quality of the target estimates.

II. Problem Description

Consider a single observer carrying a bearings-only sensor with limited field-of-view (FoV), multiple maneuvering targets, and a fixed region of interest within a two-dimensional world. The number of targets and their states are unknown to the observer; the only information available to the observer is its own state and imperfect bearing measurements of the targets. The objective is to minimize the error in its estimates of target positions within the region of interest. At each time step, the observer receives measurements from the sensor, updates its belief of the target states, and chooses the trajectory to travel until the next step.

Specifically, the state of a target consists of its 2-D position, its 2-D velocity, and a label for its current motion model. The motion model is nearly constant velocity or a coordinated turn to port or starboard, as described in [37]. At each time step, the transition between models is Markovian, using the probabilities from equation (30) of [37]. On the other hand, the state $\boldsymbol{o}_k$ of the observer at step k consists of its 2-D position and its heading. During each time step, it travels exactly straight or along a circular arc. The state of the observer is always known. Finally, we assume that the sensor has a limited FoV, with noise, clutter, and imperfect detection. For each detected target, the measurement is the relative bearing, as illustrated in Fig. 1a, with additive Gaussian noise. The detection probability within the FoV is a function of range, as illustrated by Fig. 1b. Clutter measurements are uniform in bearing within the FoV, and the number of clutter measurements is Poisson. Target measurements and clutter measurements are independent.

III. Filter

In all planning algorithms proposed in this paper, the observer uses the same probability hypothesis density (PHD) particle filter to estimate the target states. The PHD filter implementation is based on [5], [37], with a modification to handle the limited field-of-view (FoV), described below. The filter assumes the target motion model and sensor model described in Section II.

A. Persistent Cluster Assignments

A common way to perform measurement updates and produce target estimates with a PHD filter is to sample clusters of particles corresponding to measurements in the current measurement set and then update those clusters separately and report them as target estimates [5]. While this updates the PHD correctly, it does not produce estimates for targets which are undetected on the current time step, such as targets outside the sensor's FoV. To resolve this, the clustering strategy for the filter used here maintains assignments of particles to clusters even for undetected targets, and those clusters continue to be candidates for target estimates. Particles may be reassigned, but only to new clusters for new measurements. Following the clustering step, the particles are updated as normal according to either the corresponding measurement or the undetected case. In other words, this change does not affect the PHD estimate; its only effect is maintaining clusters for undetected targets and including them in the set of estimated targets. The primary limitation of this approach is that targets close to the edge of the FoV can result in two clusters representing a single target. However, this occurs rarely enough that it's not a significant issue in practice.

B. Target Estimates

Each cluster is assumed to approximately represent a single target located at the weighted mean of the particles, with probability of existence equal to the total weight of the particles. Only clusters with probability of existence of at least $1/2$ are included in the set of target estimates used for planning and evaluating the planners.

C. Target Births and Deaths

While targets do not actually appear or disappear from existence, the PHD filter implementation assumes birth and death models in order to handle previously-unobserved and no-longer-relevant targets. The filter handles births by adding particles in circular sectors conditioned on measurements, assuming a Poisson-distributed number of births, as described in [5]. While this fails to account for targets which have not yet been detected, for the problem considered, a target is unlikely to reach the region of interest without being detected. The filter assumes multi-Bernoulli target survival between time steps.

IV. Planning

At each time step, the observer uses the information provided by the PHD filter to select an action. This section starts by describing exhaustive search of every possible sequence of actions over the planning horizon, and it presents approximations to make this approach computationally feasible. It then describes how Monte Carlo tree search (MCTS) can be used to avoid evaluating all the possible action sequences.

A. Exhaustive Search

One approach for planning is to exhaustively evaluate every possible sequence of actions for the finite planning horizon, using approximations to make this computationally feasible, and then choose the first action in the sequence which has the largest expected reward. Following [6], [11], Rényi divergence is used as a heuristic reward which describes the anticipated information gain due to future measurements. When planning with a horizon of H steps into the future, the problem is

$$(\boldsymbol{a}_k, \ldots) = \underset{(\boldsymbol{a}_k, \ldots, \boldsymbol{a}_{k+H-1})}{\operatorname{argmax}} \mathbb{E}\left[R\left(D_{k+H|k}, D_{k+H|k+H}\right)\right],$$

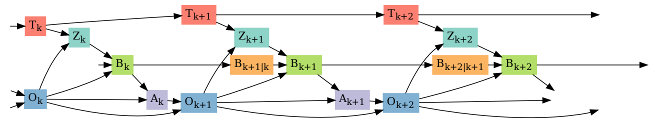
where $\boldsymbol{a}_k$ denotes the action at time step k; $Z_{1:k}$ denotes the measurement sets from step 1 to k; R is Rényi divergence with parameter $\alpha = \frac{1}{2}$; $D_{k+H|k}$ is the predicted PHD of the targets (conditioned on $Z_{1:k}, \boldsymbol{o}_{1:k}$) inside the region of interest at the planning horizon; and $D_{k+H|k+H}$ is the future posterior PHD of the targets (conditioned on $Z_{1:k+H}, \boldsymbol{o}_{1:k+H}$) inside the region of interest at the planning horizon. The objective function is an expectation, conditioned on the current belief from the filter, with respect to the unknown future measurements $Z_{k+1:k+H}$.

If the states of the particles approximating the PHD are the same between $D_{k+H|k}$ and $D_{k+H|k+H}$, i.e. only the weights differ, then the Rényi divergence can be computed efficiently using only the weights [6], [11]. To exploit this, the particle states can be preserved during planning by skipping the birth of new particles, the elimination of low-weight particles, and particle resampling. While eliminating the birth step is an approximation, for the problem of interest, the birth process has little effect over the short planning horizons considered. Similarly, the particle resampling step is dispensable, since particle collapse is limited over the short horizons considered.

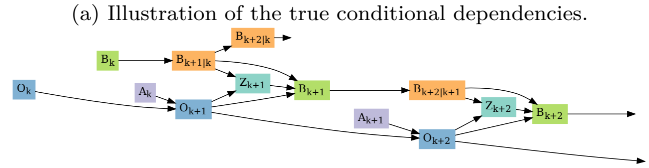
For the planner to be practical, further approximations are necessary, as described below. Note, however, that all of the approximations described in this section are applied only during planning; the full PHD filter is applied to the real measurements outside of planning.

1) Generating Measurements Based on the Ideal Measurement Set Approximation: Figure 2a shows the true conditional dependencies between the target states, measurement sets, filter states, etc. Based on this diagram, the expected reward for an action sequence could be estimated using a Monte Carlo approach: sample sets of target states according to the filter state, apply the motion model to predict the target trajectories, sample measurements using those predicted targets, perform the filter updates using those measurements, and then average together the rewards computed for these samples. Unfortunately, this approach is impractically expensive.

So, for each action sequence, the planner instead computes a single estimate of $D_{k+H|k+H}$ and uses the corresponding Rényi divergence as a rough approximation of the expected reward. For each filter update to the horizon, the planner generates a measurement set based on



(a) Illustration of the true conditional dependencies.



(b) Illustration of the conditional dependencies while planning.

Fig. 2: Bayesian networks illustrating conditional dependencies. $\mathrm{T_k}$ is the target states at step k; $\mathrm{O_k}$ is the observer state at step k; $\mathrm{B_k}$ is the filter state (containing $D_{k|k}$) approximating $\mathrm{T_k}$, after the measurement update; $\mathrm{B_{k|j}}$ is the filter state (containing $D_{k|j}$) approximating $\mathrm{T_k}$, by predicting starting from $\mathrm{B_j}$; $\mathrm{A_k}$ is the action at step k; and $\mathrm{Z_k}$ is the measurement set at step k.

the ideal measurement set approximation [6], [11]. Given the target estimates and observer state, the generated measurement set is the set of relative bearings to the target estimates, for target estimates with probability of detection greater than 1/2. Figure 2b illustrates the conditional dependencies between the states of the filter and the generated measurement sets. Each measurement set is generated using the target estimates conditioned on the previous measurement sets, since the measurement sets are not independent between time steps.

2) Exploiting Known Association of Measurements to Particle Clusters: It is possible to perform a full filter update during planning, including sampling the partitioning of particles into clusters. However, when planning with generated measurements, the association between each cluster of particles and the corresponding generated measurement (or no measurement) is known. This is exploited to reduce the cost of planning by avoiding the partitioning step. Instead, each cluster of particles is updated using the corresponding generated measurement (or no measurement). This reduces the asymptotic cost of the update from $O(|\mathcal{P}||Z|)$ to $O(|\mathcal{P}|)$, where $|\mathcal{P}|$ is the particle count and $|Z|$ is the number of measurements. For targets which are well-separated in bearing, this approximation is indistinguishable from a full update, while substantially reducing the computational cost.

3) Removing Unimportant Particles when Planning: Many of the particles have little influence on the rankings of the action sequences, so they are ignored during planning to reduce the computational cost. In particular, particles unassigned to clusters and clusters for which the total predicted weight is less than 1/2 for each step to the horizon are ignored, since they would be updated the same way for all action sequences. Additionally, any cluster for which only a very small fraction of the cluster's weight is in the region of interest for each step to the horizon is ignored, since the particle weights in the region of interest for this cluster are unlikely to change much.

B. Monte Carlo Tree Search

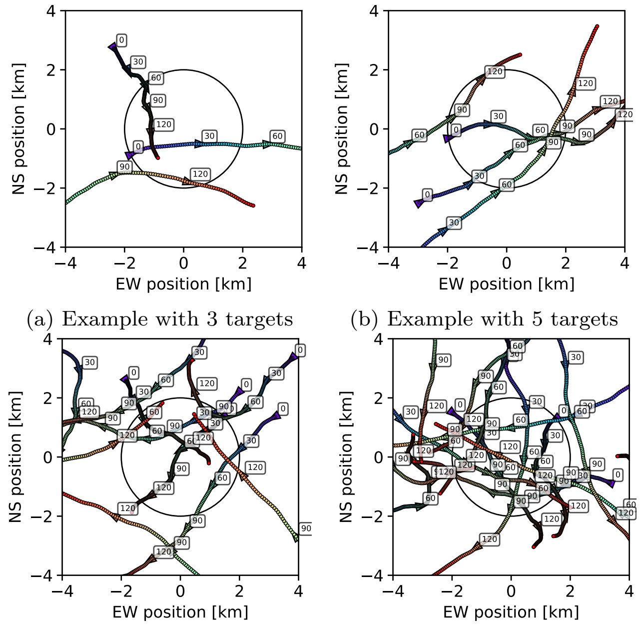
Monte Carlo tree search (MCTS) is an approach to avoid the cost of exhaustively evaluating every possible action sequence. MCTS strategically selects which actions to evaluate, based on previous evaluations, to prioritize good actions without neglecting exploration. Algorithm 1 describes the MCTS algorithm used here. For this implementation, the exploration–exploitation trade-off is handled using the polynomial upper confidence tree (PUCT) heuristic [38]. The value of a node is the mean of the rewards of its descendant leaves. After performing the desired number of iterations, the action with the largest value is chosen. Aside from the strategy for expanding the tree and updating the nodes' values, all of the other aspects of the planner are the same as the exhaustive search approach.

Algorithm 1: Monte Carlo Tree Search (MCTS).

function *mcts* **is**
  input : Observer state $\boldsymbol{o}_k$, PHD filter state $\mathcal{B}_{k|k}$.
  output : Selected action.
  $\tilde{\mathcal{B}}_{k|k} \leftarrow$ RemoveUnimportant($\mathcal{B}_{k|k}$), Sec. IV-A.3.
  $\tilde{\mathcal{B}}_{k:k+H|k} \leftarrow$ Predict($\tilde{\mathcal{B}}_{k|k}$), Sec. IV-A.
  $\tilde{\mathcal{B}}_{k:k+H|k} \leftarrow$ RemoveMoreUnimportant($\tilde{\mathcal{B}}_{k:k+H|k}$), Sec. IV-A.3.
  $\mathcal{T} \leftarrow$ initialize MCTS tree with $(\boldsymbol{o}_k, \tilde{\mathcal{B}}_{k|k})$.
  foreach iteration **do**
    for i in $k+1:k+H$ **do**
      $\boldsymbol{a}_{i-1} \leftarrow$ SelectActionPUCT($\mathcal{T}$).
      if this will be a new node **then**
        $\boldsymbol{o}_i \leftarrow$ UpdateObserver($\boldsymbol{o}_{i-1}, \boldsymbol{a}_{i-1}$), Sec. II.
        $\tilde{\mathcal{B}}_{i|i-1} \leftarrow$ PredictSameStates($\tilde{\mathcal{B}}_{i-1|i-1}, \tilde{\mathcal{B}}_{i-1|k}$), Sec. IV-A.
        $Z_i \leftarrow$ Measurements($\tilde{\mathcal{B}}_{i|i-1}, \boldsymbol{o}_i$), Sec. IV-A.1.
        $\tilde{\mathcal{B}}_{i|i} \leftarrow$ Update($\tilde{\mathcal{B}}_{i|i-1}, \boldsymbol{o}_i, Z_i$), Sec. IV-A.2.
        $\mathcal{T} \leftarrow$ AddNode($\mathcal{T}, \tilde{\mathcal{B}}_{i-1|i-1}, \boldsymbol{a}_{i-1}, \boldsymbol{o}_i, \tilde{\mathcal{B}}_{i|i}$).
      else
        $(\boldsymbol{o}_i, \tilde{\mathcal{B}}_{i|i}) \leftarrow$ GetNode($\mathcal{T}, \tilde{\mathcal{B}}_{i-1|i-1}, \boldsymbol{a}_{i-1}$).
      end if
    end for
    $r \leftarrow$ RényiDivergence($\tilde{\mathcal{B}}_{k+H|k}, \tilde{\mathcal{B}}_{k+H|k+H}$), Sec. IV-A.
    Update value of the node and its ancestors with r.
  end foreach
  return Action for the child of the root node with the largest value.
end

V. Results and Discussion

For comparison, a baseline policy of random actions, the exhaustive search planner with horizon 1 (which is most similar to existing literature, e.g. [6]), and the longer-horizon planners proposed in this paper are simulated in randomly-generated scenarios containing 3, 5, 10, or 15 targets. Note that the PHD filter makes poor estimates when targets are aligned in bearing, so it should not be used for very high target densities. For each target count, the trajectories of the targets are sampled to form 400 scenarios of 150 time steps each. Figure 3 shows some examples. The planners are used to select the observer's actions; the speed of the observer is 6 m/s, and its angular velocity is one of −4.5 °/s, 0 °/s, or 4.5 °/s. The pseudorandom number generators are seeded independently between scenarios, but the planners are applied to the same scenarios so they can be compared in a pairwise fashion. The initial states of the targets are sampled in such a way that the targets are likely to enter the region of interest, and then the motions of the targets are sampled according to Section II. The time step is 10 s. The region of interest is a circle centered at the origin with radius 2 km.



(a) Example with 3 targets (b) Example with 5 targets

(c) Example with 10 targets (d) Example with 15 targets

Fig. 3: Trajectories of the targets for some example scenarios. Small circular markers indicate the positions of the targets. Triangular markers show the positions and headings of the targets every 30 time steps, and text labels show the corresponding time indices.

Measurements are simulated according to Section II. The probability of detection is given by $p_{\mathrm{D}}(r,\theta) = 1 - \left(1 + \mathrm{e}^{-\left((r/r_{\mathrm{ref}})^2 - 3\right)}\right)^{-1}$ if $30° \leq |\theta| \leq 150°$, or 0 otherwise, where $r_{\mathrm{ref}} = 4000\,\mathrm{m}$, r is the range, and θ is the relative bearing in $[-180°, 180°)$. The standard deviation of the noise in bearing measurements is 0.5°. The mean number of clutter measurements per step is 1.

For the birth process, the filter assumes 0.01 expected births per step, a maximum range of 8000 m, a beta distribution with parameters $\alpha = 6$ and $\beta = 5$ for the speed divided by 10 m/s, a uniformly-distributed heading, and a motion model distribution equal to the stationary distribution of the Markov transition process. The filter assumes a target survival probability of 0.98.

A. Computational Cost

First, the planners are compared by computational cost. The cost of planning is dominated by the predictions, the filter updates using the generated measure-

TABLE I: Number of times each expensive planning step is performed per time step. "P" refers to predictions, "F" refers to filter updates, and "R" refers to Rényi divergence calculations. For exhaustive search, the action sequences are arranged into a tree to avoid redundant filter updates. For MCTS, the number of filter updates is not constant, since it depends on the action selections while expanding the tree.

Method	P	F	R
Exhaustive, horizon 1	1	3	3
Exhaustive, horizon 2	2	12	9
Exhaustive, horizon 3	3	39	27
MCTS, horizon 4, 9 iterations	4	$\in [24, 30]$	9

TABLE II: Geometric mean over all scenarios for each target count (specified in the column heading) of the ratio of the mean runtime during planning for exhaustive search to that for MCTS with a horizon of 4 and 9 iterations.

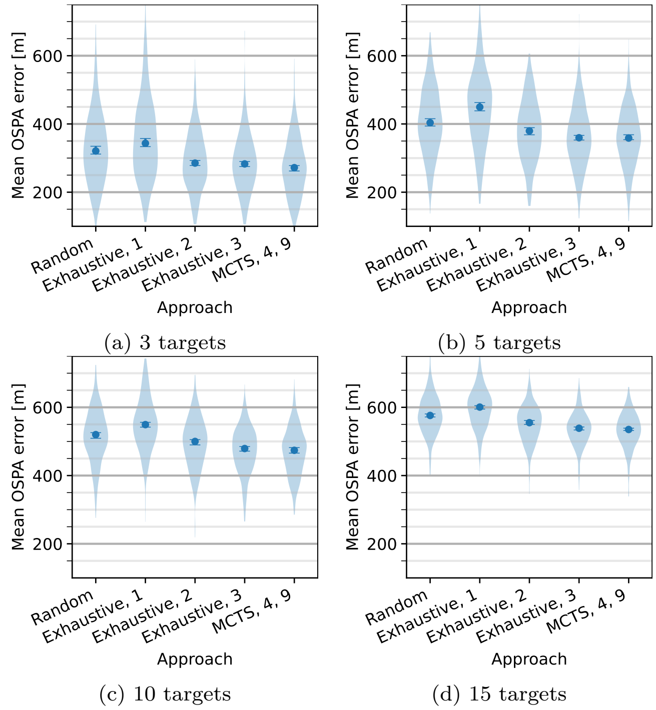
	Geometric mean of ratios			
Method	3	5	10	15
Exhaustive, horizon 1	0.15	0.15	0.16	0.16
Exhaustive, horizon 2	0.45	0.45	0.45	0.45
Exhaustive, horizon 3	1.13	1.15	1.09	1.19

ments, and the Rényi divergence evaluations, since these are only portions which operate on every particle. Each of these steps has asymptotic cost $O(\text{number of particles})$, although the number of particles is larger for the predictions, since some particles are discarded after computing the predictions, as described in Section IV-A.3. Table I lists the number of times these steps are performed. It shows that the cost of exhaustive search increases exponentially with the planning horizon. To compare the overall costs more directly, Table II shows the ratio of the runtime for each exhaustive search planner to the runtime for MCTS with a planning horizon of 4 and 9 iterations. The runtime of the MCTS approach is noticeably less than the runtime for exhaustive search with a horizon of 3.

B. OSPA Error

This section compares the quality of the estimated target positions for each planner, using the optimal subpattern assignment (OSPA) metric [39] between the subset of targets within the region of interest and the subset of target estimates within the region of interest. The parameters for the OSPA metric are $p = 2$ and $c = 1\,\text{km}$, and the distance function is the Euclidean distance. Figure 4 illustrates the empirical distribution of the mean OSPA error for each planning approach. It shows that exhaustive search with a horizon of 1 is actually worse than purely random actions, while planning over longer horizons is beneficial.

The horizon 1 planner performs poorly because it tends to myopically select actions which maximize the number of targets in the field-of-view (FoV) on the next time step; Fig. 5 shows that it maintains a large fraction of the targets of interest in the FoV. In contrast, the longer-horizon planners presented in this work more



(a) 3 targets (b) 5 targets (c) 10 targets (d) 15 targets

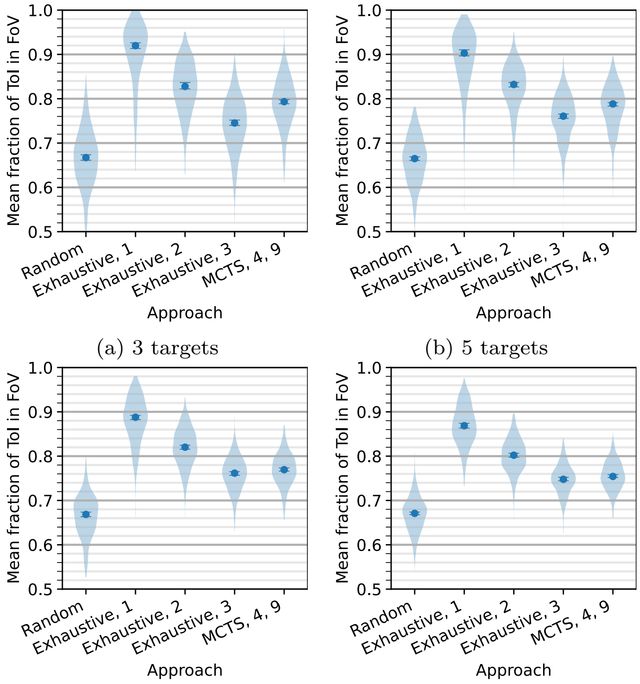
Fig. 4: Violin plots of the OSPA distance averaged over all time steps in each scenario. The circular markers indicate the medians over scenarios, and the error bars indicate 90% two-sided or 95% one-sided confidence intervals.

frequently choose actions with targets outside the FoV in exchange for better positioning the sensor for later time steps in order to obtain a better overall reward.

The remaining figures provide a pairwise comparison of the MCTS approach with a horizon of 4 and 9 iterations to the other methods. Figure 6 illustrates the empirical CDF of the pairwise difference between the mean OSPA error for each planner and that for MCTS. The curves show that the error for the MCTS approach is less than that for random actions and for exhaustive search with a horizon of 1 or 2, and it's no worse than the error for exhaustive search with a horizon of 3. There are two points on each curve to note in particular.

One point of interest is the vertical intercept; this is the fraction of the scenarios for which the given planner has lower error than MCTS. For random actions, this value is 0.24 to 0.32, so MCTS is better approximately 68% to 76% of the time. Similarly, MCTS is better than exhaustive search with a horizon of 1 approximately 81% to 84% of the time and better than exhaustive search with a horizon of 2 approximately 59% to 64% of the time. For exhaustive search with a horizon of 3, the only significant difference from 50% is for 3 targets, for which MCTS is slightly better.

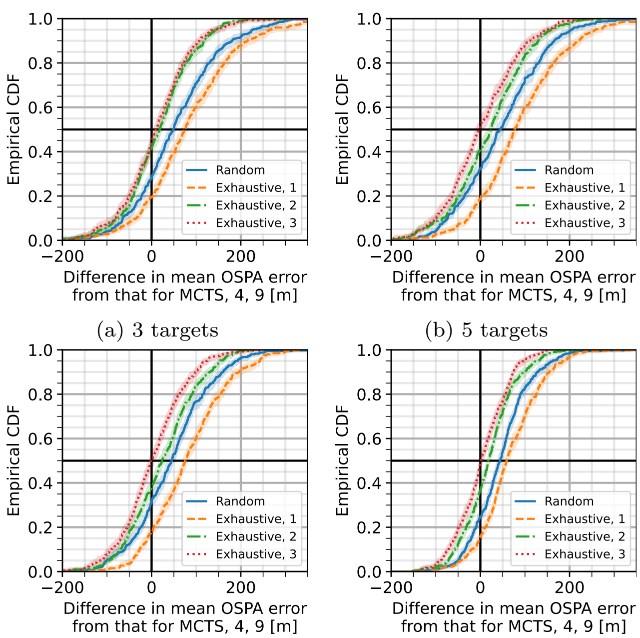
The other point of interest is horizontal intercept; this is the median difference in mean OSPA error for the planner from that for MCTS. These values are also summarized in Fig. 7, which shows that the median difference for random actions is 44 m to 48 m, that for



(a) 3 targets (b) 5 targets

(c) 10 targets (d) 15 targets

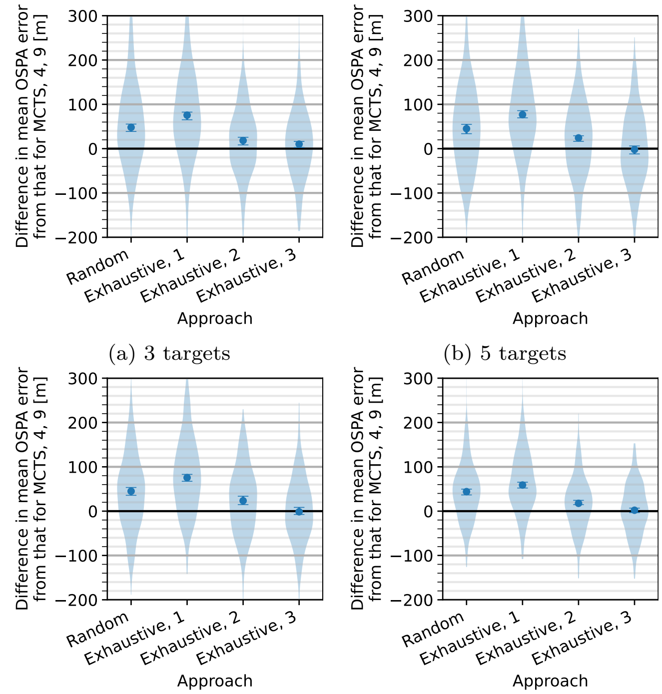
Fig. 5: Violin plots of the mean (over the time steps in the scenario) fraction of the targets of interest which were in the field-of-view. The circular markers indicate the medians over scenarios, and the error bars indicate 90% two-sided or 95% one-sided confidence intervals.



(a) 3 targets (b) 5 targets

(c) 10 targets (d) 15 targets

Fig. 6: Emprical CDF (where each scenario is a sample) of the difference in mean (over time steps in the scenario) OSPA error relative to MCTS with a horizon of 4 and 9 iterations. The shaded regions indicate the pointwise Wilson score 90% two-sided or 95% one-sided confidence intervals.



(a) 3 targets (b) 5 targets

(c) 10 targets (d) 15 targets

Fig. 7: Violin plots of the difference in mean OSPA distance from that for MCTS with horizon 4 and 9 iterations. The circular markers show the medians over scenarios; the error bars show 90% two-sided or 95% one-sided confidence intervals.

exhaustive search with horizon 1 is 59 m to 77 m, and that for exhaustive search with horizon 2 is 18 m to 24 m. In all cases, this is significantly greater than zero, so MCTS is better than random actions or exhaustive search with a horizon of 1 or 2. For exhaustive search with a horizon of 3, there is no significant difference except for 3 targets, for which MCTS is slightly better.

Based on these results, for the lowest overall OSPA error, exhaustive search with a horizon of 3 or the MCTS approach should be chosen. Since the MCTS approach has lower computational cost without increased OSPA error, it should be preferred.

VI. Conclusion

This paper demonstrates planning the motion of an observer over horizons greater than one step using a Rényi divergence reward and a PHD particle filter for bearings-only estimation of an unknown and varying number of indistinguishable, maneuvering targets of interest. This work may also be applicable to control of range-only sensors, which have similar limitations. It describes approximations to make this approach computationally feasible despite the particle representation of the filter's belief. It presents simulation results which show that planning with a horizon greater than one step reduces the OSPA error in the target estimates. Finally, it demonstrates that Monte Carlo tree search can reduce the cost of planning compared to exhaustive search without sacrificing the quality of the target estimates.

References

[1] W. Luo, J. Xing, A. Milan, X. Zhang, W. Liu, and T.-K. Kim, "Multiple object tracking: A literature review," *Artificial Intelligence*, vol. 293, p. 103448, 2021.

[2] A. Wolek, J. McMahon, B. R. Dzikowicz, and B. H. Houston, "Tracking multiple surface vessels with an autonomous underwater vehicle: Field results," *IEEE Journal of Oceanic Engineering*, pp. 1–14, 2020.

[3] Z. Wang, M. Li, Y. Lu, Y. Bao, Z. Li, and J. Zhao, "Effective multiple pedestrian tracking system in video surveillance with monocular stationary camera," *Expert Systems with Applications*, vol. 178, p. 114992, 2021.

[4] R. P. S. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, 2003.

[5] B. Ristic, M. Beard, and C. Fantacci, "An overview of particle methods for random finite set models," *Information Fusion*, vol. 31, pp. 110–126, 2016.

[6] B. Ristic, B.-N. Vo, and D. Clark, "A note on the reward function for PHD filters with sensor control," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 1521–1529, 2011.

[7] M. Beard, B.-T. Vo, B.-N. Vo, and S. Arulampalam, "Sensor control for multi-target tracking using Cauchy-Schwarz divergence," in *2015 18th International Conference on Information Fusion*, 2015, pp. 937–944.

[8] P. Dames and V. Kumar, "Autonomous localization of an unknown number of targets without data association using teams of mobile sensors," *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 3, pp. 850–864, 2015.

[9] A. El-Fallah, A. Zatezalo, R. Mahler, R. K. Mehra, and D. Donatelli, "Dynamic sensor management of dispersed and disparate sensors for tracking resident space objects," in *Signal Processing, Sensor Fusion, and Target Recognition XVII*, I. Kadar, Ed., vol. 6968, International Society for Optics and Photonics. SPIE, 2008, pp. 285–295.

[10] B. Ristic and B.-N. Vo, "Sensor control for multi-object state-space estimation using random finite sets," *Automatica*, vol. 46, no. 11, pp. 1812–1818, 2010.

[11] B. Ristic, *Particle Filters for Random Set Models*. Springer, 2013.

[12] A. K. Gostar, R. Hoseinnezhad, and A. Bab-Hadiashar, "Multi-Bernoulli sensor control via minimization of expected estimation errors," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 3, pp. 1762–1773, 2015.

[13] A. El-Fallah, A. Zatezalo, R. Mahler, R. K. Mehra, and M. Alford, "Unified Bayesian situation assessment sensor management," in *Signal Processing, Sensor Fusion, and Target Recognition XIV*, I. Kadar, Ed., vol. 5809, International Society for Optics and Photonics. SPIE, 2005, pp. 253–264.

[14] ——, "Regularized multitarget particle filter for sensor management," in *Signal Processing, Sensor Fusion, and Target Recognition XV*, I. Kadar, Ed., vol. 6235, International Society for Optics and Photonics. SPIE, 2006, pp. 224–234.

[15] P. Dames and V. Kumar, "Cooperative multi-target localization with noisy sensors," in *2013 IEEE International Conference on Robotics and Automation*, 2013, pp. 1877–1883.

[16] P. Dames, P. Tokekar, and V. Kumar, "Detecting, localizing, and tracking an unknown number of moving targets using a team of mobile robots," *The International Journal of Robotics Research*, vol. 36, no. 13–14, pp. 1540–1553, 2017.

[17] P. Dames, "Distributed multi-target search and tracking using the PHD filter," *Autonomous Robots*, vol. 44, no. 3, pp. 673–689, 2020.

[18] H. G. Hoang, B.-N. Vo, B.-T. Vo, and R. Mahler, "The Cauchy–Schwarz divergence for Poisson point processes," *IEEE Transactions on Information Theory*, vol. 61, no. 8, pp. 4475–4485, 2015.

[19] M. E. Gomes-Borges, D. Maltese, P. Vanheeghe, and E. Duflos, "A risk-based sensor management using random finite sets and POMDP," in *20th International Conference on Information Fusion*. IEEE, 2017.

[20] E. Delande, "Multi-sensor PHD filtering with application to sensor management," Ph.D. dissertation, Ecole Centrale de Lille, 2012. [Online]. Available: https://tel.archives-ouvertes.fr/tel-00688304/

[21] H. G. Hoang and B. T. Vo, "Sensor management for multi-target tracking via multi-Bernoulli filtering," *Automatica*, vol. 50, no. 4, pp. 1135–1142, 2014.

[22] M. Beard, B.-T. Vo, B.-N. Vo, and S. Arulampalam, "Void probabilities and Cauchy–Schwarz divergence for generalized labeled multi-Bernoulli models," *IEEE Transactions on Signal Processing*, vol. 65, no. 19, pp. 5047–5061, 2017.

[23] M. Jiang, W. Yi, and L. Kong, "Multi-sensor control for multi-target tracking using Cauchy-Schwarz divergence," Mar. 2016. [Online]. Available: arXiv:1603.08355v1

[24] A. K. Gostar, R. Hoseinnezhad, and A. Bab-Hadiashar, "Multi-Bernoulli sensor control using Cauchy-Schwarz divergence," in *2016 19th International Conference on Information Fusion*, 2016, pp. 651–657.

[25] A. K. Gostar, R. Hoseinnezhad, T. Rathnayake, X. Wang, and A. Bab-Hadiashar, "Constrained sensor control for labeled multi-Bernoulli filter using Cauchy–Schwarz divergence," *IEEE Signal Processing Letters*, vol. 24, no. 9, pp. 1313–1317, 2017.

[26] A. K. Gostar, R. Hoseinnezhad, A. Bab-Hadiashar, and W. Liu, "Sensor-management for multitarget filters via minimization of posterior dispersion," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 6, pp. 2877–2884, 2017.

[27] X. Wang, R. Hoseinnezhad, A. K. Gostar, T. Rathnayake, B. Xu, and A. Bab-Hadiashar, "Multi-sensor control for multi-object Bayes filters," *Signal Processing*, vol. 142, pp. 260–270, 2018.

[28] S. Panicker, A. K. Gostar, A. Bab-Hadiashar, and R. Hoseinnezhad, "Tracking of targets of interest using labeled multi-Bernoulli filter with multi-sensor control," *Signal Processing*, vol. 171, 2020.

[29] S. Papaioannou, P. Kolios, T. Theocharides, C. G. Panayiotou, and M. M. Polycarpou, "Jointly-optimized searching and tracking with random finite sets," *IEEE Transactions on Mobile Computing*, vol. 19, no. 10, pp. 2374–2391, 2020.

[30] A. J. Newman, S. R. Martin, J. T. DeSena, J. C. Clarke, J. W. McDerment, W. O. Preissler, and C. K. Peterson, "Receding horizon controller using particle swarm optimization for closed-loop ground target surveillance and tracking," in *Signal Processing, Sensor Fusion, and Target Recognition XVIII*, I. Kadar, Ed., vol. 7336, International Society for Optics and Photonics. SPIE, 2009, pp. 285–296.

[31] C. Freundlich, Y. Zhang, A. Z. Zhu, P. Mordohai, and M. M. Zavlanos, "Controlling a robotic stereo camera under image quantization noise," *The International Journal of Robotics Research*, vol. 36, no. 12, pp. 1268–1285, 2017.

[32] A. Mavrommati, E. Tzorakoleftherakis, I. Abraham, and T. D. Murphey, "Real-time area coverage and target localization using receding-horizon ergodic exploration," *IEEE Transactions on Robotics*, vol. 34, no. 1, pp. 62–80, 2018.

[33] C. Freundlich, Y. Zhang, and M. M. Zavlanos, "Distributed hierarchical control for state estimation with robotic sensor networks," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 4, pp. 2023–2035, 2018.

[34] C. Freundlich, S. Lee, and M. M. Zavlanos, "Distributed active state estimation with user-specified accuracy," *IEEE Transactions on Automatic Control*, vol. 63, no. 2, pp. 418–433, 2018.

[35] H. V. Nguyen, H. Rezatofighi, B.-N. Vo, and D. C. Ranasinghe, "Online UAV path planning for joint detection and tracking of multiple radio-tagged objects," *IEEE Transactions on Signal Processing*, vol. 67, no. 20, pp. 5365–5379, 2019.

[36] L. Calkins, P. Baldoni, J. McMahon, C. Wilhelmi, and M. M. Zavlanos, "Bearing-only active sensing under merged measurements," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4544–4551, 2021.

[37] J. Jimenez, A. Wolek, D. J. Stilwell, J. McMahon, and B. Dzikowicz, "Experimental results in bearings-only tracking using the sequential Monte-Carlo probability hypothesis density filter," in *Proc. SPIE 11018, Signal Processing, Sensor/Information Fusion, and Target Recognition XXVIII*, 2019.

[38] D. Auger, A. Couëtoux, and O. Teytaud, "Continuous upper confidence trees with polynomial exploration – consistency," in *Machine Learning and Knowledge Discovery in Databases*, ser. Lecture Notes in Computer Science, H. Blockeel, K. Kersting, S. Nijssen, and F. Železný, Eds., vol. 8188. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 194–209.

[39] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3447–3457, 2008.