

# Motion Planning for Active Data Association and Localization in Non-Gaussian Belief Spaces

Saurav Agarwal, Amirhossein Tamjidi, and Suman Chakravorty

Dept. of Aerospace Engineering,  
Texas A&M University,  
`sauravag,ahtamjidi,schakrav@tamu.edu`

**Abstract.** This paper presents a method for motion planning under uncertainty to resolve situations where ambiguous data associations result in a multimodal hypothesis on the robot state. Simultaneous localization and planning for a lost (or kidnapped) robot requires that given little to no a priori pose information, a planner should generate actions such that future observations allow the localization algorithm to recover the correct pose of a mobile robot with respect to a global reference frame. We present a Receding Horizon approach, to plan actions that sequentially disambiguate a multimodal belief to achieve tight localization on the correct pose in finite time. In our method, disambiguation is achieved through active data associations by picking target states in the map which allow distinctive information to be observed for each belief mode and creating local feedback controllers to visit the targets. Experimental results are presented for a kidnapped physical ground robot operating in an artificial maze-like environment.

## 1 Introduction

In practical mobile robot motion planning problems, situations may arise where data association between what is observed and the robot's map leads to a multimodal hypothesis on the state, for example a kidnapped robot with no a priori information or a mobile robot operating in a symmetric environment (see Fig. 1). Figure 1 depicts a problem wherein belief (the probability distribution over all possible robot states) modes are widely separated in an environment with symmetry. In such cases if a robot begins with an equal likelihood on all hypothesis, it is difficult to ascertain the true hypothesis as local sensing may result in identical information for all belief modes. Thus in practice a robot often has to seek information that helps to disambiguate its belief.

Simply relying on randomized actions to correctly recover robot pose is known to be unreliable and inefficient in practice [1]. Further, existing methods to disambiguate multimodal hypothesis [1, 10, 21] rely on heuristics-based strategies (e.g., picking random targets, wall following etc.) to seek disambiguating information. As opposed to [1, 10, 21], our approach disambiguates, i.e., rejects incorrect hypothesis in a multimodal belief by actively seeking maximally disambiguating information in the map for each mode, and recovers the robot pose with a higher certainty threshold than current state-of-the-art.

Our Multi-Modal Motion Planner (M3P) achieves disambiguation in a multimodal belief by first finding a neighboring location (referred to as target state) for each belief mode and then creating a candidate action to guide the belief mode to its target state such that these actions lead to information gathering behavior. The target states are chosen such that different modes of the robot's belief are expected to observe distinctive information at the target locations, thus accepting or rejecting hypotheses in the belief. We represent a multimodal hypothesis with a Gaussian Mixture Model (GMM) and use an Extended Kalman filter (EKF) based Multi-Hypothesis Tracking (MHT) approach to propagate the belief [9–11]. The main contributions of this work can be summarized as follows; (i) we develop a novel method for picking target states and creating candidate trajectories for a multimodal belief, our method then chooses the optimal candidate such that maximum disambiguating information is observed which helps in rejecting incorrect hypotheses, (ii) we prove that under certain realistic assumptions, through a process of iterative hypothesis elimination, our method can localize to the true robot pose, (iii) we demonstrate an application in which a kidnapped ground robot is tasked to recover its pose.

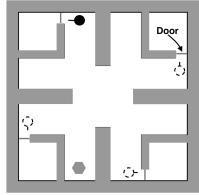


Fig. 1: A scenario depicting a multi-hypothesis localization problem with widely separated modes in a world with 4 rooms with identical doors. The true hypothesis is depicted by the solid black disk, whereas others are depicted by dashed circles. As the robot cannot distinguish between the doors, all hypotheses are equally likely.

We motivate this work with the kidnapped robot scenario since it is one of the hardest localization problems and measures the ability of an algorithm to recover from global localization failures [12]. However, the method proposed is general and can be extended to any planning situation where a multimodal belief arises in the robot state due to ambiguous data associations (a common practical issue in robot localization [12]). In the proceeding section, we present relevant related work, and discuss how our approach compares with them. In Section 3 we state some preliminaries followed by the problem description. In Section 4 we present our method followed by experimental results in Section 5.

## 2 Related Work

Recent work in sampling-based methods for belief space planning has shown promising results. Gaussian (unimodal) belief space planning methods such as [2–6, 8] provide solutions that depend on the initial belief. Recent developments in [7, 13] extend Gaussian belief space planning to multi-query settings (cases where multiple planning requests are made sequentially) by creating a belief space variant of a Probabilistic RoadMap (PRM) [14]. We note that the aforementioned methods assume that data associations between observations and information

sources (e.g., landmarks) are known and unambiguous. In contrast, we do not assume that data associations are unambiguous or that belief is unimodal. In our problem ambiguous data associations lead to a multimodal belief where the modes are widely separated (see Fig. 1), this violates the underlying Gaussian unimodal belief assumption in previously mentioned methods.

Recent work in [15, 16] extends belief space planning to non-Gaussian beliefs where the belief modes are not widely separated. The authors investigate a grasping problem with a multimodal hypothesis on the gripper’s state. Their method picks the most-likely hypothesis and a fixed number of samples from the belief distribution, then using an RHC approach, belief space trajectories are found that maximize the observation gap between the most-likely hypothesis and the drawn samples, which helps to accept or reject the most-likely hypothesis. The method in [17] builds upon the work in [15] wherein the author transposes the non-convex trajectory planning problem in belief space to a convex problem. Among other recent works, [18] reduces the computational complexity of planning for a non-Gaussian hypothesis but also assumes distributions without widely separated modes. Compared to [15–18], our method is better suited to deal with more severe cases of non-Gaussian belief space planning such as the kidnapped robot scenario. Such scenarios may not be possible to address using the trajectory optimization based techniques of [15–18] in their current form, due to the difficulty of generating an initial feasible plan for the widely separated modes in the presence of obstacles (as shown in Fig. 1).

To the extent of our knowledge, a limited number of methods approach the problem of recovering global robot pose for a mobile robot with an initial multimodal hypothesis. The analysis in [19] showed that finding the optimal (shortest) plan to re-localize a robot with multiple hypotheses in a deterministic setting (no sensing or motion uncertainty) is NP-hard. At best a greedy localization strategy can be developed whose plan length is upper bounded by a factor of the optimal plan. In a symmetric environment, [20] showed that for a robot equipped with only perfect odometry, no sequence of actions can disambiguate a pair of symmetric configurations. Compared to [19, 20], we do not assume perfect sensing or actuation. In [1], the authors develop an active localization method in a grid based scheme for a known map. Their planning method considers arbitrary targets in the robot’s local coordinate frame as atomic actions (e.g., move 1m right and 4m forward). The optimal action is selected based on the path cost and the expected decrease in entropy at the target. Compared to [1], our target selection methodology (Section 4.2) is active, i.e., M3P uses the a priori map information to select targets such that by visiting them, belief modes expect to see maximally disambiguating information (e.g., seeing a landmark with a distinctive appearance can immediately confirm or reject a hypothesis, see Fig. 2). In [10], the authors present a greedy heuristic-based planning strategy to disambiguate a multimodal hypothesis for a kidnapped robot. The method of [21] plans safe trajectories by picking a point in the vicinity of obstacles to disambiguate the hypothesis. Compared to [10, 21], we present a planning approach that explicitly reasons about the belief evolution as a result of actions in the planning stage and picks an optimal policy from a set of candidates.

### 3 Preliminaries and Problem

Let  $C$  be the configuration space and  $C_{free} \subset C$  be the set of collision free configurations. Let  $x_k \in \mathbb{X}$ ,  $u_k \in \mathbb{U}$ , and  $z_k \in \mathbb{Z}$  represent the system state, control input, and observation at time step  $k$  respectively.  $\mathbb{X}$ ,  $\mathbb{U}$ , and  $\mathbb{Z}$  denote the state, control, and observation spaces respectively. It should be noted that in our work, the state  $x_k$  refers to the state of the mobile robot, i.e., we do not model the environment and obstacles in it as part of the state. The non-linear state evolution model  $f$  and measurement model  $h$  are denoted as  $x_{k+1} = f(x_k, u_k, w_k)$  and  $z_k = h(x_k, v_k)$ , where  $w_k \sim \mathcal{N}(0, Q_k)$  and  $v_k \sim \mathcal{N}(0, R_k)$  are zero-mean Gaussian process and measurement noise, respectively. The belief  $b_k$  at time  $t_k$  can be represented by a Gaussian Mixture Model (GMM) as a weighted linear summation over Gaussian densities. Let  $w_{i,k}$ ,  $\mu_{i,k}$  and  $\Sigma_{i,k}$  be the weight, mean vector, and covariance matrix associated to the  $i^{th}$  Gaussian  $m_{i,k}$  respectively at time  $t_k$ , then  $b_k = \sum_{i=1}^{M_k} w_{i,k} m_{i,k}$ ,  $m_{i,k} \sim \mathcal{N}(\mu_{i,k}, \Sigma_{i,k})$ , where  $M_k$  is the number of modes at time  $t_k$ . We state our problem as follows:

*Given an a priori map, system dynamics and observation models, construct a belief space planner  $G(b_k)$  such that under the planner  $G$ , given an initial multimodal belief  $b_0$ , the sequence of future observations allow a robot to localize about its true pose.*

Note that there may be degenerate cases, where the map may not allow actions that lead to hypothesis elimination such that the belief converges to a unimodal distribution (e.g., in a map with two identical closed rooms, if a robot is kidnapped and placed in either room, it cannot distinguish which room it is in). In such cases, M3P attempts to minimize the number of modes  $M_k$  (by design), but it is not possible to pre-compute what this minimum value of  $M_k$  is without explicit knowledge of the true hypothesis [19] in a multimodal belief.

### 4 Methodology

We begin by defining certain key concepts used in the M3P planner.

*Uniqueness Graph:* A graph  $U_g$ , whose nodes are states sampled from the collision free space  $C_{free}$  and whose edges relate the similarity of information observed at the sampled locations.

*Target State:* A target state  $v_i^{tt} \in U_g$  for mode  $m_i$  is a node of the uniqueness graph which belongs to some neighborhood of radius  $R$  of the mode's mean  $\mu_i$  such that if each mode were to visit its target, the observations at the target would lead to disambiguation in the belief.

*Candidate Policy:* A candidate policy  $\pi_i$  for mode  $m_i$  is a local feedback controller that guides the mode to its target  $v_i^{tt}$ .

The M3P methodology has two phases, an offline phase in which we generate  $U_g$  and an online phase in which we use the offline computations and plan in a receding horizon manner to disambiguate the belief.

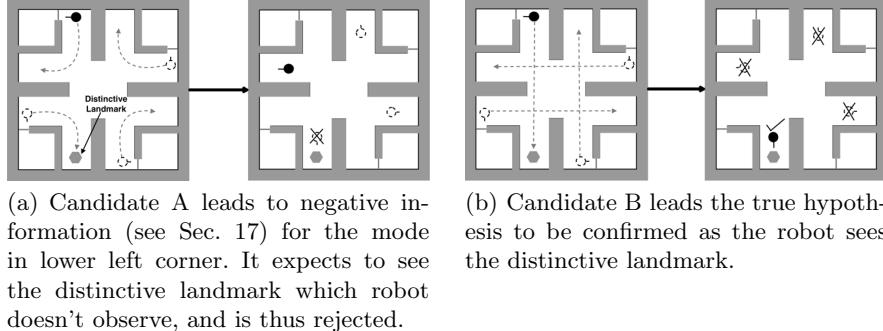


Fig. 2: Extending the example in Fig. 1, we depict how M3P creates candidate trajectories and picks the optimal one. For clarity we show only two candidates A & B and the effect of their execution. Candidate B results in complete disambiguation and is clearly a better choice.

#### 4.1 Computing the Uniqueness Graph: Offline Phase

The uniqueness graph  $U_g$  is constructed by uniformly sampling the configuration space and adding these samples as nodes of  $U_g$ . Once a node is added, we simulate the observation for the state represented by that node. Let  $v_\alpha$  be one such node and  $z^{v_\alpha}$  be the observation if the robot were to be in state  $v_\alpha$ . We add an edge  $E_{\alpha\beta}$  (undirected) between two nodes  $v_\alpha$  and  $v_\beta$  if the simulated observations at both nodes are similar. Further, the edges are weighted and the weight is dependent on the similarity in information observed, i.e., for edge  $E_{\alpha\beta}$  the weight  $\omega_{\alpha\beta} = \tau(z^{v_\alpha}, z^{v_\beta})$  where  $\tau : \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{R}$  computes a measure of similarity between two observations. Note that the form of  $\tau$  is general and can be changed to suit the problem domain (perception model). Figure 3 explains this concept visually for a landmark based observation model, where each landmark has some discrete signature (identifier) that a robot can detect. In Fig. 3 state  $v_\alpha$  observes  $z^{v_\alpha}$  with signatures  ${}^s z^{v_\alpha} = \{s_1, s_2, s_3\}$ , i.e., the landmarks with signature  $s_1, s_2$  and  $s_3$  and at  $v_\beta$  observes  ${}^s z^{v_\beta} = \{s_1, s_2, s_4\}$ , the edge weight  $\omega_{\alpha\beta}$  for edge  $E_{\alpha\beta}$  is  $\omega_{\alpha\beta} = \tau(z^{v_\alpha}, z^{v_\beta}) = |{}^s z^{v_\alpha} \cap {}^s z^{v_\beta}| = |\{s_1, s_2\}| = 2$ . A higher edge weight signifies that the states represented by the vertices of that edge are more likely to observe similar information. The lack of an edge between two nodes means that if a robot were to make an observation at those two states, it would see distinctly different information.

The complexity for the construction of  $U_g$  is  $\mathcal{O}(n^2)$  (where  $n$  is the number of samples) as each sample (node) is checked with every other for information overlap. Due to its random nature, sampling may often occur in regions of low information density (e.g., regions where there are few or no landmarks). One can often circumvent this issue by increasing the number of samples. As  $U_g$  is computed offline, the online performance is not significantly affected. Recent work

in [22] suggests a localization aware sampling strategy which may be explored in future work.

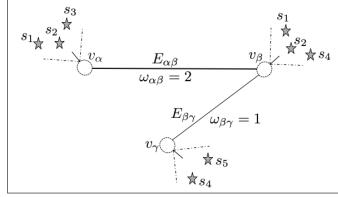


Fig. 3: Simple example of a uniqueness graph with 3 nodes  $\{v_\alpha, v_\beta, v_\gamma\}$  and 2 edges  $\{E_{\alpha\beta}, E_{\beta\gamma}\}$ . The nodes  $v_\alpha$  and  $v_\gamma$  do not see any similar landmark hence there is no edge between them. Here  $\tau(z^{v_i}, z^{v_j}) = |s_z^{v_i} \cap s_z^{v_j}|$  for  $i, j \in \{\alpha, \beta, \gamma\}$ .

## 4.2 RHC based Planning: Online Phase

In a multimodal scenario, we claim that the best action to take is one that guides a robot without collision through a path that results in information gain such that a disambiguation occurs (one or more hypotheses are rejected, see Fig. 2). Algorithm 1 describes the online planning process. In step 3, the planner picks target states for each belief mode such that visiting a target can either prove or disprove the hypothesis. In step 4, the planner generates a set of candidate policies to drive each mode to its target. In step 5, the expected information gain for each policy is computed and we pick the best one, and in step 7, the multimodal belief is propagated according to the action and observations. We proceed to describe steps 3, 4, 5 and 7 of Algorithm 1 below.

---

### Algorithm 1: M3P: MultiModal Motion Planner

---

```

1 Input:  $b$ 
2 while  $b \neq \mathcal{N}(\mu, \Sigma)$  do
3    $\{v^{tt}\} \leftarrow$  Pick target states for belief modes (see Alg. 2);
4    $\Pi \leftarrow$  Generate candidate policies to connect each mode to its target;
5    $\pi^* \leftarrow$  Pick optimal policy from  $\Pi$ ;
6   forall the  $u \in \pi^*$  do
7      $b \leftarrow$  Apply action  $u$  and update belief (see Alg. 3 for weight update
      calculation);
8     if Change in number of modes or Expect a belief mode to violate
      constraints then
9       break;
10  return  $b$ ;

```

---

**Picking the target state for a mode** Algorithm 2 describes in detail the steps involved to pick a target state for a belief mode. Let us pick a mode

$m_{i,k} \sim \mathcal{N}(\mu_{i,k}, \Sigma_{i,k})$  from the belief. To find the target  $v_{i,k}^{tt}$  for  $m_{i,k}$ , we first choose the set of nodes  $N_{i,k} \in U_g$  (Section 4.1) which belong to the neighborhood of the mean  $\mu_{i,k}$  at time  $t_k$  (steps 3 and 4, Alg. 2). Then, we find the target node  $v_{i,k}^{tt} \in N_{i,k}$  which observes information that is least similar in appearance to that observed by nodes in the neighborhoods  $N_{j,k}$  of all other modes  $m_{j,k}$  where  $j \neq i$ . To do this, after computing  $N_{i,k}$ , we calculate the total weight of the outgoing edges from every node  $v_{i,k} \in N_{i,k}$  to nodes in all other neighborhoods  $N_{j,k}$  where  $j \neq i$  (steps 7-13, Alg. 2). The node which has the smallest outgoing edge weight (steps 14-16, Alg. 2), is the target candidate  $v_{i,k}^{tt}$  for  $m_{i,k}$  as the observation  $z^{v_{i,k}^{tt}}$  would be least similar to the information observed in the neighborhood of all other modes  $m_j$  where  $j \neq i$ .

---

**Algorithm 2:** Finding the target for  $i$ -th mode

---

```

1 Input:  $b_k, i, U_g$ 
2 Output:  $v_{i,k}^{tt}$ 
3 forall the  $l \in [1, M_k]$  do
4    $N_{l,k} \leftarrow$  Find neighborhood nodes for  $\mu_{l,k}$  in  $U_g$ ;
5    $minWeight \leftarrow$  Arbitrarily large value;
6    $v_{i,k}^{tt} \leftarrow -1$ ;
7 forall the  $v \in N_{i,k}$  do
8    $w \leftarrow 0$ ;
9   for  $N_{j,k} \in \{N_{1,k}, \dots, N_{M_k,k}\} \setminus N_{i,k}$  do
10    forall the  $e \in$  Edges connected to  $v$  do
11     forall the  $p \in N_{j,k}$  do
12       if  $p$  is a target of edge  $e$  then
13          $w \leftarrow w + \text{edgeWeight}(e)$ ;
14   if  $w < minWeight$  then
15      $minWeight \leftarrow w$ ;
16      $v_{i,k}^{tt} \leftarrow v$ ;
17 return  $v_{i,k}^{tt}$ ;

```

---

**Generating candidate policies for belief modes** Once the targets corresponding to each mode have been picked, we need to find the control action that can take a mode from its current state to the target state. We generate the candidate trajectory that takes each mode to its target using the RRT\* planner [23]. Once an open loop trajectory is computed, we generate a local policy  $\pi_i$  (feedback controller) for the  $i$ -th mode, which drives the  $i$ -th mode along this trajectory. Let  $\Pi$  be the set of all such policies for the different modes.

**Picking the Optimal Policy** After generating the set  $\Pi$  of candidate policies, we evaluate the expected information gain  $\Delta I_i$  for each policy  $\pi_i$  and pick the

optimal policy  $\pi^*$  that maximizes this information gain. We model this information gain as the discrete change in the number of modes. To compute the expected change in the number of belief modes, we simulate the most-likely belief trajectory, i.e., approximating noisy observations and actions with their most-likely values [8, 24–26]. The steps to calculate the expected information gain for a candidate policy  $\pi_i \in \Pi$  are as follows:

1. For every belief mode  $m_{j,k} \in b_k$ .
  - (a) Assume that robot is at  $m_{j,k}$ .
  - (b) Simulate  $\pi_i$  and propagate all the modes.
  - (c) Compute information gain  $\Delta I_{i,m_{j,k}}$  for  $\pi_i$ .
2. Compute the weighted information gain  $\Delta I_i = \sum_{j=1}^{M_k} w_{j,k} \Delta I_{i,m_{j,k}}$ .

After computing the expected information gain for each policy, we pick the gain maximizing policy. The computational complexity of this step is  $\mathcal{O}(M_k^3 L_{max})$  (where  $M_k$  is the number of belief modes and  $L_{max}$  is the maximum candidate trajectory length). This is due to the fact that each policy is simulated for each mode for the length of policy, where at every step of policy execution, there are  $M_k$  filter updates. Figure 2 depicts the process of picking the optimal candidate trajectory in a hypothetical scenario.

**Belief Propagation Using GMM** We first discuss our decision to use EKF based MHT over a particle filtering approach. In practical localization problems, a relatively small number of Gaussian hypotheses are sufficient for maintaining the posterior over the robot state, secondly the filtering complexity grows linearly in the number of hypotheses and finally due to the computational complexity of picking the optimal policy (see previous section), the number of samples required for a particle filter would make re-planning significantly harder.

Now, we proceed to describe the weight update step which determines how likely each mode is in the belief. In a standard implementation, the weights  $w_{i,k}$ 's are updated based on the measurement likelihood function as

$$w_{i,k+1} = w_{i,k} e^{-\frac{1}{2} D_{i,k+1}^2}, \quad (1)$$

where  $D_{i,k+1}$  is the Mahalanobis distance between the sensor observation and most-likely observation for mode  $m_i$  such that

$$D_{i,k+1}^2 = (z_{k+1} - h(\mu_{i,k+1}, 0))^T R_k^{-1} (z_{k+1} - h(\mu_{i,k+1}, 0)). \quad (2)$$

The weights are normalized such that  $\sum_{i=1}^{M_k} w_{i,k+1} = 1$ . A known issue with EKF-based MHT is that it is unable to process negative information [12]. Negative information refers to the lack of information which one may expect to see and can certainly help in disproving a hypothesis (see Fig. 2(a)). We now proceed to describe how negative information is factored into the weight update.

*Factoring Negative Information:* Depending on the state of the robot, individual hypotheses and data association results, we might have several cases. We discuss this issue in the context of a landmark based measurement model.

At time  $t_{k+1}$ , let  $n_{z_{k+1}}$  be the number of landmarks observed by the robot and  $n_{z_{i,k+1}^p}$  be the number of landmarks that we predict to see for  $m_i$  where  $z_{i,k+1}^p = h(\mu_{i,k+1}, 0)$  is the predicted observation. Then  $n_{z_{k+1}} = n_{z_{i,k+1}^p}$  means that the  $i$ -th mode expected to see as many landmarks as the robot observed;  $n_{z_{k+1}} > n_{z_{i,k+1}^p}$  implies the robot observes more landmarks than predicted for the mode;  $n_{z_{k+1}} < n_{z_{i,k+1}^p}$  implies the robot observes less landmarks than predicted for the mode. Also, we can have the number of data associations to be less than the number of predicted or measured observations or both. This means that we may not be able to make a unique association between each predicted and observed landmark. At time  $t_{k+1}$ , we estimate the Mahalanobis distance  $D_{i,k+1}$  (Eq. 2) for mode  $m_i$  between the predicted and observed landmarks that are matched by the data association module and update weight according to Eq. 1. Then we multiply the updated weight by a factor  $\gamma$ , which models the effect of duration  $\beta_{i,k+1}$  for which the robot observes different landmarks than the  $i$ -th mode's prediction; and the discrepancy  $\alpha$  in the number of data associations. When a belief mode is initialized, we set  $\beta_{i,0} = 0$ . The weight update procedure is described in Algorithm 3. After each weight update step, we remove modes with negligible contribution to the belief, i.e., when  $w_{i,k+1} \leq \delta_w$  where  $\delta_w$  is user defined.

---

**Algorithm 3:** GMM Weight Update

---

```

1 Input:  $w_{i,k}, \mu_{i,k+1}, \beta_{i,k}, \delta t$ 
2 Output:  $w_{i,k+1}, \beta_{i,k+1}$ 
3  $z_{k+1}, n_{z_{k+1}} \leftarrow$  Get sensor observations;
4  $z_{i,k+1}^p, n_{z_{i,k+1}^p} \leftarrow$  Get predicted observations for  $\mu_{i,k+1}$ ;
5  $n_{z_{k+1} \cap z_{i,k+1}^p} \leftarrow$  Do data association;
6  $w'_{i,k+1} \leftarrow$  Update and normalize weight according to likelihood function;
7  $\gamma \leftarrow 1$ ;
8 if  $n_{z_{i,k+1}^p} \neq n_{z_{k+1}}$  or  $n_{z_{i,k+1}^p} \neq n_{z_{k+1} \cap z_{i,k+1}^p}$  then
9    $\alpha \leftarrow \max(1 + n_{z_{k+1}} - n_{z_{k+1} \cap z_{i,k+1}^p}, 1 + n_{z_{i,k+1}^p} - n_{z_{k+1} \cap z_{i,k+1}^p})$ ;
10   $\beta_{i,k+1} \leftarrow \beta_{i,k} + \delta t$ ;
11   $\gamma \leftarrow e^{-\alpha \beta_{i,k+1} 10^{-4}}$ ;
12 else
13    $\beta_{i,k+1} \leftarrow 0$ ;
14  $w_{i,k+1} \leftarrow w'_{i,k+1} \gamma$ ;
15 return  $w_{i,k+1}, \beta_{i,k+1}$ ;

```

---

### 4.3 Analysis

In this section, we show that under certain assumptions on the structure of the environment, the receding horizon planner M3P can guarantee that an initial

multimodal belief is driven into a unimodal belief in finite time. We now proceed to state our assumptions.

**Assumption 1** *For every mode  $m_i$ , the environment allows for the existence of some target state  $v_i^{tt}$  and some homotopy class of paths through which the robot can visit  $v_i^{tt}$  if the robot is actually at mode  $m_i$ .*

**Assumption 2** *If the robot is actually at mode  $m_i$ , and its associated target state is  $v_i^{tt}$ , let  $B_r(v_i^{tt})$  to be a neighborhood of radius  $r > 0$  centered at the target  $v_i^{tt}$  such that if the robot state  $x \in B_r(v_i^{tt})$ , exteroceptive observations can confirm that  $m_i$  is the true hypothesis.*

Due to the uncertain nature of the actuation and sensing process, the existence of a path to visit a target location does not guarantee that a robot can drive its belief along this path or that on reaching neighborhood  $B_r(v_i^{tt})$ , localization uncertainty will be sufficiently low so as to make a disambiguating data association. Let the true belief be mode  $m_i$ . Let  $F \subset C \setminus C_{free}$  be the set of failure states, and let  $L$  be the finite stopping time for policy  $\pi_i$  defined as the time at which collision occurs or the belief mean  $\mu_i$  reaches the neighborhood  $B_r(v_i^{tt})$ . Denote  $P^{\pi_i}(x_L \in F|m_i)$  as the probability that policy  $\pi_i$  drives the underlying state  $x$  into a collision given the initial belief is  $m_i$ .

**Assumption 3** *Given Assumption 1, let mode  $m_i \sim \mathcal{N}(\mu_i, \Sigma_i)$ , with  $\|\Sigma_i\| < \bar{P} < \infty$  (initial covariance is bounded) be the true hypothesis. We assume that under the feedback policy  $\pi_i$ , the failure probability  $P^{\pi_i}(x_L \in F|m_i)$  is sufficiently low such that we can drive the robot state  $x$  into the neighborhood  $B_r(v_i^{tt})$  with a high probability  $\int_{B_r(v_i^{tt})} p^{\pi_i}(x_L|m_i, \neg F) dx > 1 - \delta$  for any  $\delta > 0$  where  $p^{\pi_i}(x_L|m_i, \neg F)$  is the terminal pdf on the state under policy  $\pi_i$  when the robot does not collide.*

**Assumption 4** *The environment (world) in which the robot operates is static.*

**Proposition 1.** *Under Assumptions 1, 2, 3 and 4, given any initial multimodal belief  $b_0 = \sum_i w_{i,0} m_{i,0}$ , the receding horizon planner M3P drives the belief process into a unimodal belief  $b_T = m_T \approx \mathcal{N}(\mu_T, \Sigma_T)$  in some finite time  $T$ .*

*Proof.* Given an initial belief  $b_0$ , let  $\pi_{i^*}$ , i.e., candidate policy for mode  $m_{i^*}$ , be the one that results in most information gain as required by M3P. We have only two possibilities; (i) Case 1: Mode  $m_{i^*}$  is the true hypothesis, or (ii) Case 2: Mode  $m_{i^*}$  is not the true hypothesis. If case 1 is true, due to Assumptions 1, 2 and 3, M3P can confirm that  $m_{i^*}$  is the true hypothesis by visiting the target location and rejecting all other hypotheses in the process (see Fig. 2(b)). If case 2 is true then the robot is at some other mode  $m_j$  where  $j \neq i^*$ . In case 2, as policy  $\pi_{i^*}$  is executed, two situations can arise, either (i)  $\pi_{i^*}$  is executed fully in which case  $m_{i^*}$  will expect to see distinctive information at its target location which the robot will not observe, leading to a disambiguation immediately due to negative information (see Fig. 2(a)) or (ii) the policy  $\pi_{i^*}$  becomes unfeasible at some point of its execution in which case we immediately know that the

robot is not at mode  $i^*$  since we know that the map did not change during the execution of  $\pi_{i^*}$  (Assumption 4) and thus, there is a disambiguation whereby mode  $i^*$  is discarded. Thus we see that either  $\pi_{i^*}$  confirms the true hypothesis or the number of modes is reduced by at least one. After this disambiguation, we restart the process as before and we are assured that at least one of the modes is going to be disambiguated and so on. Thus, it follows given that we had a finite number of modes to start with, the belief eventually converges to a unimodal belief. Further, since each of the disambiguation epochs takes finite time, a finite number of such epochs also takes a finite time, thereby proving the result.

**Remarks:** The above result shows that the M3P algorithm will stabilize the belief process to a unimodal belief under Assumptions 1, 2, 3 and 4. In the case that Assumption 1 is violated we are either (i) unable to find a target which allows the robot to observe distinctive information (e.g., trivial case of a robot operating in a world with identical infinite corridors) or (ii) we may find such a target but the environment geometry does not allow for any path to visit it (e.g., robot stuck in one of many identical rooms and the doors are closed). These violations refer to degenerate cases that rarely occur in practical motion planning problems. Assumptions 2 and 3 can be violated when all candidate trajectories pass through regions lacking enough information, either because the region is unknown or featureless. In such a case the localization uncertainty on each mode may grow so high that we cannot make data associations at the target location to disambiguate the multimodal belief. Thus these two assumptions imply that the known map has enough information sources (see Fig. 4). Handling the issue of maps that are either unknown, partially known or sparse in information sources is beyond the scope of this paper and presents an important direction for future research. Assumption 4 (static world) is common in localization literature, though it may be violated in certain scenarios. In such cases, if the map is not changing rapidly, one may use sensory observations to incorporate new constraints into the map and trigger replanning.

## 5 Experimental Results

We present experimental results for two motion planning scenarios wherein the robot is placed randomly at a location in an environment which is identical to other locations in appearance<sup>1</sup>. Thus the initial belief is multimodal, the goal of the experiment is to use the non-Gaussian planner M3P described in Section 4 to localize the robot pose. We first describe the system setup to motivate the experiment followed by the results.

### 5.1 System Description

We used a low-cost Arduino based differential drive robot (shown in Fig. 5(a)) equipped with an Odroid U3 computer running ROS on Ubuntu 14.04 and an off-the-shelf Logitech C-310 webcam for sensing. The onboard computer uses a wifi

---

<sup>1</sup> Due to paucity of space we only present one experiment here, a supplementary video is provided that clearly depicts every stage of both our experiments.

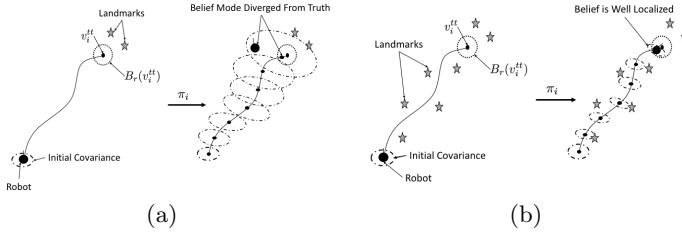


Fig. 4: Evolution of the true belief mode in environments with and without sufficient information. (a) No landmarks present along the candidate trajectory, leading to high uncertainty at the end. The belief mode has diverged from the robot pose and it is no more possible to make an accurate data association for the landmarks at the target. (b) Sufficient information along the candidate trajectory leads the belief mode to be well localized at the end, allowing unambiguous data association for the landmarks at the target.

link to communicate with the ground control station (laptop running ROS on Ubuntu 14.04). The ground station runs the planner and image processing algorithms while communicating with the robot via wifi. The kinematics of the robot are represented by a standard unicycle motion model. The observation model is a vision-based range bearing sensor augmented with appearance information (see [12] Sec. 6.6.2) such that  $z_k = h(x_k, v_k) = [(r_1, \phi_1, s_1)^T, (r_2, \phi_2, s_2)^T, \dots]$  where  $r_l, \phi_l, s_l$  are the range, bearing and signature for the  $l$ -th observed landmark. The signature is an integer value and identical landmarks have the same signature<sup>2</sup>. For this observation model, the function  $\tau$  (compute information overlap between two observations, see Sec. 4.1) is identical to that described in Fig. 3. In the real world, landmark appearances may change due to environmental conditions (e.g., lighting), perspective etc., which may adversely affect detection, such issues require more complex perception models and map representations which are outside the scope of this work.

## 5.2 Scenario

We constructed a symmetrical maze that has 8 identical rooms (R1-8) as shown in Fig. 5(b). Augmented reality (AR) markers were placed on the walls which act as the landmarks detected by the vision-based sensing system of the robot [28]. When the robot sees a landmark, it can detect the range, bearing as well as its signature. To create ambiguity in the data association, we placed multiple AR markers with the same signature in different parts of the environment. For example, one of the symmetries in our experiment is the inside of each room. Each room in the maze appears identical to the robot as markers with the same appearance are placed on each room's walls with an identical layout. Thus, if

<sup>2</sup> A detailed description of the motion and observation model parameters is omitted in the interest of space, we refer the reader to our pre-print version [27] of this paper.

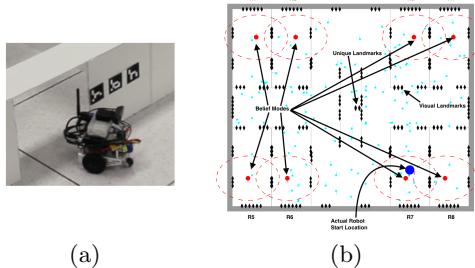
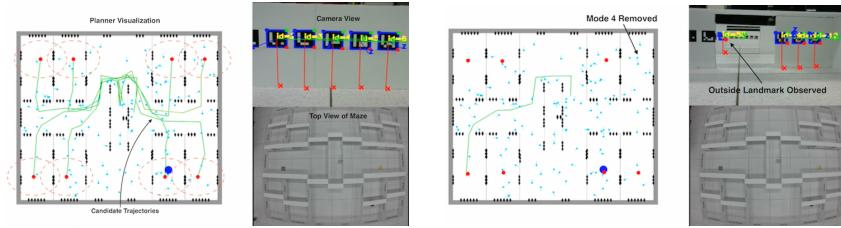


Fig. 5: (a) The robot exiting a room in the maze. It has an 11.5 cm wheelbase and measures 18 cm and 15 cm in height and length respectively. (b) The environment with 8 rooms marked R1-R8 and belief at the start of first run. Robot is placed in room R7 (blue disk), initial sampling leads to 8 belief modes, one in each room. The black diamonds mark the locations of augmented reality markers in the environment. The unique landmarks are placed inside the narrow passage, such that if the robot enters the passage from either side, it sees distinctive information.

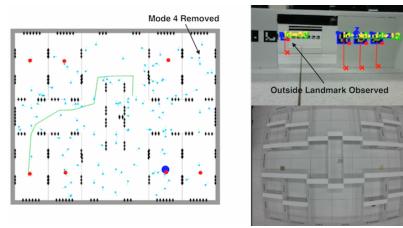
the robot is placed in a location with markers similar to another part of the environment, the data associations lead the robot to believe it could be in one of these many locations, which leads to a multimodal belief on the state. We also place four unique markers in a narrow passage in the center of maze as marked in Fig. 5(b). To successfully localize, the robot must visit this location in order to converge to its true belief.

The robot is initially placed in room R7 and is not given any prior information of its state. To estimate  $b_0$ , we uniformly sample the configuration space and set these samples as the means  $\mu_{i,k}$  of the modes of the Gaussian mixture components and assign identical covariance and uniform weight to each mode. After this, the robot remains stationary and the sensory measurements are used to update the belief state and remove the unlikely modes with weight  $w \leq \delta_w = 0.01$ . This process of elimination continues until we converge to a fixed number of modes. Figure 5(b) shows the initial belief. The robot plans its first set of candidate actions as shown in Fig. 6(a). After the candidates are evaluated, the policy based on mode  $m_5$  in room R5 is chosen and executed. As the robot turns, it sees a landmark on the wall outside the room (shown in Fig. 6(b)). This causes mode  $m_4$  to be deleted. Immediately, replanning is triggered and a new set of candidate trajectories is created. In successive steps, we see that first modes  $m_3$  and  $m_5$  are deleted and then after the next two replanning steps, modes  $m_8$ ,  $m_1$  and  $m_6$  are deleted. We notice that the robot does not move till only the 2 most-likely modes are remaining. The reason for this is that seeing the marker on the outside wall has the effect of successively lowering the weights of the unlikely modes. As the mode weights fall below the threshold, they are deleted, which triggers the replanning condition. Once the belief has converged to the two most-likely modes  $m_2, m_7$  (as expected by the

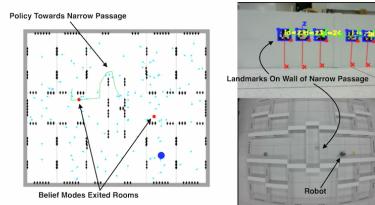
symmetry) a new set of candidate policies is created and the policy based on mode  $m_2$  is chosen. This policy leads the modes out of the rooms, and towards the narrow passage. Figure 6(c) shows both belief modes executing the policy based on mode  $m_2$ . While executing this policy, replanning is triggered as the robot exceeds maximum horizon (60 secs) for policy execution. The final policy drives the robot into the narrow passage and the unique landmarks are observed (Fig. 6(d)) which leads the belief to converge to the robot pose.



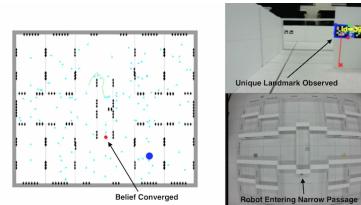
(a) The planner visualization showing the candidate trajectories (green). The top right image shows the view from the onboard camera, with the detected marker information overlaid. The bottom-right image shows the top-view of the maze in which the robot is run.



(b) The robot observes landmark ID 55 on the door of the opposite room causing the weights of modes  $m_1, m_3, m_4, m_5, m_6, m_8$  to gradually decrease which leads to these modes being removed from the belief.



(c) The robot has exited the room and is looking at the outside wall of the narrow passage. The two modes  $m_2$  and  $m_7$  are symmetrically located in the map, due to the information in the map that is observed by the robot.



(d) The belief mode has converged to the true belief as the robot enters the narrow passage and observes the unique landmark (ID 39).

Fig. 6: Snapshots of first run of the experiment at different times.

### 5.3 Discussion

Our approach results in a behavior which guides the robot to seek disambiguating information. The candidate trajectories are regenerated every time a belief mode is rejected or a constraint violation is foreseen and the time to re-plan reduces drastically as the number of modes reduce. Thus, the first few actions are the hardest which is to be expected as we start off with a large number of hypotheses. Finally, the planner is able to localize the robot safely. In [1], the authors showed that random motion is inefficient and generally incapable of localizing a robot within reasonable time horizons especially in cases with symmetry (e.g., office environments with long corridors and similar rooms). In [10] the authors consider the robot localized when one of the modes gets a weight  $\geq 0.8$ , in contrast our approach is more conservative in that we only consider the robot localized when a mode has weight  $\geq 0.99$ . We can afford to be more conservative as our localization strategy actively seeks disambiguating information using prior map knowledge as opposed to a heuristic based strategy. While our experiment acts as a proof of concept, there are certain phenomenon such as cases where the belief modes split into child modes, or dynamic environments which were not covered and will be addressed in future work.

## 6 Conclusion

In this work, we studied the problem of mobile robot motion planning for active data association in order to correctly localize a robot when the initial underlying belief is multimodal (non-Gaussian). Our main contribution in this work is a planner M3P that generates a sequentially disambiguating policy through active data association, which leads the belief to converge to the true hypothesis. We are able to show in practice that the robot is able to recover from a kidnapped state and localize in environments that present ambiguous data associations such that the underlying belief modes are widely separated. Compared to previous works, we take a non-heuristic approach to candidate policy generation and selection, while remaining conservative in accepting the true hypothesis.

A current limitation may be the computational cost for the policy selection step in large maps which lead to a high number of hypotheses. Future work will look at reducing this cost and experiments will be extended to larger problems (e.g., symmetric office environments), with more complex perception models and drastic localization failures (e.g., sequential kidnappings). Finally, there may be tasks which are feasible with a multimodal distribution on the belief. Such cases present an interesting area for future motion planning research.

## References

- [1] Fox, D., Burgard, W., Thrun, S.: Active markov localization for mobile robots. *Robotics and Autonomous Systems* **25**(34) (1998) 195 – 207 Autonomous Mobile Robots.
- [2] Prentice, S., Roy, N.: The belief roadmap: Efficient planning in belief space by factoring the covariance. *International Journal of Robotics Research* **28**(11-12) (October 2009)
- [3] Bry, A., Roy, N.: Rapidly-exploring random belief trees for motion planning under uncertainty. In: ICRA. (2011) 723–730
- [4] Chaudhari, P., Karaman, S., Hsu, D., Frazzoli, E.: Sampling-based algorithms for continuous-time POMDPs. In: the American Control Conference (ACC), Washington DC (2013)
- [5] van den Berg, J., Abbeel, P., Goldberg, K.: LQG-MP: Optimized path planning for robots with motion uncertainty and imperfect state information. In: Proceedings of Robotics: Science and Systems (RSS). (June 2010)
- [6] Kurniawati, H., Bandyopadhyay, T., Patrikalakis, N.: Global motion planning under uncertain motion, sensing, and environment map. *Autonomous Robots* **33**(3) (2012) 255–272
- [7] Agha-mohammadi, A., Chakravorty, S., Amato, N.: FIRM: Sampling-based feedback motion planning under motion uncertainty and imperfect measurements. *International Journal of Robotics Research* **33**(2) (2014) 268–304
- [8] Platt, R., Tedrake, R., Kaelbling, L., Lozano-Perez, T.: Belief space planning assuming maximum likelihood observations. In: Proceedings of Robotics: Science and Systems (RSS). (June 2010)
- [9] Reuter, J.: Mobile robot self-localization using pdab. In: Robotics and Automation, 2000. Proceedings. ICRA '00. IEEE International Conference on. Volume 4. (2000) 3512–3518 vol.4
- [10] Jensfelt, P., Kristensen, S.: Active global localization for a mobile robot using multiple hypothesis tracking. *Robotics and Automation, IEEE Transactions on* **17**(5) (Oct 2001) 748–760
- [11] Roumeliotis, S.I., Bekey, G.A.: Bayesian estimation and kalman filtering: A unified framework for mobile robot localization. In: Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on. Volume 3.. IEEE (2000) 2985–2992
- [12] Thrun, S., Burgard, W., Fox, D.: Probabilistic robotics. MIT press (2005)
- [13] Agha-mohammadi, A., Agarwal, S., Mahadevan, A., Chakravorty, S., Tomkins, D., Denny, J., Amato, N.: Robust online belief space planning in changing environments: Application to physical mobile robots. In: IEEE Int. Conf. Robot. Autom. (ICRA), Hong Kong, China (2014)
- [14] Kavraki, L., Svestka, P., Latombe, J., Overmars, M.: Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *IEEE Transactions on Robotics and Automation* **12**(4) (1996) 566–580
- [15] Platt, R., Kaelbling, L., Lozano-Perez, T., , Tedrake, R.: Efficient planning in non-Gaussian belief spaces and its application to robot grasping. In: Proc. of International Symposium of Robotics Research, (ISRR). (2011)
- [16] Platt, R., Kaelbling, L., Lozano-Perez, T., Tedrake, R.: Non-gaussian belief space planning: Correctness and complexity. In: ICRA. (2012)
- [17] Platt, R.: Convex receding horizon control in non-Gaussian belief space. In: Workshop on the Algorithmic Foundations of Robotics (WAFR). (2012)
- [18] Rafieisakhaei, M., Tamjidi, A., Chakravorty, S., Kumar, P.: Feedback motion planning under non-gaussian uncertainty and non-convex state constraints. In: 2016 IEEE International Conference on Robotics and Automation (ICRA), IEEE (2016) 4238–4244
- [19] Dudek, G., Romanik, K., Whitesides, S.: Localizing a robot with minimum travel. *SIAM Journal on Computing* **27**(2) (1998) 583–604
- [20] O'Kane, J.M., LaValle, S.M.: Localization with limited sensing. *IEEE Transactions on Robotics* **23**(4) (Aug 2007) 704–716
- [21] Gasparri, A., Panzieri, S., Pascucci, F., Ulivi, G.: A hybrid active global localisation algorithm for mobile robots. In: Robotics and Automation, 2007 IEEE International Conference on. (April 2007) 3148–3153
- [22] Pilania, V., Gupta, K.: A localization aware sampling strategy for motion planning under uncertainty. In: Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on. (Sept 2015) 6093–6099
- [23] Karaman, S., Frazzoli, E.: Sampling-based algorithms for optimal motion planning. *International Journal of Robotics Research* **30**(7) (June 2011) 846–894
- [24] Bertsekas, D.: Dynamic Programming and Optimal Control: 3rd Ed. Athena Scientific (2007)
- [25] Chakravorty, S., Erwin, R.S.: Information space receding horizon control. In: IEEE Symposium on Adaptive Dynamic Programming And Reinforcement Learning (ADPRL). (April 2011)
- [26] He, R., Brunsell, E., Roy, N.: Efficient planning under uncertainty with macro-actions. *Journal of Artificial Intelligence Research* **40** (February 2011) 523–570
- [27] Agarwal, S., Tamjidi, A., Chakravorty, S.: Motion planning for global localization in non-gaussian belief spaces. (2015) arXiv:1511.04634 [cs.RO].
- [28] Garrido-Jurado, S., Muoz-Salinas, R., Madrid-Cuevas, F., Marn-Jimnez, M.: Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition* **47**(6) (2014) 2280 – 2292