Robust Exploration with Multiple Hypothesis Data Association

Jinkun Wang and Brendan Englot

***Abstract*— We study the ambiguous data association problem confronting simultaneous localization and mapping (SLAM), specifically for the autonomous exploration of environments lacking rich features. In such environments, a single false positive assignment might lead to catastrophic failure, which even robust back-ends may be unable to resolve. Inspired by multiple hypothesis tracking, we present a novel approach to effectively manage multiple hypotheses (MH) of data association inherited from traditional joint compatibility branch and bound (JCBB), which entails the generation, ordering and elimination of hypotheses. We analyze the performance of MHJCBB in two particular situations, one applying it to SLAM over a predefined trajectory and the other showing its applicability in exploring unknown environments. Statistical results demonstrate that MHJCBB's maintenance of diverse hypotheses under ambiguous conditions significantly improves map accuracy.**

I. Introduction

Simultaneous localization and mapping (SLAM) has been applied successfully for state estimation and map-building with data collected on various sensing platforms. Although accurate and robust solutions have been proposed to tackle SLAM problems using feature-rich sensors, for example monocular SLAM (ORB-SLAM [1]) and lidar odometry (LOAM [2]), it is still a key challenge for an autonomous vehicle to manage the quality of its state estimation and map in environments with sparse features and severe disturbances, e.g. in unknown subsea environments [3]. A hurdle remains in the *data association*, or the front-end of SLAM, which is responsible for recognizing features that have been observed before. A lack of redundancy of features combined with uncertainty caused by odometry drift makes it difficult to detect the correct loop closure, thus leading to errors.

Recently, the ambiguity in data association has been addressed by numerous robust back-ends. In [4], dynamic covariance scaling (DCS) was introduced to remove outliers in a factor graph generated by imperfect front-ends. Similarly, given a set of odometry and loop-closing constraints, the RRR algorithm ([5]) is able to correct loop closures by performing a consistency check on clustered constraints. A generalized graph SLAM was proposed in [6], where ambiguous data associations are represented as mixtures of Gaussian constraints, and a generalized prefilter is used to resolve ambiguous graphs before graph optimization.

However, the above solutions all rely on front-ends that are capable of detecting *global* associations. Here global associations correspond to mappings from measurements to existing features far from the current pose, and as a result, a robust back-end can delete outliers and recover correct associations despite a robot incurring critical drift from an odometry-based trajectory estimate. While this is feasible for ground vehicles equipped with laser rangefinders or cameras (e.g. using a bag-of-words model), it is generally hard for applications under sparse and featureless conditions. A workaround solution in semantic SLAM was proposed [7], where the geometric features are endowed with semantic labels, reducing the quantity of incorrect associations.

Representing a robot's state using a multi-modal probability distribution has also been used to manage data association ambiguity. Using particle filters, uncertain data associations are naturally handled by assigning each particle different measurement pairings [8]. In addition, multi-modal loop-closures may be incorporated into a factor graph, and non-parametric belief propagation is applied to achieve multiple hypothesis inference [9]. However, computation is prohibitive for large-scale problems. In the target tracking community, maintaining multiple association hypotheses over time is used in multiple hypothesis tracking (MHT) [10]. Our algorithm shares the same spirit with MHT, except in SLAM, the *vehicle* is moving rather than independent objects.

Failure to secure state estimation against ambiguous data association is also a major concern for decision-making in autonomous navigation. A corrupted map generated by incorrect loop-closure proposals will pose a threat for navigation. Provided multiple hypotheses of poses (e.g. mixtures of Gaussians), active planning strategies in perceptual aliasing environments have been discussed in [11], [12].

In this paper, we provide a multiple hypothesis joint compatibility branch and bound (MHJCBB) solution for SLAM subject to ambiguity in data association, and we apply it to an autonomous exploration framework. Based on ideas shared with tracking algorithms, we hold off decisive associations until unlikely hypotheses can be safely deleted after acquiring sufficient re-observation. We manage limited numbers of hypotheses by performing expansion and elimination based on a designated optimality ordering, and a set of rules for hypothesis elimination. Furthermore, instead of actively choosing actions to disambiguate multiple hypotheses, our proposal is to minimize map uncertainty during exploration. We show that MHJCBB can substantially mitigate data association ambiguity in SLAM. In the next section, the data association problem is introduced and we also give a brief review of JCBB. In Section 3 we describe the details of MHJCBB's techniques for hypothesis and track ordering, hypothesis elimination and an efficient traversal method. We present its application to autonomous exploration in Section 4. The proposed algorithm is validated

J. Wang and B. Englot are with the Department of Mechanical Engineering, Stevens Institute of Technology, Castle Point on Hudson, Hoboken NJ 07030, USA, {jwang92, benglot}@stevens.edu.

over simulated environments in Section 5.

II. The Data Association Problem

A. Problem Definition

We first give the definition of data association that will be assumed throughout this paper. A robot navigating throughout the environment repeatedly collects observations from a set of m features. Let $\{\mathbf{z}_1, ..., \mathbf{z}_m\}$ be the set of measurements of the landmarks $\{\mathbf{l}_1, ..., \mathbf{l}_m\}$ at state $\mathbf{x}$. The data association problem is to determine the hypothesis $H \triangleq \{j_1, ..., j_m\}$, each of which associates one measurement $\mathbf{z}_i$ with one non-repeated landmark $\mathbf{l}_{j_i}$. A measurement from a new landmark, or a null-pairing, is denoted as $j_i = 0$. Subsequently, the back-end of SLAM updates the system's estimate of the robot state and landmark locations, incorporating this set of new measurements.

The selection criterion is defined by the Mahalanobis distance between actual measurements and predicted measurements given a noisy observation model. Mathematically, the joint measurement model under hypothesis H is

$$\mathbf{z}_H = \mathbf{h}_H(\mathbf{x}, \mathbf{l}_H) + \mathbf{v}, \tag{1}$$

where $\mathbf{h}_H \triangleq [\mathbf{h}_{1j_i}, ..., \mathbf{h}_{mj_m}]^T$ is the collection of independent measurement models with zero-mean Gaussian noise $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_H)$, and $\mathbf{l}_H \triangleq [\mathbf{l}_{1j_1}, ..., \mathbf{l}_{mj_m}]^T$ is the tested landmark vector corresponding to $\mathbf{z}_H = \{\mathbf{z}_1, ..., \mathbf{z}_m\}$. The distance d_H^2 is given by

$$\mathbf{e}_H = \mathbf{z}_H - \mathbf{h}_H(\hat{\mathbf{x}}, \hat{\mathbf{l}}_H) \tag{2}$$

$$\mathbf{C}_H = \mathbf{H}_H \hat{\mathbf{P}}_H \mathbf{H}_H^T + \mathbf{R}_H \tag{3}$$

$$d_H^2 = \mathbf{e}_H^T \mathbf{C}_H^{-1} \mathbf{e}_H, \tag{4}$$

where $\mathbf{H}_H$ is the Jacobian matrix with respect to robot pose and landmark positions, and $\hat{\mathbf{P}}_H$ is the joint covariance of the estimates. One set of pairings can be accepted, or is *jointly compatible*, if the predicted measurements lie in the validation gate based on the chi-squared distribution,

$$\text{jc}(H) \text{ if } d_H^2 \leq \chi_{d,\alpha}^2, \tag{5}$$

where $d \triangleq \dim(\mathbf{h}_H)$, and α is the confidence level.

B. Joint Compatibility Branch and Bound

In general, only one "optimal" solution is derived from data association algorithms, and the problem is challenging due to the exponential growth of the *interpretation tree* [13]. Therefore, obtaining the hypothesis with the minimum Mahalanobis distance is computationally intractable.

Joint compatibility branch and bound (JCBB) [14] instead searches for the hypothesis which has the maximum number of non-null pairings $N(H) = \sum_{i=1}^{m} \mathbb{I}_{j_i \neq 0}$,

$$H^* = \underset{H}{\text{argmax}}\, N(H), \text{ s.t. jc}(H). \tag{6}$$

The underlying idea is that the probability of accepting a spurious pairing decreases as we increase the number of jointly compatible non-null pairings. With this redefinition, the combinatorial problem becomes tractable using a branch and bound algorithm. While generating the interpretation tree, we obtain an incomplete hypothesis $H_i \triangleq \{j_1, ..., j_{i<m}\}$ at every leaf. Given the best H^* so far, the branching of a leaf, or insertion of a new pairing $H_{i+1} = H_i \cup \{j_{i+1}\}$, is bounded by the lower bound $\underline{\text{N}}(H^*)$. More specifically, H_{i+1} will be considered as a candidate only if $\overline{N}(H_{i+1}) > \underline{\text{N}}(H^*) \wedge \text{jc}(H_{i+1})$. The upper bound of the number of non-null pairings $\overline{N}$ can be estimated by assuming all future independently compatible pairings are also jointly compatible. Thus a good lower bound will significantly reduce the computation involved in the compatibility check.

III. Multiple Hypothesis JCBB

The myopic pairings derived from JCBB can be inaccurate. For example, when faced with the situation in the top left of Fig. 4, the solution that greedily incorporates more pairings with existing landmarks will distort the robot's state estimate. The assumptions of JCBB are typically capable of rejecting false positive pairings with a dense feature cloud, such as that from a camera image. However, these assumptions do not hold when dealing with fewer features. In addition, errors can result from the linearization of priors [15] and from the fact that maximum a posteriori estimation converges to local minima.

In this section, we introduce our proposed approach to addressing the limitations of JCBB. Similar to multiple hypothesis tracking (MHT) [10], the key idea is to defer decision-making about data association and ultimately to pick the association "track" with the maximum number of non-null pairings. Consequently, the need for interpretation at a single ambiguous moment is avoided by accumulating measurements over a longer time horizon.

A. Problem Definition

Let $T^{(t-1)}$ be the hypothesis track associated with measurements $Z^{(t-1)}$ over the trajectory $X^{(t-1)}$,

$$T^{(t-1)} \triangleq \{H_1, ..., H_{t-1}\},$$
$$Z^{(t-1)} \triangleq \{\mathbf{z}_{H^{(1)}}, ..., \mathbf{z}_{H^{(t-1)}}\},$$
$$X^{(t-1)} \triangleq \{\mathbf{x}_1, ..., \mathbf{x}_{t-1}\}.$$

Here, we use the superscript to indicate a time step and the subscript for indexing, e.g., $T_i^{(j)}$ means the ith track at time j. In the notation that follows, the superscript will occasionally be omitted for the sake of brevity.

We seek to resolve ambiguity by maintaining the K-best solutions of the data association problem, which are represented as $\{T_1^{(t-1)}, ..., T_{k \leq K}^{(t-1)}\}$. Given the latest measurements, the goal is to correctly populate a compact set of the most probable association histories $\{T_k^{(t)}\}$ after pruning. However, without intentionally planning to eliminate ambiguity, there is no guarantee that all tracks will eventually be reduced to one.

B. Hypothesis Orderings

A good measure is needed to select the K-best solutions, and a common approach is to leverage the joint probability

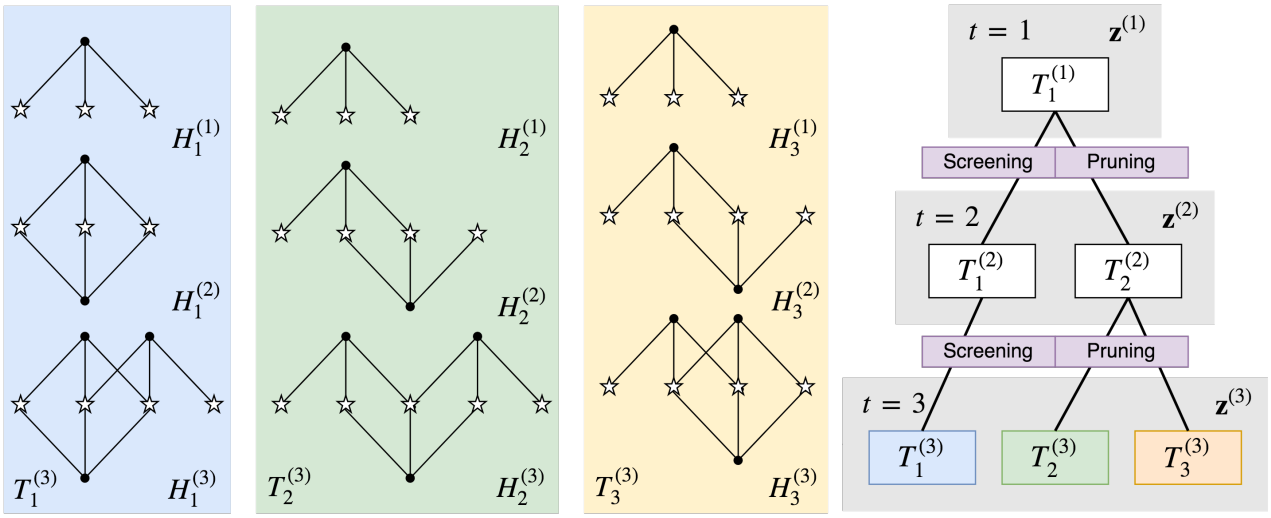


Fig. 1: A simple example of MHJCBB. The left three panels visualize three tracks resulting from ambiguous data associations after taking three observations (from top to bottom). At each time instant, tracks diverge, taking into account possible hypotheses. The formation and ordering of tracks are discussed in Sec. III.B. Redundant hypotheses are screened and the resulting tracks are pruned to keep the K-best solutions (Sec. III.C).

model [16], or the sum of residual error d_T^2 introduced by measurement and odometry constraints in its logarithmic form [17],

$$d_T^2 = \sum_t ||\mathbf{f}_t(\mathbf{x}_{t-1}, \mathbf{u}_{t-1}) - \mathbf{x}_t||^2_{\mathbf{Q}_t} + ||\mathbf{h}_{H_t}(\mathbf{x}_t, \mathbf{l}_{H_t}) - \mathbf{z}_{H_t}||^2_{\mathbf{R}_{\mathbf{H}_t}},$$

where $\mathbf{x}_t = \mathbf{f}_t(\mathbf{x}_{t-1}, \mathbf{u}_t) + \mathbf{w}$ is the process model with zero-mean Gaussian noise $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_t)$. Unlike d_H^2, the posterior residual error d_T^2 associated with a track of associations is computed after optimization.

However, a small d_T^2 doesn't necessarily indicate a good mapping result, because the optimization may easily become trapped in local minima [18]. For instance, consider two mapping results produced from the same trajectory (Fig. 2), using JCBB and using known data associations respectively. The former result, which is severely distorted, turns out to have a smaller posterior error. Additionally, a data association algorithm that assumes all measurements to be from new landmarks can drive the d_T^2 error to zero.

Therefore, we propose two total orderings for intermediate hypothesis H and association track T. After new measurements are introduced into the least-squares problem, the first ordering is used for sorting association tracks. Inspired by the objective function used in JCBB, an association track is measured in terms of the total number of non-null pairings over the entire trajectory,

$$N(T) = \sum_t N(H_t). \tag{7}$$

Thus, the tracks are sorted in the following way,

$$T_i < T_j \;\Rightarrow\; N(T_i) < N(T_j) \vee N(T_i) = N(T_j) \wedge d^2_{T_i} > d^2_{T_j}. \tag{8}$$

By maximizing the overall non-null pairings, populating new landmarks is discouraged. However, tracks that have false associations are likely to diverge from the correct trajectory, and so maintaining maximum non-null pairings is less probable. In the example in Fig. 1, there are three tracks with $N(T_1) = 0{+}3{+}2, N(T_2) = 0{+}2{+}2, N(T_3) = 0{+}2{+}3$.

Based on the above definition, we could follow the same search procedure as in JCBB, using the lower bound computed by $N(T^{*(t-1)}) + N(H^{*(t)})$, where $H^{*(t)}$ is generated from the track $T^{*(t-1)}$. However, if we consider a loop-closure observing existing landmarks, true tracks with even smaller $N(T)$ will produce more non-null pairings in the short-term. Thus, the second ordering is designed by valuing short-term hypotheses, solely incorporating $N(H)$.

The conventional JCBB method is used to search the interpretation forest consisting of interpretation trees rooted at multiple tracks. In the example shown in Fig. 1 at time $t = 2$, two hypotheses $\{H_1^{(2)}, H_2^{(2)} = H_3^{(2)}\}$ from T_1 are jointly compatible, which consequently form two tracks $\{T_1^{(2)}, T_2^{(2)}\}$. Similarly, at $t = 3$, one $\{H_1^{(3)}\}$ from $T_1^{(2)}$ and two $\{H_2^{(3)}, H_3^{(3)}\}$ from $T_2^{(2)}$ split into three tracks $\{T_1^{(3)}, T_2^{(3)}, T_3^{(3)}\}$. A priority queue with maximum size K is maintained to house jointly compatible hypotheses from any track that could be one of the K-best. The priority is defined using the following total ordering:

$$H_i < H_j \;\Rightarrow\; N(H_i) < N(H_j) \vee N(H_i) = N(H_j) \wedge d^2_{H_i} > d^2_{H_j}. \tag{9}$$

In this way, the last entry in the queue serves as the lower bound while traversing the forest. A hypothesis will be inserted into the priority queue as long as it is superior to the lower bound either in the number of non-null pairings or in the chi-squared distance, both of which are non-decreasing with respect to node level. It is also worth noting that

until now we do not distinguish hypotheses from different association tracks, thus the outcome is comprised of the i-th best H_{ki} from association track T_k.

C. Hypothesis Elimination

Hypotheses and association tracks grow exponentially (due to the fact that we must perform JCBB over K association tracks), so poor management of the various hypotheses will increase the computational burden, and more importantly, it will neglect correct associations. This happens in situations where hypotheses with more false positives have occupied the majority of spots in the priority queue, leaving no space for correct hypotheses with more true negatives. Thus, we propose a number of techniques to eliminate unlikely associations that fall into two categories: screening and pruning (similar to concepts in [19]).

Screening refers to eliminating unnecessary branches before generation. Suppose a jointly compatible hypothesis is $H = \{j_1 = 1, j_2 = 2\}$. Then, there are three trivial hypotheses that are also jointly compatible: $\{j_1 = 1, j_2 = 0\}$, $\{j_1 = 0, j_2 = 2\}$, and $\{j_1 = 0, j_2 = 0\}$, which become redundant if they result in little to no ambiguity during the ensuing data association. Thus we formulate two screening rules by examining the posterior estimate as follows:

$$\text{delete } H_{ki} \text{ if } H_{ki} < H_{kj} \wedge ||\hat{\mathbf{x}}_{t,H_{ki}} - \hat{\mathbf{x}}_{t,H_{kj}})||^2_{\mathbf{P}} < \alpha_1, \quad \text{(Rule 1)}$$

$$\text{delete } H_i \text{ if } N(H_i) = 0 \wedge \ (\exists k \ (N(H_{kj}) \geq 1 \wedge ||\hat{\mathbf{x}}_{t,H_i} - \hat{\mathbf{x}}_{t,H_{kj}}||^2_{\mathbf{P}} < \alpha_1)), \quad \text{(Rule 2)}$$

where $\hat{\mathbf{x}}_H$ is the updated pose estimate and $\mathbf{P}$ is the updated covariance estimate from any of the two hypotheses. It is worth noting that in practice, we use an extended Kalman filter (EKF) update step to predict the posterior pose and covariance estimates after incorporating a new hypothesis H, instead of performing the full least-squares optimization that is applied elsewhere. The first rule selectively discards any hypothesis that is close to a high-ranking hypothesis derived from the same association track T_k, and the second rule ignores the all-null hypothesis if the posterior state is in the proximity of any hypothesis with non-null pairings. Both rules restrain the addition of new landmarks.

After the generation of hypotheses, *pruning* is used for the elimination of redundant tracks, since all states are updated with new measurements. The rules are as follows:

$$\text{delete } T_i \text{ if } d^2_{T_i} \text{ is an outlier} \vee d^2_{T_i} > \alpha_2, \quad \text{(Rule 3)}$$

$$\text{delete } T_i \text{ if } N(T_j) - N(T_i) > \alpha_3, \quad \text{(Rule 4)}$$

$$\text{delete } T_i \text{ if } T_i < T_j \wedge \ (\forall n \leq t \ (||\mathbf{x}_{n,T_i} - \mathbf{x}_{n,T_j}||^2_{\mathbf{P}} < \alpha_4). \quad \text{(Rule 5)}$$

The third and fourth rules delete unlikely tracks with regard to the chi-squared error and the number of landmarks that are observed at least twice. However, the fourth pruning rule is not executed frequently, considering that the observation of existing landmarks doesn't occur until loop-closure (observing the same landmarks at consecutive poses usually will not eliminate ambiguity). Therefore, we consider tracks with many new landmarks to be less probable after a certain number of re-observations are performed. The fifth rule inspects the closeness of two estimated trajectories by calculating the maximum Mahalanobis distance between two poses at every step, and we consider the track of the highest order to represent the others in its close proximity.

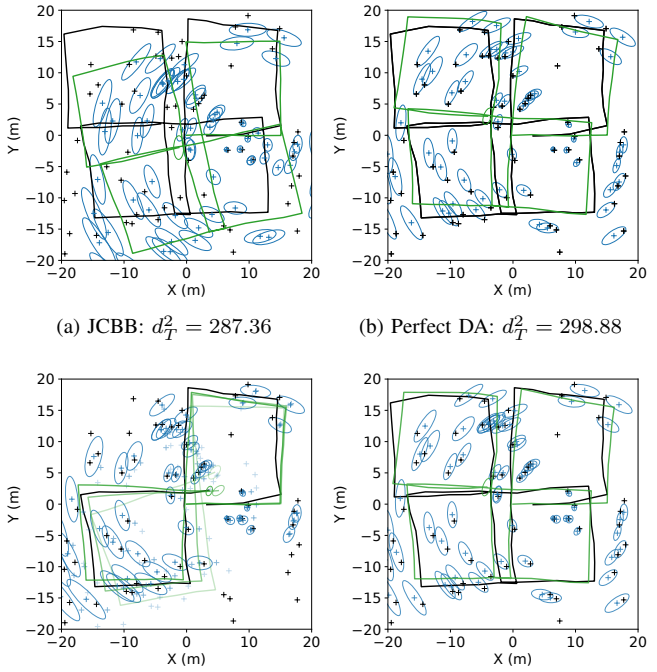
Algorithm 1: MHJCBB

Result: Return best estimates $\hat{X}, \hat{L}$
Given measurements Z, priors $T^{(1)}$;
for $\{\mathbf{z}_i^{(t)}\}$ **do**
  Priority queue $PQ \leftarrow \emptyset$ ordered by Eq. 9;
  for *MSDBF* $H_{ki}^{(t)}$ **do**
    if $H_{ki}^{(t)} > PQ.top()$ and jc(H) **then**
      Delete or insert $H_{ki}^{(t)}$ to PQ by rules 1-2;
    end
  end
  for $H_{ki}^{(t)} \in PQ$ **do**
    Add $\{\mathbf{z}_i^{(t)}\}, H_{ki}^{(t)}$ to $T_k^{(t)}$;
    Optimize states $\hat{X}^{(t)}_{T_k^{(t)}}$, $\hat{L}^{(t)}_{T_k^{(t)}}$;
    Delete $T_k^{(t)}$ by rules 3-5;
  end
end
Return $\hat{X}^{(t)}_{T_1^{(t)}}, \hat{L}^{(t)}_{T_1^{(t)}}$ with $T_1^{(t)}$ ordered by Eq. 8;

D. Traversal Order

The branch and bound algorithm relies on the quick computation of lower bounds such that suboptimal branches are not expanded. Typically, it is solved by using a depth-first search (DFS) strategy to obtain one jointly compatible hypothesis at the leaf node. However, in the context of multiple hypothesis data association, the K-best solutions are generally distributed across different "track trees". Thus exhaustive traversal of a forest, tree by tree, is less efficient. In [20], a novel traversal strategy, mixed stacked depth-breadth first (MSDBF) search, was proposed to speed up the search for an optimal solution.

Firstly, a non-recursive traversal with stacks replaces function recursion. After visiting the first leaf, instead of visiting its siblings, mixed depth-breadth visits the first leaf traced back to the other children from the root by using a sequence of stacks (see [20] for implementation details). In principle, leaves in a tree or in a forest are explored in parallel. For example, two binary trees with two levels each will have leaves $\{l_1^1, l_2^1, l_3^1, l_4^1, l_1^2, l_2^2, l_3^2, l_4^2\}$, which will be explored in MSDBF in the order $(l_1^1, l_1^2, l_3^1, l_3^2, l_2^1, l_2^2, l_4^1, l_4^2)$. In this manner, the number of visited nodes is reduced, and if the computation time is limited, stopping the search early



(a) JCBB: $d_T^2 = 287.36$ (b) Perfect DA: $d_T^2 = 298.88$

(c) MHJCBB($K = 5$): Step 43 (d) MHJCBB($K = 5$): $d_T^2 = 291.04$

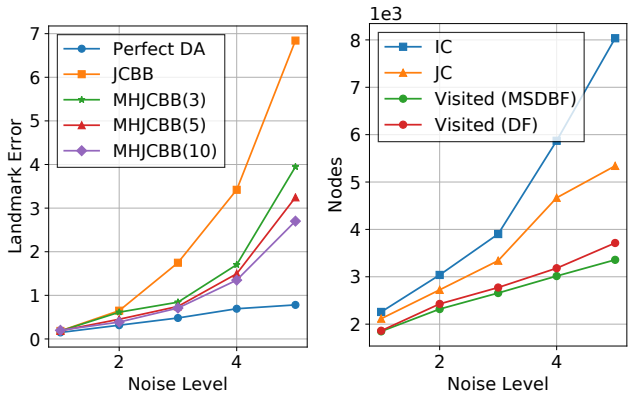
Fig. 2: Mapping examples in a simulated environment (noise level = 4) using (a) JCBB, (b) known data associations, and (c-d) MHJCBB($K = 5$). JCBB distorts the trajectory but iSAM2 reports a relatively small error. Tracks using MHJCBB diverge when the robot revisits its start location, but they converge to the optimal trajectory after collecting enough observations. **Black**: ground truth, **green**: multiple estimated trajectories, **blue**: multiple estimated landmarks and error ellipses (2 std. deviations). Darker-colored trajectories and landmarks in (c) indicate the favored hypotheses.

has a less detrimental effect on the outcome of MSDBF, as shown in the experiment to follow.

We describe the MHJCBB data association procedure as a whole in Algorithm 1. The process proceeds whenever new measurements arrive by populating association hypotheses (the first inner loop) and appending them to current tracks (the second inner loop). The MSDBF search is used to iterate over leaves in the interpretation tree rooted at its corresponding track, and the search is bounded by the least optimal case in PQ using hypothesis ordering 9 (or it ignores the bound if PQ hasn't reach its capacity). Given a jointly compatible hypothesis, Rules 1-2 are used to determine whether it will result in uncertain state estimation. Hypothesis generation is followed by optimization of the entire trajectory and map using a smoothing and mapping method, and unnecessary tracks are removed following Rules 3-5. The algorithm outputs an estimate selected from the remaining tracks using "track ordering" per Equation 8.

IV. Data Association in Autonomous Exploration

The purpose of exploration is to autonomously discover an unknown environment, where the objective guiding exploration is to produce a map of a robot's surroundings accurately and efficiently. In environments with sparse features and severe noise, spurious data associations, specifically false positive ones, can result in catastrophic failure during exploration. A false positive pairing can lead to an overconfident and divergent estimate of robot pose, which brings about consequential errors in future data associations.



(a) (b)

Fig. 3: Mean landmark error and interpretation forest node counts (using MHJCBB(10)) are plotted with respect to the noise level over 50 trials of the example shown in Fig. 2 with randomized landmark locations. MHJCBB effectively manages errors at low noise levels, and MSDBF reduces the number of forest nodes visited (IC: total independently compatible node count, JC: total jointly compatible node count, Visited: nodes actually visited during the search).

Max Nodes	Noise:	1	2	3	4	5
30	DFS	-0.03	**+0.46**	+0.39	+0.40	+0.59
	MSDBF	**-0.04**	+0.57	**+0.16**	**+0.12**	**+0.39**
50	DFS	-0.02	**+0.14**	+0.20	+0.19	+0.43
	MSDBF	**-0.04**	+0.22	**+0.18**	**+0.03**	**+0.24**
70	DFS	+0.03	**-0.03**	**+0.06**	+0.19	+0.33
	MSDBF	**-0.03**	+0.05	+0.15	**+0.12**	**+0.24**

TABLE I: The increase/decrease in landmark error when limiting the max. number of visited nodes is compared over different node limits using MHJCBB($K = 10$) (baselines without restricting visited nodes are set to zero). Mixed depth-breadth first search generally obtains optimal hypotheses faster than depth first search.

As shown in Fig. 3(a), map accuracy deteriorates as the level of additive noise increases. This relationship requires an exploration algorithm to take localization and map uncertainties into account, and we employ our previous algorithm for exploration with expectation maximization (EM) [21]. Here, we briefly introduce the EM exploration algorithm that is used in concert with MHJCBB data association.

We seek a sequence of actions $\mathbf{a}$ to minimize a utility function composed of action cost $C(\mathbf{a})$ (e.g., distance traveled) and exploration utility. Unlike entropy-based exploration frameworks, we seed finely and evenly distributed *virtual*

landmarks $\mathcal{V} = \{\mathbf{v}_k\}$ throughout the environment as a prior, and the discovery of unknown space and the minimization of uncertainty are successfully combined into one term:

$$\mathbf{a}^* = \underset{\mathbf{a}}{\operatorname{argmin}}\, U(\mathbf{a}) = \underset{\mathbf{a}}{\operatorname{argmin}} \sum_{\mathbf{v}_k \in \mathcal{V}^*} \phi(\Sigma_{\mathbf{v}_k}) + \alpha C(\mathbf{a}). \quad (10)$$

This exploration approach alternates between an expectation step and a maximization step. In the E-step, the contents of the environment are estimated using the accumulated measurements, and they are represented as an expected occupancy map. Furthermore, we threshold the map to obtain the set of landmarks $\mathcal{V}^* \subset \mathcal{V}$ that includes true landmarks, and the virtual landmarks that do not represent free space. Next, in the M-step, the covariance of the landmarks in $\mathcal{V}^*$ after executing actions $\Sigma_{\mathbf{v}_k}$ is predicted by forward-simulating maximum likelihood measurements. These are used to minimize an uncertainty criterion $\phi(\Sigma) = \text{tr}(\Sigma)$ over path candidates generated using a sampling-based planner (e.g. Rapidly Exploring Random Trees (RRTs)).

In considering multiple map hypotheses resulting from uncertain data associations, it would be problematic to perform exploration using only the favored hypothesis, which may lead to a distorted and incomplete map. Therefore, we modify the objective function to combine the individual utility contributed by different tracks,

$$\mathbf{a}^* = \underset{\mathbf{a}}{\operatorname{argmin}} \sum_{k=1}^{K} \omega_k U(\mathbf{a}|T_k^{(t)}). \quad (11)$$

During planning, only one RRT is constructed, and it is transformed into the frame at pose $\mathbf{x}^{(t)}_{T_k^{(t)}}$. A path is favorable if it leads to greater expected uncertainty reduction of the virtual landmarks over the weighted combination of tracks. We have found uniform weights $\omega_k = 1/K$ to perform best, where K is the total number of tracks. We have observed instabilities when instead using K-best tracks with small K.

V. Experiments and Results

We analyze the performance of the proposed multiple hypothesis JCBB in two simulated scenarios: SLAM given a predefined trajectory (Figs. 1, 2, and Table 1), and autonomous exploration of unknown environments (Figs. 3-5). We use the iSAM2 [22] implementation in the GTSAM library [23] for state estimation, and simulations are built upon [21]. In describing the parameterizations of MHJCBB investigated hereafter, we will use the notation MHJCBB(K) to refer to an implementation of the method that maintains at most K hypothesis tracks.

A. SLAM Over a Predefined Trajectory

Our simulation of a predefined trajectory employs an environment with 60 uniformly distributed random point features (Fig. 2). The robot is equipped with a sensor with a limited field of view (5m, 120°) that is capable of measuring the relative range and bearing to a landmark. Zero mean Gaussian noise is added to both process and measurement models. The robot is commanded to travel in four square

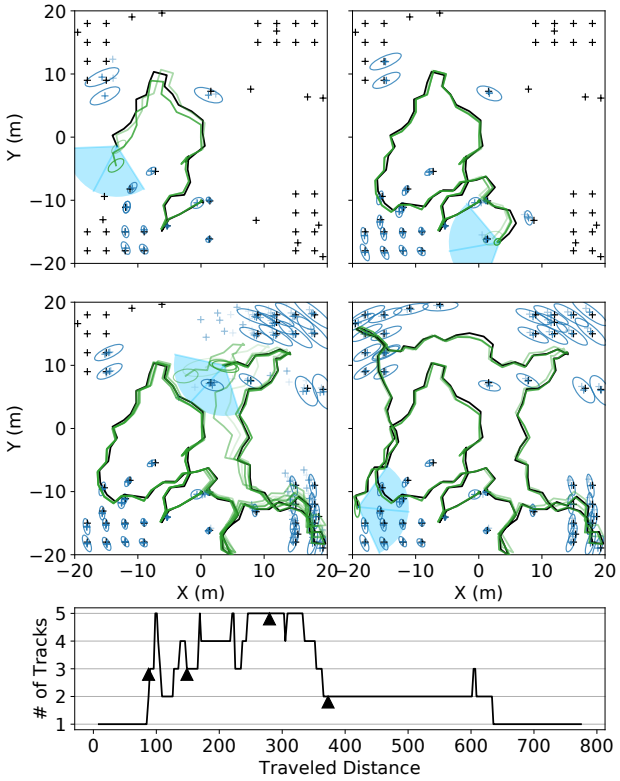


Fig. 4: Four representative steps from EM exploration using our proposed MHJCBB($K = 5$) algorithm. The estimated trajectories (*lines*) and maps (*crosses*) correspond to the four instants along the robot's trajectory that are marked in the plot at bottom, which shows the evolving number of tracks.

patterns ($15m \times 15m$). To demonstrate association error, we gradually multiplied the standard deviation ($0.05°$ for rotation and bearing measurement, $0.05m$ for translation and range measurement) by factors of $1, 2, 3, 4, 5$ (this factor is the "noise level" plotted along the x-axes of Fig. 2). All methods included 50 trials with different random seeds.

The outcomes across one representative trial are illustrated in Fig. 2. JCBB generates a misleading map as a result of a few incorrect data associations, whereas by using MHJCBB, tracks diverge when ambiguity occurs and converge to one track as soon as the robot observes enough existing landmarks. The average performance is shown in Fig. 3 by analyzing root-mean-square error with respect to true landmark positions. Given a mapping from ground truth to landmark estimates ($\mathbf{l}_i \rightarrow \mathbf{z}_j \rightarrow \hat{\mathbf{l}}_k$), we define $\hat{L}(\mathbf{l}_i) = \{\hat{\mathbf{l}}_k | \mathbf{l}_i \rightarrow \cdot \rightarrow \hat{\mathbf{l}}_k\}$ to be the set of estimates corresponding to the same landmark. Then we calculate the error as follows,

$$\text{RMSE}(\hat{L}) = \sqrt{\frac{1}{|L|} \sum_{\mathbf{l}_i \in L} \text{MSE}(\mathbf{l}_i)} \quad (12)$$

$$\text{MSE}(\mathbf{l}_i) = \frac{1}{|\hat{L}(\mathbf{l}_i)|} \sum_{\hat{\mathbf{l}}_k \in \hat{L}(\mathbf{l}_i)} ||\mathbf{l}_i - \hat{\mathbf{l}}_k||_2^2. \quad (13)$$

It is clear that landmark error rises dramatically as we increase the noise level, and our proposed method recovers

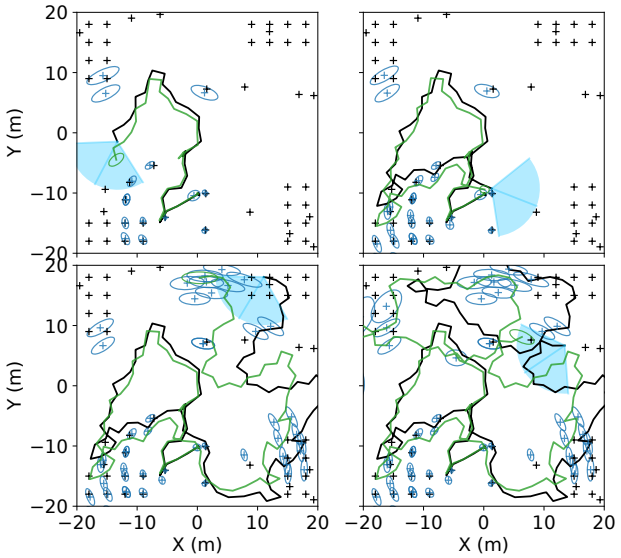


Fig. 5: Snapshots at the same travel distances shown in Fig. 4 are now shown using a single-hypothesis JCBB algorithm to support EM exploration. In the first panel, an erroneous association with an existing landmark corrupts state estimation, and hinders the subsequent exploration process.

accurate landmark positions under high uncertainty. However, although allowing more estimation tracks improves performance, especially in extremely uncertain environments, the performance drop compared with SLAM under perfect knowledge of data association is still substantial. This can be explained by the fact that the error introduced by an incorrect pairing is exacerbated by increased noise, thus more hypotheses even with false positive pairings are considered jointly compatible. Consequently, even if we increase the quantity of hypotheses, the subtle differences in the posterior error makes the hypothesis pruning process error-prone.

In addition, we show the number of visited nodes in the interpretation forest during the entire trajectory, along with all independently and jointly compatible nodes, in Fig. 3(b). As mentioned above, the numbers of visited nodes along with jointly compatible nodes undergo an approximately linear increase as the noise level increases. However, the required computation is still demanding in contrast to single hypothesis JCBB, which visits fewer than 250 nodes (this JCBB visit count is not shown in Fig. 2 due to the scaling of the plot). Mixed depth-breadth first search aids MHJCBB by reducing the number of joint compatibility checks required. Further investigation shows that when limiting the maximum number of nodes the search is allowed to visit, using MSDBF search to spread out visited leaves among multiple interpretation trees yields less accuracy loss, as shown in Table 1.

B. Autonomous Exploration

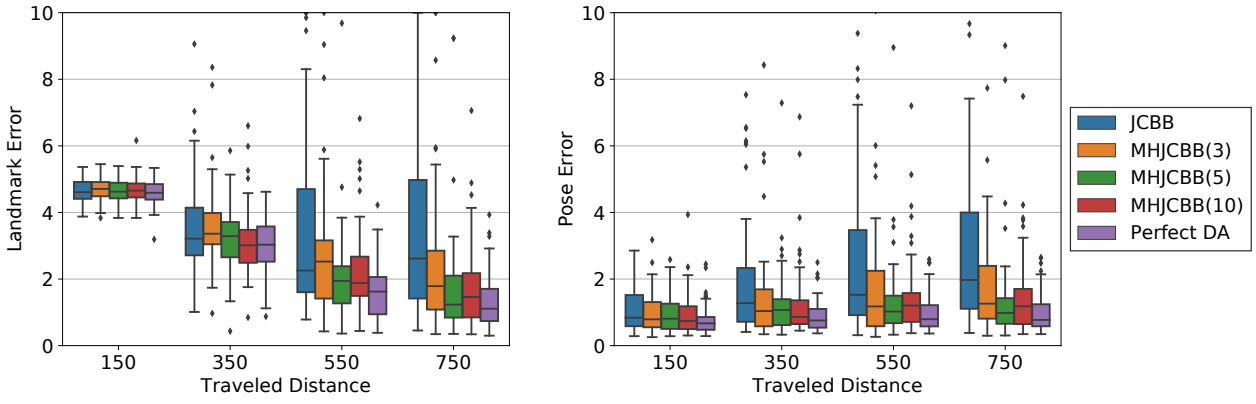
We further demonstrate the limitations of JCBB in a scenario where a robot incurs critical drift while exploring an unknown environment that is sparsely populated with features. We have designed a series of environments as shown in Fig. 4, where there are 4 “piers" at the corners, each consisting of 8 landmarks, and another 20 landmarks distributed uniformly at random in the environment. The robot has the same sensor with a fixed noise level $(1.5°, 0.15m$ for bearing and range measurement, $2.0°, 0.20m$ for rotation and translation). Similarly, all competing methods were run over 50 trials with different random landmarks selected.

From the representative example trial compared across Fig. 4 and Fig. 5, we can observe two factors contributing to the success of multi-hypothesis exploration with MHJCBB. First, when faced with ambiguous data association, the belief splits into multiple hypotheses, which has occurred in the top left panel of Fig. 4 when the robot tries to relocalize itself by revisiting its starting point, and in the bottom left panel when the robot travels through a featureless region into unknown space. Second, the EM exploration algorithm prevents uncertainty from increasing by carrying out loop-closing consistently, thereby eliminating unlikely hypotheses.

From simulated exploration conducted over a variety of random features, we show the performance of our proposed framework with respect to landmark error and localization error in Fig. 6. As the prior uncertainty is small (the first group of bars in Fig. 6(b)), all methods produce almost the same maps (first group of bars in Fig. 6(a)). However, incorrect data association has a negative influence on mapping and localization (second group of bars). Mistakes will not be rectified because the robot still keeps a consistent map by adding a new landmark whenever there exists a measurement that disagrees with the current constraints. As a result, a small divergence causes more error in data association, which can be observed in the growing error bars (last two groups).

We can also conclude that on one hand, increasing the number of hypotheses improves robustness to ambiguity (from MHJCBB($K = 3$) to MHJCBB($K = 5$)), and on the other hand, using more hypotheses than necessary does not bring about any further improvement (MHJCBB($K = 10$)). The cause is as follows. As described above, the elimination process is not performed regularly. Therefore, the utility function in the planning step must take into account more improbable trajectory and map hypotheses, such that decisions are adjusted away from optimal ones. Despite MHJCBB's ability to distinguish correct hypotheses among multiple tracks, information is still lost. A number of pairings that do not necessarily result in ambiguity are considered to be from new landmarks, which in Fig. 6 is evident from the presence of slightly higher bars from MHJCBB than those from perfect associations.

When applied to EM exploration, the most costly aspect of using MHJCBB is the need to propagate its hypotheses over the *multiple* candidate paths being considered by the robot at every decision-making step. This is required to select the path whose multiple tracks give rise to the best expected reduction in map uncertainty relative to distance traveled. When implemented with our C++ code, and run on an Intel i7 6950 in Ubuntu 14.04, MHJCBB($K = 3$) runs in real-time over the examples explored here, and MHJCBB($K = 10$) requires about 10 seconds of computation time per exploration decision.



(a) (b)

Fig. 6: Results showing landmark errors and robot pose errors with respect to distance traveled over 50 trials of EM exploration. Errors accumulate as the robot explores using JCBB, and in contrast, MHJCBB with different maximum numbers of hypotheses ($K = 3, 5, 10$) produces comparable results to the trials with perfect knowledge of landmark associations.

VI. Conclusion

In this paper, we have presented a multiple hypothesis data association framework for SLAM in ambiguous environments. Multiple trajectory and map tracks are maintained using a series of association hypotheses generated from JCBB. Efficient hypothesis management is organized by limiting hypothesis generation and pruning unlikely tracks, which is enabled by providing two orderings. The most important application we demonstrated is in online exploration, which necessitates recovery from bad associations. Simulations show that our algorithm effectively explores the entire environment, more importantly without diverging from the ground truth. However, the computational complexity is substantial, and further investigation on how to efficiently abstract ambiguity is our next step. Also, we do not consider the active data association problem [11], [12], where the robot is able to generate actions to diminish ambiguity.

Acknowledgments

This research was supported in part by the National Science Foundation, grants IIS-1652064 and IIS-1723996, and by a grant from Schlumberger Technology Corporation.

References

[1] R. Mur-Artal, J. M. M. Montiel, J. D. Tardos, "ORB-SLAM: a versatile and accurate monocular SLAM system," *IEEE Transactions on Robotics,* 31(5), pp. 1147-1163, 2015.

[2] J. Zhang, and S. Singh, "LOAM: Lidar odometry and mapping in real-time," *Robotics: Science and Systems*, Vol. 2, 2014.

[3] J. Wang, S. Bai and B. Englot, "Underwater localization and 3D mapping of submerged structures with a single-beam scanning sonar," *IEEE International Conference on Robotics and Automation,*, pp. 4898-4905, 2017.

[4] P. Agarwal, G. D. Tipaldi, L. Spinello, C. Stachniss, W. Burgard, "Robust map optimization using dynamic covariance scaling," *IEEE International Conference on Robotics and Automation,* pp. 62-69, 2013.

[5] Y. Latif, C. Cadena, J. Neira, "Robust loop closing over time for pose graph SLAM," *The International Journal of Robotics Research,* 32(14), pp. 1611-1626, 2013.

[6] M. Pfingsthorn, and A. Birk, "Generalized graph SLAM: Solving local and global ambiguities through multimodal and hyperedge constraints," *The International Journal of Robotics Research*, 35(6), pp. 601-630, 2016.

[7] S. L. Bowman, N. Atanasov, K. Daniilidis, G. J. Pappas, "Probabilistic data association for semantic slam," *IEEE International Conference on Robotics and Automation,*, pp. 1722-1729, 2017.

[8] M. Montemerlo, S. Thrun, D. Koller, B. Wegbreit, "FastSLAM: A factored solution to the simultaneous localization and mapping problem," *Aaai/iaai*, 593598, 2002.

[9] D. Fourie, J. Leonard, and M. Kaess, "A nonparametric belief solution to the Bayes tree," *IEEE/RSJ International Conference on Intelligent Robots and Systems,*, pp. 2189-2196, 2016.

[10] S.S. Blackman, "Multiple hypothesis tracking for multiple target tracking," *IEEE Aerospace and Electronic Systems Magazine,* pp. 5-18, 2004.

[11] S. Agarwal, A. Tamjidi, S. Chakravorty, "Motion planning for active data association and localization in non-Gaussian belief spaces," *Workshop on the Algorithmic Foundations of Robotics*, 2016.

[12] V. Indelman, S. Pathak, A. Thomas, "Nonmyopic data association aware belief space planning for robust active perception," *IEEE International Conference on Robotics and Automation,* pp. 4487-4494, 2017.

[13] W.E.L. Grimson, *Object Recognition by Computer: The Role of Geometric Constraints*, The MIT Press, Cambridge, Mass., 1990.

[14] J. Neira and J.D. Tardós, "Data association in stochastic mapping using the joint compatibility test," *IEEE Transactions on Robotics and Automation,* 17(6), pp. 890-897, 2001.

[15] E.B. Olson and Y. Li, "IPJC: The incremental posterior joint compatibility test for fast feature cloud matching," *IEEE/RSJ International Conference on Intelligent Robots and Systems,*, 2012.

[16] D. Hähnel, S. Thrun, B. Wegbreit, W. Burgard, "Towards lazy data association in SLAM," *International Symposium on Robotics Research,* pp. 421-431, 2005.

[17] F. Dellaert, and M. Kaess, "Square Root SAM: Simultaneous localization and mapping via square root information smoothing," *International Journal of Robotics Research,* pp. 1181-1203, 2006.

[18] E.B. Olson, "Robust and efficient robotic mapping," 2008.

[19] I.J. Cox, and J.J. Leonard, "Modeling a dynamic environment using a Bayesian multiple hypothesis approach," *Artificial Intelligence,* pp. 311-344, 1994.

[20] H.A. Fayed and A.F. Amir, "A mixed breadth-depth first strategy for the branch and bound tree of Euclidean k-center problems," *Computational Optimization and Applications*, pp. 675-703, 2013.

[21] J. Wang and B. Englot, "Autonomous exploration with Expectation-Maximization," *International Symposium on Robotics Research,* 16 pp., 2017.

[22] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. J. Leonard and F. Dellaert, "iSAM2: Incremental smoothing and mapping using the Bayes tree," *International Journal of Robotics Research*, pp. 216-235, 2012.

[23] GTSAM, `https://bitbucket.org/gtborg/gtsam/`