Towards Cooperative Multi-Robot Belief Space Planning in Unknown Environments

Vadim Indelman

Abstract We investigate the problem of cooperative multi-robot planning in unknown environments, which is important in numerous applications in robotics. The research community has been actively developing belief space planning approaches that account for the different sources of uncertainty within planning, recently also considering uncertainty in the environment observed by planning time. We further advance the state of the art by reasoning about future observations of environments that are *unknown at planning time*. The key idea is to incorporate within the belief indirect multi-robot constraints that correspond to these future observations. Such a formulation facilitates a framework for active collaborative state estimation while operating in unknown environments. In particular, it can be used to identify best robot actions or trajectories among given candidates generated by existing motion planning approaches, or to refine nominal trajectories into locally optimal trajectories using direct trajectory optimization techniques. We demonstrate our approach in a multi-robot autonomous navigation scenario and show that modeling future multi-robot interaction within the belief allows to determine robot trajectories that yield significantly improved estimation accuracy.

1 Introduction

Autonomous operation under uncertainty is essential in numerous problem domains, including autonomous navigation, object manipulation, multi-robot localization and tracking, and robotic surgery. As the robot state is never accurately known due to motion uncertainty and imperfect state estimation obtained from partial and noisy sensor measurements, planning future actions should be performed in the belief space - a probability distribution function (pdf) over robot states and additional states of interest.

V. Indelman
Department of Aerospace Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel. e-mail: vadim.indelman@technion.ac.il

Belief space planning has been investigated extensively in the last two decades. While the corresponding problem can be described in the framework of partially observable Markov decision process (POMDP), which is known to be computationally intractable for all but the smallest problems [17], several approaches that tradeoff optimal performance with computational complexity have been recently developed. These approaches can be segmented into several categories: point-based value iteration methods, simulation based approaches, sampling based approaches and direct trajectory optimization approaches.

Point-based value iteration methods (e.g. [14, 19]) select a number of representative belief points and calculate a control policy over belief space by iteratively applying value updates to these points. *Simulation-based approaches* (e.g. [23, 24]) generate a few potential plans and select the best policy according to a given metric. They are referred to as simulation-based approaches, since they simulate the evolution of the belief for each potential plan to quantify its quality.

Sampling based approaches (e.g. [1, 6, 21]) discretize the state space using randomized exploration strategies to explore the belief space in search of an optimal plan. While many of these approaches, including probabilistic roadmap (PRM) [13], rapidly exploring random trees (RRT) [15] and RRT* [12], assume perfect knowledge of the state, deterministic control and a known environment, efforts have been devoted in recent years to alleviate these restricting assumptions. These include, for example, the belief roadmap (BRM) [21] and the rapidly-exploring random belief trees (RRBT) [1], where planning is performed in the belief space, thereby incorporating the predicted uncertainties of future position estimates. We note that similar strategies are used to address also informative planning problems (see, e.g., [6]).

Direct trajectory optimization methods (including [9, 18, 20, 25]) calculate locally optimal trajectories and control policies, starting from a given nominal path. Approaches in this category perform planning over a continuous state and action spaces, which is often considered more natural as the robot states (e.g., poses) and controls (e.g., steering angles) are not constrained to few discrete values. For example, Platt et al. [20] apply linear quadratic regulation (LQR) to compute locally optimal policies, while Van den Berg et al. [25] develop a related method using optimization in the belief space and avoiding assuming maximum likelihood observations in predicting the belief evolution. These approaches reduce computational complexity to polynomial at the cost of guaranteeing only locally optimal solutions.

While typically, belief space planning approaches consider the environment is known, in certain scenarios of interest (e.g. navigation in unknown environments) this is not a feasible assumption. In these cases, the environment is either a priori unknown, uncertain or changes dynamically, and therefore should be appropriately modeled as part of the inference and decision making processes. Such a concept was recently developed in [8, 9], where random variables representing the observed environment have been incorporated into the belief and locally optimal motion plans were calculated using a direct trajectory optimization approach. In [7], the approach was extended to a multi-robot belief space planning centralized framework and was used to facilitate active collaborative estimation in unknown environments. Simulation- and sampling-based approaches that consider a priori unknown environments have also been recently developed in the context of active SLAM (see, e.g. [4, 24]). A limitation of these approaches is that the belief only considers the environment observed by planning time and does not reason, in the context of uncertainty reduction, about new environments to be observed in the future as the robot continues exploration.

In this work we alleviate this limitation, considering the problem of cooperative multi-robot autonomous navigation in unknown environments. While it is well known that collaboration between robots can significantly improve estimation accuracy, existing approaches (e.g. [3, 10, 22]) typically focus on the inference part, considering robot actions to be determined externally. On the other hand, active multi-robot SLAM approaches (e.g. [2]) typically focus on coordination aspects and on the trade-off between exploring new regions and reducing uncertainty by re-observing previously mapped areas (performing loop closures). In contrast, in this paper we consider the question - how should the robots act to collaboratively improve state estimation while autonomously navigating to individual goals and operating in unknown environments?

Addressing this question requires incorporating multi-robot collaboration aspects into belief space planning. To that end, we present an approach to evaluate the probability distributions of multiple robot states while modeling future observations of mutual areas that are unknown at planning time (Figure 1a). The key idea is that although the environment may be unknown a priori, or has not been mapped yet, it is still possible to reason in terms of robot actions that will result in the same unknown environments to be observed by multiple robots, possibly at different future time instances. Such observations can be used to formulate non-linear constraints between appropriate robot future states. Importantly, these constraints allow collaborative state estimation without the need for the robots to actually meet each other, in contrast to the commonly used direct relative pose observations that require rendezvous between robots (e.g. [22]). We show how such constraints can be incorporated within a multi-robot belief, given candidate paths that can be generated by any motion planning method. One can then identify the best path with respect to a user-defined objective function (e.g. reaching a goal with minimum uncertainty), and also refine best alternatives using direct trajectory optimization techniques (e.g. [9, 18, 25]).

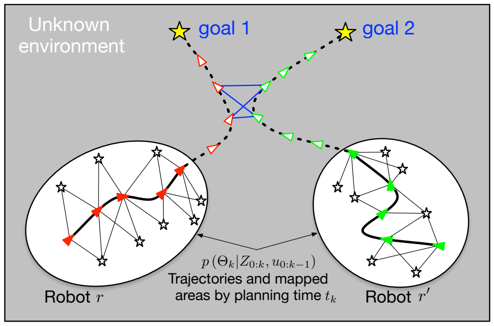
2 Notations and Problem Formulation

Let x_i^r represent the pose of robot r at time t_i and denote by L_i^r the perceived environment by that robot, e.g. represented by 3D points, by that time. We let Z_i^r represent the local observations of robot r at time t_i, i.e. measurements acquired by its onboard sensors, and define the joint state Θ^r over robot past and current poses and observed 3D points as

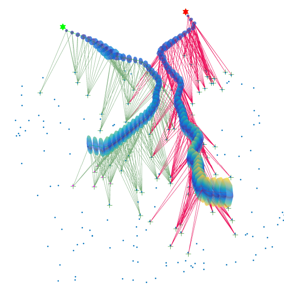
$$\Theta_k^r \doteq X_k^r \cup L_k^r \ , \ X_k^r \doteq \{x_0^r, \dots, x_k^r\} \, . \tag{1}$$

The joint probability distribution function (pdf) over this joint state given local observations $Z_{0:k}^r \doteq \left\{Z_0^r, \dots, Z_k^r\right\}$ and controls $u_{0:k-1}^r \doteq \left\{u_0^r, \dots, u_{k-1}^r\right\}$ is given by

$$p\left(\Theta_k^r | Z_{0:k}^r, u_{0:k-1}^r\right) \propto p\left(x_0^r\right) \prod_{i=1}^{k} \left[p\left(x_i^r | x_{i-1}^r, u_{i-1}^r\right) p\left(Z_i^r | \Theta_i^{ro}\right) \right] , \tag{2}$$



(a)



(b)

Fig. 1: (a) Illustration of the proposed concept. Multi-robot indirect constraints representing mutual future observations of unknown environments are shown in blue. (b) 3D view of the scenario from Figure 4b: Robots operate in an unknown environment and follow paths generated by PRM that have been identified by the proposed approach to provide the best estimation accuracy upon reaching the goals. One can observe the mutually-observed 3D points that induce indirect multi-robot constraints involving different time instances; these constraints have been accounted for in the planning phase. Robot initial positions are denoted by $\star$ marks (at the *top* of the figure); uncertainty covariances of robot poses are represented by ellipsoids.

where $\Theta_i^{ro} \subseteq \Theta_i^r$ are the involved random variables in the measurement likelihood term $p(Z_i^r|\Theta_i^{ro})$, which can be further expanded in terms of individual measurements $z_{i,j}^r \in Z_i^r$ representing observations of 3D points l_j: $p(Z_i^r|\Theta_i^{ro}) = \prod_j p\left(z_{i,j}^r|x_i^r, l_j\right)$. The motion and observation models in Eq. (2) are assumed to be with additive Gaussian noise,

$$x_{i+1}^r = f(x_i^r, u_i^r) + w_i^r \quad , \quad z_{i,j}^r = h(x_i^r, l_j) + v_i^r \tag{3}$$

where $w_i \sim N(0, \Sigma_w^r), v_i \sim N(0, \Sigma_v^r)$, with Σ_w^r and Σ_v^r representing the process and measurement noise covariance matrices, respectively.

We consider now a group of R collaborating robots, and denote by Θ_k the corresponding joint state

$$\Theta_k \doteq X_k \cup L_k \ , \ X_k \doteq \{X_k^r\}_{r=1}^R \tag{4}$$

comprising the past and current poses X_k of all robots, and where L_k represents the perceived environment by the entire group. Assuming a common reference frame between the robots is established, L_k includes all the 3D points in L_k^r for each r, expressed in that reference frame.

The joint pdf over Θ_k, the *belief* at planning time t_k, can now be written as

$$b(\Theta_k) \doteq p(\Theta_k|Z_{0:k}, u_{0:k-1}) \propto \prod_{r=1}^R p\left(\Theta_k^r|Z_{0:k}^r, u_{0:k-1}^r\right), \tag{5}$$

where $u_{0:k-1}$ represents the controls of all robots and is defined as $u_{0:k-1} \doteq \left\{u^r_{0:k-1}\right\}_{r=1}^{R}$.

The joint belief at a future time t_{k+l} can now be similarly defined as

$$b\left(\Theta_{k+l}\right) \doteq p\left(\Theta_{k+l}|Z_{0:k+l}, u_{0:k+l-1}\right), \tag{6}$$

where $u_{k:k+l-1}$ are future actions for a planning horizon of l steps and $Z_{k+1:k+l}$ are the corresponding observations to be obtained. We will discuss in detail how such a belief can be formulated in the sequel (Sections 3.1 and 3.2).

We can now define a general multi-robot objective function

$$J\left(u_{k:k+L-1}\right) \doteq \mathbb{E}\left[\sum_{l=0}^{L} c_l\left(b\left(\Theta_{k+l}\right), u_{k+l}\right) + c_L\left(b\left(\Theta_{k+L}\right)\right)\right], \tag{7}$$

that involves L future steps for all robots, and where c_l is the immediate cost function for the lth step. The expectation operator accounts for all the possible future observations $Z_{k+1:k+l}$. While for notational convenience the same number L of future steps is assumed for all robots in Eq. (7), this assumption can be easily relaxed.

Our objective is to find the optimal controls $u^{\star}_{k:k+L-1}$ for all R robots:

$$u^{\star}_{k:k+L-1} = \underset{u_{k:k+L-1}}{\arg\min}\, J\left(u_{k:k+L-1}\right). \tag{8}$$

3 Approach

In this work we show how to incorporate into belief space planning multi-robot collaboration aspects such that estimation accuracy is significantly improved while operating in unknown environments. Our approach extends the state of the art by incorporating into the belief (6) multi-robot constraints induced by multiple robots observing, possibly at different *future* time instances, environments that are *unknown* at planning time. In lack of sources of absolute information (such as reliable GPS, beacons, and known 3D points), these constraints are the *key* for collaboratively improving estimation accuracy.

One can then identify best robot actions or motion plans, according to Eq. (8), among those generated by existing motion planning approaches (e.g. sampling based approaches), or resort to direct optimization techniques to obtain locally optimal solutions in a timely manner. In this work, we focus on the former case, and consider we are given candidate paths for different robots (generated, e.g. by PRM or RRT). A schematic illustration of the proposed approach is shown in Figure 1a.

We start with a recursive formulation of the multi-robot belief (Section 3.1) and then discuss in Section 3.2 our approach to incorporate into the multi-robot belief future constraints that correspond to mutual observations of unknown scenes. Evaluating the objective function (7) involves simulating belief evolution along candidate robots paths.

3.1 Recursive Formulation of a Multi-Robot Belief

We begin with a recursive formulation of the multi-robot belief (6), considering future controls $u_{0:k+l-1}$ for all robots to be given. These are determined from candidate robot paths that are being evaluated, or alternatively in the case of direct trajectory optimization approaches, the controls are determined from either nominal or perturbed robot paths (see, e.g. [9] for further details).

Given future controls for all robots, the multi-robot belief $b(\Theta_{k+l})$ at the lth future step can be written recursively as follows (see also Eq. (2)):

$$b(\Theta_{k+l}) \doteq p(\Theta_{k+l}|Z_{0:k+l}, u_{0:k+l-1}) \\ = \eta\, b(\Theta_{k+l-1}) \prod_{r=1}^{R} p\left(x^r_{k+l}|x^r_{k+l-1}, u^r_{k+l-1}\right) p\left(Z^r_{k+l}|\Theta^{ro}_{k+l}\right), \tag{9}$$

where η is a normalization constant, and $p\left(x^r_{k+l}|x^r_{k+l-1}, u^r_{k+l-1}\right)$ and $p\left(Z^r_{k+l}|\Theta^{ro}_{k+l}\right)$ are respectively the motion model and measurement likelihood terms.

We now focus on the measurement likelihood term $p\left(Z^r_{k+l}|\Theta^{ro}_{k+l}\right)$, noting that it appears recursively in Eq. (9), for each look ahead step. As earlier, this term represents sensor observations of the environment (represented e.g. by 3D points), see Eq. (2). However, now, these are future observations of the environment to be made according to robot r's planned motion. It therefore makes sense to distinguish between the following two cases: (a) observation of 3D points from $L_k \subset \Theta_k$ representing environments already mapped by planning time t_k, and (b) observation of new areas that were not previously explored by any of the robots.

The former case allows to plan single- and multi-robot loop closures (e.g. as in [9]), i.e. to quantify the expected information gain due to re-observation of previously mapped areas by any of the robots.

We focus on the latter case, which has not been investigated, to the best of our knowledge, in the context of collaborative active state estimation and uncertainty reduction. Since environments that are unknown at planning time t_k are considered, the key question is how to quantify the corresponding measurement likelihood term.

3.2 Incorporating Future Multi-Robot Constraints

Despite the fact that the environments (or objects) to be observed are unknown at planning time, it is still possible to reason in terms of mutual observations of these unknown environments to be made by different robots, possibly at different future time instances. We can then formulate constraints relating appropriate robot states while marginalizing out the corresponding random variables representing the unknown environments.

More specifically, let us consider robots r and r' mutually observing at future times t_{k+l} and t_{k+j}, respectively, an unknown environment represented, e.g., by 3D points $L^{r,r'}_{k+l,k+j}$, with $1 \leq j \leq l$. The joint pdf involving the corresponding states and these 3D points can be written as

$$p\left(x^r_{k+l}, x^{r'}_{k+j}, L^{r,r'}_{k+l,k+j} | z^r_{k+l}, z^{r'}_{k+j}\right) \propto p\left(z^r_{k+l} | x^r_{k+l}, L^{r,r'}_{k+l,k+j}\right) p\left(z^{r'}_{k+j} | x^{r'}_{k+j}, L^{r,r'}_{k+l,k+j}\right)$$

We can now marginalize out the unknown 3D points $L^{r,r'}_{k+l,k+j}$ to get

$$p\left(z^r_{k+l}, z^{r'}_{k+j} | x^r_{k+l}, x^{r'}_{k+j}\right) \propto p\left(x^r_{k+l}, x^{r'}_{k+j} | z^r_{k+l}, z^{r'}_{k+j}\right) = \tag{10}$$

$$= \int p\left(x^r_{k+l}, x^{r'}_{k+j}, L^{r,r'}_{k+l,k+j} | z^r_{k+l}, z^{r'}_{k+j}\right) dL^{r,r'}_{k+l,k+j}, \tag{11}$$

which corresponds to a multi-robot constraint involving different time instances.

In the passive problem setting, i.e. controls and measurements are given, this constraint is typically a nonlinear function that involves the robot poses, say x^r_i and $x^{r'}_j$, and the measured constraint $z^{r,r'}_{i,j}$ which is obtained by matching the measurements z^r_i and $z^{r'}_j$. Typical examples include matching laser scans or images using standard techniques (e.g. ICP, vision-based motion estimation). The corresponding measurement likelihood term can thus be written as

$$p(z^{r,r'}_{i,j} | x^r_i, x^{r'}_j) \propto \exp\left(-\frac{1}{2}\|z^{r,r'}_{i,j} - g\left(x^r_i, x^{r'}_j\right)\|^2_{\Sigma^{MR}_v}\right) \tag{12}$$

where Σ^{MR}_v is the corresponding measurement noise covariance matrix, and g is an appropriate measurement function. For example, this function could represent a nonlinear relative pose constraint.

Coming back to Eq. (10), while in our case the *future* observations are *not* given, the reasoning is very similar: we can denote by $z^{r,r'}_{k+l,k+j}$ the measured constraint that would be obtained by matching z^r_{k+l} and $z^{r'}_{k+j}$ if these were known, and considering, as before, the match is successful (i.e. not outlier), it is possible to quantify the measurement likelihood (10) as

$$p\left(z^{r,r'}_{k+l,k+j} | x^r_{k+l}, x^{r'}_{k+j}\right) \propto \exp\left(-\frac{1}{2}\|z^{r,r'}_{k+l,k+j} - g\left(x^r_{k+l}, x^{r'}_{k+j}\right)\|^2_{\Sigma^{MR}_v}\right) \tag{13}$$

Note the above assumes robots r and r' will observe the same unknown scene from future states x^r_{k+l} and $x^{r'}_{k+j}$. How to determine if two future measurements (e.g. images, laser scans), to be captured from robot poses x^r_{k+l} and $x^{r'}_{k+j}$, will be overlapping, i.e. represent a mutually observed a scene? The answer to this question is scenario specific. For example, in an aerial scenario with robots equipped with downward looking cameras, it is possible to assess if the images are overlapping given robot poses and a rough estimate of height above ground. Ground scenarios allow similar reasoning, however here it is more likely that the same (unknown) scene is observed from multiple views (e.g. autonomous driving with a forward looking camera), and moreover, obstacles, that are unknown at planning time, may prevent two adjacent views to observe a mutual scene in practice.

In this paper we assume one is able to predict if two future poses will mutually observe a scene. Specifically, in Section 4 we consider aerial robots with downward facing cameras and take a simplified approach, considering two future poses

x^r_{k+l} and $x^{r'}_{k+j}$ to overlap if they are "sufficiently" nearby, quantified by a relative distance below a threshold d. Naturally, more advanced approaches can be considered (e.g. account also for viewpoint variation) and be encapsulated by an indicator function as in [16] - we leave the investigation of these aspects to future research.

Given candidate robot paths it is possible to determine using the above method which future views (poses) will overlap and formulate the corresponding multi-robot constraints (13). In particular, multi-robot constraints between robot r at time t_{k+l} and other robots r' at time t_{k+j} with $0 \leq j \leq l$ can be enumerated as

$$\prod_j p\left(z^{r,r'}_{k+l,k+j} | x^r_{k+l}, x^{r'}_{k+j}\right). \tag{14}$$

We can now write the measurement likelihood term $p\left(Z^r_{k+l} | \Theta^{ro}_{k+l}\right)$ from Eq. (9) as:

$$\begin{aligned} p\left(Z^r_{k+l} | \Theta^{ro}_{k+l}\right) = & \prod_{l_j \in \Theta^{ro}_{k+l}} p\left(z^r_{k+l,j} | x^r_{k+l}, l_j\right) p\left(z^r_{k+l,k+l-1} | x^r_{k+l}, x^r_{k+l-1}\right) \cdot \\ & \cdot \prod_j p\left(z^{r,r'}_{k+l,k+j} | x^r_{k+l}, x^{r'}_{k+j}\right). \end{aligned} \tag{15}$$

The first product represents observations of previously mapped 3D points $l_j \in L_k$, with Θ^{ro}_{k+l} including those 3D points that are actually visible from x^r_{k+l}. The second term $p\left(z^r_{k+l,k+l-1} | x^r_{k+l}, x^r_{k+l-1}\right)$ denotes a constraint stemming from robot r observing a mutual unknown scene from adjacent views, while the last product represents multi-robot constraints (14) that correspond to different robots observing common areas that have not yet been mapped by planning time t_k. See schematic illustration in Figure 1a, where these future constraints are shown in blue.

Substituting Eq. (15) into Eq. (9) yields the final expression for $b\left(\Theta_{k+l}\right)$:

$$\begin{aligned} b\left(\Theta_{k+l}\right) = \eta\, b\left(\Theta_{k+l-1}\right) \prod_{r=1}^{R} \Bigg[& p\left(x^r_{k+l} | x^r_{k+l-1}, u^r_{k+l-1}\right) \prod_{l_j \in \Theta^{ro}_{k+l}} p\left(z^r_{k+l,j} | x^r_{k+l}, l_j\right) \\ & p\left(z^r_{k+l,k+l-1} | x^r_{k+l}, x^r_{k+l-1}\right) \cdot \prod_j p\left(z^{r,r'}_{k+l,k+j} | x^r_{k+l}, x^{r'}_{k+j}\right) \Bigg]. \end{aligned} \tag{16}$$

Several remarks are in order at this point. First, observe that direct multi-robot constraints, where a robot measures its pose relative to another robot, are naturally supported in the above formulation by considering the same (future) time index, i.e. $p\left(z^{r,r'}_{k+l,k+l} | x^r_{k+l}, x^{r'}_{k+l}\right)$. Of course, being able to formulate constraints involving also different future time instances, as in Eq. (16), provides enhanced flexibility since planning rendezvous between robots is no longer required. Second, observe the formulation (14) is an approximation of the underlying joint pdf of *multiple* views X making observations Z of an unknown scene L, since it only considers *pairwise* potentials. More concretely, marginalizing L out, $p(X|Z) = \int p\left(X, L|Z\right) dL$, introduces mutual information between all views in X, i.e. any two views in X become correlated. Thus, a more accurate formulation than (14) would consider all robot poses observing a mutual scene together. Finally, one could also incorporate reasoning regarding (robust) data association, i.e. whether a match $z^{r,r'}_{k+l,k+j}$ from raw measurements (images, laser scans) z^r_{k+l} and $z^{r'}_{k+j}$ is expected to be an inlier, as for example done in [11] for the passive case. These aspects are left to future research.

3.3 Inference Over Multi-Robot Belief Given Controls

Having described in detail the formulation of a multi-robot belief $b(\Theta_{k+l-1})$ at each future time t_{k+l}, this section focuses on simulating belief evolution over time given robot controls or paths. As discussed in Section 3, this calculation is required both for sampling based motion planning and direct trajectory optimization approaches.

Thus, we are interested in evaluating the belief $b(\Theta_{k+l})$ from Eq. (16)

$$b(\Theta_{k+l}) \equiv p(\Theta_{k+l}|Z_{0:k+l}, u_{0:k+l-1}) = N\left(\Theta^\star_{k+l}, I_{k+l}\right). \tag{17}$$

which is required for evaluating the objective function (7). Observe that for conciseness we are using here $I_{k+l} \equiv I_{k+l|k+l}$ and $\Theta^\star_{k+l} \equiv \hat{\Theta}_{k+l|k+l}$.

This process involves a maximum a posteriori (MAP) inference

$$\Theta^\star_{k+l} = \underset{\Theta_{k+l}}{\arg\max}\, b(\Theta_{k+l}) = \underset{\Theta_{k+l}}{\arg\min}\, [-\log b(\Theta_{k+l})], \tag{18}$$

which also determines the corresponding information matrix $I_{k+l} = \Sigma^{-1}_{k+l}$.

To perform this inference, recall the recursive formulation (9) and denote the MAP inference of the belief at a previous time by $b(\Theta_{k+l-1}) = N\left(\Theta^\star_{k+l-1}, I_{k+l-1}\right)$. The belief at time t_{k+l} can therefore be written as

$$\begin{aligned}-\log b(\Theta_{k+l}) = &\left\|\Theta_{k+l-1} - \Theta^\star_{k+l-1}\right\|^2_{\Sigma_{k+l-1}} + \\ &+ \sum_{r=1}^{R}\left[\left\|x^r_{k+l} - f(x^r_{k+l-1}, u^r_{k+l-1})\right\|^2_{\Sigma_Q} - \log p\left(Z^r_{k+l}|\Theta^{ro}_{k+l}\right)\right]\end{aligned} \tag{19}$$

We now focus on the term $-\log p\left(Z^r_{k+l}|\Theta^{ro}_{k+l}\right)$. Recalling the discussion from Section 3.2 and Eq. (15), this term can be written as

$$\begin{aligned}-\log p\left(Z^r_{k+l}|\Theta^{ro}_{k+l}\right) = &\sum_{l_j \in \Theta^{ro}_{k+l}} \left\|z^r_{k+l,j} - h(x^r_{k+l}, l_j)\right\|^2_{\Sigma_v} + \\ + \left\|z^r_{k+l,k+l-1} - g(x^r_{k+l}, x^r_{k+l-1})\right\|^2_{\Sigma_v} &+ \sum_j \left\|z^{r,r'}_{k+l,k+j} - g(x^r_{k+l}, x^{r'}_{k+j})\right\|^2_{\Sigma^{MR}_v},\end{aligned} \tag{20}$$

where the motion and measurement models f and h are defined in Section 2, and the nonlinear function g was introduced in Eqs. (12) and (13). We note that while here we consider the measurement noise covariance Σ^{MR}_v to be constant, one could go further and model also accuracy deterioration, e.g. as the relative distance between robot poses increases.

We now proceed with the MAP inference (18), which, if the future observations Z^r_{k+l} were known, could be solved using standard iterative non-linear optimization techniques (e.g. Gauss-Newton and Levenberg-Marquardt): in each iteration the system is linearized, the delta vector $\Delta\Theta_{k+l}$ is recovered and used to update the linearization point, and the process is repeated until convergence.

Let us first describe in more detail this fairly standard approach, considering for a moment the future measurements Z^r_{k+l} are known. The linearization point $\bar{\Theta}_{k+l}$ is discussed first. Recalling that we are to evaluate belief evolution given robot paths, these paths can be considered as the linearization point for robot poses. On the other hand, in the case of direct trajectory optimization approaches, the nominal controls over the planning horizon can be used to generate the corresponding nominal trajectories according to (similar to the single robot case, see, e.g. [9])

$$\bar{x}^r_{k+l} = \begin{cases} f(\bar{x}^r_{k+l-1}, u^r_{k+l-1}), & l > 1 \\ f(\hat{x}^r_k, u^r_k), & l = 1 \end{cases} \tag{21}$$

The linearizaiton point for the landmarks $L_k \subset \Theta_{k+l}$ (see Section 2) is taken as their most recent MAP estimate. We first linearize Eq. (19)

$$\begin{aligned} -\log b\left(\Theta_{k+l}\right) = & \left\|B_{k+l}\Delta\Theta_{k+l}\right\|^2_{\Sigma_{k+l-1}} + \\ & + \sum_{r=1}^{R}\left[\left\|F^r_{k+l}\Delta\Theta_{k+l} - b^r_{k+l}\right\|^2_{\Sigma_Q} - \log p\left(Z^r_{k+l}|\Theta^{ro}_{k+l}\right)\right] \end{aligned} \tag{22}$$

and then linearize the term $-\log p\left(Z^r_{k+l}|\Theta^{ro}_{k+l}\right)$ from Eq. (20):

$$-\log p\left(Z^r_{k+l}|\Theta^{ro}_{k+l}\right) = \sum_{l_j\in\Theta^{ro}_{k+l}}\left\|H^r_{k+l,j}\Delta\Theta_{k+l} - b^r_{k+l,j}\right\|^2_{\Sigma_v} + \tag{23}$$

$$+\left\|G^r_{k+l,k+l-1}\Delta\Theta_{k+l} - b^r_{k+l,k+l-1})\right\|^2_{\Sigma_v} + \sum_j\left\|G^{r,r'}_{k+l,k+j}\Delta\Theta_{k+l} - b^{r,r'}_{k+l,k+j})\right\|^2_{\Sigma^{MR}_v},$$

where the matrices F, H and G and the vectors b are the appropriate Jacobians and right-hand-side (rhs) vectors. The binary matrix B_{k+l} in Eq. (22) is conveniently defined such that $B_{k+l}\Delta\Theta_{k+l} = \Delta\Theta_{k+l-1}$.

Using the relation $\Sigma^{-1} \equiv \Sigma^{-\frac{T}{2}}\Sigma^{-\frac{1}{2}}$ to switch from $\|a\|^2_\Sigma$ to $\|\Sigma^{-\frac{1}{2}}a\|^2$ and stacking all the Jacobians and rhs vectors into $\mathscr{A}_{k+l}$ and $\breve{b}_{k+l}$, respectively, we get

$$\Delta\Theta^\star_{k+l} = \underset{\Delta\Theta_{k+l}}{\arg\min}\left\|\mathscr{A}_{k+l}\Delta\Theta_{k+l} - \breve{b}_{k+l}\right\|^2. \tag{24}$$

The a posteriori information matrix I_{k+l} of the joint state vector Θ_{k+l} can thus be calculated as $I_{k+l} = \mathscr{A}^T_{k+l}\mathscr{A}_{k+l}$.

This constitutes the first iteration of the nonlinear optimization. Recalling again that the future observations Z^r_{k+l} are unknown, it is not difficult to show [9] that, while the a posteriori information matrix I_{k+l} is not a function of these observations, the equivalent rhs vector $\breve{b}_{k+l}$ from Eq. (24) does depend on Z^r_{k+l}. This presents difficulties in carrying out additional iterations as the linearization point itself becomes a function of the unknown random variables Z^r_{k+l}.

As common in related works (e.g. [9, 18, 20, 25]), we assume a single iteration sufficiently captures the impact of a candidate action(s). Alternatively, to better predict uncertainty evolution, one could resort to using the unscented transformation, as in [5], or to particle filtering techniques. Furthermore, for simplicity in this paper we also make the maximum-likelihood measurement assumption, according to which a future measurement z is assumed equal to the predicted measurement using the most recent state estimate. As a result, it can be shown that the rhs vector $\breve{b}_{k+l}$ becomes zero and thus $\Theta^\star_{k+l} = \bar{\Theta}_{k+l}$. We note one could avoid making this assumption altogether at the cost of more complicated expressions, see, e.g. [9, 25].

To summarize, the output of the described inference procedure is a Gaussian that models the multi-robot belief as in Eq. (17): $b\left(\Theta_{k+l}\right) = N\left(\Theta^\star_{k+l}, I_{k+l}\right)$.

3.4 Evaluation of Candidate Paths

Given candidate paths for robots in the group, one can identify the best candidates by evaluating the objective function J from Eq. (7) for different path combinations. Such a process involves simulating belief evolution along the candidate paths of different robots in the group, as discussed in Section 3.3, while accounting for multi-robot collaboration in terms of mutual observations of unknown environments (as discussed in Section 3.2).

4 Simulation Results

In this section we demonstrate the proposed approach considering the problem of multi-robot autonomous navigation while operating in unknown GPS-deprived environments. We consider an aerial scenario, where each robot has its own goal and the objective is to reach these goals in minimum time but also with highest accuracy. This can be quantified by the follwing objective function:

$$J = \sum_{r=1}^{R} \left[\kappa^r t^r_{goal} + (1-\kappa^r) tr\left(\Sigma^r_{goal}\right)\right], \tag{25}$$

where Σ^r_{goal} and t^r_{goal} represent, respectively, the covariance upon reaching the goal and time of travel (or path length) for robot r. The parameter $\kappa^r \in [0, 1]$ weights the importance of each term.

As the environment is unknown and there are no beacons, radio sources or any other means to reset estimation error, the robots can only rely on onboard sensing capabilities and collaboration with each other to reduce drift as much as possible. We assume each robot is equipped with camera and range sensors and can observe natural landmarks in the environment, which are used to estimate robot pose within a standard SLAM framework. However, since the environment is unknown ahead of time, these landmarks are discovered on the fly while the planning process has access only to environments observed by planning time (Section 3). Initial relative poses between the robots are assumed to be known, such that the robots have a common reference frame - approaches that relax this assumption do exist (e.g. [11]).

In this basic study we use a state of the art sampling based motion planning approach, a probabilistic roadmap (PRM) [13], to discretize the environment and generate candidate paths for different robots over the generated roadmap. Figure 2 shows some of these candidate paths considering a scenario of two robots starting operating from different locations. In each case we also show the belief evolution (in terms of uncertainty covariance) along each path, calculated as described in Section 3.3, and the multi-robot constraints that have been incorporated into the appropriate beliefs (denoted by cyan color). In the current implementation, these constraints, possibly involving different future time instances, are formulated between any two poses with relative distance closer than d meters. We use $d = 300$ meters for this threshold parameter (in the considered scenario the aerial robots height is about 500 meters). More advanced methods could be implemented of course, considering also viewpoint variation and incorporating statistical knowledge.

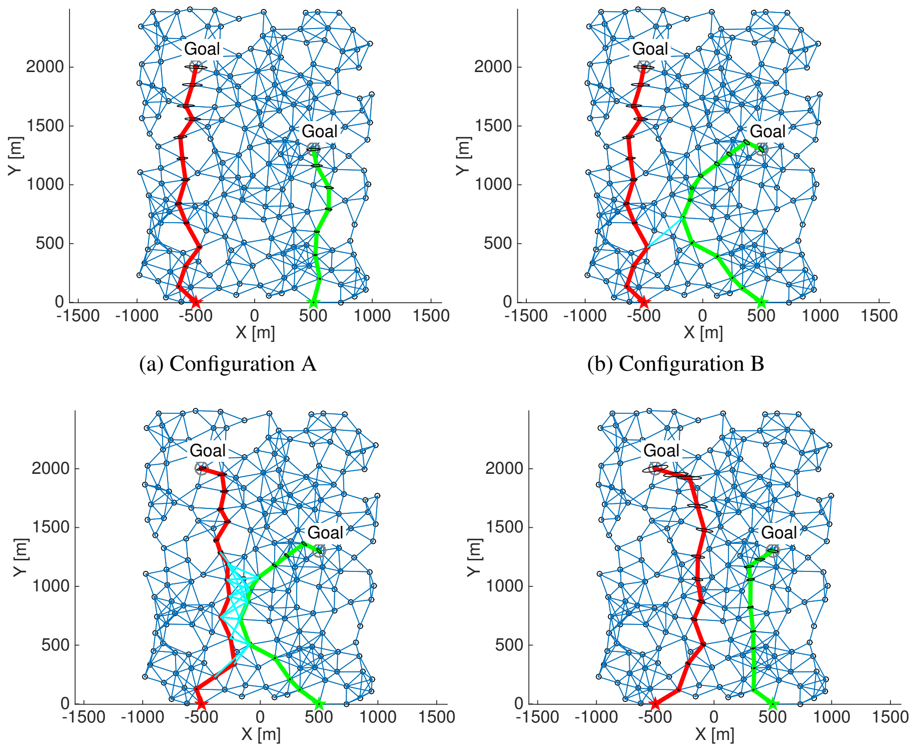
As seen in Figure 2, only in two of the considered cases (Figures 2b and 2c), robot paths were sufficiently close to facilitate multi-robot constraints within belief space planning. In practice, however, only in the latter case numerous informative constraints have been incorporated. Figure 3 compares between the two terms in the considered objective function (25), path length and uncertainty upon reaching the goal, for the candidate paths shown in Figure 2.

The lowest predicted uncertainty covariances are obtained for candidate paths with identified multi-robot constraints as shown in Figure 3b. In particular, the predicted uncertainty is reduced by about 40% from 35 meters to below 20 meters for the first (red) robot. There is a price to pay, however, in terms of path lengths (or time of arrival): as shown in Figure 3a, to attain these levels of uncertainty, the path of the second (gree) robot is not the shortest among the considered candidate paths. The decision what solution is the best therefore depends on the parameter κ from Eq. (25) that weights the importance of each term in the objective function.

Next, we consider actual performance while navigating to pre-defined goals in unknown environments using as controls the identified robot paths in the planning phase described above. The results are shown in Figure 4 for two alternatives from Figures 2a and 2c. Only the latter included multi-robot constraints within planning. One can observe that also in practice, using controls from Configuration C drives the robots sufficiently close to make mutual observations of 3D points (that were unknown at planning time) and as a result significantly improve estimation accuracy for both robots (see Figures 4c and 4d, and Figure 1b for a 3D view).

5 Discussion and Future Work

Results from the previous section indicate estimation accuracy can be significantly improved by modeling multi-robot mutual observations of unknown areas within belief space planning. More generally, we believe similar reasoning can be used to improve multi-robot collaboration aspects while operating also in uncertain, possibly dynamic, environments.



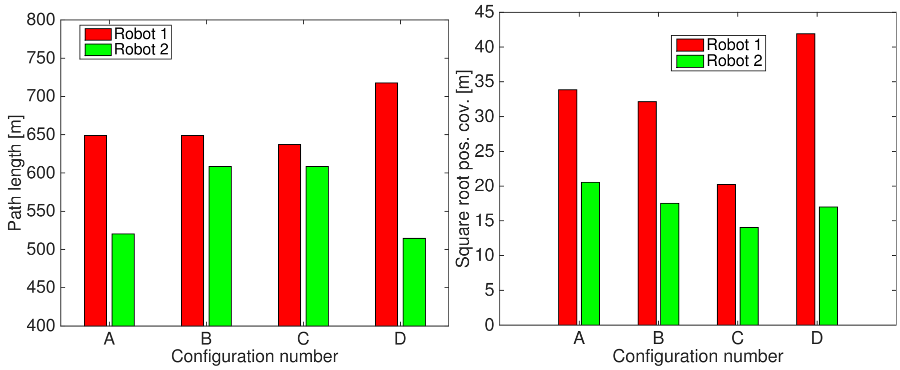
(a) Configuration A (b) Configuration B

(c) Configuration C (d) Configuration D

Fig. 2: Different candidate paths for red and green robots calculated over a PRM. Robot initial positions are denoted by ⋆ marks; each robot has to navigate to a different goal, while operating in an unknown environment. The figures show the covariance evolution along each path. Multi-robot constraints have been incorporated (denoted by cyan color) whenever robot poses are sufficiently close, which happens mainly in (c); as a result, uncertainty covariances are drastically reduced. Note these constraints involve different *future* time instances. Covariances were artificially inflated by a constant factor for visualization - actual values are shown in Figure 3.

In this basic study we have made several simplifying assumptions and did not address some of the challenges that are expected to arise in practical applications.

- Obstacles: While initially the environment is unknown, it may be that after some time obstacles are identified as the robots continue in exploration. These obstacles can be efficiently avoided upon discovery by discarding appropriate paths, as commonly done in sampling based approaches.
- Scalability: Although current implementation uses PRM, our approach can be formulated within any motion planning algorithm. The combinatorial problem associated with evaluating candidate trajectories of different robots is a topic of future research. We note approaches addressing related problems have been actively developed in recent years (e.g. [16]). An interesting direction is to also



(a) Path length (b) Trace of square root final position covariance

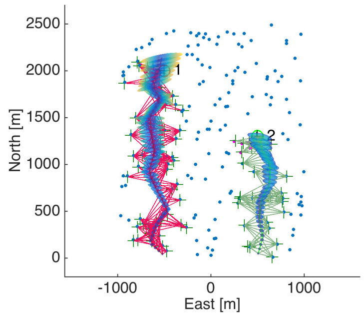
Fig. 3: Quantitative comparison between the four alternatives shown in Figure 2: (a) Path length; (b) covariance upon reaching the goals. Multi-robot constraints lead to lowest predicted uncertainty represented by Configuration C from Figure 2c.

consider generalization of BRM and RRBT to the multi-robot case. A complimentary aspect is to consider direct trajectory optimization approaches, which could allow reducing sampling resolution.

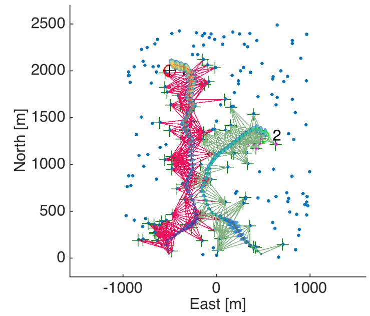
- Belief consistency: While here we consider a centralized approach, decentralized or distributed approaches are often more suitable in practice for numerous reasons. Resorting to these architectures requires the beliefs maintained by different robots to be consistent with each other.

6 Conclusions

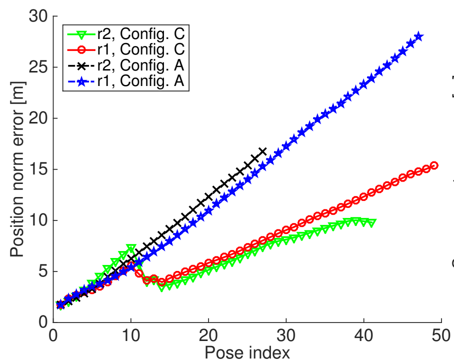
We presented an approach for collaborative multi-robot belief space planning while operating in unknown environments. Our approach advances the state of the art in belief space planning by reasoning about observations of environments that are unknown at planning time. The key idea is to incorporate within the belief constraints that represent multi-robot observations of unknown mutual environments. These constraints can involve different future time instances, thereby providing enhanced flexibility to the group as rendezvous are no longer necessary. The corresponding formulation facilitates an active collaborative state estimation framework. Given candidate robot actions or trajectories, it allows to determine best trajectories according to a user-defined objective function, while modeling future multi-robot interaction and its impact on the belief evolution. Candidate robot trajectories can be generated by existing motion planning algorithms, and most promising candidates could be further refined into locally optimal solutions using direct trajectory optimization approaches. The approach was demonstrated in simulation considering the problem of cooperative autonomous navigation in unknown environments, yielding significantly reduced estimation errors.



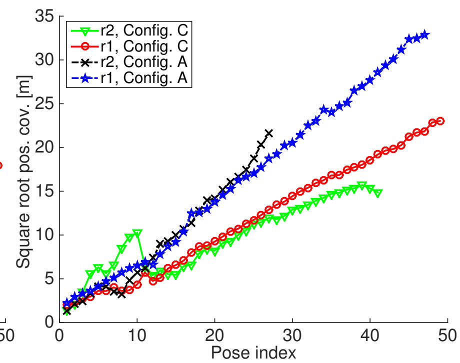
(a) Configuration A



(b) Configuration C



(c) Position estimation errors



(d) Trace of square root position covariance

Fig. 4: Autonomous navigation to goals according to identified robot paths in the planning phase. The environment, represented by a sparse set of landmarks, is initially unknown and only gradually discovered. Figures (a) and (b) show robot trajectories and landmark observations using paths defined, respectively, by Configuration A and C (see Figure 2). The latter involves numerous mutual observations of landmarks, that induce indirectly multi-robot constraints. A 3D view is also shown in Figure 1b. Figures (c) and (d) show the corresponding estimation errors and developing covariance over time, which, in overall, agree with the predicted belief evolution from Figure 3b.

7 Acknowledgments

This work was partially supported by the Technion Autonomous Systems Program.

References

1. A. Bry and N. Roy. Rapidly-exploring random belief trees for motion planning under uncertainty. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 723–730, 2011.
2. W. Burgard, M. Moors, C. Stachniss, and F. Schneider. Coordinated multi-robot exploration. *IEEE Trans. Robotics*, 2005.

3. L. Carlone, M. Kaouk Ng, J. Du, B. Bona, and M. Indri. Rao-Blackwellized particle filters multi robot SLAM with unknown initial correspondences and limited communication. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 243–249, 2010.
4. S. M. Chaves, A. Kim, and R. M. Eustice. Opportunistic sampling-based planning for active visual slam. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, pages 3073–3080. IEEE, 2014.
5. R. He, S. Prentice, and N. Roy. Planning in information space for a quadrotor helicopter in a gps-denied environment. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 1814–1820, 2008.
6. G. A. Hollinger and G. S. Sukhatme. Sampling-based robotic information gathering algorithms. *Intl. J. of Robotics Research*, pages 1271–1287, 2014.
7. V. Indelman. Towards multi-robot active collaborative state estimation via belief space planning. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, September 2015.
8. V. Indelman, L. Carlone, and F. Dellaert. Towards planning in generalized belief space. In *The 16th International Symposium on Robotics Research*, Singapore, December 2013.
9. V. Indelman, L. Carlone, and F. Dellaert. Planning in the continuous domain: a generalized belief space approach for autonomous navigation in unknown environments. *Intl. J. of Robotics Research*, 34(7):849–882, 2015.
10. V. Indelman, P. Gurfil, E. Rivlin, and H. Rotstein. Distributed vision-aided cooperative localization and navigation based on three-view geometry. *Robotics and Autonomous Systems*, 60(6):822–840, June 2012.
11. V. Indelman, E. Nelson, N. Michael, and F. Dellaert. Multi-robot pose graph localization and data association from unknown initial relative poses via expectation maximization. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2014.
12. S. Karaman and E. Frazzoli. Sampling-based algorithms for optimal motion planning. *Intl. J. of Robotics Research*, 30(7):846–894, 2011.
13. L.E. Kavraki, P. Svestka, J.-C. Latombe, and M.H. Overmars. Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *IEEE Trans. Robot. Automat.*, 12(4):566–580, 1996.
14. H. Kurniawati, D. Hsu, and W. S. Lee. Sarsop: Efficient point-based pomdp planning by approximating optimally reachable belief spaces. In *Robotics: Science and Systems (RSS)*, volume 2008, 2008.
15. S. M. LaValle and J. J. Kuffner. Randomized kinodynamic planning. *Intl. J. of Robotics Research*, 20(5):378–400, 2001.
16. D. Levine, B. Luders, and J. P. How. Information-theoretic motion planning for constrained sensor networks. *Journal of Aerospace Information Systems*, 10(10):476–496, 2013.
17. C. Papadimitriou and J. Tsitsiklis. The complexity of markov decision processes. *Mathematics of operations research*, 12(3):441–450, 1987.
18. S. Patil, G. Kahn, M. Laskey, J. Schulman, K. Goldberg, and P. Abbeel. Scaling up gaussian belief space planning through covariance-free trajectory optimization and automatic differentiation. In *Intl. Workshop on the Algorithmic Foundations of Robotics*, 2014.
19. J. Pineau, G. J. Gordon, and S. Thrun. Anytime point-based approximations for large pomdps. *J. of Artificial Intelligence Research*, 27:335–380, 2006.
20. R. Platt, R. Tedrake, L.P. Kaelbling, and T. Lozano-Pérez. Belief space planning assuming maximum likelihood observations. In *Robotics: Science and Systems (RSS)*, pages 587–593, 2010.
21. S. Prentice and N. Roy. The belief roadmap: Efficient planning in belief space by factoring the covariance. *Intl. J. of Robotics Research*, 2009.
22. S.I. Roumeliotis and G.A. Bekey. Distributed multi-robot localization. *IEEE Trans. Robot. Automat.*, August 2002.
23. C. Stachniss, G. Grisetti, and W. Burgard. Information gain-based exploration using rao-blackwellized particle filters. In *Robotics: Science and Systems (RSS)*, pages 65–72, 2005.
24. R. Valencia, M. Morta, J. Andrade-Cetto, and J.M. Porta. Planning reliable paths with pose SLAM. *IEEE Trans. Robotics*, 2013.
25. J. Van Den Berg, S. Patil, and R. Alterovitz. Motion planning under uncertainty using iterative local optimization in belief space. *Intl. J. of Robotics Research*, 31(11):1263–1278, 2012.