iX−BSP: Belief Space Planning through Incremental Expectation

Elad I. Farhi and Vadim Indelman

***Abstract*— Belief space planning (BSP) is a fundamental problem in robotics. Determining an optimal action quickly grows intractable as it involves calculating the expected accumulated cost (reward), where the expectation accounts for all future measurement realizations. State of the art approaches therefore resort to simplifying assumptions and approximations to reduce computational complexity. Importantly, while in robotics re-planning is essential, these approaches calculate each planning session from scratch. In this work we contribute a novel approach, `iX−BSP`, that is based on the *key insight* that calculations in consecutive planning sessions are similar in nature and can be thus re-used. Our approach performs incremental calculation of the expectation by appropriately re-using computations already performed in a precursory planing session while accounting for the information obtained in inference between the two planning sessions. The formulation of our approach considers general distributions and accounts for data association aspects. We evaluate `iX−BSP` in statistical simulation and show that incremental expectation calculations significantly reduce runtime without impacting performance.**

I. Introduction

Decision making under uncertainty and belief space planning (BSP) approaches are entrusted with providing the next optimal action sequence given a certain objective. The aforementioned is accomplished by reasoning about belief evolution for different candidate actions while taking into account different sources of uncertainty. The corresponding problem is an instantiation of a Partially Observable Markov Decision Process (POMDP) problem, which is known to be computationally intractable [1] for all but the smallest problems, i.e. no more than few dozen states [2].

The intractability of the BSP problem originates mainly from the use of expectation in the objective function, $J(\mathcal{U}) = \mathbb{E}_z[\sum_i c_i(b_i, u_{i-1})]$. The objective over a candidate action sequence $\mathcal{U}$, is obtained by calculating the expected value of all possible costs (rewards) c received from following $\mathcal{U}$. Since the cost (reward) function is a function of the belief b and the action led to it u, in practice the objective considers all future beliefs obtained from following $\mathcal{U}$, i.e. all future measurements z. We refer to this general problem as the full solution of BSP, denoted by `X−BSP`, expectation based BSP.

The exponential growth of possible measurements and candidate actions, usually denoted as the *curse of history*, is the key aspect targeted by a lot of research efforts. As in any computational problem, one can either streamline the solution process or change the problem, i.e. take simplifying assumptions or approximations.

Elad I. Farhi is with the Technion Autonomous Systems Program (TASP), Technion - Israel Institute of Technology, Haifa 32000, Israel, `eladf@campus.technion.ac.il`. Vadim Indelman is with the Department of Aerospace Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel, `vadim.indelman@technion.ac.il`.

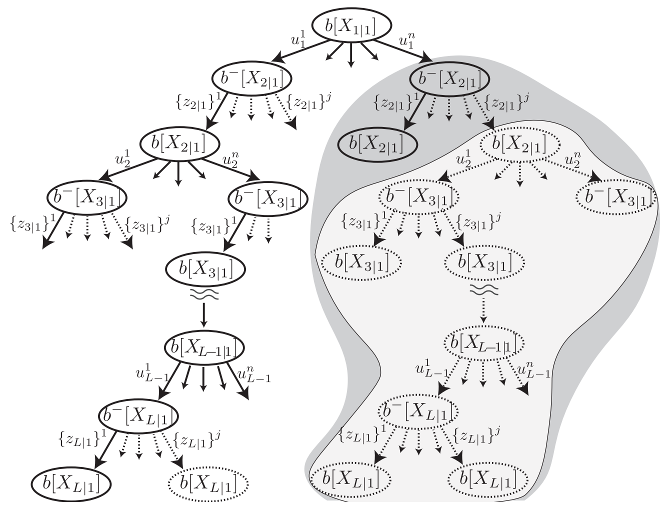


Fig. 1: `X−BSP` performs lookahead search on a tree with depth L. Each belief tree node represents a belief. For each node, the tree branches either for a candidate action or a sampled measurement. The corresponding belief tree for `ML−BSP` is marked with solid lines, while the dashed lines represent the parts of `X−BSP` that relate to sampled measurements. Under `iX−BSP`, the gray-marked parts of the tree are being re-used for the succeeding planning session.

Indeed, over the years, numerous approaches have been developed to trade-off suboptimal performance with reduced computational complexity of POMDP, see e.g. [3]–[6]. Sampling based approaches, e.g. [4], [5], [7]–[9], discretize the state space using randomized exploration strategies to locate the belief's optimal strategy. While many sampling based approaches (e.g. [10]–[12]) assume perfect knowledge of the state (i.e. MDP framework), along with deterministic control and known environment, efforts have been made to assuage these simplifying assumptions. These efforts vary in the alleviated-assumptions, from the belief roadmap (BRM) [9] and the rapidly exploring random belief trees (RRBT) [8], through, Partially Observable Monte-Carlo Planning (POMCP) [13], Determinized Sparse Partially Observable Tree (DESPOT) [14], [15] and up to active full SLAM in discrete [16] and continuous [17] domains accounting for uncertainties in the environment mapped thus far as part of the decision making process (e.g. [16]–[19]) at the price of increased state dimensionality.

While all the aforementioned research efforts tackle the curse of history through providing various approximations to the `X−BSP` problem, the common denominator for most of them is the Maximum Likelihood (ML) assumption [20], which allows to prune `X−BSP` by considering only the maximum likelihood measurements rather than all possible ones. We denote the use of ML in BSP as `ML−BSP`.

In strike contrast to the vast amount of research invested in approximating the `X−BSP` problem, only few tried re-using calculations. Although under simplifying assumptions, amongst them ML, both [21] and [22] re-use computationally expensive calculations during planning. In their work, Chaves and Eustice [21], consider a Gaussian belief under ML−BSP in a Bayes tree [23] representation. All candidate action sequences consider a shared location (entrance pose), thus enabling to re-use a lot of the calculations through state ordering constrains. That work enables to efficiently evaluate a single candidate action across multiple time steps, and is conceptually applicable to multiple candidate actions at a single time step. While Kopitkov and Indelman [22], also consider a Gaussian belief under ML−BSP, they utilize a factor graph representation of the belief while considering an information theoretic cost. Using an (augmented) determinant lemma, they are able to avert from belief propagation while re-using calculations throughout the planning session. Although they consider calculation re-use within the same planning session, their work can be augmented to consider re-use also between planning sessions.

To the best of our knowledge, in-spite of aforementioned research efforts, calculation re-use has only been done over ML−BSP, with restricting assumptions. While ML−BSP is widely used, the pruning of X−BSP by considering only the most likely measurements, might mean choosing a sub-optimal action in case the biggest available reward is not the most likely one, in particular in presence of significant estimation uncertainty. As for today, X−BSP approaches do not re-use calculations between consecutive planning sessions, and regard each planning session for its own.

Our *key observation* is that the similarity between two successive planning sessions can be utilized to re-use calculations, thus salvaging valuable computation time. In this paper we provide a novel paradigm for Incremental eXpectation BSP, or iX−BSP, which incrementally updates the expectation related calculations in X−BSP, by re-using the measurements sampled in a precursory planning session. Instead of re-calculating the planning session each time from scratch, we create it by incrementally updating the precursory session with newly received information, using our previous work on efficient belief update [24], [25].

Figure 1 illustrates a planning session of X−BSP at planning time $t = 1$ for a horizon of L steps, which determined action u_1^n as optimal. In the succeeding planning session, instead of re-calculating from scratch, we take the dark-gray-marked part from precursory planning and update it with information received between the successive planning sessions. First, we locate the predicted measurement $\{z^j_{2|1}\}$ closest to the one received in inference, and prune our selection accordingly, i.e. light-gray-marked part. Now we go over all previously sampled measurements, update their impact over current objective (i.e. importance sampling [26]), which includes updating relevant beliefs while accounting for possibly different data association (DA), and re-sampling.

To summarize, our contributions in this paper are as follows: (a) We introduce a novel approach for incremental expectation belief space planning (iX−BSP) that avoids the common assumption of maximum likelihood (ML) observations. Our approach calculates incrementally the expectation over future observations by appropriately re-using sampled measurements and additional calculations from the previous planning session. (b) We incorporate within our approach data association (DA) aspects while accounting for the potentially changing DA across planning sessions. (c) We evaluate iX−BSP in simulation and compare it to X−BSP, which calculates expectation form scratch, while considering the problem of autonomous navigation in unknown environments. (d) We provide a statistical study that highlights BSP with expectation is superior compared to ML−BSP, and show iX−BSP significantly reduces runtime of X−BSP.

II. Background and Problem Formulation

This section provides the theoretical background for belief space planning (BSP), starting with belief definition, followed by the BSP formulation. While the formulation, as well as the suggested paradigm, are impartial to a specific belief distribution, throughout this paper we also provide the conventional case which deals with Gaussian distributions.

A. Belief Definition

Let x_t denote the agent's state at time instant t and $\mathcal{L}$ represent the mapped environment thus far. The joint state, up to and including time k, is defined as $X_k = \{x_0, ..., x_k, \mathcal{L}\}$. We shall be using the notation $t|k$ to refer to some time instant t while considering information up to and including time k. The unique time notation is required since this paper makes use of both current and future time indices in the same equations. Let $z_{t|k}$ and $u_{t|k}$ denote, respectively, measurements and the control action at time t, while the current time is k. The measurements and controls up to time t given current time is k, are represented by

$$z_{1:t|k} \doteq \{z_{1|k}, ..., z_{t|k}\}\ ,\ u_{0:t-1|k} \doteq \{u_{0|k}, ..., u_{t-1|k}\}, \quad (1)$$

The posterior probability density function (pdf) over the joint state, denoted as the *belief*, is given by

$$b[X_{t|k}] \doteq \mathbb{P}(X_t|z_{1:t|k}, u_{0:t-1|k}) = \mathbb{P}(X_t|H_{t|k}). \quad (2)$$

where $H_{t|k} \doteq \{u_{0:t-1|k}, z_{1:t|k}\}$ represents history at time t given current time k. The propagated belief at time t, i.e. belief $b[X_{t|k}]$ lacking the measurements of time t, is denoted by $b^-[X_{t|k}] \doteq b[X_{t-1|k}] \cdot \mathbb{P}(x_t|x_{t-1}, u_{t-1|k}) = \mathbb{P}(X_t|H^-_{t|k})$, where $H^-_{t|k} \doteq H_{t-1|k} \cup \{u_{t-1|k}\}$.

Using Bayes rule, Eq. (2) can be rewritten as $b[X_{t|k}] \propto \mathbb{P}(X_0) \prod_{i=1}^{t} \Big[\mathbb{P}(x_i|x_{i-1}, u_{i-1|k}) \prod_{j \in \mathcal{M}_{i|k}} \mathbb{P}(z_{i,j|k}|x_i, l_j)\Big]$, where $\mathbb{P}(X_0)$ is the prior on the initial joint state, and $\mathbb{P}(x_i|x_{i-1}, u_{i-1|k})$ and $\mathbb{P}(z_{i,j|k}|x_i, l_j)$ denote, respectively, the motion and measurement likelihood models. Here, $z_{i,j|k}$ represents an observation of landmark l_j from robot pose x_i, while the set $\mathcal{M}_{i|k}$ contains all landmark indices observed at time i, i.e. it denotes data association (DA). The DA of few time steps is denoted by $\mathcal{M}_{1:i|k} \doteq \{\mathcal{M}_{1|k}, \cdots, \mathcal{M}_{i|k}\}$.

B. Belief Space Planning

The purpose of BSP is to determine an optimal action given an objective function J, belief $b[X_{k|k}]$ at planning time instant k and, considering a discrete action space, a set of candidate actions $\mathcal{U}_k$. While these actions can be with different planning horizons, we consider for simplicity the same horizon of L look ahead steps for all actions, i.e. $\mathcal{U}_k = \{u_{k:k+L-1}\}$. The optimal action is given by $u^\star_{k:k+L-1|k} = \arg\min_{u_{k:k+L-1|k} \in \mathcal{U}_k} J(u_{k:k+L-1|k})$, where the general objective function $J(.)$ is defined as

$$J(u) \doteq \mathbb{E}\left[\sum_{i=k+1}^{k+L} c_i\left(b[X_{i|k}], u_{i-1|k}\right)\right], \tag{3}$$

with $u \doteq u_{k:k+L-1|k}$, immediate costs (or rewards) c_i and where the expectation is with respect to future observations $z_{k+1:k+L|k}$. The expectation can be written explicitly as

$$J(u) = \int_{z_{k+1|k}} \mathcal{P}_{k+1|k}(c_{k+1} + \ldots \int_{z_{i|k}} \mathcal{P}_{i|k}(c_i + \ldots)\ldots) \tag{4}$$

where each integral accounts for all possible measurement realizations from an appropriate look ahead step, $\mathcal{P}_{i|k} \doteq \mathbb{P}(z_{i|k}|H^-_{i|k})$ with $i \in [k+1, k+L]$ and $H^-_{i|k}$ is a function of a specific sequence of measurement realization, i.e.

$$H^-_{i|k} = H_{k|k} \cup \{z_{k+1:i-1|k}, u_{k:i-1|k}\}. \tag{5}$$

Above, we also used $c_i \doteq c_i\left(b[X_{i|k}], u_{i-1|k}\right)$, where $b[X_{i|k}] = \mathbb{P}(X_{i|k}|H^-_{i|k}, z_{k+i|k})$.

Evaluating the objective for each candidate action in $\mathcal{U}_k$ involves calculating (4), considering different measurement realizations. As solving these integrals analytically is typically not feasible, in practice these are approximated by sampling future measurements from appropriate distributions.

Specifically, consider the i-th future step and corresponding $H^-_{i|k}$ to some realization of measurements from the previous steps. In order to sample from $\mathbb{P}(z_{i|k}|H^-_{i|k})$, we should marginalize over the future robot pose x_i and landmarks $\mathcal{L}$

$$\mathbb{P}(z_{i|k}|H^-_{i|k}) = \int_{x_i}\int_{\mathcal{L}} \mathbb{P}(z_{i|k}|x_i, \mathcal{L}) \cdot \mathbb{P}(x_i, \mathcal{L}|H^-_{i|k}) dx_i d\mathcal{L}, \tag{6}$$

where $\mathbb{P}(x_i, \mathcal{L}|H^-_{i|k})$ can be calculated from the belief $b^-[X_{i|k}] \doteq \mathbb{P}(X_{i|k}|H^-_{i|k})$. We approximate the above integral via sampling as summarized in Algorithm 1. One can also choose to approximate further by considering only landmark estimates $\hat{\mathcal{L}}$ (i.e. without sampling $\mathcal{L}$).

Algorithm 1 Sampling $z_{i|k} \sim \mathbb{P}(z_{i|k}|H^-_{i|k})$

1: $\chi_i \doteq \{x_i, \mathcal{L}\} \sim \mathbb{P}(x_i, \mathcal{L}|H^-_{i|k})$
2: Determine data association $\mathcal{M}_{i|k}(x_i, \mathcal{L})$
3: $z_{i|k} = \{z_{i,j|k}\}_{j \in \mathcal{M}_{i|k}(\chi_i)}$ with $z_{i,j|k} \sim \mathbb{P}(z_{i,j|k}|x_i, l_j)$
4: **return** $z_{i|k}$ and χ_i

Each sample χ_i and the determined DA (lines 1-2 of Alg. 1) define a measurement likelihood $\mathbb{P}(z_{i|k}|\chi_i, \mathcal{M}_{i|k}(\chi_i)) = \prod_{j \in \mathcal{M}_{i|k}(\chi_i)} \mathbb{P}(z_{i,j|k}|x_i, l_j)$ from which observations are sampled in line 3. Considering n_x samples, $\{\chi_i^n\}_{n=1}^{n_x}$, we can approximate Eq. (6) by $\mathbb{P}(z_{i|k}|H^-_{i|k}) \approx \eta_i \sum_{n=1}^{n_x} w_i^n \cdot \mathbb{P}(z_{i|k}|\chi_i^n, \mathcal{M}_{i|k}(\chi_i^n))$, where w_i^n represents n-th sample weight, χ_i^n, and $\eta_i^{-1} \doteq \sum_{n=1}^{n_x} w_i^n$. Here, since all samples are generated from their original distribution (corresponding to the proposal distribution in importance sampling), see line 1, we have identical weights.

For each sample $\chi_i^n \in \{\chi_i^n\}_{n=1}^{n_x}$, we can generally consider n_z measurement samples (line 3), providing the set $\{z_{i|k}^{n,m}\}_{m=1}^{n_z}$. In other words, Algorithm 1 yields $n_x \cdot n_z$ sampled measurements, denoted by $\{z_{i|k}\}$, for a given realization of $z_{k+1:i-1|k}$. Thus, considering all such possible realizations, we get $(n_x \cdot n_z)^i$ sampled measurements for the i-th look ahead step. Figure 1 illustrates this conceptually.

We can now write an unbiased estimator for (4), considering the $(n_x \cdot n_z)^i$ sampled measurements. In particular, for the i-th look ahead step, we get

$$\mathbb{E}_{z_{k+1:i|k}}[c_i] \approx \eta_{k+1} \sum_{\{z_{k+1|k}\}} w_{k+1}^n(\cdots(\eta_i \sum_{\{z_{i|k}\}} w_i^n \cdot c_i)\cdots) \tag{7}$$

where $H^-_{i|k}$ varies with each measurement realization.

The above exponential complexity makes the described calculations quickly infeasible, due to both curse of dimensionality and history. In practice, approximate approaches, e.g. Monte-Carlo tree search [13], must be used. However, in this work we prefer to present our paradigm considering the above formulation, without any further approximations, referring to it as X−BSP. We believe our proposed concept can be applied in conjunction with existing approximate approaches, however leave this endeavor for future research.

Before proceeding further, we mention another common approximation to Eq. (3), ML−BSP, which corresponds to $n_x = n_z = 1$ (and considering the mean of the propagated belief in the above generative model). The belief propagation illustrated in Figure 1 is reduced to contain only the beliefs marked with a solid line, thus drastically reducing complexity at the expense of sacrificing performance.

C. Problem Statement

We are now in a position to formulate the problem addressed in this work. Consider the planning session at time instant k has been solved by evaluating the objective (3) via appropriate measurement sampling for each action in $\mathcal{U}_k$ and subsequently choosing the optimal action $u^\star_{k:k+L-1|k}$. A subset of this action, $u^\star_{k:k+l-1|k} \in u^\star_{k:k+L-1|k}$ with $l \in [1, L)$, is now executed, new measurements $z_{k+1:k+l|k+l}$ are obtained and the posterior belief $b[X_{k+l|k+l}]$ in inference is calculated, upon which a new planning session is initiated.

Determining the optimal action sequence at time instant $k+l$ involves evaluating the objective function for each candidate action $u' \doteq u_{k+l:k+l+L-1|k+l} \in \mathcal{U}_{k+l}$

$$J(u') \doteq \mathbb{E}\left[\sum_{i=k+l+1}^{k+l+L} c_i\left(b[X_{i|k+l}], u'_{i-1|k+l}\right)\right], \tag{8}$$

where the expectation is with respect to future observations $z_{k+l+1:k+l+L|k+l}$. Existing approaches perform these costly evaluations from scratch for each candidate action.

In contrast, our goal in this work is to develop an approach for evaluating the objective function (8) more efficiently by re-using calculations from the previous planning session.

At this point, we summarize our assumptions in this work.

Assumption 1: Calculations from a precursory planning session are accessible from the current planning session.

Assumption 2: The planning horizon of current time $k+l$, overlaps the planning horizon of the precursory planning time k, i.e. $l \in [1, L)$.

Assumption 3: Action sets $\mathcal{U}_{k+l}$ and $\mathcal{U}_k$ overlap in the sense that actions in $\mathcal{U}_k$ which overlap in the executed portion of the optimal action also partially reside in $\mathcal{U}_{k+l}$. In other words, $\forall u \in \mathcal{U}_k$ with $u \doteq \{u_{k:k+l-1|k}, u_{k+l:k+L-1|k}\}$ and $u_{k:k+l|k} \equiv u^\star_{k:k+l-1|k}$, $\exists u' \in \mathcal{U}_{k+l}$ such that $u' \doteq \{u'_{k+l:k+L-1}, u'_{k+L:k+l+L-1}\}$ and $u'_{k+l:k+L-1} \cap u_{k+l:k+L-1|k} \notin \emptyset$.

III. Approach

Our *key observation* is that expectation related calculations from two successive X-BSP planning sessions at times k and $k+l$ are similar and, often, calculations related to sampled measurements from the former planning session can be appropriately re-used in the planning session at time $k+l$. Based on this observation we develop incremental expectation BSP (iX-BSP) approach that saves valuable computation time, while at the same time preserving the benefits of the expectation solution obtained by X-BSP.

A. Approach Overview

Let us compare between the objective functions for two planning sessions (3) and (8), with $u \in \mathcal{U}_k$ and $u' \in \mathcal{U}_{k+l}$, and consider, for simplicity, u and u' overlap, i.e. $u \ni u_{k+l:k+L-1|k} = u_{k+l:k+L-1|k+l} \in u'$, and $u = u^\star$ was the optimal action determined at time k, where a sub-sequence $u_{k:k+l-1} \in u$ has been executed by the current time $k+l$. We note our approach is applicable also to the more general case where actions only partially overlap (see Assumption 3), as discussed in the sequel.

Along the overlapping planning horizon (between $k+l+1$ and $k+L$), calculations performed for u at time k and those to be performed for u' at time $k+l$ are similar - in both cases, measurements are sampled from *similar* distributions. The key difference resides in the information up to time $k+l$: While at time k, different measurement realizations $\{z_{k+1:k+l|k}\}$ were considered (via sampling), at time $k+l$ a particular measurement realization $z_{k+1:k+l|k+l}$ was captured in inference, in practice.

For example, consider the look ahead time step $k+l+1$. At the first planning session, the expectation is with respect to $\mathbb{P}(z_{k+l+1|k}|H^-_{k+l+1|k})$, where $H^-_{k+l+1|k}$ is a function of a measurement realization $z_{k+1:k+l|k} \in \{z_{k+1:k+l|k}\}$. On the other hand, at the second planning session, the expectation is with respect to $\mathbb{P}(z_{k+l+1|k}|H^-_{k+l+1|k+l})$, where $H^-_{k+l+1|k+l}$ involves the actually obtained measurements $z_{k+1:k+l|k+l}$.

Considering again the sampled different measurement realizations from time k, $\{z_{k+1:k+l|k}\}$ for each action in $\mathcal{U}_k$, it is possible to identify a realization, in terms of action sequence $u_{k:k+l-1|k}$ and observations $z_{k+1:k+l|k}$, that is closest to actions taken $u_{k:k+l-1:k+l}$ and obtained measurements $z_{k+1:k+l|k+l}$, such that the corresponding beliefs $b[X_{k+l|k}]$ and $b[X_{k+l|k+l}]$ are closest. We formalize this notion and discuss our method to do so in Sec. III-B.

The identified closest realization $z_{k+1:k+l|k}$, corresponds to its own $H_{k+l|k}$, which has been used, at time k, for sampling measurements in the next look ahead steps, e.g. $\{z_{k+l+1|k}\}$, $\{z_{k+l+2|k}\}$ etc., as described in Sec. II-B. Importantly, the costs c_i for these sampled measurements have been already calculated at time k. Assuming these calculations are available at time $k+l$, we now discuss how these computations can be appropriately re-used.

To this end, our approach considers two aspects, considering the fact that, in the general case, $z_{k+1:k+l|k}$ and $z_{k+1:k+l|k+l}$ are not identical. First, the sampled measurements $z_{i|k}$ for each look ahead step $i \in [k+l+1, k+L]$, that were originally sampled from $\mathbb{P}(z_{i|k}|H^-_{i|k})$, see Eq. (6), should be re-weighted such that these samples could be considered as samples from the correct distribution $\mathbb{P}(z_{i|k}|H^-_{i|k+l})$. This step is discussed in Sec. III-C. Second, we re-use calculations of costs c_i that were performed at planning time k, while accounting for the difference between $b[X_{k+l|k}]$ and $b[X_{k+l|k+l}]$. This is discussed in Sec. III-D. The planning steps that cannot be re-used from the precursory planning session, e.g. $k+L+1 : k+l+L$, are calculated from scratch (as in X-BSP).

Next, we go over each step in iX-BSP: selecting calculations for re-use (Sec. III-B), re-using the samples (Sec. III-C) and updating information from current time (Sec. III-D).

B. Selecting Closest Predictions for Re-Use

While the precursory planning session contains several candidate actions $u_{k:k+l-1|k}$ and for each of which several future beliefs $b[X_{k+l|k}]$, the current planning session contains a single given action sequence $u_{k:k+l-1|k+l}$ resulting in a single posterior $b[X_{k+l|k+l}]$. In order to utilize calculations from precursory planning we need to first choose from it, the candidate action sequence, and resulting belief which are closest to those in current time $k+l$.

We start the selection process by considering all action sequences $u_{k:k+L-1|k}$ meeting Assumption 3. Each of these actions propagated the precursory posterior $b[X_{k|k}]$ into several candidate future beliefs $b[X_{k+l|k}]$ using different sets of sampled measurements $z_{k+1:k+l|k}$. We wish to prune the calculations from precursory planning further and remain with a single belief $b[X_{k+l|k}]$, which is equal to the current posterior ($b[X_{k+l|k+l}]$) in the sense of sampled measurements. Because the actions are already identical the measurements are the only difference between both beliefs.

It is highly unlikely to find a belief such that $z_{k+1:k+l|k} \equiv z_{k+1:k+l|k+l}$, hence we look for the closest match in the sense of DA and measurement values. First we

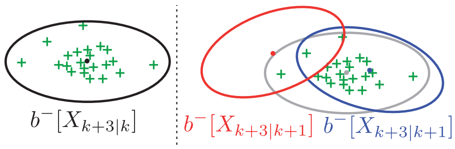


Fig. 2: Illustration for adequate and inadequate representative sample. Samples in green, beliefs mean and covariance are represented by ellipse and dot respectivly.

prune according to DA, we keep only the beliefs closest to $b[X_{k+l|k+l}]$ in the sense of DA, i.e. the set of beliefs corresponding all DA such that $\mathcal{M}_{k+1:k+l|k} \in \underset{\mathcal{M}_{k+1:k+l|k}}{\arg\max}(\mathcal{M}_{k+1:k+l|k} \cap \mathcal{M}_{k+1:k+l|k+l})$.

In case we are not left with a single belief, we continue pruning seeking minimal measurement value difference. We choose the set of beliefs that correspond to the sampled measurements $z_{k+1:k+l|k}$ which are closest to $z_{k+1:k+l|k+l}$ by value. This comparison is crucial since the sampled measurements that we are forcing in current planning session, i.e. $\{z_{k+l+1:k+L|k}\}$, were sampled as a function of $z_{k+1:k+l|k}$ (see Sec. II-B), hence directly affects the relevance of the re-used samples (as discussed in the sequel).

At the end of the process we are either left with a single belief $b[X_{k+l|k}]$ closest to $b[X_{k+l|k+l}]$, or we are left with few beliefs which are equally close to $b[X_{k+l|k+l}]$; in case of the latter, we just pick one and prune the others.

C. Re-using Samples

After Sec. III-B, we have a set of sampled measurements, from the precursory planning session, for each of the overlapping horizon steps $k+l+1 : k+L$.

Since we are forcing the measurement samples, the values used for the estimation are of $z_{i|k}$, rather than of $z_{i|k+l}$. Moreover, the old samples were generated from the set of sampled states $\chi_{i|k}$, hence the measurement model should consider $\chi_{i|k}$ as the given sampled location (Algorithm 1, line 1), rather than $\chi_{i|k+l}$. Under importance sampling the approximation of Eq. (6) is given by $\mathbb{P}(z_{i|k}|H^-_{i|k+l}) \approx \eta_i \sum_{n=1}^{n_x} w_i^n \cdot \mathbb{P}(z_{i|k}|\chi^n_{i|k}, \mathcal{M}_{i|k+l}(\chi^n_{i|k}))$, where w_i^n are no longer uniform, but instead equal the probability $\mathbb{P}(\chi = \chi_{i|k}|H^-_{i|k+l})$.

Since we are re-using previously sampled measurements (a.k.a. importance sampling), we need to assure they constitute a representative sample of the true measurement likelihood. Figure 2 illustrates the problem with re-using samples from precursory planning sessions. Consider propagated belief $b^-[X_{k+3|k}]$, colored in black. During planning time k we sampled (Alg. 1, line 1) from the propagated belief $b^-[X_{k+3|k}]$. We now consider planning time $k+1$, and in it, two propagated beliefs $b^-[X_{k+3|k+1}]$, colored red and blue in Figure 2, each of which consider different realizations at inference time $k+1$, hence with different posterior. While the samples provide adequate representative sample for the blue propagated belief, they lack doing so for the red propagated belief. To avoid this problem, we examine re-used samples coverage by judging weight values. When needed, we re-sample measurements from scratch to provide with an adequate representative sample for the distribution.

D. Incremental Belief Update: from $b[X_{k+l|k}]$ to $b[X_{k+l|k+l}]$

The last remaining aspect in `iX-BSP` approach is re-using the costs c_i. Specifically, for each sampled measurement $z_{i|k} \sim \mathbb{P}(z_{i|k}|H^-_{i|k})$, the corresponding posterior belief $b[X_{i|k}]$ and cost c_i have already been calculated. However, in-spite re-using the sampled measurements $z_{k+l+1:i|k}$, the posterior belief $b[X_{i|k}]$ still needs to be updated to recover $b[X_{i|k+l}]$. In particular, these two beliefs can be written as

$$b[X_{i|\alpha}] \propto b[X_{k+l|\alpha}] \prod_{s=k+l+1}^{i} [\mathbb{P}(x_s|x_{s-1}, u_{s-1|\alpha}) \prod_{j \in \mathcal{M}_{s|\alpha}} \mathbb{P}(z^j_{s|\alpha}|x_s, l_j)],$$

with $\alpha \in \{k, k+l\}$ and $\mathcal{M}_{s|k} = \mathcal{M}_{s|k+l}$, $\forall s$. Thus, to recover $b[X_{i|k+l}]$, $b[X_{i|k}]$ has to be updated to account for the difference between $b[X_{k+l|k}]$ and $b[X_{k+l|k+l}]$. In case u and u' partially overlap (see Assumption 3), the motion model terms for $b[X_{i|k+l}]$ and $b[X_{i|k}]$ will also partially differ.

Given two beliefs, $b[X_{k+l|k}]$ and $b[X_{k+l|k+l}]$, the former can be used to update the latter to the point of algebraic equality. Luckily, this update can be done *incrementally*, building upon our previous work [24], [25], which considered multivariate Gaussian distributions. We believe similar concepts are also applicable to more general distributions, which should be investigated in future work.

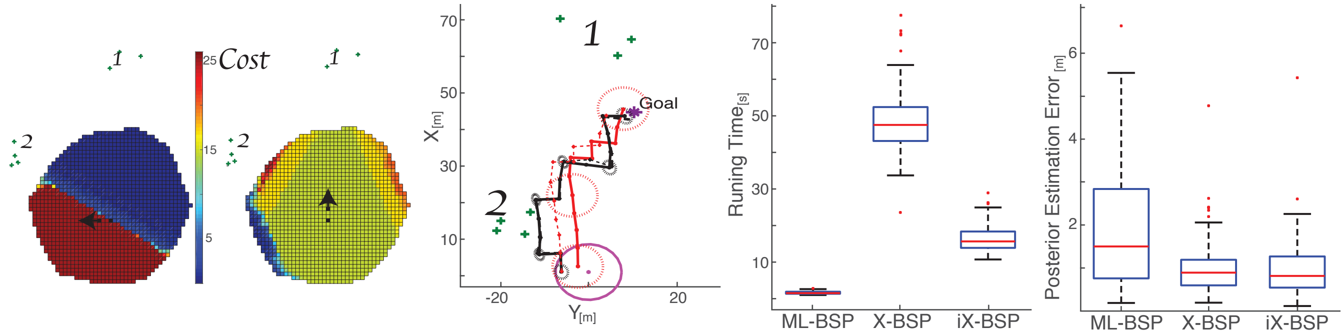
Briefly, for the special case of consistent DA between two beliefs, i.e. $\mathcal{M}_{k+l|k} = \mathcal{M}_{k+l|k+l}$, such update can be performed using the `DU-OO` method (see details in [24]). In case of inconsistent DA, as presented in our recent work (see [25], one can update the DA before using the aforementioned to update the belief. In particular, the first step in updating the DA is to mark the inconsistency between both beliefs, meaning which associations in $b[X_{k+l|k}]$ do not appear in $b[X_{k+l|k+l}]$ and vice versa. Once the inconsistencies are flagged, we incrementally remove the unnecessary associations from $b[X_{k+l|k}]$, and add the correct ones from $b[X_{k+l|k+l}]$ using `iSAM2` efficient methodologies (see [27]).

IV. Results

In this section we provide statistical comparison between `X-BSP`, `ML-BSP` and `iX-BSP`, using Active-SLAM as a test-bed under Model Predictive Control (MPC) framework. For simplicity we consider the following: (a) Motion and observation models with additive zero-mean Gaussian noise. (b) All landmarks are already part of the joint state. (c) All previously sampled measurements provide adequate representative sample in current planning time, i.e. all samples are re-used. Code implemented in MATALB, and executed on a Linux machine, with Xeon E3-1241v3 3.5GHz processor with 64GB of memory. In the sequel we present a single-action statistical comparison of `X-BSP` to `ML-BSP`, and then compare between `iX-BSP`, `X-BSP`, and `ML-BSP`.

A. The ML Assumption

In this section we provide a glimpse behind the curtains of `X-BSP` and `ML-BSP`. We show the results of a single



(a) (b) (c) (d) (e)

Fig. 3: (a) and (b) Spatial sensitivity to the ground truth location in respect to the objective value when considering "left" and "forward" actions accordingly. While `X-BSP` considers the weighted average of different possible measurements, denoted by colored area, `ML-BSP` considers only the most likely measurement, denoted by the black arrow. (c) Testing scenario, landmarks denoted by green "+", prior state and uncertainty in solid purple, `ML-BSP` denoted by red, `X-BSP` and `iX-BSP` denoted by black, see Sec. IV-B for more details. (d) and (e) Box plots of 100 rollouts for planning session timing (d) and posterior estimation error (e) upon reaching the goal.

planning session, for which expectation and ML produced different optimal actions. Consider a robot with initial estimated location and covariance, given two candidate actions. The world consists out of two types of landmarks, the first with high covariance and the second with low. Figures 3a-3b present the spatial cost values which are the result of choosing "left" or "forward" actions accordingly, and where warm colors denote higher cost values. While ML considers only the cost value where the arrow is placed, expectation considers multiple samples from different spatial locations. As a result expectation favored the "left" action while ML favored "forward". For $20k$ inference rollouts, each with a different ground truth location, choosing left is favorable in the sense of minimizing cost (uncertainty), 74% of the times.

B. *Re-Using The Precursory Planning Session*

In this section we compare all three methods in the sense of planning-session computation time and the posterior estimation error upon reaching the goal. For comparison we perform 100 rollouts (entire mission run), each with a different sampled ground-truth for the prior state. For each rollout, we time the planning sessions of all three methods. Figure 3c presents the scenario on which all rollouts were performed. Considering the same world and same landmarks as in Sec. IV-A. A robot equipped with a stereo camera, is required to reach the goal whilst not crossing a covariance threshold, i.e. cost consisting of distance to goal and a covariance penalty above a certain value. Figure 3c shows one of the 100 rollouts that were calculated, in which the estimated trajectory by each method is denoted by a solid line, the ground truth by a dashed line and the posterior covariance by a dashed ellipse. In Figure 3c both `X-BSP` and `iX-BSP`, in black, chose the same optimal actions along the mission, while `ML-BSP`, in red, chose differently. We can also see the effect of this difference over each method's covariance, `X-BSP` and `iX-BSP` action choice led to a smaller covariance along the entire path.

Figure 3d presents the statistical data of the planning session running time. Since in this example we follow an MPC framework, the last step of each horizon is required to be calculated from scratch. Since doing so is identical to the course of action in `X-BSP`, we present the computation time of the entire horizon, excluding the last horizon step. As expected, for average timing as well as for each separate rollout, both `ML-BSP` and `iX-BSP` timings are lower than that of `X-BSP`. By re-using previous planning session, instead of calculating it from scratch we save valuable computation time, theoretically without effecting the planning solution. We examine the effect on the planning solution in Figure 3e, by comparing the posterior estimation error upon reaching the goal. As expected, the statistical results of 100 rollouts presented in Figure 3e, shows that `X-BSP` is statistically superior to `ML-BSP`: in 63% of the rollouts it has a smaller estimation error while in 10% they are equal. Importantly, we can also see that `iX-BSP` is statistically similar to `X-BSP`, with 41% of the rollouts with smaller estimation error and 15% equal. We note that relaxing the simplifying assumption that all samples provide an adequate representative sample, would result with an even better match between `X-BSP` and `iX-BSP`.

V. Conclusions

State of the art approaches under `X-BSP` paradigm (BSP with expectation) calculate each planning session from scratch. In this paper we presented an alternative paradigm, `iX-BSP`, that utilizes similarities and appropriately re-uses calculations between two consecutive planning sessions. Performing importance sampling using samples from precursory planning session enables to incrementally update the expectation calculations for the current planning sessions. Thus, planning calculations can be efficiently re-used in order to save valuable computation time. After recalling the strengths of expectation over the common maximum likelihood observations assumption, we showed our proposed `iX-BSP` is statistically equal to `X-BSP` whilst providing shorter computation time.

Since our paradigm changes the solution approach of the original, un-approximated, problem (`X-BSP`), we believe it can be utilized to also reduce computation time of existing approximations of `X-BSP`.

References

[1] D. Bernstein, R. Givan, N. Immerman, and S. Zilberstein, "The complexity of decentralized control of markov decision processes," *Mathematics of operations research*, vol. 27, no. 4, pp. 819–840, 2002.

[2] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Artificial intelligence*, vol. 101, no. 1, pp. 99–134, 1998.

[3] J. Pineau, G. J. Gordon, and S. Thrun, "Anytime point-based approximations for large POMDPs." *J. of Artificial Intelligence Research*, vol. 27, pp. 335–380, 2006.

[4] 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)*, vol. 2008, 2008.

[5] G. A. Hollinger and G. S. Sukhatme, "Sampling-based robotic information gathering algorithms," *Intl. J. of Robotics Research*, pp. 1271–1287, 2014.

[6] M. Toussaint, "Robot trajectory optimization using approximate inference," in *Intl. Conf. on Machine Learning (ICML)*. ACM, 2009, pp. 1049–1056.

[7] A.-A. Agha-Mohammadi, S. Chakravorty, and N. M. Amato, "FIRM: Sampling-based feedback motion planning under motion uncertainty and imperfect measurements," *Intl. J. of Robotics Research*, vol. 33, no. 2, pp. 268–304, 2014.

[8] A. Bry and N. Roy, "Rapidly-exploring random belief trees for motion planning under uncertainty," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2011, pp. 723–730.

[9] S. Prentice and N. Roy, "The belief roadmap: Efficient planning in belief space by factoring the covariance," *Intl. J. of Robotics Research*, vol. 28, no. 11-12, pp. 1448–1465, 2009.

[10] L. Kavraki, P. Svestka, J.-C. Latombe, and M. Overmars, "Probabilistic roadmaps for path planning in high-dimensional configuration spaces," *IEEE Trans. Robot. Automat.*, vol. 12, no. 4, pp. 566–580, 1996.

[11] S. M. LaValle and J. J. Kuffner, "Randomized kinodynamic planning," *Intl. J. of Robotics Research*, vol. 20, no. 5, pp. 378–400, 2001.

[12] S. Karaman and E. Frazzoli, "Sampling-based algorithms for optimal motion planning," *Intl. J. of Robotics Research*, vol. 30, no. 7, pp. 846–894, 2011.

[13] D. Silver and J. Veness, "Monte-carlo planning in large pomdps," in *Advances in Neural Information Processing Systems (NIPS)*, 2010, pp. 2164–2172.

[14] N. Ye, A. Somani, D. Hsu, and W. S. Lee, "Despot: Online pomdp planning with regularization," *JAIR*, vol. 58, pp. 231–266, 2017.

[15] Y. Luo, H. Bai, D. Hsu, and W. S. Lee, "Importance sampling for online planning under uncertainty," *IJRR*, p. 0278364918780322, 2016.

[16] C. Stachniss, D. Haehnel, and W. Burgard, "Exploration with active loop-closing for FastSLAM," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2004.

[17] 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*, vol. 34, no. 7, pp. 849–882, 2015.

[18] R. Valencia, M. Morta, J. Andrade-Cetto, and J. Porta, "Planning reliable paths with pose SLAM," *IEEE Trans. Robotics*, pp. 1050–1059, 2013.

[19] A. Kim and R. M. Eustice, "Active visual SLAM for robotic area coverage: Theory and experiment," *Intl. J. of Robotics Research*, vol. 34, no. 4-5, pp. 457–475, 2014.

[20] R. Platt, R. Tedrake, L. Kaelbling, and T. Lozano-Pérez, "Belief space planning assuming maximum likelihood observations," in *Robotics: Science and Systems (RSS)*, Zaragoza, Spain, 2010, pp. 587–593.

[21] S. M. Chaves and R. M. Eustice, "Efficient planning with the Bayes tree for active SLAM," in *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*. IEEE, 2016, pp. 4664–4671.

[22] D. Kopitkov and V. Indelman, "No belief propagation required: Belief space planning in high-dimensional state spaces via factor graphs, matrix determinant lemma and re-use of calculation," *Intl. J. of Robotics Research*, vol. 36, no. 10, pp. 1088–1130, August 2017.

[23] M. Kaess, V. Ila, R. Roberts, and F. Dellaert, "The Bayes tree: An algorithmic foundation for probabilistic robot mapping," in *Intl. Workshop on the Algorithmic Foundations of Robotics*, Dec 2010.

[24] E. I. Farhi and V. Indelman, "Towards efficient inference update through planning via jip - joint inference and belief space planning," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2017.

[25] E. Farhi and V. Indelman, "ix-bsp: Belief space planning through incremental expectation," Technion - Israel Institute of Technology, Tech. Rep. ANPL-2019-01, 2019. [Online]. Available: http://vindelman.technion.ac.il/Publications/Farhi19icra_Supplementary.pdf

[26] P. W. Glynn and D. L. Iglehart, "Importance sampling for stochastic simulations," *Management Science*, vol. 35, no. 11, pp. 1367–1392, 1989.

[27] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. Leonard, and F. Dellaert, "iSAM2: Incremental smoothing and mapping using the Bayes tree," *Intl. J. of Robotics Research*, vol. 31, pp. 217–236, Feb 2012.