Efficient Informative Planning with High-dimensional Non-Gaussian Beliefs by Exploiting Structure

Gilad Rotman

Efficient Informative Planning with High-dimensional Non-Gaussian Beliefs by Exploiting Structure

Research Thesis

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Autonomous Systems and Robotics

Gilad Rotman

Submitted to the Senate
of the Technion — Israel Institute of Technology
Adar alef 5782 Haifa February 2022

This research was carried out under the supervision of Associate Professor Vadim Indelman, in the interdepartmental program of Autonomous Systems and Robotics.

Some results in this thesis have been published or submitted as articles by the author and research collaborators in conferences and journals during the course of the author's research period, the most up-to-date versions of which being:

G. Rotman and V. Indelman. Towards involve-MI: Informative Planning with High-Dimensional Non-Parametric Beliefs. In *61st Israel Annual Conference on Aerospace Sciences (IACAS)*, 2022.

ACKNOWLEDGEMENTS

This period in my life has been one of the most challenging I have faced yet, both in terms of professional aspects and in terms of personal, family aspects, together with the Covid-19 outbreak which has been hard for almost everyone. Yet, I could not have been in a more supportive environment.

First of all, I would like to thank my supervisor, Assoc. Prof. Vadim Indelman. There are not enough words to express my appreciation for him. The combination between his knowledge, involvement, enthusiasm, kindness and patience is what makes him the ultimate supervisor from my perspective.

I would also like to thank the administrating staff of TASP for the highly supportive approach towards the students.

And last but not least, I would like to thank my family. In particular my wife Noya, which I got married with and had two adorable little children during this period. She is the one who believed in me the most, and allowed me to take this journey to the unknown, which has been challenging for her no less than it was for me.

The generous financial help of the Technion is gratefully acknowledged.

Contents

List of Figures

Abstract

One of the most complex tasks of decision making and planning is to gather information. This task becomes even more complex when the state is high-dimensional and its belief cannot be expressed with a parametric distribution. Although the state is high-dimensional, in many problems only a small fraction of it might be involved in transitioning the state and generating observations. We exploit this fact to calculate an information-theoretic expected reward, mutual information (MI), over a much lower-dimensional subset of the state, to improve efficiency and without sacrificing accuracy. A similar approach was used in previous works, yet specifically for Gaussian distributions, and we here extend it for general distributions. Moreover, we apply the dimensionality reduction for cases in which the new states are augmented to the previous, yet again without sacrificing accuracy. We then continue by developing an estimator for the MI which works in a Sequential Monte Carlo (SMC) manner, and avoids the reconstruction of future belief's surfaces. Finally, we show how this work is applied to the informative planning optimization problem. This work is then evaluated in a simulation of an active SLAM problem, where the improvement in both accuracy and timing is demonstrated.

Abbreviations and Notations

POMDP	:	Partially Observable Markov Decision Process
SLAM	:	Simultaneous Localization and Mapping
SMC	:	Sequential Monte Carlo
IG	:	Information Gain
MI	:	Mutual Information
X_0	:	prior state vector (when planning starts); also referred to as X
X_t	:	state vector at future time t; also referred to as X'
x_t	:	new state augmented to X_t at time t
X^{in}	:	subset of X_0 with variables *involved* in future factors
$X^{\neg in}$	:	subset of X_0 with variables *not involved* in future factors
X_{new}	:	subset of X_t representing *new augmented* variables
$z_{1:t}$	:	realization of future observation sequence up to time t
$Z_{1:t}$	:	future observation sequence up to time t as a random variable
$a_{0:t-1}$	:	candidate future action sequence up to time t
h_t	:	history up to time t
$\mathbb{P}_T$	:	probabilistic transition model
$\mathbb{P}_Z$	:	probabilistic observation model
$\mathcal{P}_T$	:	sequential transition model
$\mathcal{P}_Z$	:	sequential observation model
η^{-1}	:	sequential normalizer
$b[X]$	:	belief of state vector X, its probability density function
$J(\cdot)$	:	objective function
ρ_t	:	belief-based reward at time t
$\mathcal{H}[\cdot]$	:	differential entropy
$IG[\cdot]$	:	information gain
$IG_{aug}[\cdot]$	:	augmented information gain
$I[\cdot]$	:	mutual information
$I_{aug}[\cdot]$	:	augmented mutual information

Chapter 1

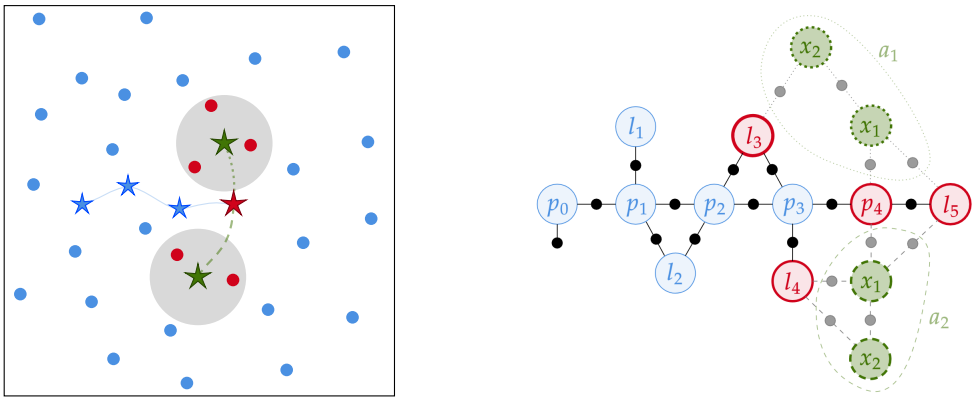
Introduction

Planning under uncertainty is of most importance for many applications. Our world is stochastic in nature, thus for every inference and planning task this stochasticity needs to be taken into account, or catastrophes may occur.

Addressing stochasticity can be done in many levels. At the simplest level, the stochasticity is only being considered at the inference phase, while at planning the state is treated as if it is completely deterministic. In one of the next levels, stochasticity is also being considered within planning, however the planner is considered to reach the goal without caring about the uncertainty of the state. In one of the top levels, the uncertainty of the state also needs to be measured. For example, if the plan is to reach a goal state with some minimum probability or if the plan is to purely gain as much information as possible about the state. Such an approach, where we select a sequence of actions based also (or only) upon the certainty of the state, is known as informative planning. For instance, a task which is considered as informative planning is search and rescue, where in the ”search” phase an exploration of unknown terrains might be done.

There are a few measures for the uncertainty of a state, which we will refer to as information-theoretic costs or rewards. A very common such cost is entropy. Yet, for many applications we wouldn't be interested in the absolute value of the uncertainty, but rather by how much we expect it to be reduced, or in other words how much information is to be gained. Such a reward is called Information Gain (IG). Also, since the state is not always directly observable, observations are used. But when evaluating the future, observations themselves are also unknown, thus the uncertainty of these should also be taken into account. Consequently, we consider Mutual Information (MI), which is the expected IG over the observations. We note that the reader might find that some parts of the literature refer to MI also as IG, but as in the artificial intelligence literature, we distinct between the two terms in this thesis.

Naturally, by increasing the dimension of the problem, the complexity of evaluating information-theoretic terms is increased as well. This relation is exponential (known as the Curse of Dimensionality), making the solution intractable.



(a) (b)

Figure 1.1: A toy example of an active SLAM problem, where a drone has to choose between two actions. **(a)** As a 2D map: stars represent drone's poses; dots represent landmarks; big circles are drone's eye of sight from future poses. **(b)** As a factor graph: circles represent variable nodes; dots represent factors (probabilistic constraints); the unity of all solid-circled variable nodes represent the prior state X_0. In both figures, colors partition between involved, uninvolved and new (future) variables.

1.1 Related Work

Reviewing state-of-the-art approaches for planning with non-Gaussian (non-parametric) beliefs, most do not attempt to address the problem of evaluating the uncertainty of a state (e.g. [8, 9, 14–16, 20, 21]). These works consider the POMDP framework, which does not support belief-dependent rewards. The recent approaches presented in e.g. [24] and [7], do support belief-dependent rewards, by using the more general ρ-POMDP framework and simulating particle sets in a MCTS fashion to construct a search tree. [17] uses GMMs to represent non-Gaussian beliefs and relaxes assumptions regarding data association. Yet, while these are concerned with improving the search over a tree or relax assumptions, the calculation of the information-theoretic terms when the state is high-dimensional remains a problem.

Many estimators exist for information-theoretic rewards. A brief overview of the most common estimators for entropy can be found in [1]. Another approach to estimate entropy, a particle-filter based approach, is presented in [2]. Furthermore, [19] suggests to use a piece-wise linear approximation of the beliefs' surface. Yet, all will face the Curse of Dimensionality. The work presented in [23] claims to support high-dimensional spaces, however it requires a minimum number of samples dependent on the dimension. Also, simulations show this approach isn't superior in terms of accuracy for *any* arbitrary belief compared to other approaches.

When changing the context a bit, and looking at an expected reward, mutual information, [3] and [25] both state that the MI over a multi-dimensional variable can be calculated over a lower dimensional subset of this variable. Yet their statements are applied specifically for the problems of feature matching and active mapping, respectively. Also, these do not address problems in which quantifying the information should also take into account the state might change.

There are two works which, at some level, construct planners with an information-gathering task and which support high-dimensional non-Gaussian beliefs with complexity not exponential with the dimension of the state, yet do not support the more general setting we are addressing in this thesis. One of which is [22], addressing the very specific active SLAM problem. It exploits

the unique structure of SLAM to use Rao-Blackwellization, which might not be useful for more general problems. Also, it approximates the joint entropy by averaging the individual entropies of the state variables, thus breaks up the correlations between them.

The second work is [18], which uses a slightly different and more specific formulation than the general POMDP, where the objective is to minimize the expected cost. In its formulation, the objective is to reach a desired region of the state space with a guaranteed minimum probability of success. This, in turn, means that it wouldn't necessarily choose the most informative path. Also, it does not support, at least directly, information-theoretic rewards. And, lastly, this approach was eventually tested only on problems with up to four dimensions.

Assuming Gaussian distributions, [11] discovered that there are cases in which the correlations between some of the state variables could be discarded for the calculation of entropy without sacrificing accuracy. This, in turn, means that efficiency is improved. This idea was extended to more general cases by [4] and [13].

To the best of our knowledge, there are no existing approaches which deal with the painful and critical issue of calculating information-theoretic terms in general ρ-POMDPs with high-dimensional non-Gaussian beliefs. Our approach aims to extend the works of [4] and [13] to general non-Gaussian distributions in order to fill this void.

1.2 Contributions

In this thesis we propose a novel approach to fill the above gap. By exploiting structure inherent in many planning problems, we break the relation between the dimension of the problem and the complexity of calculating the information-theoretic terms. Fig. 1.1a shows a toy example of a 2D active SLAM problem to illustrate the structure our approach exploits. An aerial drone observes landmarks from above with a certain sensing range. At current time it needs to choose between two actions. While the state is high-dimensional because it contains many landmarks, we observe that only a few landmarks might be observed at future time, and we will denote these landmarks as the *involved* variables. The *key idea* of our approach is to discard the uninvolved landmarks at planning time for the calculation of the expected reward, MI, although these are impacted as well. We prove that by discarding these variables, the solution is still exact, yet the dimension (and thus the complexity) of the problem is reduced dramatically. Furthermore, a naive approach entails also explicit estimation of future beliefs' surfaces, which adds yet another level of errors. Another key contribution of our approach is that it allows skipping this step by having access to the problem's models. We finalize by showing how these contributions can be used together with existing belief tree solvers.

1.3 Organization

This thesis is organized as follows.

1. Chapter 2 introduces the informative planning problem, and gives a formal statement of it.

2. Chapter 3 describes our approaches `involve-MI` and `MI-SMC` and their relation to the informative planning problem.

3. Chapter 4 presents experimental results, evaluating the proposed approach and comparing it against current approaches.

4. Conclusions are drawn in Chapter 5.

5. For purpose of simplicity, proofs are given in the Appendix in Chapter 6.

Chapter 2

Notations and Problem Formulation

2.1 POMDPs

A Partially Observable Markov Decision Process (POMDP) is modeled as a tuple $\langle \mathcal{X}, \mathcal{A}, \mathcal{Z}, b\left[X_0\right], \mathbb{P}_T, \mathbb{P}_Z, r \rangle$, where $\mathcal{X}$ is the state space, $\mathcal{A}$ is the action space, and $\mathcal{Z}$ is the observation space; $b\left[X_0\right] \triangleq \mathbb{P}(X_0)$ is the prior belief over the state, $\mathbb{P}_T$ is the probabilistic transition model, and $\mathbb{P}_Z$ is the probabilistic observation model; $r\left(X_t, a_t\right)$ is an immediate reward function.

The prior belief and probabilistic models are used in the inference layer for determining future posterior beliefs. The belief over the state X_t at time t is defined by

$$b\left[X_t\right] \triangleq \mathbb{P}(X_t \mid h_t), \tag{2.1}$$

where $h_t = \{a_{0:t-1}, z_{1:t}\}$ is the history, containing all actions $a_{0:t-1}$ and observations $z_{1:t}$ acquired up to time t. The transition model $\mathbb{P}_T \triangleq \mathbb{P}(x_t \mid X_{t-1}^{tr}, a_{t-1})$ defines the distribution of the successor state x_t, given a subset of its previous state $X_{t-1}^{tr} \subseteq X_{t-1}$ participating in the transition, and the chosen action a_{t-1}. In this paper, we use the *smoothing* formulation, where the successor state is *augmented* to the previous, thus constructing the *joint* successor state $X_t = \{X_{t-1}, x_t\}$. This formulation means that the state's dimension increases in time.

The observation model $\mathbb{P}_Z \triangleq \mathbb{P}(z_t \mid X_t^{obs})$ defines the conditional distribution of receiving an observation z_t, given a subset of the state $X_t^{obs} \subseteq X_t$ which participates in generating the observation.

For example, in the case of active (full) SLAM, the state is defined as the union of all poses and landmarks. The transition model can be formulated between two consecutive poses, and the observation model can be formulated such that an observation is generated given the last pose and a specific landmark. This example is illustrated in the factor graph in Figure 1.1b, where p refers to poses, l to landmarks and the factor nodes represent the probabilistic constraints between the variables, given with the motion and observation models.

Using Bayes' rule and the chain rule, the belief can be defined recursively as

$$b\left[X_t\right] = \frac{\mathbb{P}(x_t \mid X_{t-1}^{tr}, a_{t-1})\mathbb{P}(z_t \mid X_t^{obs})}{\mathbb{P}(z_t \mid h_t)} b\left[X_{t-1}\right], \tag{2.2}$$

and, more generally, for $0 \leq l < t$ the belief can be formulated as

$$b\left[X_t\right] = \frac{\prod_{i=l+1}^{t} \mathbb{P}(x_i \mid X_{i-1}^{tr}, a_{i-1}) \prod_{j=l+1}^{t} \mathbb{P}(z_j \mid X_j^{obs})}{\prod_{k=l+1}^{t} \mathbb{P}(z_k \mid h_k)} b\left[X_l\right]. \tag{2.3}$$

Denoting the sequential transition model, observation model and normalizer, respectively, as

$$\begin{aligned} \mathcal{P}_T &\triangleq \prod_{i=l+1}^{t} \mathbb{P}(x_i \mid X_{i-1}^{tr}, a_{i-1}) \\ \mathcal{P}_Z &\triangleq \prod_{j=l+1}^{t} \mathbb{P}(z_j \mid X_j^{obs}) \\ \eta^{-1} &\triangleq \prod_{k=l+1}^{t} \mathbb{P}(z_k \mid h_k^-), \end{aligned} \tag{2.4}$$

where $h_k^- \triangleq h_k \setminus z_k$. The belief is then formulated, in short, as

$$b\left[X_t\right] = \eta \mathcal{P}_T \mathcal{P}_Z b\left[X_l\right]. \tag{2.5}$$

The tasks of the planning scheme are represented with the immediate reward function $r\left(X_t, a_t\right)$. Obviously, each planning task might involve multiple different (and sometimes contradicting) tasks, such as energy consumption and time to reach a goal, thus the reward functions might be shaped with multiple different terms. As will further be explained, the reward in our formulation must also express information-gathering tasks, which in turn means that the reward has to be belief-dependent. Thus, we use ρ-POMDP, an extension of POMDP, whose reward at each time t is $\rho_t \triangleq \rho\left(b\left[X_t\right], a_t\right)$.

Planning T steps into the future, the objective is then to find an action sequence $a_{0:T-1}$ which maximizes the expected sum of rewards, denoted as the objective function

$$J\left(b\left[X_0\right], a_{0:T-1}\right) = \mathop{\mathbb{E}}_{\mathcal{Z}_{1:T}} \left[\sum_{t=0}^{T-1} \rho_t + \rho_T\right], \tag{2.6}$$

where $\mathcal{Z}_{1:T}$ is the sequential space of future observations and $\rho_T \triangleq \rho\left(b\left[X_T\right]\right)$ is a terminal reward. As will further be explained, our approach is based upon evaluation of the expected rewards rather than the rewards themselves. Due to commutativity, the objective function can also be defined as the sum of expected rewards

$$J\left(b\left[X_0\right], a_{0:T-1}\right) = \sum_{t=0}^{T-1} \left[\mathop{\mathbb{E}}_{\mathcal{Z}_{1:t}} \left[\rho_t\right]\right] + \mathop{\mathbb{E}}_{\mathcal{Z}_{1:T}} \left[\rho_T\right]. \tag{2.7}$$

We note that while the problem is formulated with an open loop formulation, our approach is not limited to it alone. Since the approach focuses on evaluating the (expected) rewards in the objective function, it also supports a close loop formulation, in which the objective is to find a policy π rather than an action sequence. This, in turn, means that each action a_t becomes a function of the belief $b[X_t]$, thus is not determined beforehand. Focusing on the objective function evaluation also means that other building blocks of the planning task, such as the inference engine and the optimization solver, can be chosen independently.

We now return to explain that the reason we have formulated our problem as ρ-POMDP is that we focus on planners which contain information-gathering tasks (among other tasks). Arguably, evaluating the information-theoretic terms involves the heaviest calculations of the objective function.

2.2 Information-gathering tasks

We start by introducing some known information-theoretic terms for the simple case in which the state is static, i.e. $X_t = \tilde{X}$, $\forall t$. A commonly used information-theoretic reward is negative (differential) entropy. The entropy of the state $\tilde{X} \in \mathcal{X}$, distributed with $\mathbb{P}(\tilde{X})$, is defined as

$$\mathcal{H}\left[\tilde{X}\right] \triangleq -\int_{\mathcal{X}} \mathbb{P}(\tilde{X}) \log \mathbb{P}(\tilde{X}) d\tilde{X}. \tag{2.8}$$

Getting an observation $\tilde{z} \in \mathcal{Z}$, without transitioning the state, yields

$$\mathcal{H}\left[\tilde{X} \mid \tilde{Z} = \tilde{z}\right] \triangleq -\int_{\mathcal{X}} \mathbb{P}(\tilde{X} \mid \tilde{Z} = \tilde{z}) \log \mathbb{P}(\tilde{X} \mid \tilde{Z} = \tilde{z}) dX, \tag{2.9}$$

where $Z_{1:t} = \tilde{Z}$ represents the observation sequence as a random variable. The expected reward, referred to as conditional entropy, is then defined as

$$\mathcal{H}\left[\tilde{X} \mid \tilde{Z}\right] = \mathop{\mathbb{E}}_{\mathcal{Z}}\Big[\mathcal{H}\left[\tilde{X} \mid \tilde{Z} = \tilde{z}\right]\Big] = -\iint_{\mathcal{X},\mathcal{Z}} \mathbb{P}(\tilde{X}, \tilde{Z}) \log \mathbb{P}(\tilde{X} \mid \tilde{Z}) d\tilde{X} d\tilde{Z}. \tag{2.10}$$

Another commonly used information-theoretic reward is information gain (IG), which quantifies the amount of information gained for a certain variable by knowing the value of another variable. It is defined as the difference between the entropy prior to this additional knowledge and the entropy afterwards

$$IG\left[\tilde{X}; \tilde{Z} = \tilde{z}\right] \triangleq \mathcal{H}\left[\tilde{X}\right] - \mathcal{H}\left[\tilde{X} \mid \tilde{Z} = \tilde{z}\right]. \tag{2.11}$$

Similarly to the entropy case, we introduce the expected reward of IG, which is referred to as the mutual information (MI). It can also be defined as the difference between the entropy of the state and the expected entropy of the state given the new observation

$$I\left[\tilde{X}; \tilde{Z}\right] \triangleq \mathop{\mathbb{E}}_{\mathcal{Z}}\Big[IG\left[\tilde{X}; \tilde{Z} = \tilde{z}\right]\Big] = \mathcal{H}\left[\tilde{X}\right] - \mathcal{H}\left[\tilde{X} \mid \tilde{Z}\right]. \tag{2.12}$$

For the case where the state changes between time steps, as presented in the previous section,

the original definitions of IG and MI are insufficient, since these do not account for the additional uncertainty obtained by changes in the state. Thus, we define the reward more generally as the difference between the entropies of the prior state X_0 and the successor state X_t. We remind that in our formulation the posterior state is augmented, i.e. $X_t = \{X_0, x_{1:t}\}$, thus we will refer to this reward as *augmented* IG, defined at each time t as

$$\begin{aligned} IG_{aug}\left[X_0 \boxplus x_{1:t}; Z_{1:t} = z_{1:t} \mid a_{0:t-1}\right] &\triangleq \mathcal{H}\left[X_0\right] - \mathcal{H}\left[X_t \mid Z_{1:t} = z_{1:t}, a_{0:t-1}\right] \\ &= \mathcal{H}\left[X_0\right] - \mathcal{H}\left[X_0, x_{1:t} \mid Z_{1:t} = z_{1:t}, a_{0:t-1}\right], \end{aligned} \tag{2.13}$$

where $Z_{1:t}$ represents the observation sequence as a random variable, and the symbol $\boxplus$ provides a distinction between the prior state X_0 and the new, augmented subset $x_{1:t}$ of the successor state X_t. This distinction is necessary since $x_{1:t}$ only appears in the posterior entropy term, while X_0 appears in both terms. The corresponding expected reward, *augmented* MI, is then defined at each time t as

$$\begin{aligned} I_{aug}\left[X_0 \boxplus x_{1:t}; Z_{1:t} \mid a_{0:t-1}\right] &\triangleq \underset{Z_{1:t}}{\mathbb{E}}\left[IG_{aug}\left[X_0 \boxplus x_{1:t}; Z_{1:t} = z_{1:t}, a_{0:t-1}\right]\right] \\ &= \mathcal{H}\left[X_0\right] - \mathcal{H}\left[X_t \mid Z_{1:t}, a_{0:t-1}\right]. \end{aligned} \tag{2.14}$$

Note that augmented IG and augmented MI converge to IG and MI, respectively, when there is no change in the state, i.e. $x_{1:t} = \emptyset$ and $X_t = X_0$. Thus, augmented IG and augmented MI can be considered as generalizations of IG and MI.

For the purposes of planning, since negative entropy and augmented IG differ only by the value $\mathcal{H}\left[X_0\right]$, which is constant for each action, using any of these as rewards is equivalent. We further continue to present our approach by choosing augmented IG as the reward, $\rho_t = IG_{aug}\left[X_0 \boxplus x_{1:t}; Z_{1:t} = z_{1:t} \mid a_{0:t-1}\right]$ (for the terminal reward as well), which dictates that the expected reward is augmented MI. The basis of our approach focuses on ion of the augmented MI, thus we will present it over the augmented MI at time t alone. Naturally, it will apply for the entire horizon. For readability, we denote the prior state as $X = X_0$, the state at time t as $X' = X_t = \{X_0, x_{1:t}\}$ and the augmented part of the state as $X_{new} = \{x_{1:t}\}$. Also, the future observations sequence up to time t is denoted as $Z = \{Z_{1:t}\}$ and its space is denoted as $\mathcal{Z} = \{\mathcal{Z}_{1:t}\}$. Lastly, we omit the conditioning over future actions. The augmented MI definition (2.14) then becomes

$$\begin{aligned} I_{aug}\left[X \boxplus X_{new}; Z\right] &\triangleq \mathcal{H}\left[X\right] - \mathcal{H}\left[X' \mid Z\right] \\ &= \mathcal{H}\left[X\right] - \mathcal{H}\left[X, X_{new} \mid Z\right]. \end{aligned} \tag{2.15}$$

The evaluation of any of the presented information-theoretic terms involves integration over the state-space, thus we might face the Curse of Dimensionality when the state is high-dimensional.

2.3 Non-parametric entropy estimation

For any of the presented information-theoretic terms, one possible calculation scheme is to go through calculation of entropy. Having a non-parametric belief over a state $X \in \mathbb{R}^D$, it is usually approximated by a weighted particle set $\{(X^{(i)}, w^{(i)})\}_{i=1}^N$ with normalized weights, where $X^{(i)}$ is the i-th particle, $w^{(i)}$ is the weight of the i-th particle, and N is the number of particles. The Curse of Dimensionality in this case means that in order to have sufficient resolution to represent the belief, the number of particles N needs to be exponential in the dimension D, i.e. $N \propto \alpha^D$ where $\alpha \geq 1$. The entropy in this case is approximated as well using this particle set, and thus will suffer from the Curse of Dimensionality as well, as will further be shown. Several entropy estimators exist in literature, each has its own advantages and disadvantages compared to the others. The most well-known estimators can be found in [1]. One of these estimators, for example, is the re-substitution estimator, for which the entropy is estimated as

$$\hat{\mathcal{H}}[X] \triangleq \sum_{i=1}^{N} w^{(i)} \log \hat{b}\left[X^{(i)}\right], \tag{2.16}$$

where $\hat{b}[X]$ is an approximation of the belief obtained by a probability distribution estimator such as Kernel Density Estimator (KDE). The computational complexity of calculating entropy with the re-substitution estimator with KDE is $O\left(N^2 D\right)$. Another estimator, presented in [23], performs k-d partitioning of the state-space, and thus achieves a complexity of $O\left(N \log N\right)$. There are many more estimators, such as estimators which are based on nearest-neighbor or m_n-spacings, however analyzing these is outside the scope of this work.

Although the complexity of the presented estimators might not seem to have an exponential relation to the state's dimension D, or any relation at all, we remind that N should be exponential with D in order to get an accurate enough estimation, thus the complexity of any estimator is exponential with the dimension D, even if not explicitly.

Chapter 3

Approach

Our work includes multiple contributions when approaching the problem of informative planning with high-dimensional non-Gaussian beliefs. These will be presented in different sections. Our first contribution is an exact mathematical derivation which shows that the dimension of the belief-space can be reduced for the calculation of MI by only exploiting the structure of the problem. We then generalize this conclusion for augmented MI. It thus provides a solution to the curse of dimensionality. We refer to this part of the approach as `involve-MI`. As our second contribution, we derive a method that allows to avoid the explicit reconstruction of the beliefs' surfaces usually required for the augmented MI calculation. We then use this to construct an estimator, which we refer to as `MI-SMC`. In the last section, we show how `involve-MI` and `MI-SMC` can be used together with existing solvers of the informative planning problem, more specifically tree-based solvers.

3.1 Dimensionality reduction for MI calculation

We aim at reducing the complexity of calculating the MI over a high-dimensional state $X \in \mathbb{R}^D$, where $D \gg 1$, distributed with a non-Gaussian belief. As the dimension increases, the number of samples required to get the same accuracy, and thus the complexity, exponentially increases. So, more specifically, the key idea of our approach is to reduce the dimensionality of the problem by exploiting its structure. As in the works [4, 13], it starts by dividing the high-dimensional *prior* state into two subsets, such that $X = \left\{X^{in}, X^{\neg in}\right\}$. We remind that only subsets of the state participate in the probabilistic transition and observation models, $\mathbb{P}_T$ and $\mathbb{P}_Z$, respectively. The same also applies for the sequential counterparts, $\mathcal{P}_T$ and $\mathcal{P}_Z$, defined in (2.4). Thus, $X^{in} \in \mathbb{R}^d$, which we will refer to as the *involved* subset of the state, is defined as a union of all variables in the prior state $X = X_0$ which participate in generating future state transitions and future observations

$$X^{in} = \left[\bigcup_{i=1}^{t} \left[X_{i-1}^{tr} \cup X_i^{obs}\right]\right] \cap X. \tag{3.1}$$

where the intersection with the prior state X is to emphasize that while the subsets X_{i-1}^{tr} and X_i^{obs} might also include new augmented variables X_{new}, which are inherently involved, our definition

of X^{in} is that it is a subset of the prior alone.

To make this definition clearer, we will explain with a specific example. Consider a multi-dimentional prior state of a robot $X = \{x_0, x_1, x_2, l_1, l_2\}$, where x and l resemble poses and landmarks, respectively. A specific action leads to generating a new pose x_3, and observing two landmarks, one of them is a previously observed landmark l_1 and the second is a new landmark l_3. The successive state then becomes $X' = \{x_0, x_1, x_2, x_3, l_1, l_2, l_3\}$. The involved subset of the state, on the other hand, is $X^{in} = \{x_2, l_1\}$. This is since x_2 directly participates in generating x_3 (motion), and l_1 directly participates in generating an observation. Both are states in the prior. Although x_3 and l_3 also directly participate in generating the new factors, these are considered as new variables, thus are not a part of the involved subset, as we have defined it.

The involved subset is usually very small compared to the dimension of the entire state, i.e. $d \ll D$, a fact which is of key importance in our approach. Determining the involved subset is done heuristically, as is naturally done when transitioning and generating future measurements. This idea is also illustrated in a SLAM example in Figure 1.1a. In this example, the involved variables are the prior pose and observed landmarks.

We note that working with the expected reward (MI) rather than the reward itself (IG) entails two underlying assumptions for our approach. The first is that data association for future observation sequences is known and perfect. The second is that all realizations of the future observation sequence share the same structure (per action), or in other words, we assume we have only one future factor graph structure per action. We conjecture these assumptions can be relaxed, yet this is outside the scope of this research.

Phase 1: MI

We will start by analyzing a degenerate problem, for which the state $\tilde{X}$ is static or already transitioned, and we only wish to quantify the expected amount of information the future observations $\tilde{Z}$ will add. We then define $\tilde{X}^{in}$ as just the union of all variables in the state which participate in generating future observations. In this case, we get back to the original MI definition in eq. (2.12).

Theorem 3.1. *Let $\tilde{X}$ be some state and $\tilde{Z}$ be an observation sequence (multivariate random variable) over a subset of this state, $\tilde{X}^{in} \subseteq \tilde{X}$, such that the observation model is $\mathbb{P}(\tilde{Z} \mid \tilde{X}) \equiv \mathbb{P}(\tilde{Z} \mid \tilde{X}^{in})$. Then,*

$$I\left[\tilde{X}; \tilde{Z}\right] = I\left[\tilde{X}^{in}; \tilde{Z}\right]. \tag{3.2}$$

The proof of this theorem, as well as the following theorems, lemmas and propositions, is given in the appendix.

In words, Theorem 3.1 states that the expected information the observations entail over the state X is exactly the expected information they entail over the involved variables X^{in} alone. The result of Theorem 3.1 is also illustrated as an information diagram in Figure 3.1. Note that discarding (marginalizing out) the subset $\tilde{X}^{\neg in} \triangleq \tilde{X} \setminus \tilde{X}^{in}$ (blue circle) does not affect the calculation since the shaded area we are calculating remains the same.

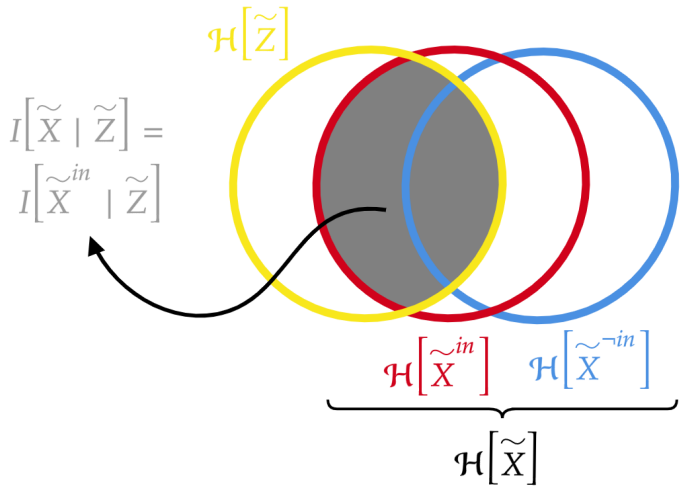


Figure 3.1: Illustration of the static (non-augmented) problem as an information diagram. The area of each circle represents the entropy value of a subset of variables. The mutual areas represent the MI values between these subsets. The key idea, obtained by Theorem 3.1, is that there is no mutual area between $\tilde{X}^{\neg in}$ and $\tilde{Z}$ given $\tilde{X}^{in}$. The gray area is the MI value we are calculating, $I\left[\tilde{X}^{in}; \tilde{Z}\right]$.

While this conclusion might remind the one in [13], which considered only the Gaussian case, it can be viewed as a non-trivial extension to an arbitrary distribution. Specifically, Theorem 3.2 states that *in expectation* the IG and its involved counterpart are exactly equal, no matter how the belief is distributed. Although the most general motivation is to reduce the complexity of calculating the reward itself, going further in this thesis, we will also show that for ρ-POMDP this issue is immaterial.

Remark: We would like to note that [3], for the more specific case of feature selection, has come to an analogous conclusion. In that work, the variables were features. The probability distribution was then approximated as a tree, in which the nodes are the features and the vertices are dependencies between these features. Then, the authors concluded that the MI a node in this tree is predicted to provide to the rest of the nodes in this tree is equivalent to the MI it is predicted to give to its neighbors alone. This conclusion can of course be generalized to any tree. Our problem can be formulated as well as a tree, in which the nodes are $\tilde{X}^{\neg in}$, $\tilde{X}^{in}$ and $\tilde{Z}$, since there is no dependence between $\tilde{X}^{\neg in}$ and $\tilde{Z}$, thus we could have used the above conclusion to prove Theorem 3.1. The works [12] and [25] has also come to an analogous conclusion, yet again the context of these works, active mapping, is very specific. We also note none of these works deal with the more general case of augmentation, about which we discuss next.

Phase 2: augmented MI

Continuing with our initial problem formulation, in which the state is transitioned (augmented, more specifically) and we wish to also take into account the uncertainty this transitioning has introduced, we return to the augmented MI definition in eq. (2.15).

Lemma 3.1.1. *Let X be some prior state, and $X' = \{X, X_{new}\}$ be a successor state, where X_{new} is the augmented subset of the successor state. Let Z be an observation sequence (multivariate*

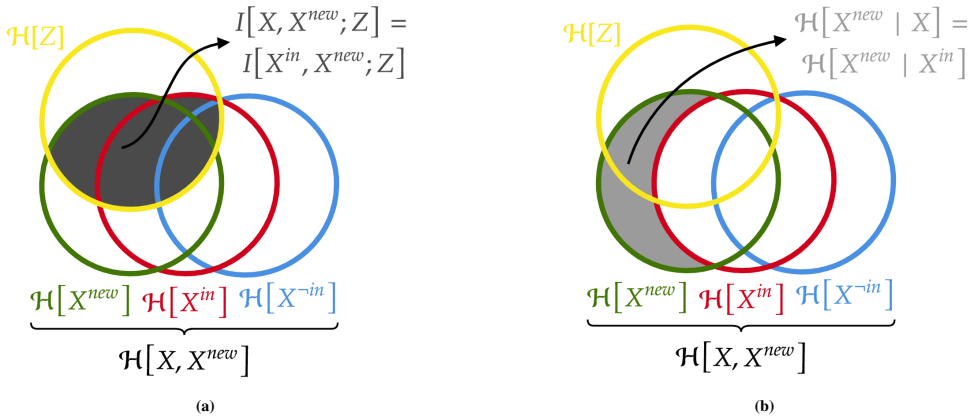


Figure 3.2: Illustration of the augmented problem as an information diagram. The area of each circle represents the entropy value of a subset of variables. The mutual areas represent the MI values between these subsets. The key idea, obtained by Theorem 3.2, is that there is no mutual area between $X^{\neg in}$ and both X_{new}, Z given X^{in}. The dark gray area in (a) is $I\left[X^{in}, X_{new}; Z\right]$, the light gray area in (b) is $\mathcal{H}\left[X_{new} \mid X^{in}\right]$. Subtraction of the dark gray area by the light gray area yields the augmented MI.

random variable) over this successor state. Then, the relation between MI and augmented MI is

$$I_{aug}\left[X \boxplus X_{new}; Z\right] = I\left[X, X_{new}; Z\right] - \mathcal{H}\left[X_{new} \mid X\right]. \tag{3.3}$$

In words, the difference between the two MI variants is the expected uncertainty obtained directly from transitioning the state. Using this Lemma, we then get a generalization to Theorem 3.1.

Theorem 3.2. *Let X be some prior state, and $X' = \{X, X_{new}\}$ be a successor state, where X_{new} is the augmented subset of the successor state. Let Z be an observation sequence (multivariate random variable) over a subset of the successor state, $\left\{X^{in}, X_{new}\right\} \subseteq X'$, such that $\mathbb{P}(Z \mid X, X_{new}) \equiv \mathbb{P}(Z \mid X^{in}, X_{new})$. Then,*

$$I_{aug}\left[X \boxplus X_{new}; Z\right] = I_{aug}\left[X^{in} \boxplus X_{new}; Z\right]. \tag{3.4}$$

Meaning that the same conclusion we have got for the unaugmented case applies also for the case of augmentation. As in the static case, the result of Theorem 3.2 is also illustrated as an information diagram in Figure 3.2. Also here, when discarding (marginalizing out) the subset $X^{\neg in} \triangleq X \setminus X^{in}$ (blue circle) it does not affect the calculation since the shaded areas we are calculating remain the same.

The involved variables depend on the specific action and its observations. Reasoning about the exact involved variables of each action and marginalizing out the uninvolved variables for each action might be costly operations which will eventually make this whole approach worthless. It is easy to show that if we choose a subset X^{in+} which contains the involved variables but might also contain some other uninvolved variables, so that $X^{in} \subseteq X^{in+} \subseteq X$, then Theorem 3.2 can be written more generally as

$$I_{aug}\left[X \boxplus X_{new}; Z\right] = I_{aug}\left[X^{in+} \boxplus X_{new}; Z\right], \tag{3.5}$$

meaning that it is sufficient to choose a subset *containing* the involved rather than *only* the

involved. This subset has larger dimensions, yet it might still be much smaller in dimensions compared to the whole state. This result will be helpful in case we have an easy way to calculate marginalized beliefs of subsets which are not exactly the involved but contain them. Also, if we have n candidate actions, denoting the involved subset for the j-th action as $X^{in(j)}$ and choosing $X^{in+} = X^{in(1)} \cup X^{in(2)} \cup \cdots \cup X^{in(n)}$, will result in a one-time marginalization rather then marginalizing for each action separately. This concept is similar to the one suggested in [13]. We will continue to refer to X^{in} rather than X^{in+} for readability reasons, yet we emphasize that the following is true for X^{in+} as well.

Using the result of Theorem 3.2, we propose an approach which we will refer to as `involve-MI`, and is concluded in Algorithm 3.1. Given the prior belief, the probabilistic models and an action sequence, this algorithm first determines the involved variables using some heuristic. Then, it calculates the marginalized prior belief over the involved variables. And, finally, it calculates the (augmented) MI value using the marginalized prior belief (while propagating future beliefs). This is instead of naively using the entire prior belief. We remind that either or both the complexity and the accuracy of any calculation scheme of the MI depend on the dimension of the entire state, D. The main contribution of our approach is that it cancels this dependence. Instead, the dependence is over the dimension of a smaller subset of the state, d, which improves either or both the complexity and the accuracy. In order to make the overall marginalization process more efficient, we can instead work with X^{in+}, and the algorithm is then slightly changed depending on how it is defined, yet the key idea remains the same.

Algorithm 3.1 `involve-MI`

Input: $b[X], \mathcal{P}_T, \mathcal{P}_Z, a$
Output: $I_{aug}[X \boxplus X_{new}; Z]$
1: $X^{in} \leftarrow \textsc{DetermineInvolved}(b[X], \mathcal{P}_T, \mathcal{P}_Z, a)$
2: $b\left[X^{in}\right] \leftarrow \textsc{Marginalize}(b[X], X^{in})$
3: $I_{aug}\left[X^{in} \boxplus X_{new}; Z\right] \leftarrow \textsc{CalcMI}(b\left[X^{in}\right], \mathcal{P}_T, \mathcal{P}_Z, a)$ ▷ any calculation scheme
4: $I_{aug}[X \boxplus X_{new}; Z] \leftarrow I_{aug}\left[X^{in} \boxplus X_{new}; Z\right]$

Note also this approach is not limited to any specific calculation scheme of the MI. The calculation, for example, might go through its relation to entropy, by calculating the entropy terms or estimating them using entropy estimators, such as the common re-substitution estimator with KDE we have briefly introduced in Section 2.3. We shall use it now to demonstrate the complexity reduction we get from using our approach. It is easy to first see that for the same number of samples, our approach reduces the complexity of estimating the entropy terms from $O\left(N^2 D\right)$ to $O\left(N^2 d\right)$. However, as we have already mentioned, the Curse of Dimensionality suggests that the number of samples required to get the same accuracy is exponential in the dimension. Thus, if we wish to preserve the accuracy, using our approach, only $n \propto \alpha^d$ samples are required, compared to $N \propto \alpha^D$ samples which were required without our approach. Since $d \ll D$, it means that $n \ll N$. This reduces the complexity even further to $O\left(n^2 d\right)$. To generalize and simplify our conclusion, we consider the complexity of any estimator is at least linear in the number of samples, thus the complexity is reduced, at least, from $O\left(\alpha^D\right)$ to $O\left(\alpha^d\right)$, where we

remind that $d \ll D$ and $\alpha \geq 1$.

3.2 Avoiding the reconstruction of future beliefs' surfaces

The estimation scheme of the involved MI might require reconstructing the surfaces of future beliefs, which adds to the estimation error or perhaps entails another level of complexity in the form of new hyperparameters. In this section, we present a theoretical derivation that allows to avoid this reconstruction step for the augmented MI calculation. We then use this derivation to construct an estimator. This estimator, however, as will shortly be explained, can only be used in cases in which the probabilistic models are explicitly given.

Theorem 3.3. *Let X be some prior state, and $X' = \{X, X_{new}\}$ be a successor state, where X_{new} is the augmented subset of the successor state. Let Z be an observation sequence (multivariate random variable) over a subset of the successor state, $\left\{X^{in}, X_{new}\right\} \subseteq X'$, such that $\mathbb{P}(Z \mid X, X_{new}) \equiv \mathbb{P}(Z \mid X^{in}, X_{new})$. Then, the augmented MI can be factorized as*

$$I_{aug}\left[X \boxplus X_{new}; Z\right] = -\mathcal{H}\left[X_{new} \mid X^{in}\right] - \mathcal{H}\left[Z \mid X^{in}, X_{new}\right] + \mathcal{H}\left[Z\right]. \tag{3.6}$$

This result suggests that the augmented MI can be expressed as a superposition of the expected entropies of the sequential transition model, observation model and normalizer. This means that the objective function can be calculated without the need to reconstruct future beliefs' surfaces, which is a key result.

Using the definition of entropy and the notations of the sequential models from (2.4), the first term on the right hand side of eq. (3.6) becomes

$$\mathcal{H}\left[X_{new} \mid X^{in}\right] = -\int_{\mathcal{X}^{in}} b\left[X^{in}\right] \left[\int_{\mathcal{X}_{new}} \mathcal{P}_T \log \mathcal{P}_T dX_{new}\right] dX^{in}. \tag{3.7}$$

Doing the same for the second term of eq. (3.6) yields

$$\mathcal{H}\left[Z \mid X^{in}, X_{new}\right] = -\int_{\mathcal{X}^{in}} b\left[X^{in}\right] \left[\int_{\mathcal{X}_{new}} \mathcal{P}_T \left[\int_{\mathcal{Z}} \mathcal{P}_Z \log \mathcal{P}_Z dZ\right] dX_{new}\right] dX^{in}. \tag{3.8}$$

And, lastly, the third term in eq. (3.6) is defined by

$$\mathcal{H}\left[Z\right] = -\int_{\mathcal{Z}} \eta^{-1} \log \eta^{-1} dZ, \tag{3.9}$$

where we can use the law of total probability to get

$$\eta^{-1} = \int_{\mathcal{X}^{in}} b\left[X^{in}\right] \left[\int_{\mathcal{X}_{new}} \mathcal{P}_T \mathcal{P}_Z dX_{new}\right] dX^{in}. \tag{3.10}$$

Plugging this back into eq. (3.9) and changing integration order yields

$$\mathcal{H}\left[Z\right] = \int_{\mathcal{X}^{in}} b\left[X^{in}\right] \left[\int_{\mathcal{X}_{new}} \mathcal{P}_T \left[\int_{\mathcal{Z}} \mathcal{P}_Z \log \eta^{-1} dZ\right] dX_{new}\right] dX^{in}. \tag{3.11}$$

The above analysis provides an exact formulation for the calculation of the augmented MI over a high-dimensional state that only involves integration over the involved variables, for any of the terms, including η^{-1}. Also, any of these terms can be expressed by only using the prior belief and the probabilistic models, assuming these are explicitly given. Thus, without explicitly requiring the posterior belief, the reconstruction step can be avoided. Plugging these terms back into eq. (3.6) and using sampling, i.e.

$$\begin{aligned} \left(x^{in(i)}, w^{(i)}\right) &\sim b\left[X^{in}\right] \\ x_{new}^{(i,j)} &\sim \mathcal{P}_T\left(X_{new} \mid x^{in(i)}\right) \\ z^{(i,j,k)} &\sim \mathcal{P}_Z\left(Z \mid x^{in(i)}, x_{new}^{(i,j)}\right), \end{aligned} \tag{3.12}$$

the augmented MI is approximated as

$$\begin{aligned} I_{aug}\left[X \boxplus X_{new}; Z\right] \approx & \sum_{i=1}^{n_1} w^{(i)} \left[\frac{1}{n_2} \sum_{j=1}^{n_2} \log \mathcal{P}_T^{(i,j)}\right] \\ & + \sum_{i=1}^{n_1} w^{(i)} \left[\frac{1}{n_2} \sum_{j=1}^{n_2} \left[\frac{1}{n_3} \sum_{k=1}^{n_3} \log \mathcal{P}_Z^{(i,j,k)}\right]\right] \\ & - \sum_{i=1}^{n_1} w^{(i)} \left[\frac{1}{n_2} \sum_{j=1}^{n_2} \left[\frac{1}{n_3} \sum_{k=1}^{n_3} \log \eta^{-1^{(i,j,k)}}\right]\right], \end{aligned} \tag{3.13}$$

where

$$\begin{aligned} \mathcal{P}_T^{(i,j)} &= \mathcal{P}_T\left(x^{in(i)}, x_{new}^{(i,j)}\right) \\ \mathcal{P}_Z^{(i,j,k)} &= \mathcal{P}_Z\left(x^{in(i)}, x_{new}^{(i,j)}, z^{(i,j,k)}\right) \\ \eta^{-1^{(i,j,k)}} &= \eta^{-1}\left(x^{in(i)}, x_{new}^{(i,j)}, z^{(i,j,k)}\right), \end{aligned} \tag{3.14}$$

and the normalizer, for each sampled instance, is then also approximated as

$$\eta^{-1^{(i,j,k)}} \approx \sum_{l=1}^{n_4} w^{(l)} \left[\frac{1}{n_5} \sum_{m=1}^{n_5} \mathcal{P}_Z^{(l,m,k)}\right]. \tag{3.15}$$

As in a particle filter, $\mathcal{P}_Z^{(l,m,k)}$ can be considered an update for the particle's weight. Thus, the approximation of $\eta^{-1^{(i,j,k)}}$ can be viewed as an average of the updated weights.

This estimation scheme is concluded in Algorithm 3.2. Since it uses particles from the prior and propagates them as in Sequential Monte Carlo (SMC) methods, we will refer to it as `MI-SMC`. Note that this formulation allows the algorithm to be an *anytime* algorithm. Given that we have a limited time for each planning step, we can start with a small set of particles and incrementally refine the calculation by increasing the number of particles until the time is up, without the need to calculate from scratch whenever more samples are added. This might not always be possible using other estimation schemes.

The presented estimator might seem at a first glance very similar to the estimator in [2], although `MI-SMC` calculates MI and [2] calculates entropy. Both use samples from the prior

belief, propagates them and calculate entropy/MI without the reconstruction of future belief's surfaces. However, the key difference is the problem formulation these two address. Whereas in our work we treat a smoothing formulation, [2] treats a filtering formulation. A filtering formulation means marginalizing out past states, which eventually makes the entire state involved, thus [2] cannot be scaled to high-dimensional problems compared to our approach.

Also note that although this estimator is formulated by already exploiting the dimensionality reduction, it is not vital. The estimator can get as an input the full prior belief $b[X]$ instead of the involved prior belief $b\left[X^{in}\right]$. The particles would then be high-dimensional, yet since these are used only for the evaluation of the probabilistic models, marginalization would automatically be done in the context of these particles. This attribute makes this estimator closely related to the `involve-MI` approach, without explicitly using it beforehand. However, to avoid the Curse of Dimensionality, it is preferred to maintain and sample from a lower-dimensional belief to begin with, which is exactly the result of using `involve-MI`.

Algorithm 3.2 `MI-SMC`

Input: $b\left[X^{in}\right], \mathcal{P}_T, \mathcal{P}_Z, a$
Output: $I_{aug}\left[X^{in} \boxplus X_{new}; Z\right]$

$sum_1 \leftarrow 0$
$sum_2 \leftarrow 0$
$sum_3 \leftarrow 0$
for $i = 1$ **to** n_1 **do**
  $\left(x^{in(i)}, w^{(i)}\right) \sim b\left[X^{in}\right]$
  for $j = 1$ **to** n_2 **do**
    $x_{new}^{(i,j)} \sim \mathcal{P}_T(X_{new} \mid x^{in(i)}, a)$
    $value_1 \leftarrow w^{(i)} \frac{1}{n_2} \log \mathcal{P}_T^{(i,j)}$
    $sum_1 \leftarrow sum_1 + value_1$
    for $k = 1$ **to** n_3 **do**
      $z^{(i,j,k)} \sim \mathcal{P}_Z(Z \mid x^{in(i)}, x_{new}^{(i,j)})$
      $value_2 \leftarrow w^{(i)} \frac{1}{n_2} \frac{1}{n_3} \log \mathcal{P}_Z^{(i,j,k)}$
      $sum_2 \leftarrow sum_2 + value_2$
      $\eta^{-1} \leftarrow 0$
      for $l = 1$ **to** n_4 **do**
        $\left(x^{in(l)}, w^{(l)}\right) \sim b\left[X^{in}\right]$
        for $m = 1$ **to** n_5 **do**
          $x_{new}^{(l,m)} \sim \mathcal{P}_T(X_{new} \mid x^{in(l)}, a)$
          $value \leftarrow w^{(l)} \frac{1}{n_5} \mathcal{P}_Z^{(l,m,k)}$
          $\eta^{-1} \leftarrow \eta^{-1} + value$
        end for
      end for
      $value_3 \leftarrow w^{(i)} \frac{1}{n_2} \frac{1}{n_3} \log \eta^{-1}$
      $sum_3 \leftarrow sum_3 + value_3$
    end for
  end for
end for
$I_{aug}\left[X^{in} \boxplus X_{new}; Z\right] \leftarrow sum_1 + sum_2 - sum_3$

In terms of complexity, the most expensive step of this approach is the estimation of $\mathcal{H}[Z]$, thus its complexity is the complexity of the entire estimator. Estimating each $\eta^{-1}{}^{(i,j,k)}$ has a complexity of $O(n_4 n_5 d)$. In turn, the complexity of estimating $\mathcal{H}[Z]$ is of $O(n_1 n_2 n_3 n_4 n_5 d)$. Considering that we have a total number of m observations instances, i.e. $n_1 n_2 n_3 = m$, and also that the total number of particles is n, i.e. $n_4 n_5 = n$, the complexity becomes $O(mnd)$. In comparison, the complexity of using a re-substitution estimator with KDE is $O\left(mn^2 d\right)$ when using `involve-MI`, which makes our estimator favorable in terms of complexity. We also again emphasize that our estimator avoids the intermediate step of belief surface reconstruction, and hence we conjecture it is expected to be more accurate. Many other entropy estimators exist in the literature, such as the nearest neighbor estimator, which can be found in [1], and the k-d partitioning estimator, presented in [23]. When estimating the MI value with these estimators, the complexity of both can get to $O\left(mn \log n\right)$, which is comparable to the complexity of `MI-SMC` when reminding again that n should be exponential in the dimension d. We leave further comparison to these additional estimators for future research.

3.3 Applicability to belief trees

In this section we wish to relate the above approaches to the informative planning optimization problem. We remind that although the following analysis considers an open-loop formulation, for which we seek an optimal action sequence, $a_{0:T-1}$, it also applies for a close-loop formulation, for which we seek for a policy, $\pi_{0:T-1}$. The solution to the ρ-POMDP is obtained by maximization of the objective function, given in eq. (2.6), and denoted shortly as $J_0 \triangleq J(b[X_0], a_{0:T-1})$

$$J_0^\star = \max_{a_{0:T-1}} \left\{ \mathop{\mathbb{E}}_{\mathcal{Z}_{1:T}} \left[\sum_{t=0}^{T-1} \rho_t + \rho_T \right] \right\}. \tag{3.16}$$

Formulating it recursively yields the Bellman optimality equation

$$J_t^\star = \max_{a_t} \left\{ \rho_t + \mathop{\mathbb{E}}_{\mathcal{Z}_{t+1}} \left[J_{t+1}^\star \right] \right\}. \tag{3.17}$$

where $J_t \triangleq J(b[X_t], a_{t:T-1})$.

A common solver to this optimization problem is to construct a search over a tree. More specifically, for ρ-POMDP, which is the case of belief-dependent rewards, a belief tree is used. In a belief tree, the beliefs $b[X_t]$ are propagated using instances of future actions and observations, then the rewards ρ_t are calculated, and the action sequence providing the maximum value for the objective function is eventually chosen. Since, in general, the action and observation spaces can be large, in order to be able to solve this optimization problem in reasonable time, it is approximated with a belief tree which propagates only a few sampled instances of future actions and observations. Dealing with continuous such spaces, a belief tree is an approximation of the problem to begin with. The planning literature contains lots of tree-based solvers (e.g. [24], [7]). However, since our analysis so far was done considering an expected reward, augmented MI, it

is not trivial to prove that our approach, `involve-MI`, and our estimator, `MI-SMC`, are able to cope with such solvers. This is the purpose of this section.

We denote the augmented IG, the augmented MI and their involved counterparts shortly as

$$\begin{aligned} IG_0^t &\triangleq IG_{aug}\left[X_0 \boxplus x_{1:t}; Z_{1:t} = z_{1:t} \mid a_{0:t-1}\right] \\ I_0^t &\triangleq I_{aug}\left[X_0 \boxplus x_{1:t}; Z_{1:t} \mid a_{0:t-1}\right] \\ {IG_0^t}^{in} &\triangleq IG_{aug}\left[X_0^{in} \boxplus x_{1:t}; Z_{1:t} = z_{1:t} \mid a_{0:t-1}\right] \\ {I_0^t}^{in} &\triangleq I_{aug}\left[X_0^{in} \boxplus x_{1:t}; Z_{1:t} \mid a_{0:t-1}\right]. \end{aligned}$$

We will also from now omit the term "augmented" while still referring to the more general case of augmentation. For readability, our analysis is done for IG as the only term of the reward, meaning $\rho_t = IG_0^t$, $\forall\, t \in [1, T]$. Yet, the conclusions will also apply when there are additional terms for the reward, state-based terms for example. Eq. (3.16) then becomes

$$J_0^\star = \max_{a_{0:T-1}} \left\{ \sum_{t=0}^{T} \mathop{\mathbb{E}}_{\mathcal{Z}_{1:T}} \left[IG_0^t \right] \right\} = \max_{a_{0:T-1}} \left\{ \sum_{t=0}^{T} I_0^t \right\}. \tag{3.18}$$

Using Theorem 3.2 over this equation yields

$$J_0^\star = \max_{a_{0:T-1}} \left\{ \sum_{t=0}^{T} {I_0^t}^{in} \right\}. \tag{3.19}$$

Theorem 3.4. *Let us define a new reward, $\rho_t^{in} = {IG_0^t}^{in}$. Solving the ρ-POMDP optimization problem with this reward is equivalent to solving it with the original reward, $\rho_t = IG_0^t$, such that*

$$J_t^\star = \max_{a_t} \left\{ \rho_t^{in} + \mathop{\mathbb{E}}_{\mathcal{Z}_{t+1}} \left[J_{t+1}^\star \right] \right\}. \tag{3.20}$$

This, in turn, means that any optimization solver suitable for the original problem, with the reward ρ_t, is also suitable when changing it to ρ_t^{in}. This is a key result, since in general these rewards are only equal in expectation, however it is much more efficient to calculate ρ_t^{in}, as already discussed for the MI case. The belief tree which resembles this new equivalent optimization problem is shown in Figure 3.3a. We can then use the IG definition given in eq. (2.13) and calculate it through the entropy terms. This result is again general, but non-Gaussian distributions might be non-parametric, which then necessitate the usage of entropy estimators, plenty of which exist in the literature, as already discussed.

Proposition 3.3.1. *Naively calculating the values I_0^t yields a degenerate belief tree, in which there are only action nodes, without observation nodes.*

This type of a tree can be seen in Figure 3.3c. We do not here analyze whether this formulation is good or bad compared to the standard formulation nor whether it would even suit a policy formulation or not. We leave it for future research. We cling to the fact that none of the state-of-the-art tree-based solvers work this way, and suggest another approach.

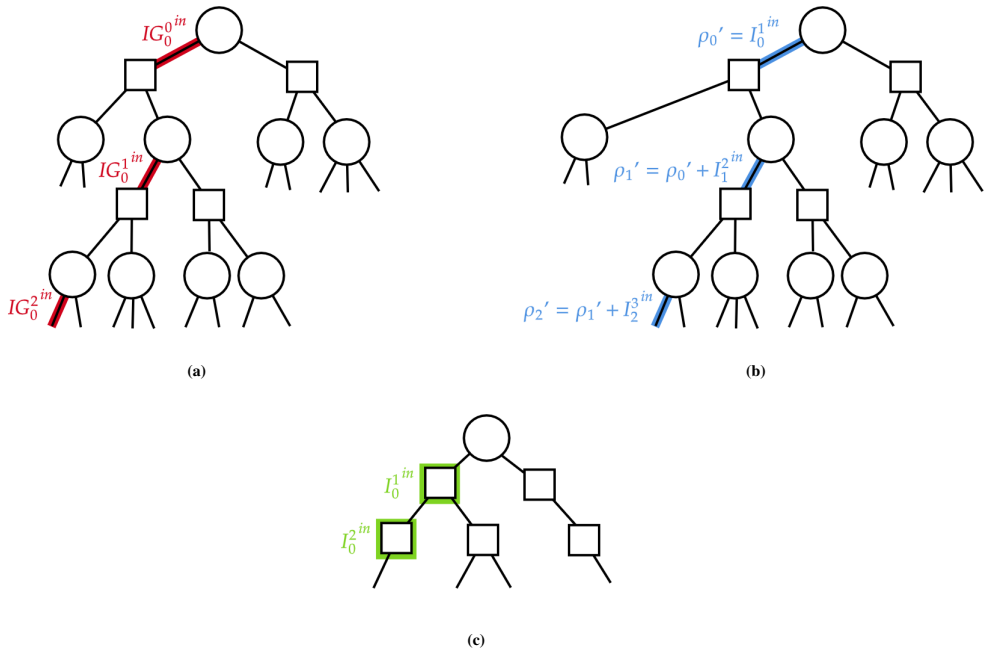


Figure 3.3: Belief trees over the involved variables only, which resemble optimization problems equivalent to the original optimization problem. Circles are observation nodes, squares are action nodes. (a) shows a belief tree where the rewards are sequential involved IGs; (b) shows a belief tree where the rewards are updated incrementally with consecutive involved MIs; (c) shows the resultant degenerate belief tree when trying to directly go through the calculation of the sequential involved MIs. It is degenerate in the sense that there are only action nodes, without observation nodes.

Lemma 3.3.2. *Let* I_0^t *denote a sequential MI between times* 0 *and* t*, and* $I_{i-1}^i = I_{aug}\left[X_{i-1} \boxplus x_i; Z_i \mid h_i^-\right]$ *denote a consecutive MI between times* $i-1$ *and* i*, where* $h_i^- = \{z_{1:i-1}, a_{0:i-1}\}$ *is the history up to time* i*, without the last observation* z_i*. The sequential MI can be decomposed into multiple consecutive MI values, such that*

$$I_0^t \triangleq \sum_{i=1}^{t} \left[\mathop{\mathbb{E}}_{\mathcal{Z}_{1:i-1}} \left[I_{i-1}^i \right] \right]. \tag{3.21}$$

The main result of Theorem 3.2 can be applied on both the sequential and the consecutive MI values by assigning the notations in a slightly different manner, such that the result of Lemma 3.3.2 is transformed into

$$I_0^{t\,in} = \sum_{i=1}^{t} \left[\mathop{\mathbb{E}}_{\mathcal{Z}_{1:i-1}} \left[I_{i-1}^{i\,in} \right] \right], \tag{3.22}$$

where $I_{i-1}^{i\,in} = I_{aug}\left[X_{i-1}^{in} \boxplus x_i; Z_i \mid h_i^-\right]$ is the consecutive MI over the involved subset of the state X_{i-1}.

Theorem 3.5. *Let us define a new reward,* $\rho_t' = \sum_{i=1}^{t+1} I_{i-1}^{i\,in}$*. Solving the* ρ*-POMDP optimization problem with this reward is equivalent to solving it with the original reward,* $\rho_t = IG_0^t$*, such that*

$$J_t^\star = \max_{a_t} \left\{ \rho_t' + \mathop{\mathbb{E}}_{\mathcal{Z}_{t+1}} \left[J_{t+1}^\star \right] \right\}. \tag{3.23}$$

This allows the usage of estimators which directly estimate MI, as our suggested estimator `MI-SMC` does, together with the usage of tree-based solvers of ρ-POMDP. However, we emphasize that instead of sequential MI values, we will calculate consecutive MI values.

We note that ${I^i_{i-1}}^{in} = \mathop{\mathbb{E}}_{\mathcal{Z}_i}\left[IG^i_{i-1}\right]$. This means that the calculation of the MI values is not limited only to the observations that are used for constructing the tree, thus the calculation can be more accurate, which is another added value of this formulation.

And, lastly, we note that $\rho'_t = \sum_{i=1}^{t+1}\left[{I^i_{i-1}}^{in}\right] = \rho'_{t-1} + {I^{t+1}_t}^{in}$. Meaning that for each node, we can calculate the reward based on the previous reward and just update the new information incrementally, without having to calculate the entire reward from scratch. The belief tree which resembles this optimization problem is shown in Figure 3.3b.

The above analysis suggests that the entire tree can be constructed considering only the marginalized beliefs rather than the entire-state beliefs. This, in turn, reduces also the complexity of constructing this tree, since we avoid maintaining and propagating the beliefs over unnecessary states. This is done by determining ahead all the involved variables, together with variables which are required for other reward functions, and marginalizing out the rest of the variables (the formulation of one-time marginalization we have introduced in section 3.1). Care should be taken when using this approach, since marginalizing out a variable which would in retrospect be found to be involved would mean that the tree should be updated from the root. Also note that this approach might prevent the usage of calculation re-use approaches (e.g. [5], [6]) since we only consider a subset of the state for the whole planning process.

Chapter 4

Results

Our approach was tested on an instance of active Simultaneous Localization and Mapping (SLAM) problem, which is a classical choice for high-dimensional problems, since the state contains past trajectory and the map. As in Figure 1.1a, an autonomous drone is flying around, observing landmarks which construct a 2D map. At each time step, it needs to decide where to move next in order to reduce its state uncertainty, i.e. the drone's trajectory and the map. In order to make a decision, the drone estimates the MI value of the different possible actions at that time. The MI value of states which are distributed with Gaussians has an analytical solution, which makes it a perfect choice as a first validation of our approach. Although Gaussian, we emphasize that the different algorithms only get samples as an input, thus work as if it is a purely non-parametric scenario. Using this simulation, we have tested (i) the impact of the dimensionality on choosing an action; and (ii) the impact of increasing the dimensionality on accuracy and timing.

We note that since we chose to deal with Gaussian noise and linear factors, the values of the states and observations do not matter for the MI calculation, but only the noise values and the graph structure. This allowed an implementation simplicity, in which the graph of the problem was created synthetically rather than by creating a specific scenario which includes landmark locations and a trajectory. It means that motion and measurements were *randomly* generated, yet the graph resembled a graph of a SLAM problem. More specifically, the graph was generated using the following process:

1. It was initialized with a prior over the first pose
2. A new pose was then added by applying motion
3. An observation was made such that a new first landmark was created
4. It then continued by randomly adding 100 factors to the graph (motion and observation), using the following approach:
    - At each time a number was sampled uniformly between 0 and 1
    - An observation factor of the last previously observed landmark was generated if the previous factor was motion and the number was between 0.6 and 1

- An observation factor of a new landmark was generated if the previous factor was motion and the number was between 0.3 and 0.6
- A motion factor was generated otherwise

4.1 Impact of the dimensionality on choosing an action

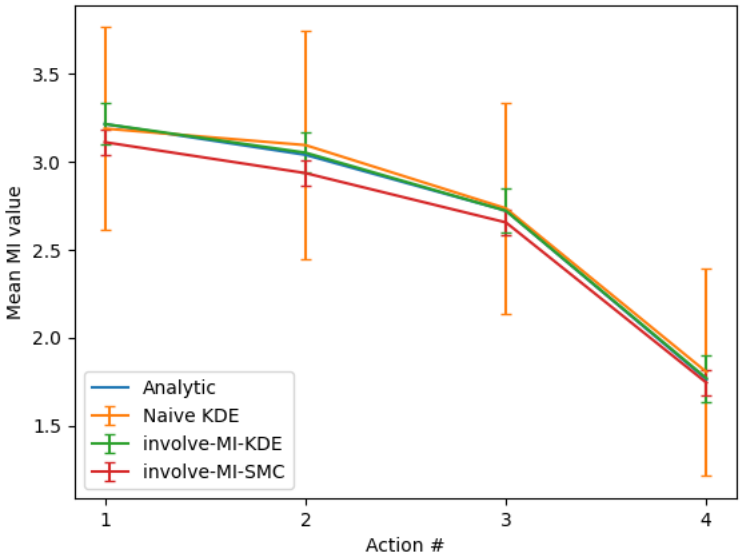
At time t, the drone needs to choose between four different actions, each involving an observation of a different landmark. In this section, we compare between the analytical solution and the estimation results obtained by three different methods, which we will refer to as `Naive KDE`, `involve-MI-KDE` and `involve-MI-SMC`. `Naive KDE` is the naive approach, which uses a re-substitution estimator with KDE over the entire state; `involve-MI-KDE` first uses the `involve-MI` approach and only then the re-substitution estimator with KDE, i.e. the estimation is only over the involved subset of the state; `involve-MI-SMC` is our suggested estimation scheme, `MI-SMC`, where we just emphasize that `involve-MI` is inherent in it. We note that for these specific tests, the methods which included KDE were implemented as if the inference engine is perfect, i.e. the posterior samples were generated from the true posterior, whereas `MI-SMC`, as shown in Algorithm 3.2, uses samples from the prior belief and propagates them in a SMC manner. This gives a slight advantage to the methods with KDE over `involve-MI-SMC`. The simulations were done in Python, where we used Scipy's KDE for the two KDE methods.

Using the graph generation discussed above, the prior state has ~ 150 dimensions, whereas each action currently involves a subset of the state with only 4 dimensions. Each method was provided with 300 particles (for `involve-MI-SMC`, $n_1 = n_4 = 300$, $n_2 = n_3 = n_5 = 1$), and the MI was calculated 100 times to evaluate its standard deviation. The mean values and standard deviations of each method and for each action are shown in Figure 4.1. As expected, `Naive KDE` has big mean errors and big variances. These errors are big enough to make the drone choose an action which isn't optimal. `involve-MI-KDE` and `involve-MI-SMC`, on the other hand, both present pretty similar and better results compared with `Naive KDE`, with the former being slightly better in terms of the mean value, probably due to the prior advantage we have provided it with. We note that also for the case of the `involve-MI` estimators, the drone might choose an action which isn't optimal due to the similarity between the MI values of this specific test, however an increase in the number of samples will solve this (a smaller increase compared with the one needed for the naive approach).

4.2 Impact of increasing the dimensionality on accuracy and timing

We have also tested the impact of increasing state dimensionality on the standard deviation of each estimator, where the action now stays the same. The original motivation is to show that as the dimension grows, an exponential number of samples would be required in order to get the same accuracy for the naive approach. However, it is not feasible with the dimensions we chose for this simulation. Instead, we approach it in a different manner, and show that as the

Figure 4.1: A comparison between the mean MI values of the different actions and calculation methods. The standard deviation of the calculation methods are shown as error bars. While the mean values, even if not close to the analytical values, maintain action consistency, the real problem is the standard deviation. For a specific trial, the actions ordering might be changed due to the overlaps between MI's possible values. The naive approach suffers the most from this problem.



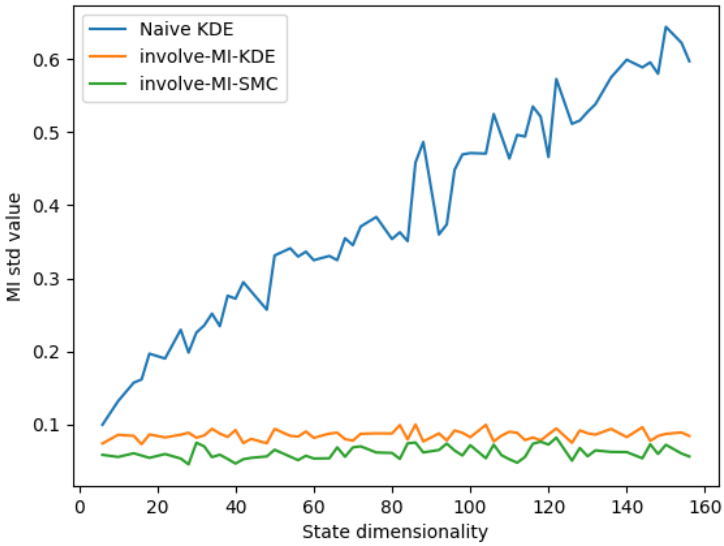


Figure 4.2: Impact of dimensionality on MSE (standard deviation) for the three methods. `Naive KDE`'s error linearly increases with the state's dimension. Using `involve-MI` keeps it roughly constant, independent of the state's dimension.

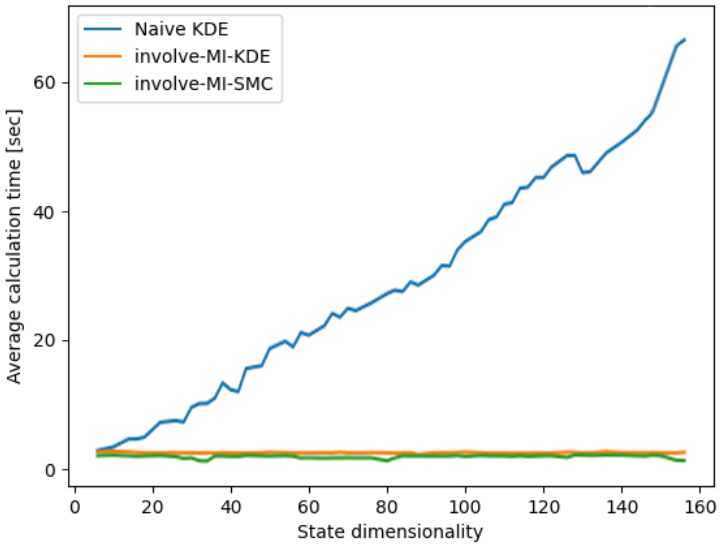


Figure 4.3: Impact of dimensionality on calculation time for the three methods. `Naive KDE`'s calculation time linearly increases with the state's dimension. Using `involve-MI` keeps it roughly constant, independent of the state's dimension.

dimension grows, for a constant number of samples, the accuracy is harmed. To achieve the increasing dimensionality, we exploited the fact that in full SLAM the graph's dimension is increased in time. During the creation of the big synthetic graph generation, at each time just before applying motion, the MI value of the current graph was calculated. Each estimator, again, was provided with 300 particles and the MI was calculated 100 times to evaluate its standard deviation and the average calculation time. The results can be seen in Figures 4.2 and 4.3. Both the standard deviation and calculation time of the MI using `Naive KDE` increase with the state's dimension, an increase which seems linear. On the other hand, for `involve-MI-KDE` and `involve-MI-SMC` both the standard deviation and calculation time are roughly constant. This is thanks to the fact that the involved subset is of the same dimension during this simulation (only one landmark is observed for each action). This demonstrates our main contribution, for which `involve-MI` is better both in terms of accuracy and time complexity, for the same number of samples. The standard deviation and timing of both `involve-MI-KDE` and `involve-MI-SMC` are comparable, with a slight advantage to `involve-MI-SMC`. We remind that `involve-MI-KDE` has a complexity of $O(mn^2d)$ whereas `involve-MI-SMC` has a complexity of $O(mnd)$, which suggests that a better performance should have been obtained for `involve-MI-SMC`. However, it is very likely that the Scipy's KDE implementation is optimized, whereas the implementation of `involve-MI-SMC` is currently very simple and straightforward. We thus conjecture that the timing could be further improved. We also conjecture that providing the KDE methods with a practical inference engine, rather than the current perfect inference it was provided with, and using more complex distributions, will increase the gap in terms of accuracy.

Chapter 5

Conclusions and Future Work

To conclude, we have identified a void in the informative planning subject. For the case of high-dimensional non-parametric beliefs, the complexity becomes too high for solving (not to mention online solutions), thus current state-of-the-art approaches either avoid the high-dimensionality, assume very specific settings or apply rough approximations which impact the accuracy.

In this work, we have filled this void, mainly by reducing the dimensionality of the state for the expected reward's (MI) calculation, while relaxing many of the former assumptions and approximations. We have then shown the same conclusion applies also for the more general and complex augmented scenario, in which the dimensionality is inherently increased at each time step. We refer to the approach of dimensionality reduction for MI calculation as `involve-MI`. Next, we have introduced an estimator, `MI-SMC`, which avoids the reconstruction of future belief's surfaces in case that we have access to the probabilistic models of the problem, which we conjecture might reduce the estimation error and timing, compared to KDE-based approaches, for example. This, in turn, also makes it an anytime algorithm. Lastly, for the completeness of this work, we have related back to the planning optimization problem and shown how both `involve-MI` and `MI-SMC` can be applied in it. We emphasize again that this work contributes specifically to the objective function calculation, and thus can be used as a black box together with many other state-of-the-art algorithms which contribute to other building blocks of the planning task.

Our work was tested on an active SLAM problem, and was compared to a naive approach. The results show that `involve-MI` is superior to the naive approach of going through the entire state, in terms of both accuracy and timing. `MI-SMC` is shown to be only slightly better than the KDE estimator, however we conjecture that code optimization for `MI-SMC` will increase the gap.

5.1 Future Work

In future work, we believe our approach can be extended in the following directions.

1. Working with the more complex focused case, for which we are only interested in

quantifying the uncertainty over a subset of the entire state, as it was done in [13] for Gaussian distributions

2. Given the latest novelties in non-parametric inference, such as the work of [10], we believe it would be very relevant for efficient and accurate non-parametric marginalization
3. Relaxation of the underlying assumptions regarding data-association and single factor graph structure per action
4. Transferring our approach to other uncertainty measures. We conjecture that, at least approximately, the behavior will remain
5. Simplification via removing the correlations between the involved variables, i.e. calculating the uncertainty of each involved variable individually

Chapter 6

Appendix

Proof of Theorem 3.1

Using the chain rule for MI, where the state is arbitrarily partitioned as $\tilde{X} = \{\tilde{X}^A, \tilde{X}^B\}$ yields

$$I\left[\tilde{X};\tilde{Z}\right] = I\left[\tilde{X}^A, \tilde{X}^B;\tilde{Z}\right] = I\left[\tilde{X}^A;\tilde{Z}\right] + I\left[\tilde{X}^B;\tilde{Z} \mid \tilde{X}^A\right]. \tag{6.1}$$

Given that $\tilde{X}^A = \tilde{X}^{in}$, we denote $\tilde{X}^B = \tilde{X}^{\neg in}$, such that the conditional MI becomes $I\left[\tilde{X}^B;\tilde{Z} \mid \tilde{X}^A\right] = I\left[\tilde{X}^{\neg in};\tilde{Z} \mid \tilde{X}^{in}\right]$. By definition, the conditional MI is

$$\begin{aligned} I\left[\tilde{X}^{\neg in};\tilde{Z} \mid \tilde{X}^{in}\right] \triangleq \int_{\tilde{Z}} \int_{\tilde{X}^{in}} \int_{\tilde{X}^{\neg in}} \mathbb{P}(\tilde{Z}, \tilde{X}^{in}, \tilde{X}^{\neg in}) \cdot \\ \cdot \log\left[\frac{\mathbb{P}(\tilde{X}^{\neg in}, \tilde{Z} \mid \tilde{X}^{in})}{\mathbb{P}(\tilde{X}^{\neg in} \mid \tilde{X}^{in})\mathbb{P}(\tilde{Z} \mid \tilde{X}^{in})}\right] d\tilde{X}^{\neg in} d\tilde{X}^{in} d\tilde{Z}. \end{aligned} \tag{6.2}$$

Using the chain rule over the numerator inside the log term, we get

$$\mathbb{P}(\tilde{X}^{\neg in}, \tilde{Z} \mid \tilde{X}^{in}) = \mathbb{P}(\tilde{Z} \mid \tilde{X}^{in}, \tilde{X}^{\neg in})\mathbb{P}(\tilde{X}^{\neg in} \mid \tilde{X}^{in}). \tag{6.3}$$

Since $\tilde{X}^{in}$ is defined such that it contains all the variables involved in generating the observations $\tilde{Z}$, we can state that $\mathbb{P}(\tilde{Z} \mid \tilde{X}^{in}, \tilde{X}^{\neg in}) = \mathbb{P}(\tilde{Z} \mid \tilde{X}^{in})$, so eq. (6.3) becomes

$$\mathbb{P}(\tilde{X}^{\neg in}, \tilde{Z} \mid \tilde{X}^{in}) = \mathbb{P}(\tilde{Z} \mid \tilde{X}^{in})\mathbb{P}(\tilde{X}^{\neg in} \mid \tilde{X}^{in}). \tag{6.4}$$

Plugging this term back into eq. (6.2) yields

$$I\left[\tilde{X}^{\neg in};\tilde{Z} \mid \tilde{X}^{in}\right] \triangleq \int_{\tilde{Z}} \int_{\tilde{X}^{in}} \int_{\tilde{X}^{\neg in}} \mathbb{P}(\tilde{Z}, \tilde{X}^{in}, \tilde{X}^{\neg in}) \log(1)\, d\tilde{X}^{\neg in} d\tilde{X}^{in} d\tilde{Z} = 0. \tag{6.5}$$

Using the above result, eq. (6.1) then transforms into

$$I\left[\tilde{X};\tilde{Z}\right] = I\left[\tilde{X}^{in};\tilde{Z}\right]. \tag{6.6}$$

Proof of Lemma 3.1.1

We remind the augmented MI definition in eq. (2.15) is

$$I_{aug}\left[X \boxplus X_{new}; Z\right] \triangleq \mathcal{H}\left[X\right] - \mathcal{H}\left[X, X_{new} \mid Z\right]. \tag{6.7}$$

Using the following known identities

$$\mathcal{H}\left[X\right] \triangleq \mathcal{H}\left[X, X_{new}\right] - \mathcal{H}\left[X_{new} \mid X\right] \tag{6.8}$$

$$I\left[X, X_{new}; Z\right] \triangleq \mathcal{H}\left[X, X_{new}\right] - \mathcal{H}\left[X, X_{new} \mid Z\right], \tag{6.9}$$

we get the relation between MI and the augmented MI

$$I_{aug}\left[X \boxplus X_{new}; Z\right] = I\left[X, X_{new}; Z\right] - \mathcal{H}\left[X_{new} \mid X\right]. \tag{6.10}$$

Proof of Theorem 3.2

By defining $\tilde{X} = X'$ and $\tilde{X}^{in} = \{X^{in}, X_{new}\}$, we can use the result from Theorem 3.1 to get

$$I\left[X, X_{new}; Z\right] = I\left[X^{in}, X_{new}; Z\right]. \tag{6.11}$$

Looking at the conditional entropy term in the result of Lemma 3.1.1, we can rewrite it as $\mathcal{H}\left[X_{new} \mid X\right] = \mathcal{H}\left[X_{new} \mid X^{in}, X^{\neg in}\right]$. By our definition of X^{in}, X_{new} is conditionally independent of $X^{\neg in}$ given X^{in}, i.e. $\mathbb{P}(X_{new} \mid X^{in}, X^{\neg in}) = \mathbb{P}(X_{new} \mid X^{in})$. Thus, one of the conditional entropy properties states that

$$\mathcal{H}\left[X_{new} \mid X\right] = \mathcal{H}\left[X_{new} \mid X^{in}\right] \tag{6.12}$$

Plugging (6.11) and (6.12) back into the result of Lemma 3.1.1 (eq. (6.10)) we get that

$$I_{aug}\left[X \boxplus X_{new}; Z\right] = I\left[X^{in}, X_{new}; Z\right] - \mathcal{H}\left[X_{new} \mid X^{in}\right]. \tag{6.13}$$

We then observe that by using the result from eq. (6.10), the right hand side in eq. (6.13) is equal to $I_{aug}\left[X^{in} \boxplus X_{new}; Z\right]$, and so we finally conclude that

$$I_{aug}\left[X \boxplus X_{new}; Z\right] = I_{aug}\left[X^{in} \boxplus X_{new}; Z\right]. \tag{6.14}$$

Proof of Theorem 3.3

We begin by using the definition of the augmented MI, given in eq. (2.15), over the involved subset, which is

$$I_{aug}\left[X^{in} \boxplus X_{new}; Z\right] \triangleq \mathcal{H}\left[X^{in}\right] - \mathcal{H}\left[X^{in}, X_{new} \mid Z\right]. \tag{6.15}$$

Using the chain rule for conditional entropy over the second term on the right hand side of eq. (6.15) yields

$$\mathcal{H}\left[X^{in}, X_{new} \mid Z\right] = \mathcal{H}\left[X^{in}, X_{new}, Z\right] - \mathcal{H}\left[Z\right]. \tag{6.16}$$

Using the same principle twice again eventually yields

$$\mathcal{H}\left[X^{in}, X_{new} \mid Z\right] = \mathcal{H}\left[X^{in}\right] + \mathcal{H}\left[X_{new} \mid X^{in}\right] + \mathcal{H}\left[Z \mid X^{in}, X_{new}\right] - \mathcal{H}\left[Z\right]. \tag{6.17}$$

Plugging back into eq. (6.15), we observe that the term $\mathcal{H}\left[X^{in}\right]$ is canceled out. Then, by using the result of Theorem 3.2, the augmented MI term over the high-dimensional state finally becomes

$$I_{aug}\left[X \boxplus X_{new}; Z\right] = -\mathcal{H}\left[X_{new} \mid X^{in}\right] - \mathcal{H}\left[Z \mid X^{in}, X_{new}\right] + \mathcal{H}\left[Z\right]. \tag{6.18}$$

Proof of Theorem 3.4

We remind eq. (3.19) is

$$J_0^\star = \max_{a_{0:T-1}} \left\{ \sum_{t=0}^{T} {I_0^t}^{in} \right\}. \tag{6.19}$$

The involved MI is by definition an expectation over the involved IG

$${I_0^t}^{in} \triangleq \mathop{\mathbb{E}}_{\mathcal{Z}_{1:T}} \left[{IG_0^t}^{in} \right]. \tag{6.20}$$

Plugging this back into eq. (6.19) yields

$$J_0^\star = \max_{a_{0:T-1}} \left\{ \sum_{t=0}^{T} \left[\mathop{\mathbb{E}}_{\mathcal{Z}_{1:t}} \left[{IG_0^t}^{in} \right] \right] \right\}. \tag{6.21}$$

Due to commutativity, we can again switch between the order of expectation and summation, which yields

$$J_0^\star = \max_{a_{0:T-1}} \left\{ \mathop{\mathbb{E}}_{\mathcal{Z}_{1:T}} \left[\sum_{t=0}^{T} \left[{IG_0^t}^{in} \right] \right] \right\}. \tag{6.22}$$

We then separate the first action a_0 from the rest of the actions $a_{1:T-1}$. We also observe that Z_1 is not a function of $a_{1:T-1}$ and that ${IG_0^0}^{in}$ is not a function of both $a_{1:T-1}$ and Z_1. This yields

$$J_0^\star = \max_{a_0} \left\{ {IG_0^0}^{in} + \mathop{\mathbb{E}}_{\mathcal{Z}_1} \left[\max_{a_{1:T-1}} \left\{ \mathop{\mathbb{E}}_{\mathcal{Z}_{2:T}} \left[\sum_{t=1}^{T} {IG_0^t}^{in} \right] \right\} \right] \right\}. \tag{6.23}$$

We then observe that the term inside the expectation over Z_1 is equal to $J_1^\star$, which yields the following recursive form

$$J_0^\star = \max_{a_0} \left\{ {IG_0^0}^{in} + \mathop{\mathbb{E}}_{\mathcal{Z}_1} \left[J_1^\star \right] \right\}, \tag{6.24}$$

and, in general, for each $t \in [1, T-1]$

$$J_t^\star = \max_{a_t} \left\{ {IG_0^t}^{in} + \mathop{\mathbb{E}}_{\mathcal{Z}_{t+1}} \left[J_{t+1}^\star \right] \right\}. \tag{6.25}$$

We observe that this is the Bellman optimality equation with a new reward, $\rho_t^{in} \triangleq {IG_0^t}^{in}$. This eventually means that Solving the ρ-POMDP optimization problem with this reward is equivalent to solving it with the original reward we have started with, $\rho_t = IG_0^t$.

Proof of Proposition 3.3.1

We remind eq. (3.19) is

$$J_0^\star = \max_{a_{0:T-1}} \left\{ \sum_{t=0}^{T} {I_0^t}^{in} \right\}. \tag{6.26}$$

Since by definition ${I_0^0}^{in} = 0$, we can start the summation from $t = 1$

$$J_0^\star = \max_{a_{0:T-1}} \left\{ \sum_{t=1}^{T} {I_0^t}^{in} \right\}. \tag{6.27}$$

We then separate the first action a_0 from the rest of the actions $a_{1:T-1}$. We also observe that ${I_0^1}^{in}$ is not a function of $a_{1:T-1}$. This yields

$$J_0^\star = \max_{a_0} \left\{ {I_0^1}^{in} + \max_{a_{1:T-1}} \left\{ \sum_{t=2}^{T-1} {I_0^t}^{in} \right\} \right\}. \tag{6.28}$$

We then observe that the term $\max_{a_{1:T-1}} \left\{ \sum_{t=2}^{T-1} {I_0^t}^{in} \right\}$ is equal to $J_1^\star$, which yields the following recursive form

$$J_0^\star = \max_{a_0} \left\{ {I_0^1}^{in} + J_1^\star \right\}, \tag{6.29}$$

and, in general, $\forall t \in [1, T-1]$

$$J_t^\star = \max_{a_t} \left\{ {I_0^{t+1}}^{in} + J_{t+1}^\star \right\}. \tag{6.30}$$

We observe that this recursive form is slightly different than the Bellman optimality equation. The Bellman optimality equation, as can be seen in eq. (3.17), includes also expectation over future observations, while in this formulation it is omitted (more specifically, it is considered at the level of calculating the values I_0^t). This, in turn, means that a corresponding tree will lack observation nodes, thus it will be a degenerate belief-tree. We note again that this is the result of naively going through direct calculations of the values I_0^t.

Proof of Lemma 3.3.2

We remind that the sequential augmented MI is defined as

$$I_0^t \triangleq I_{aug}\left[X_0 \boxplus x_{1:t}; Z_{1:t} \mid a_{0:t-1}\right] \triangleq \mathop{\mathbb{E}}_{\mathcal{Z}_{1:t}}\Big[IG_{aug}\left[X_0 \boxplus x_{1:t}; Z_{1:t} = z_{1:t} \mid a_{0:t-1}\right]\Big]. \tag{6.31}$$

Detaching the observations $Z_{l+1:t}$, where $0 < l < t$, and expressing the augmented IG with entropies, we get

$$I_{aug}\left[X_0 \boxplus x_{1:t}; Z_{1:t} \mid a_{0:t-1}\right] = \mathop{\mathbb{E}}_{\mathcal{Z}_{1:l}}\Bigg[\mathop{\mathbb{E}}_{\mathcal{Z}_{l+1:t}}\Big[\mathcal{H}\left[X_0\right] - \mathcal{H}\left[X_t \mid h_t\right]\Big]\Bigg]. \tag{6.32}$$

Adding and subtracting the term $\mathcal{H}\left[X_l \mid h_l\right]$, it becomes

$$\begin{aligned} I_{aug}\left[X_0 \boxplus x_{1:t}; Z_{1:t} \mid a_{0:t-1}\right] = \mathop{\mathbb{E}}_{\mathcal{Z}_{1:l}}\Bigg[\mathop{\mathbb{E}}_{\mathcal{Z}_{l+1:t}}\Big[\big\{\mathcal{H}\left[X_0\right] - \mathcal{H}\left[X_l \mid h_l\right]\big\} + \\ + \big\{\mathcal{H}\left[X_l \mid h_l\right] - \mathcal{H}\left[X_t \mid h_t\right]\big\}\Big]\Bigg]. \end{aligned} \tag{6.33}$$

Observing that both new differences are augmented IGs as well, and that the first difference is not a function of the last observation, we get

$$\begin{aligned} I_{aug}\left[X_0 \boxplus x_{1:t}; Z_{1:t} \mid a_{0:t-1}\right] = \mathop{\mathbb{E}}_{\mathcal{Z}_{1:l}}\Bigg[IG_{aug}\left[X_0 \boxplus x_{1:l}; z_{1:l} \mid a_{0:l-1}\right] + \\ + \mathop{\mathbb{E}}_{\mathcal{Z}_{l+1:t}}\Big[IG_{aug}\left[X_l \boxplus x_{l+1:t}; z_{l+1:t} \mid a_{0:t-1}, z_{1:l}\right]\Big]\Bigg], \end{aligned} \tag{6.34}$$

The expectation over the augmented IG is the augmented MI, and so we get the following recursive form

$$\begin{aligned} I_{aug}\left[X_0 \boxplus x_{1:t}; Z_{1:t} \mid a_{0:t-1}\right] = & I_{aug}\left[X_0 \boxplus x_{1:l}; Z_{1:l} \mid a_{0:l-1}\right] + \\ & + \mathop{\mathbb{E}}_{\mathcal{Z}_{1:l}}\Big[I_{aug}\left[X_l \boxplus x_{l+1:t}; Z_{l+1:t} \mid a_{0:t-1}, z_{1:l}\right]\Big]. \end{aligned} \tag{6.35}$$

The specific case of choosing $l = t - 1$ yields

$$\begin{aligned} I_{aug}\left[X_0 \boxplus x_{1:t}; Z_{1:t} \mid a_{0:t-1}\right] = & I_{aug}\left[X_0 \boxplus x_{1:t-1}; Z_{1:t-1} \mid a_{0:t-2}\right] + \\ & + \mathop{\mathbb{E}}_{\mathcal{Z}_{1:t-1}}\Big[I_{aug}\left[X_{t-1} \boxplus x_t; Z_t \mid h_t^-\right]\Big], \end{aligned} \tag{6.36}$$

where $h_t^- = \{z_{1:t-1}, a_{0:t-1}\}$ is the history up to time t, without the last observation z_t. Opening the recursive form of the sequential augmented MI in eq. (6.36) yields

$$\begin{aligned} I_{aug}\left[X_0 \boxplus x_{1:t}; Z_{1:t} \mid a_{0:t-1}\right] = I_{aug}\left[X_0 \boxplus x_1; Z_1 \mid h_1^-\right] + \\ + \mathop{\mathbb{E}}_{\mathcal{Z}_1}\Big[I_{aug}\left[X_1 \boxplus x_2; Z_2 \mid h_2^-\right]\Big] + \cdots + \mathop{\mathbb{E}}_{\mathcal{Z}_{1:t-1}}\Big[I_{aug}\left[X_{t-1} \boxplus x_t; Z_t \mid h_t^-\right]\Big], \end{aligned} \tag{6.37}$$

which can more compactly be written as

$$I_{aug}\left[X_0 \boxplus x_{1:t}; Z_{1:t} \mid a_{0:t-1}\right] = \sum_{i=1}^{t} \left[\mathop{\mathbb{E}}_{\mathcal{Z}_{1:i-1}} \left[I_{aug}\left[X_{i-1} \boxplus x_i; Z_i \mid h_i^-\right] \right] \right]. \tag{6.38}$$

Returning to the short notations, we finally get

$$I_0^t = \sum_{i=1}^{t} \left[\mathop{\mathbb{E}}_{\mathcal{Z}_{1:i-1}} \left[I_{i-1}^i \right] \right]. \tag{6.39}$$

Proof of Theorem 3.5

We remind eq. (3.19) is

$$J_0^\star = \max_{a_{0:T-1}} \left\{ \sum_{t=0}^{T} {I_0^t}^{in} \right\}. \tag{6.40}$$

Since by definition ${I_0^0}^{in} = 0$, we can start the summation from $t = 1$

$$J_0^\star = \max_{a_{0:T-1}} \left\{ \sum_{t=1}^{T} {I_0^t}^{in} \right\}. \tag{6.41}$$

Plugging the result from eq. (3.22) into the above yields

$$J_0^\star = \max_{a_{0:T-1}} \left\{ \sum_{t=1}^{T} \left[\sum_{i=1}^{t} \left[\mathop{\mathbb{E}}_{\mathcal{Z}_{1:i-1}} \left[{I_{i-1}^i}^{in} \right] \right] \right] \right\}. \tag{6.42}$$

Due to commutativity, we can switch between the order of expectation and summation, which yields

$$J_0^\star = \max_{a_{0:T-1}} \left\{ \mathop{\mathbb{E}}_{\mathcal{Z}_{1:T-1}} \left[\sum_{t=1}^{T} \left[\sum_{i=1}^{t} \left[{I_{i-1}^i}^{in} \right] \right] \right] \right\}. \tag{6.43}$$

We then denote $\rho'_{t-1} \triangleq \sum_{i=1}^{t} \left[{I_{i-1}^i}^{in} \right]$, and get

$$J_0^\star = \max_{a_{0:T-1}} \left\{ \mathop{\mathbb{E}}_{\mathcal{Z}_{1:T-1}} \left[\sum_{t=1}^{T} \rho'_{t-1} \right] \right\} = \max_{a_{0:T-1}} \left\{ \mathop{\mathbb{E}}_{\mathcal{Z}_{1:T-1}} \left[\sum_{t=0}^{T-1} \rho'_t \right] \right\}. \tag{6.44}$$

We then separate the first action a_0 from the rest of the actions $a_{1:T-1}$. We also observe that Z_1 is not a function of $a_{1:T-1}$, and that $\rho'_0 = {I_0^1}^{in}$ is not a function of both $a_{1:T-1}$ and Z_1 (since I_0^1 is already an expectation over Z_1). This yields

$$J_0^\star = \max_{a_0} \left\{ \rho'_0 + \mathop{\mathbb{E}}_{\mathcal{Z}_1} \left[\max_{a_{1:T-1}} \left\{ \mathop{\mathbb{E}}_{\mathcal{Z}_{2:T-1}} \left[\sum_{t=1}^{T-1} \rho'_t \right] \right\} \right] \right\}. \tag{6.45}$$

We then observe that the term inside the expectation over Z_1 is equal to $J_1^\star$, which yields the following recursive form

$$J_0^\star = \max_{a_0} \left\{ \rho_0' + \mathop{\mathbb{E}}_{\mathcal{Z}_1} \left[J_1^\star \right] \right\}, \tag{6.46}$$

and, in general, $\forall t \in [1, T-1]$

$$J_t^\star = \max_{a_t} \left\{ \rho_t' + \mathop{\mathbb{E}}_{\mathcal{Z}_{t+1}} \left[J_{t+1}^\star \right] \right\}. \tag{6.47}$$

We observe that this is the Bellman optimality equation with the new reward, ρ_t'. This eventually means that solving the ρ-POMDP optimization problem with this reward is equivalent to solving it with the original reward we have started with, $\rho_t = IG_0^t$. We note another slight difference between the formulations, for which the latter formulation does not include a terminal reward.

Bibliography

[1] J. Beirlant, E. J. Dudewicz, L. Györfi, E. C. Van der Meulen, et al. Nonparametric entropy estimation: An overview. *International Journal of Mathematical and Statistical Sciences*, 6(1):17–39, 1997.

[2] Y. Boers, H. Driessen, A. Bagchi, and P. Mandal. Particle filter based entropy. In *2010 13th International Conference on Information Fusion*, pages 1–8, 2010.

[3] Margarita Chli. *Applying Information Theory to Efficient SLAM*. PhD thesis, Imperial College London, 2009.

[4] K. Elimelech and V. Indelman. Simplified decision making in the belief space using belief sparsification. *Intl. J. of Robotics Research*, 2021. Accepted.

[5] 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.

[6] E. I. Farhi and V. Indelman. ix-bsp: Belief space planning through incremental expectation. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, May 2019.

[7] Johannes Fischer and Omer Sahin Tas. Information particle filter tree: An online algorithm for pomdps with belief-based rewards on continuous domains. In *Intl. Conf. on Machine Learning (ICML)*, Vienna, Austria, 2020.

[8] Neha P Garg, David Hsu, and Wee Sun Lee. Despot-α: Online pomdp planning with large state and observation spaces. In *Robotics: Science and Systems (RSS)*, 2019.

[9] Marcus Hoerger, Hanna Kurniawati, and Alberto Elfes. Multilevel monte-carlo for solving pomdps online. In *Proc. International Symposium on Robotics Research (ISRR)*, 2019.

[10] Q. Huang, C. Pu, D. Fourie, K. Khosoussi, J. P. How, and J. J. Leonard. Nf-isam: Incremental smoothing and mapping via normalizing flows. *arXiv preprint arXiv:2105.05045*, 2021.

[11] V. Indelman. No correlations involved: Decision making under uncertainty in a conservative sparse information space. *IEEE Robotics and Automation Letters (RA-L)*, 1(1):407–414, 2016.

[12] Brian J Julian, Sertac Karaman, and Daniela Rus. On mutual information-based control of range sensing robots for mapping applications. *The International Journal of Robotics Research*, 33(10):1375–1392, 2014.

[13] 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*, 36(10):1088–1130, August 2017.

[14] H. Kurniawati, , and V. Yadav. An online pomdp solver for uncertainty planning in dynamic environment. In *Robotics Research*, pages 611–629. Springer, 2016.

[15] H. Kurniawati, D. Hsu, and W. S. Lee. SARSOP: Efficient point-based POMDP planning by approximating optimally reachable belief spaces., 2008.

[16] Michael H. Lim, Claire Tomlin, and Zachary N. Sunberg. Sparse tree search optimality guarantees in pomdps with continuous observation spaces. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, Jul 2020.

[17] S. Pathak, A. Thomas, and V. Indelman. A unified framework for data association aware robust belief space planning and perception. *Intl. J. of Robotics Research*, 32(2-3):287–315, 2018.

[18] R. Platt, L. Kaelbling, T. Lozano-Perez, and R. Tedrake. Efficient planning in non-gaussian belief spaces and its application to robot grasping. In *Proc. of the Intl. Symp. of Robotics Research (ISRR)*, 2011.

[19] Allison Ryan. Information-theoretic tracking control based on particle filter estimate. In *AIAA Guidance, Navigation and Control Conference*, pages 1–15, 2008.

[20] David Silver and Joel Veness. Monte-carlo planning in large pomdps. In *Advances in Neural Information Processing Systems (NIPS)*, pages 2164–2172, 2010.

[21] A. Somani, N. Ye, D. Hsu, and W. S. Lee. Despot: Online pomdp planning with regularization. *Advances in neural information processing systems*, 26:1772–1780, 2013.

[22] 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.

[23] Dan Stowell and Mark D Plumbley. Fast multidimensional entropy estimation by k-d partitioning. *IEEE Signal Processing Letters*, 16(6):537–540, 2009.

[24] Zachary N Sunberg and Mykel J Kochenderfer. Online algorithms for pomdps with continuous state, action, and observation spaces. In *Twenty-Eighth International Conference on Automated Planning and Scheduling*, 2018.

[25] Zhengdong Zhang, Theia Henderson, Sertac Karaman, and Vivienne Sze. Fsmi: Fast computation of shannon mutual information for information-theoretic mapping. *The International Journal of Robotics Research*, 39(9):1155–1177, 2020.

יתר על כן, גישה נאיבית עשויה להצריך שיערוך מפורש של כל משטחי ההתפלגויות העתידיות האפשריות, מה שמוסיף עוד נדבך למורכבות החישובית ומהווה מקור נוסף לשגיאות שיערוך. אנו מציגים במחקר זה תרומה נוספת – שיטה המאפשרת לפסוח על שלב זה, תחת הנחה שקיימת גישה למודלי הרעש של הבעיה – ובכך מאפשרים את הקטנת מורכבות החישוב ושגיאות השיערוך. לשיטה זו אנו קוראים 'MI-SMC'.

לטובת שלמות השיטות המדוברות, אנו מסיימים בניתוחן על בעיית התכנון האינפורמטיבי, ומציגים כיצד ניתן להשתמש בשיטות אלה ביחד עם פותרני עצים קיימים. האתגר נובע מכך שפותרני עצים עובדים ברמת פונקציית התגמול (IG), בעוד שהשיטות שלנו נכונות ברמת הציפייה על פונקצייה זו (MI) בלבד. מצאנו שניתן להמיר את פונקצית התגמול המקורית לשתי פוקנציות תגמול אחרות, אשר מאפשרות לנצל את השיטות היעילות המוצגות במחקר זה, ואינן משנות את הבעיה המקורית.

אנו מסיימים בבחינת השיטות שפיתחנו על בעיית 'לוקליזציה ומיפוי בו-זמנית' (SLAM) אל מול הגישות הנאיביות בהן נעשה שימוש היום. אנו מראים שהשיטות שלנו, עבור כמות דגימות זהה, בעלות יתרון משמעותי הן מבחינת זמן החישוב והן מבחינת דיוק הפתרון.

תקציר

תכנון תחת אי-ודאות הוא בעל חשיבות רבה עבור יישומים רבים. העולם שלנו הוא סטוכסטי באופיו, לכן, בכל מקרה שבו אנחנו נדרשים לבצע פעולות של הסקה או תכנון לעתיד, יש לקחת בחשבון את הסטוכסטיות הזו, אחרת עלולות להתרחש קטסטרופות.

לקיחת הסטוכסטיות בחשבון יכולה להיעשות במספר רמות. ברמה הפשוטה ביותר, הסטוכסטיות נלקחת בחשבון רק בשלב ההסקה, בעוד שבשלב התכנון מפשטים את הבעיה ומתייחסים אליה כאילו היא דטרמיניסטית לחלוטין. באחת הרמות הבאות, הסטוכסטיות נלקחת בחשבון גם בשלב התכנון, אולם הדרישה היא להגיע למצב סופי מסוים בלבד, מבלי לדאוג לאי-הוודאות על מצב זה. באחת הרמות הגבוהות ביותר, אנו נדרשים גם לכמת את אי-הוודאות על המצב. לדוגמה, אם המטרה היא להגיע למצב יעד כלשהו עם הסתברות מינימלית כלשהי או אם המטרה היא לצבור כמה שיותר מידע על מצב נתון כלשהו (להקטין את אי-הוודאות עליו). בעיה כזו, שבה יש לבחור רצף של פעולות על בסיס אי-וודאותו של מצב כלשהו, מוכרת כתכנון אינפורמטיבי. משימה שניתן לסווגה כתכנון אינפורמטיבי, לדוגמה, היא חיפוש והצלה, כאשר בשלב ה"חיפוש" עשויה להתבצע סקירה של שטחים לא מוכרים.

ישנם כמה מדדים לכימות אי-הוודאות, אשר יכולים לשמש כפונקציות מחיר או פונקציות תגמול בבעיית התכנון האינפורמטיבי. 'אנטרופיה' היא בין המדדים המוכרים ביותר. עם זאת, עבור יישומים רבים אין הכרח בערך המוחלט של אי-הוודאות, אלא רק בכמה היא תפחת עקב הפעולה שתבוצע, או במילים אחרות – כמה מידע נוכל להשיג מפעולה זו. למדד זה נקרא 'אינפורמציה מושגת' (information gain – IG). מכיוון שהמצב לא תמיד אובזרבבילי, נעשה שימוש במדידות, אך כאשר מסיקים על העתיד, גם המדידות עצמן אינן ידועות, ולכן גם אי-הוודאות עליהן צריכה להילקח בחשבון. כדי להתמודד עם זה, אנו מציגים מדד נוסף, 'אינפורמציה הדדית' (mutual information – MI), שהוא למעשה הציפייה שלנו על 'האינפורמציה המושגת' בהינתן שאיננו יודעים מה יהיו המדידות, ואיתו אנו עובדים במחקר זה. נציין כי חלקים מהספרות מתייחסים ל-MI ו-IG כאותו מדד, אך כמו בספרות של בינה מלאכותית, אנו מבדילים בין שני אלו.

מטבע הדברים, כאשר גדל מימד המצב, כך גם גדלה המורכבות של חישוב מדדי אי-הוודאות. במקרה הזה היחס הוא אקספוננציאלי. אנו מכנים מצב זה כ'קללת המימדיות' (Curse of Dimensionality). למיטב ידיעתנו, לא קיימות גישות העוסקות בסוגיה הכואבת והקריטית של חישוב מדדים אלו עבור בעיות תכנון אינפורמטיבי כלליות, שהמצב בהן הוא רב-מימדי והפילוג ההסתברותי שלו אינו בהכרח גאוסי.

אנו מציעים שיטה חדשה למילוי החלל הזה. על-ידי ניצול מבנה הגלום בבעיות תכנון רבות, ניתן לשבור את היחס בין מימד הבעיה למורכבות חישוב מדדי אי-הוודאות. הרעיון המרכזי של השיטה שלנו הוא בזמן התכנון "להשליך" את המשתנים הלא מעורבים עבור חישוב 'האינפורמציה ההדדית' (בין המצב למדידות), על אף שגם משתנים אלו מושפעים. אנו מוכיחים שעל ידי כך שאיננו מתחשבים במשתנים הללו בזמן חישוב מדד זה, כלומר שמימד המצב קטן, מורכבות החישוב מצמטמצמת גם היא באופן דרמטי, תוך שהפתרון עדיין מדויק. אנו קוראים לשיטה זו 'involve-MI'.

המחקר בוצע בהנחייתו של פרופסור חבר ואדים אינדלמן, במסגרת התוכנית הבין-יחידתית למערכות אוטונומיות ורובוטיקה.

חלק מן התוצאות בחיבור זה פורסמו או הוגשו כמאמרים מאת המחבר ושותפיו למחקר בכנסים ובכתבי-עת במהלך תקופת מחקר המאסטר של המחבר, אשר גרסאותיהם העדכניות ביותר הינן:

G. Rotman and V. Indelman. Towards involve-MI: Informative planning with high-dimensional non-parametric beliefs. In *61st Israel Annual Conference on Aerospace Sciences (IACAS),* 2022.

תודות

תקופת התואר השני הייתה מהמאתגרות ביותר בחיי, הן מבחינה מקצועית והן מבחינה אישית, בייחוד עם התפרצות מגיפת הקורונה אשר היוותה קושי לכולם ללא יוצא מן הכלל. ולמרות זאת, לא יכולתי להיות בסביבה תומכת יותר.

ראשית, אני רוצה להודות למנחה שלי, פרופ׳ חבר ואדים אינדלמן. קצרה היריעה מלתאר את ההערכה שלי אליו. השילוב בין הידע, המקצועיות, המעורבות, התשוקה, האדיבות והסבלנות שלו הופכות אותו למנחה המושלם מנקודת הראייה שלי.

כמו כן, אני רוצה להודות לצוות האדמיניסטרטיבי של התוכנית הבין-יחידתית למערכות אוטונומיות ורובוטיקה, אשר נתנו תמיכה ואוזן קשבת לסטודנטים.

ולסיום, אני רוצה להודות למשפחתי, ובפרט לאשתי נויה, אשר איתה התחתנתי והבאתי 2 ילדים מתוקים במהלך התקופה הזאת. היא האמינה בי מההתחלה ונתנה לי את ברכת הדרך לצאת למסע הזה, אשר היה מאתגר לה לא פחות מאשר לי.

אני מודה לטכניון על התמיכה הכספית הנדיבה במשך השתלמותי.

תכנון אינפורמטיבי יעיל עם פילוגים הסתברותיים רב-מימדיים לא-גאוסיים תוך ניצול מבנה

חיבור על מחקר

לשם מילוי חלקי של הדרישות לקבלת התואר
מגיסטר למדעים במערכות אוטונומיות ורובוטיקה

גלעד רוטמן

הוגש לסנט הטכניון – מכון טכנולוגי לישראל
אדר א׳ ה׳תשפ״ב חיפה פברואר 2022

תכנון אינפורמטיבי יעיל עם פילוגים הסתברותיים רב-מימדיים לא-גאוסיים תוך ניצול מבנה

גלעד רוטמן