On Decision Making and Planning in the *Conservative* Information Space - Is the Concept Applicable to Active SLAM?

Vadim Indelman*

Decision making under uncertainty is essential in numerous problems involving robot autonomy and artificial intelligence, including informative planning, active sensing, sensor deployment and active simultaneous localization and mapping (SLAM). These problems are typically addressed using information-theoretic decision making and belief-space planning approaches, where one looks for the most informative action(s) according to a given objective function.

In many cases, however, decision making should be performed over a high-dimensional information space that captures spatial and, perhaps also temporal relations between the random variables. For example, in active SLAM, the state represents the robot trajectories and the observed environment, while in the context of environment modeling it represents a multidimensional field which can be dynamically changing over time. In these cases, evaluating the impact of a *single* candidate action (or sequence of actions), for example considering entropy as an information-theoretic objective function, is by itself an expensive operation. Considering an n dimensional state and without making any assumptions on the sparsity of the information (covariance) matrix, the computational complexity is $O\left(n^3\right)$. The complexity can be significantly reduced in the case the involved matrices are sparse. In the limit, *if* the information matrix *was* diagonal, i.e. no correlations between states, the complexity of evaluating a single candidate action would be $O\left(n\right)$.

We suggest a new paradigm [1] for decision making under uncertainty: The basic idea is to resort to conservative information fusion techniques for information-theoretic decision making. Such a concept is motivated by the fact that conservative information fusion techniques, e.g. covariance intersection [2], allow the correlations between different states to be unknown, while yielding consistent (but sub-optimal) state estimates. What happens if similar ideas are used for information-theoretic decision making and belief space planning?

In general, regardless of the sparsity pattern of the original information space, one can always resort to a *sparse* conservative information space, where some of the correlations are *appropriately* dropped, and perform decision making considering that information space. Clearly, this would result in a greatly reduced computational complexity. The question, however, is whether the performance is sacrificed with respect to using the original information space, i.e. whether the same decisions would be made in both cases. The best one could hope for, is that such a concept would allow evaluating a candidate action *without accounting for any* correlations while guaranteeing no sacrifice in performance.

In recent research we began exploring this concept, and obtained promising results [1]. In particular, it was proved that decision making with unary observation models possibly involving arbitrary states can be performed considering a conservative pdf that decouples the states (i.e. diagonal information matrix) and without any sacrifice in performance compared to using the original information space. As a result, the complexity of evaluating a single candidate action is reduced from $O\left(n^3\right)$ to $O\left(n\right)$, regardless if the original information space is sparse or not.

This motivates further research on the applicability of this concept to additional scenarios and problems of interest, including belief space planning and active SLAM. In this short paper we summarize current status and outline key questions within this new paradigm.

Basic Concept

Consider an information matrix I, representing the second moment of the pdf over a multidimensional state X, and let I_c denote some conservative information matrix with respect to I. Suppose we have access to different sensor modalities and are to decide what parts of the state space to measure. Specifically, consider different measurement likelihoods $p_i\left(z_i|x_i\right)$ involving arbitrary states $x_i \in X$, with the corresponding observation models of the form

$$z_i = h_i\left(x_i\right) + v_i \quad , \quad v_i \sim N\left(0, \Sigma_{vi}\right), \tag{1}$$

where both h_i and Σ_{vi} can vary. These models can represent, for example, different types of sensors with varying accuracy.

The concept is formulated in the following theorem considering, for now, a one-dimensional state $X \in \mathbb{R}$ and some two candidate actions a and b.

Theorem 1:

$$\left|I^{a+}\right| \leq \left|I^{b+}\right| \text{ iff } \left|I_c^{a+}\right| \leq \left|I_c^{b+}\right|. \tag{2}$$

Here, I^{a+} and I_c^{a+} represent the a posteriori information matrices after performing action a

$$I^{a+} = I + A^T \Sigma_v^{-1} A \quad , \quad I_c^{a+} = I_c + A^T \Sigma_v^{-1} A, \tag{3}$$

where $A \doteq \frac{\partial h_a}{\partial x}$ is the measurement Jacobian corresponding to the measurement model (1) of action a. The information matrices I^{b+} and I_c^{b+} are similarly defined for action b with a measurement Jacobian $B \doteq \frac{\partial h_b}{\partial x}$.

Theorem 1 states the impact of any two candidate actions on entropy, or other information-theoretic utility measures involving the determinant operator, has the same trend regardless if it is calculated based on the original pdf $p\left(x\right)$ or based on the conservative pdf $p_c\left(x\right)$. The proof of this theorem for a single dimensional state is trivial [1].

*Department of Aerospace Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel.

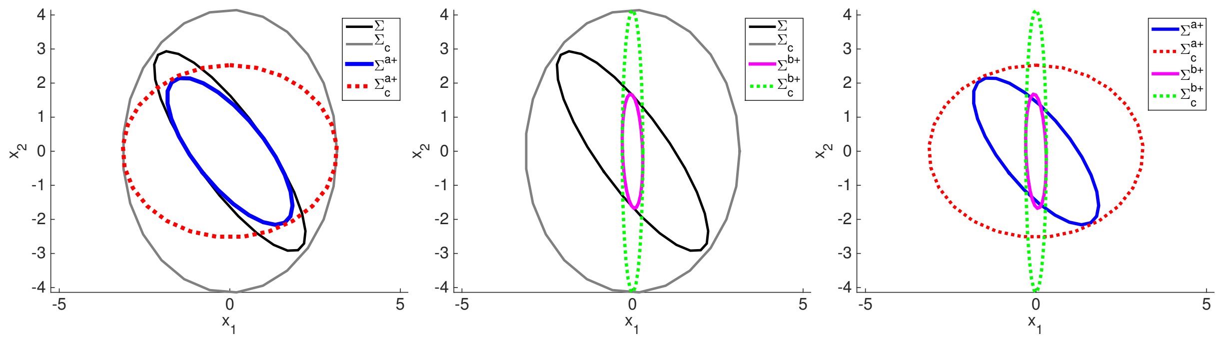

Fig. 1. Concept illustration for a two dimensional case, $X = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$, considering two candidate actions corresponding to measuring either the first or the second state. Left and middle figures show an a priori covariance Σ and the conservative covariance Σ_c with no correlations between states, and the a posteriori covariances corresponding to two actions a and b that are equivalent to choosing the appropriate measurement model (1). Right figure shows a posteriori covariances for each action using original and conservative a priori covariances: $\Sigma^{a+}, \Sigma^{b+}, \Sigma_c^{a+}, \Sigma_c^{b+}$. Theorem 2 states that if the area enclosed by Σ^{a+} is smaller (larger) than the area enclosed by Σ^{b+}, then necessarily, the same relation holds also between Σ_c^{a+} and Σ_c^{b+}.

High-Dimensional State Spaces

Theorem 1 provides a mechanism to calculate the impact of candidate actions using *some* conservative information matrix. While formulated for a single dimensional state, the question is whether it is valid also for high-dimensional states $X \in \mathbb{R}^n$. Intuitively, if Theorem 1 is indeed valid, then one may construct a *sparse* conservative approximation of the information matrix I, which will yield significant reduction in computational complexity (e.g. when calculating the determinant), especially in problems that involve many states and a dense I. Importantly, Theorem 1 states this will be obtained for free, without any sacrifice in performance.

In recent work [1], we took this idea to the extreme and considered a conservative approximation that decouples the states in X, leading to a *diagonal* matrix I_c. This decoupled conservative approximation can be obtained as

$$p_c\left(X\right) \doteq \eta \prod_i p^{w_i}\left(x_i\right), \tag{4}$$

with x_i being the ith component in X, η a normalization constant, and w_i are weights such that $\sum_i w_i = 1$. Here, $p\left(x_i\right)$ is the marginal distribution over x_i: $p\left(x_i\right) = \int_{\neg x_i} p\left(X\right)$. For Gaussian distributions, this is equivalent to scaling the covariance of each component of Σ by a factor of $1/w$, and can be considered as a special instance of covariance intersection [2]. As a result, the (square root) information matrix I_c and the covariance Σ_c, representing the second moment of the pdf (4), become *diagonal* matrices.

We can then formulate the following theorem, which is illustrated in Figure 1 and proved in [1].

Theorem 2: Theorem 1 holds for high dimensional state spaces for unary measurement models (1) and the decoupled conservative approximation (4).

According to Theorem 2, decision making involving unary observation models (1) can be performed over the decoupled conservative information space, represented by the diagonal information matrix I_c. This framework admits decision making without accounting for any correlations between the states while guaranteeing to *yield exactly the same decisions* as would be obtained by using the original information space. As a result, the computational complexity of evaluating a single candidate action is reduced from $O\left(n^3\right)$ to $O\left(n\right)$ without sacrificing performance and without making any assumptions on the sparsity pattern of the original information matrix.

Applicability to Active SLAM

Is the concept of decision making over a conservative information space applicable also to active single- and multi-robot SLAM and belief space planning problems? To answer this question, a number of aspects should be further considered, including:

- More general observation models: e.g. pairwise observation models, where each observation typically involves robot pose and either past pose or a landmark. We note, it was demonstrated in [1] that Theorem 1 in conjunction with the decoupled belief (4) is valid in certain cases also for pairwise observation models.
- Active SLAM can often be considered as active *focused* inference [3], where only a subset of states is of interest. Is the concept valid also for active focused inference?
- Multi-step planning horizon: while the proposed concept can be applied in a greedy fashion, an open question is whether it is also valid to non-myopic planning.

We note the decoupled conservative belief (4) is only one possible (and extreme) alternative and additional conservative approximations exist that could be better suited for the above aspects. Furthermore, since the general problem is NP-complete, it might make sense to consider a tradeoff between suboptimal performance and computational complexity. Ongoing research aims to address these aspects.

References

[1] V. Indelman. Towards information-theoretic decision making in a conservative information space. In *American Control Conference*, July 2015.

[2] S.J. Julier and J.K. Uhlmann. A non-divergent estimation algorithm in the presence of unknown correlations. In *American Control Conference*, pages 2369–73, 1997.

[3] D. Levine and J. P. How. Sensor selection in high-dimensional gaussian trees with nuisances. In *Advances in Neural Information Processing Systems (NIPS)*, pages 2211–2219, 2013.