Efficient Belief Space Planning via Factor-Graph Propagation Action Tree and Incremental Covariance Update
Supplementary Material

Dmitry Kopitkov and Vadim Indelman

This document provides supplementary material to the paper [1]. Therefore, it should not be considered a self-contained document, but instead regarded as an appendix of [1].

This document is organized as follows: Appendix A provides a derivation for updating covariance entries after applying action a, that does not depends on state dimension n; Appendix B provides a proof for information gain of action a to be equal to sum of information gains of a's sub-actions.

1 Appendix A: Covariance Update

Problem definition: Given a priori belief $b[X_-]$ with prior information matrix $\Lambda_- \in \mathbb{R}^{n\times n}$, the candidate action a with an *increment* $I(a) \doteq \{F_{new}, X_{new}\}$ is applied and the posterior belief $b[X_+]$ is obtained. Consider the set of variables $Y \subset X_+$ whose marginal covariance $\Sigma_+^{M,Y}$ from $b[X_+]$ we would like to calculate. We are looking for function $\Sigma_+^{M,Y} = f(\Sigma_-^{M,W})$ where $\Sigma_-^{M,W}$ is prior marginal covariance of set $W \doteq \{Y_{old}, X^I\}$. The variables set Y_{old} is intersection $X_- \cap Y$, or in other words the old variables inside Y. The $X^I \subset X_-$ is the set of *involved* variables in action's newly introduced factors F_{new}. Moreover, we are looking for function $f()$ whose complexity does not depend on state dimension n.

First, let's separate all actions into different categories accordingly to action's *increment* properties.

In case X_{new} is empty, we will call such action as *not-augmented*. Such action does not change the state vector ($X_- \equiv X_+$) and only introduces new information through new factors. The information matrix of such action can be propagated through:

$$\Lambda_+ = \Lambda_- + A^T \cdot A \tag{A1}$$

D. Kopitkov is with the Technion Autonomous Systems Program (TASP), Technion - Israel Institute of Technology, Haifa 32000, Israel, dimkak@tx.technion.ac.il . V. Indelman is with the Department of Aerospace Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel, vadim.indelman@technion.ac.il. This work was supported by the Israel Science Foundation.

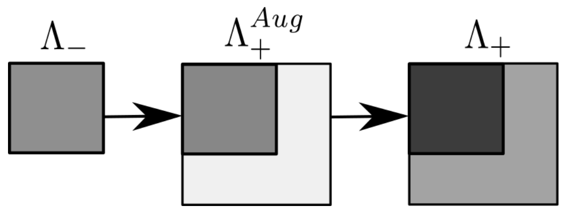

Figure 1: Illustration of Λ_+'s construction for a given candidate action in Augmented BSP case. First, Λ_+^{Aug} is created by adding zero rows and columns representing the new state variables. Then, the new information of belief is computed through $\Lambda_+ = \Lambda_+^{Aug} + A^T A$.

where matrix $A \in \mathbb{R}^{m \times n}$ is a noise-weighted Jacobian of newly introduced factors F_{new} w.r.t. state variables X_+, and A's height m is dimension of all new factors within F_{new}.

In case X_{new} is not empty, we will call such action as *rectangular*. Such action is augmenting the state vector to be $X_+ = \{X_-, X_{new}\}$ and also introduces new information through the new factors. Here the information matrix can be propagated through:

$$\Lambda_+ = \Lambda_+^{Aug} + A^T \cdot A \tag{A2}$$

where $\Lambda_+^{Aug} \in \mathbb{R}^{N \times N}$ is constructed by first augmenting the prior information matrix Λ_- with zero rows and columns representing the new state variables X_{new}, as illustrated in Figure 1; $N = |X_+| = n + |X_{new}|$ is posterior state dimension; A here will be $m \times N$ matrix.

Finally, for case when X_{new} is not empty and total dimension of new factors m is equal to number of newly introduced variables $|X_{new}|$, we will call such action as *squared*. Clearly, the *squared* action is specific case of *rectangular* action which for instance can represent the new robot poses of candidate trajectory and the new motion model factors. The reason for this specific action case to be dealt with in special way is due to the fact that its $f()$ function is much simpler than $f()$ function of more general *rectangular* action, as we will show below. Thus, when $m = |X_{new}|$ it would be advisable to use function $f()$ of *squared* action.

Next, in Section 1.1 we will present the function $f()$ separately for each one of the *not-augmented*, *rectangular* and *squared* cases. Although function $f()$ has an intricate form (especially in *rectangular* case), all matrix terms involved in it have dimensions m, $|X_{new}|$ or $|X^I|$ and overall calculation of posterior $\Sigma_+^{M,Y}$ does not depend on state dimension n. In Section 1.2 we will show how this function was derived in each one of the cases.

1.1 Final Expression for $f()$

1.1.1 *Not-augmented* Action

For the *not-augmented* action the posterior marginal covariance $\Sigma_+^{M,Y}$ can be calculated as:

$$\Sigma_+^{M,Y} = \Sigma_-^{M,Y} - B \cdot C^{-1} \cdot B^T, \quad B \triangleq \Sigma_-^C \cdot ({}^I\!A)^T, \quad C \triangleq I_m + {}^I\!A \cdot \Sigma_-^I \cdot ({}^I\!A)^T \tag{A3}$$

where Σ_{-}^{I}, Σ_{-}^{Y} and Σ_{-}^{C} are parts of prior marginal covariance $\Sigma_{-}^{M,W}$ partitioned through $W = \{Y, X^I\}$:

$$\Sigma_{-}^{M,W} = \begin{pmatrix} \Sigma_{-}^{Y} & \Sigma_{-}^{C} \\ (\Sigma_{-}^{C})^T & \Sigma_{-}^{I} \end{pmatrix} \tag{A4}$$

and where ${}^{I}\!A$ consists of A's columns belonging to *involved* variables X^I.

1.1.2 *Rectangular* Action

For the *rectangular* action the prior marginal covariance $\Sigma_{-}^{M,W}$ and the posterior marginal covariance $\Sigma_{+}^{M,Y}$ have forms:

$$\Sigma_{-}^{M,W} = \begin{pmatrix} \Sigma_{-}^{Y_{old}} & \Sigma_{-}^{C} \\ (\Sigma_{-}^{C})^T & \Sigma_{-}^{I} \end{pmatrix} \tag{A5}$$

$$\Sigma_{+}^{M,Y} = \begin{pmatrix} \Sigma_{+}^{M,Y_{old}} & \Sigma_{+}^{(Y_{old},Y_{new})} \\ (\Sigma_{+}^{(Y_{old},Y_{new})})^T & \Sigma_{+}^{M,Y_{new}} \end{pmatrix} \tag{A6}$$

where we partition Y variables into two subsets $Y_{old} \doteq X_{-} \cap Y$ and $Y_{new} \doteq X_{new} \cap Y$, and where $W = \{Y_{old}, X^I\}$.

Using parts of $\Sigma_{-}^{M,W}$ we can calculate parts of $\Sigma_{+}^{M,Y}$ as:

$$\Sigma_{+}^{M,Y_{old}} = \Sigma_{-}^{M,Y_{old}} - B \cdot G^{-1} \cdot B^T \tag{A7}$$

$$\Sigma_{+}^{M,Y_{new}} = P^{(Y_{new},:)} \tag{A8}$$

$$C \triangleq I_m + {}^{I}\!A \cdot \Sigma_{-}^{I} \cdot ({}^{I}\!A)^T \tag{A9}$$

$$P \triangleq [(A_{new}^T \cdot C^{-1} \cdot A_{new})^{-1}]^{(:,Y_{new})} \tag{A10}$$

$$F \triangleq (A_{new}^T \cdot A_{new})^{-1} \tag{A11}$$

$$K \triangleq I_m - A_{new} \cdot F \cdot A_{new}^T \tag{A12}$$

$$K_1 \triangleq K \cdot {}^{I}\!A \tag{A13}$$

$$B \triangleq \Sigma_{-}^{C} \cdot K_1^T \tag{A14}$$

$$G \triangleq I_m + K_1 \cdot \Sigma_{-}^{I} \cdot K_1^T \tag{A15}$$

where ${}^{I}\!A$ consists of A's columns belonging to *involved* variables X^I, and A_{new} - columns belonging to newly introduced variables X_{new}.

For $\Sigma_{+}^{(Y_{old},Y_{new})}$ from Eq. (A6) there are two methods to calculate it.

Method 1:

$$\Sigma_{+}^{(Y_{old},Y_{new})} = \Sigma_{-}^{C} \cdot ({}^{I}\!A)^T \cdot [C^{-1} \cdot {}^{I}\!A \cdot \Sigma_{-}^{I} \cdot ({}^{I}\!A)^T - I_m] \cdot A_{new} \cdot P \tag{A16}$$

Method 2:

$$\Sigma_{+}^{(Y_{old},Y_{new})} = \Sigma_{-}^{C} \cdot [K_1^T \cdot G^{-1} \cdot K_1 \cdot \Sigma_{-}^{I} - I_k] \cdot ({}^{I}\!A)^T \cdot A_{new} \cdot F^{(:,Y_{new})} \tag{A17}$$

Empirically we have found that method 2 is the fastest option.

1.1.3 *Squared* Action

For the *squared* action the prior marginal covariance $\Sigma_{-}^{M,W}$ and the posterior marginal covariance $\Sigma_{+}^{M,Y}$ have forms:

$$\Sigma_{-}^{M,W} = \begin{pmatrix} \Sigma_{-}^{Y_{old}} & \Sigma_{-}^{C} \\ (\Sigma_{-}^{C})^T & \Sigma_{-}^{I} \end{pmatrix} \tag{A18}$$

$$\Sigma_{+}^{M,Y} = \begin{pmatrix} \Sigma_{+}^{M,Y_{old}} & \Sigma_{+}^{(Y_{old},Y_{new})} \\ (\Sigma_{+}^{(Y_{old},Y_{new})})^T & \Sigma_{+}^{M,Y_{new}} \end{pmatrix} \tag{A19}$$

where we partition Y variables into two subsets $Y_{old} \doteq X_{-} \cap Y$ and $Y_{new} \doteq X_{new} \cap Y$, and where $W = \{Y_{old}, X^I\}$.

Using parts of $\Sigma_{-}^{M,W}$ we can calculate parts of $\Sigma_{+}^{M,Y}$ as:

$$\Sigma_{+}^{M,Y_{old}} = \Sigma_{-}^{M,Y_{old}} \tag{A20}$$

$$\Sigma_{+}^{M,Y_{new}} = A_{iv} \cdot C \cdot A_{iv}^T \tag{A21}$$

$$\Sigma_{+}^{(Y_{old},Y_{new})} = -\Sigma_{-}^{C} \cdot ({}^{I}\!A)^T \cdot (A_{iv})^T \tag{A22}$$

$$A_{iv} \triangleq [A_{new}^{-1}]^{(Y_{new},:)} \tag{A23}$$

$$C \triangleq I_m + {}^{I}\!A \cdot \Sigma_{-}^{I} \cdot ({}^{I}\!A)^T \tag{A24}$$

where ${}^{I}\!A$ consists of A's columns belonging to *involved* variables X^I, and A_{new} - columns belonging to newly introduced variables X_{new}. We can see that in case of *squared* action the covariances of old variables X_{-} do not change.

1.2 Derivation of $f()$

1.2.1 *Not-augmented* Action

The variables set W in this case is $\{Y_{old}, X^I\} = \{Y, X^I\}$. Define next prior marginal covariance matrices: $\Sigma_{-}^{I} \equiv \Sigma_{-}^{M,X^I}$, $\Sigma_{-}^{Y} \equiv \Sigma_{-}^{M,Y}$. Also notate prior cross covariance between Y and X^I as Σ_{-}^{C}. The $\Sigma_{-}^{M,W}$ then will have following form:

$$\Sigma_{-}^{M,W} = \begin{pmatrix} \Sigma_{-}^{Y} & \Sigma_{-}^{C} \\ (\Sigma_{-}^{C})^T & \Sigma_{-}^{I} \end{pmatrix} \tag{A25}$$

Additionally, let's separate prior (old) state variables X_{-} into *involved* X^I (in new factors F_{new}) and *not involved* $X^{\neg I}$. Similarly let's partition the Jacobian matrix A into:

$$A = \begin{pmatrix} {}^{\neg I}\!A & {}^{I}\!A \end{pmatrix} = \begin{pmatrix} 0 & {}^{I}\!A \end{pmatrix} \tag{A26}$$

where ${}^{\neg I}\!A$ contains noise-weighted Jacobians w.r.t. $X^{\neg I}$, and ${}^{I}\!A$ - w.r.t. X^I. From its definition we can conclude that ${}^{\neg I}\!A$ contains only zeros.

Next, using the Woodbury matrix identity and Eq. (A1) the posterior covariance matrix is:

$$\begin{aligned}\Sigma_+ = (\Lambda_+)^{-1} = (\Lambda_- + A^T \cdot A)^{-1} = \\ = \Sigma_- - \Sigma_- \cdot A^T \cdot [I_m + A \cdot \Sigma_- \cdot A^T]^{-1} \cdot A \cdot \Sigma_- = \\ = \Sigma_- - \Sigma_- \cdot A^T \cdot [I_m + {}^{I}\!A \cdot \Sigma_-^I \cdot ({}^{I}\!A)^T]^{-1} \cdot A \cdot \Sigma_- \end{aligned} \tag{A27}$$

where $A \cdot \Sigma_- \cdot A^T = {}^{I}\!A \cdot \Sigma_-^I \cdot ({}^{I}\!A)^T$ because of A's sparsity structure.

Then Σ_+ can be calculated as:

$$\Sigma_+ = \Sigma_- - \Sigma_- \cdot A^T \cdot C^{-1} \cdot A \cdot \Sigma_- \tag{A28}$$

$$C = I_m + {}^{I}\!A \cdot \Sigma_-^I \cdot ({}^{I}\!A)^T \tag{A29}$$

Then $\Sigma_+^{M,Y}$ can be calculated by retrieving from Σ_+ rows and columns that belong to variables Y:

$$\begin{aligned}\Sigma_+^{M,Y} = \Sigma_-^{M,Y} - \Sigma_-^{(Y,:)} \cdot A^T \cdot C^{-1} \cdot A \cdot \Sigma_-^{(:,Y)} = \\ = \Sigma_-^{M,Y} - [A \cdot \Sigma_-^{(:,Y)}]^T \cdot C^{-1} \cdot [A \cdot \Sigma_-^{(:,Y)}] = \\ = \Sigma_-^{M,Y} - [{}^{I}\!A \cdot (\Sigma_-^C)^T]^T \cdot C^{-1} \cdot [{}^{I}\!A \cdot (\Sigma_-^C)^T] = \\ = \Sigma_-^{M,Y} - [\Sigma_-^C \cdot ({}^{I}\!A)^T] \cdot C^{-1} \cdot [\Sigma_-^C \cdot ({}^{I}\!A)^T]^T \end{aligned} \tag{A30}$$

where using the Matlab syntax the used matrices are $\Sigma_-^{(Y,:)} = \Sigma_-(Y,:)$ and $\Sigma_-^{(:,Y)} = \Sigma_-(:,Y)$. ■

Note that the variables inside information matrices do not have to be ordered in any particular way, and that the provided above proof is correct for any ordering whatsoever.

1.2.2 *Rectangular* Action

In this case we can partition variables set Y into two subsets $Y_{old} \doteq X_- \cap Y$ and $Y_{new} \doteq X_{new} \cap Y$, or in other words into old and new state variables. The posterior marginal covariance matrix $\Sigma_+^{M,Y}$ will have then next form:

$$\Sigma_+^{M,Y} = \begin{pmatrix} \Sigma_+^{M,Y_{old}} & \Sigma_+^{(Y_{old},Y_{new})} \\ (\Sigma_+^{(Y_{old},Y_{new})})^T & \Sigma_+^{M,Y_{new}} \end{pmatrix} \tag{A31}$$

and we are looking for efficient way to calculate matrices $\Sigma_+^{M,Y_{old}}$, $\Sigma_+^{M,Y_{new}}$ and $\Sigma_+^{(Y_{old},Y_{new})}$.

The variables set W in this case is $\{Y_{old}, X^I\}$. Define next prior marginal covariance matrices: $\Sigma_-^I \equiv \Sigma_-^{M,X^I}$, $\Sigma_-^{Y_{old}} \equiv \Sigma_-^{M,Y_{old}}$. Also notate prior cross covariance between Y_{old} and X^I as Σ_-^C. The $\Sigma_-^{M,W}$ then will have following form:

$$\Sigma_-^{M,W} = \begin{pmatrix} \Sigma_-^{Y_{old}} & \Sigma_-^C \\ (\Sigma_-^C)^T & \Sigma_-^I \end{pmatrix} \tag{A32}$$

Additionally, let's separate prior (old) state variables X_- into *involved* X^I (in new factors F_{new}) and *not involved* $X^{\neg I}$. The posterior state vector is then $X_+ = \{X^I, X^{\neg I}, X_{new}\}$. Similarly let's partition the Jacobian matrix A into:

$$A = \begin{pmatrix} A_{old} & A_{new} \end{pmatrix}, \quad A_{old} = \begin{pmatrix} {}^{\neg I}\!A & {}^{I}\!A \end{pmatrix} = \begin{pmatrix} 0 & {}^{I}\!A \end{pmatrix} \tag{A33}$$

where A_{old} contains noise-weighted Jacobians w.r.t. old variables X_-, A_{new} - w.r.t. new variables X_{new}, ${}^{\neg I}\!A$ - w.r.t. $X^{\neg I}$, and ${}^{I}\!A$ - w.r.t. X^I. From its definition we can conclude that ${}^{\neg I}\!A$ contains only zeros.

Following Eq. (A2) the posterior information matrix can be partitioned using separation $X_+ = \{X_-, X_{new}\}$ as:

$$\Lambda_- = \begin{pmatrix} \Lambda_- + A_{old}^T \cdot A_{old} & A_{old}^T \cdot A_{new} \\ A_{new}^T \cdot A_{old} & A_{new}^T \cdot A_{new} \end{pmatrix} \tag{A34}$$

Now, let us partition the posterior covariance matrix Σ_+ in a similar way:

$$\Sigma_+ = \begin{pmatrix} \Sigma_+^{old} & \Sigma_+^{cross} \\ (\Sigma_+^{cross})^T & \Sigma_+^{new} \end{pmatrix} \tag{A35}$$

Giving the setup till now, we will derive each of the matrices $\Sigma_+^{M,Y_{old}}$, $\Sigma_+^{M,Y_{new}}$ and $\Sigma_+^{(Y_{old},Y_{new})}$ from Eq. (A31) using parts from $\Sigma_-^{M,W}$ defined in Eq. (A32).

$\Sigma_+^{M,Y_{new}}$

By using block-wise matrix inversion (which is based on notion of Schur Complements), the Σ_+^{new} is equal to:

$$\begin{aligned} \Sigma_+^{new} = (A_{new}^T \cdot A_{new} - A_{new}^T \cdot A_{old} \cdot (\Lambda_- + A_{old}^T \cdot A_{old})^{-1} \cdot A_{old}^T \cdot A_{new})^{-1} = \\ = (A_{new}^T \cdot (I_m - A_{old} \cdot (\Lambda_- + A_{old}^T \cdot A_{old})^{-1} \cdot A_{old}^T) \cdot A_{new})^{-1} \end{aligned} \tag{A36}$$

Now, let's define matrix C as following:

$$C \triangleq I_m + A_{old} \cdot \Sigma_- \cdot A_{old}^T = I_m + {}^{I}\!A \cdot \Sigma_-^I \cdot ({}^{I}\!A)^T \tag{A37}$$

Through Woodbury matrix identity it can easily shown that C's inverse is:

$$C^{-1} = I - A_{old} \cdot (\Lambda_- + A_{old}^T \cdot A_{old})^{-1} \cdot A_{old}^T \tag{A38}$$

Therefore, Σ_+^{new} is equal to:

$$\Sigma_+^{new} = (A_{new}^T \cdot C^{-1} \cdot A_{new})^{-1} \tag{A39}$$

and $\Sigma_+^{M,Y_{new}}$ can be calculated in state dimension independent way as:

$$\Sigma_+^{M,Y_{new}} = P^{(Y_{new},:)}, \quad P \triangleq [(A_{new}^T \cdot C^{-1} \cdot A_{new})^{-1}]^{(:,Y_{new})} \tag{A40}$$

where in brackets we are using Matlab syntax to index relevant rows/columns. Note that $[(A_{new}^T \cdot C^{-1} \cdot A_{new})^{-1}]^{(:,Y_{new})}$ can be calculated without calculation of full $(A_{new}^T \cdot C^{-1} \cdot A_{new})^{-1}$, by using backslash operator in Matlab:

$$P = [A_{new}^T \cdot C^{-1} \cdot A_{new}] \backslash I^{(:,Y_{new})} \tag{A41}$$

where $I^{(:,Y_{new})}$ are particular columns of identity matrix.

$\boldsymbol{\Sigma_{+}^{M,Y_{old}}}$

Using the block-wise matrix inversion again we know that the Σ_{+}^{old} from Eq. (A35) is equal to:

$$
\begin{aligned}
\Sigma_{+}^{old} &= (\Lambda_{-} + A_{old}^{T} \cdot A_{old} - A_{old}^{T} \cdot A_{new} \cdot (A_{new}^{T} \cdot A_{new})^{-1} \cdot A_{new}^{T} \cdot A_{old})^{-1} = \\
&= (\Lambda_{-} + A_{old}^{T} \cdot (I_{m} - A_{new} \cdot (A_{new}^{T} \cdot A_{new})^{-1} \cdot A_{new}^{T}) \cdot A_{old})^{-1} = \\
&= (\Lambda_{k} + A_{old}^{T} \cdot K \cdot A_{old})^{-1}
\end{aligned}
\tag{A42}
$$

where

$$
K \triangleq I_{m} - A_{new} \cdot (A_{new}^{T} \cdot A_{new})^{-1} \cdot A_{new}^{T} = I_{m} - A_{new} \cdot F \cdot A_{new}^{T}, \quad F \triangleq (A_{new}^{T} \cdot A_{new})^{-1} \tag{A43}
$$

where K is singular, symmetric, idempotent projection matrix, with properties $K = K^{2}$ and $K = K^{T}$.

The Σ_{+}^{old} can be then rewritten as:

$$
\begin{aligned}
\Sigma_{+}^{old} &= (\Lambda_{-} + A_{old}^{T} \cdot K^{T} \cdot K \cdot A_{old})^{-1} = \\
&= \Lambda_{-}^{-1} - \Lambda_{-}^{-1} \cdot A_{old}^{T} \cdot K^{T} \cdot (I_{m} + K \cdot A_{old} \cdot \Lambda_{-}^{-1} \cdot A_{old}^{T} \cdot K^{T})^{-1} \cdot K \cdot A_{old} \cdot \cdot \Lambda_{-}^{-1} = \\
&= \Sigma_{-} - \Sigma_{-} \cdot A_{old}^{T} \cdot K^{T} \cdot (I_{m} + K \cdot A_{old} \cdot \Sigma_{-} \cdot A_{old}^{T} \cdot K^{T})^{-1} \cdot K \cdot A_{old} \cdot \cdot \Sigma_{-} = \\
&= \Sigma_{-} - \Sigma_{-} \cdot A_{old}^{T} \cdot K^{T} \cdot G^{-1} \cdot K \cdot A_{old} \cdot \cdot \Sigma_{-}
\end{aligned}
\tag{A44}
$$

where

$$
G \triangleq I_{m} + K \cdot A_{old} \cdot \Sigma_{-} \cdot A_{old}^{T} \cdot K^{T} = I_{m} + K \cdot {}^{I}\!A \cdot \Sigma_{-}^{I} \cdot ({}^{I}\!A)^{T} \cdot K^{T} = I_{m} + K_{1} \cdot \Sigma_{-}^{I} \cdot K_{1}^{T} \tag{A45}
$$

$$
K_{1} \triangleq K \cdot {}^{I}\!A \tag{A46}
$$

where K_1 are non-zero columns from A_{old} projected outside of vector space that is spanned by columns in A_{new}. In other words, K_1 contains information from A_{old} that isn't contained within A_{new}.

Then $\Sigma_{+}^{M,Y_{old}}$ can be calculated by retrieving from Σ_{+}^{old} rows and columns that belong to variables Y_{old}:

$$
\begin{aligned}
\Sigma_{+}^{M,Y_{old}} &= \Sigma_{-}^{M,Y_{old}} - \Sigma_{-}^{(Y_{old},:)} \cdot A_{old}^{T} \cdot K^{T} \cdot G^{-1} \cdot K \cdot A_{old} \cdot \Sigma_{-}^{(:,Y_{old})} = \\
&= \Sigma_{-}^{M,Y_{old}} - [K \cdot A_{old} \cdot \Sigma_{-}^{(:,Y_{old})}]^{T} \cdot G^{-1} \cdot [K \cdot A_{old} \cdot \Sigma_{-}^{(:,Y_{old})}] = \\
&= \Sigma_{-}^{M,Y_{old}} - [K \cdot {}^{I}\!A \cdot \Sigma_{-}^{(X^{I},Y_{old})}]^{T} \cdot G^{-1} \cdot [K \cdot {}^{I}\!A \cdot \Sigma_{-}^{(X^{I},Y_{old})}] = \\
&= \Sigma_{-}^{M,Y_{old}} - [K_{1} \cdot (\Sigma_{-}^{C})^{T}]^{T} \cdot G^{-1} \cdot [K_{1} \cdot (\Sigma_{-}^{C})^{T}] = \\
&= \Sigma_{-}^{M,Y_{old}} - B \cdot G^{-1} \cdot B^{T}
\end{aligned}
\tag{A47}
$$

where

$$
B \triangleq \Sigma_{-}^{C} \cdot K_{1}^{T} = \Sigma_{-}^{C} \cdot ({}^{I}\!A)^{T} \cdot K \tag{A48}
$$

$\Sigma_+^{(Y_{old},Y_{new})}$ - Method 1

Using the block-wise matrix inversion again we know that the Σ_+^{cross} from Eq. (A35) is equal to:

$$\begin{aligned}
\Sigma_+^{cross} &= -(\Lambda_- + A_{old}^T \cdot A_{old})^{-1} \cdot A_{old}^T \cdot A_{new} \cdot (A_{new}^T \cdot C^{-1} \cdot A_{new})^{-1} = \\
&= -(\Sigma_- - \Sigma_- \cdot A_{old}^T \cdot C^{-1} \cdot A_{old} \cdot \Sigma_-) \cdot A_{old}^T \cdot A_{new} \cdot (A_{new}^T \cdot C^{-1} \cdot A_{new})^{-1} = \\
&= -\Sigma_- \cdot A_{old}^T \cdot A_{new} \cdot (A_{new}^T \cdot C^{-1} \cdot A_{new})^{-1} + \Sigma_- \cdot A_{old}^T \cdot C^{-1} \cdot A_{old} \cdot \Sigma_- \cdot A_{old}^T \cdot A_{new} \cdot (A_{new}^T \cdot C^{-1} \cdot A_{new})^{-1} = \\
&= \Sigma_- \cdot A_{old}^T \cdot [-I_m + C^{-1} \cdot A_{old} \cdot \Sigma_- \cdot A_{old}^T] \cdot A_{new} \cdot (A_{new}^T \cdot C^{-1} \cdot A_{new})^{-1} = \\
&= \Sigma_- \cdot A_{old}^T \cdot [C^{-1} \cdot {}^{I}\!A \cdot \Sigma_-^I \cdot ({}^{I}\!A)^T - I_m] \cdot A_{new} \cdot (A_{new}^T \cdot C^{-1} \cdot A_{new})^{-1}
\end{aligned} \tag{A49}$$

where matrix C is defined in Eq. (A37)

Then $\Sigma_+^{(Y_{old},Y_{new})}$ can be calculated by retrieving from Σ_+^{cross} the entries that are corresponding to Y_{old} rows and Y_{new} columns:

$$\begin{aligned}
\Sigma_+^{(Y_{old},Y_{new})} &= \Sigma_-^{(Y_{old},:)} \cdot A_{old}^T \cdot [C^{-1} \cdot {}^{I}\!A \cdot \Sigma_-^I \cdot ({}^{I}\!A)^T - I_m] \cdot A_{new} \cdot [(A_{new}^T \cdot C^{-1} \cdot A_{new})^{-1}]^{(:,Y_{new})} = \\
&= \Sigma_-^C \cdot ({}^{I}\!A)^T \cdot [C^{-1} \cdot {}^{I}\!A \cdot \Sigma_-^I \cdot ({}^{I}\!A)^T - I_m] \cdot A_{new} \cdot P
\end{aligned} \tag{A50}$$

where matrix P is defined in Eq. (A40).

$\Sigma_+^{(Y_{old},Y_{new})}$ - Method 2

Using another form of block-wise matrix inversion, the Σ_+^{cross} from Eq. (A35) is equal to:

$$\begin{aligned}
\Sigma_+^{cross} &= \\
&= -[\Sigma_- - \Sigma_- \cdot A_{old}^T \cdot K^T \cdot G^{-1} \cdot K \cdot A_{old} \cdot \cdot \Sigma_-] \cdot A_{old}^T \cdot A_{new} \cdot (A_{new}^T \cdot A_{new})^{-1} = \\
&= [\Sigma_- \cdot A_{old}^T \cdot K^T \cdot G^{-1} \cdot K \cdot A_{old} \cdot \cdot \Sigma_- \cdot A_{old}^T - \Sigma_- \cdot A_{old}^T] \cdot A_{new} \cdot (A_{new}^T \cdot A_{new})^{-1} = \\
&= [\Sigma_- \cdot A_{old}^T \cdot K^T \cdot G^{-1} \cdot K \cdot A_{old} \cdot \cdot \Sigma_- \cdot A_{old}^T - \Sigma_- \cdot A_{old}^T] \cdot A_{new} \cdot F = \\
&= [\Sigma_- \cdot A_{old}^T \cdot K^T \cdot G^{-1} \cdot K \cdot {}^{I}\!A \cdot \Sigma_-^I \cdot ({}^{I}\!A)^T - \Sigma_- \cdot A_{old}^T] \cdot A_{new} \cdot F
\end{aligned} \tag{A51}$$

where matrix F is defined in Eq. (A43).

Then $\Sigma_+^{(Y_{old},Y_{new})}$ can be calculated by retrieving from Σ_+^{cross} the entries that are corresponding to Y_{old} rows and Y_{new} columns:

$$\begin{aligned}
\Sigma_+^{(Y_{old},Y_{new})} &= [\Sigma_-^{(Y_{old},:)} \cdot A_{old}^T \cdot K^T \cdot G^{-1} \cdot K \cdot {}^{I}\!A \cdot \Sigma_-^I \cdot ({}^{I}\!A)^T - \Sigma_-^{(Y_{old},:)} \cdot A_{old}^T] \cdot A_{new} \cdot F^{(:,Y_{new})} = \\
&= [\Sigma_-^C \cdot ({}^{I}\!A)^T \cdot K^T \cdot G^{-1} \cdot K \cdot {}^{I}\!A \cdot \Sigma_-^I \cdot ({}^{I}\!A)^T - \Sigma_-^C \cdot ({}^{I}\!A)^T] \cdot A_{new} \cdot F^{(:,Y_{new})} = \\
&= \Sigma_-^C \cdot [({}^{I}\!A)^T \cdot K^T \cdot G^{-1} \cdot K \cdot {}^{I}\!A \cdot \Sigma_-^I - I_k] \cdot ({}^{I}\!A)^T \cdot A_{new} \cdot F^{(:,Y_{new})} = \\
&= \Sigma_-^C \cdot [K_1^T \cdot G^{-1} \cdot K_1 \cdot \Sigma_-^I - I_k] \cdot ({}^{I}\!A)^T \cdot A_{new} \cdot F^{(:,Y_{new})}
\end{aligned} \tag{A52}$$

where matrix K_1 is defined in Eq. (A46) and identity matrix I_k has dimension $|X^I|$. ■

Note that the variables inside information matrices do not have to be ordered in any particular way, and that the provided above proof is correct for any ordering whatsoever.

1.2.3 *Squared* Action

The *squared* action is special case of *rectangular* action, and thus we will use here the same setup as for the *rectangular* action. In other words, we will use the partitioning that was defined in Eq. (A31), (A32) and (A33).

In case of *squared* action we have that $m = |X_{new}|$ from which we can conclude that matrix A_{new} from Eq. (A33) is squared matrix. Then matrix K from Eq. (A12) is equal to zero matrix:

$$K = I_m - A_{new} \cdot (A_{new}^T \cdot A_{new})^{-1} \cdot A_{new}^T = I_m - A_{new} \cdot A_{new}^{-1} \cdot (A_{new}^T)^{-1} \cdot A_{new}^T = 0 \quad \text{(A53)}$$

Then matrices K_1 and B from Eq. (A13) and Eq. (A14) contain only zeros, and $\Sigma_+^{M,Y_{old}}$ is equal to:

$$\Sigma_+^{M,Y_{old}} = \Sigma_-^{M,Y_{old}} - B \cdot G^{-1} \cdot B^T = \Sigma_-^{M,Y_{old}} \quad \text{(A54)}$$

The Σ_+^{new} from Eq. (A39) can be calculated as:

$$\Sigma_+^{new} = (A_{new}^T \cdot C^{-1} \cdot A_{new})^{-1} = A_{new}^{-1} \cdot C \cdot (A_{new}^T)^{-1} = A_{new}^{-1} \cdot C \cdot (A_{new}^{-1})^T \quad \text{(A55)}$$

and $\Sigma_+^{M,Y_{new}}$ is equal to:

$$\begin{aligned} \Sigma_+^{M,Y_{new}} &= [A_{new}^{-1}]^{(Y_{new},:)} \cdot C \cdot [(A_{new}^{-1})^T]^{(:,Y_{new})} = \\ &= [A_{new}^{-1}]^{(Y_{new},:)} \cdot C \cdot ([A_{new}^{-1}]^{(Y_{new},:)})^T = A_{iv} \cdot C \cdot A_{iv}^T \end{aligned} \quad \text{(A56)}$$

where

$$A_{iv} \triangleq [A_{new}^{-1}]^{(Y_{new},:)} \quad \text{(A57)}$$

and can be efficiently calculated through Matlab backslash operator:

$$A_{iv} = A_{new} \backslash I^{(:,Y_{new})} \quad \text{(A58)}$$

Next, we can reduce Eq. (A17) to:

$$\begin{aligned} \Sigma_+^{(Y_{old},Y_{new})} &= \Sigma_-^C \cdot [K_1^T \cdot G^{-1} \cdot K_1 \cdot \Sigma_-^I - I_k] \cdot ({}^I\!A)^T \cdot A_{new} \cdot F^{(:,Y_{new})} = \\ &= -\Sigma_-^C \cdot ({}^I\!A)^T \cdot A_{new} \cdot F^{(:,Y_{new})} = -[\Sigma_-^C \cdot ({}^I\!A)^T \cdot A_{new} \cdot F]^{(:,Y_{new})} = \\ &= -[\Sigma_-^C \cdot ({}^I\!A)^T \cdot A_{new} \cdot A_{new}^{-1} \cdot (A_{new}^T)^{-1}]^{(:,Y_{new})} = -[\Sigma_-^C \cdot ({}^I\!A)^T \cdot (A_{new}^T)^{-1}]^{(:,Y_{new})} = \\ &= -\Sigma_-^C \cdot ({}^I\!A)^T \cdot [(A_{new}^T)^{-1}]^{(:,Y_{new})} = -\Sigma_-^C \cdot ({}^I\!A)^T \cdot ([A_{new}^{-1}]^{(Y_{new},:)})^T = \\ &= -\Sigma_-^C \cdot ({}^I\!A)^T \cdot (A_{iv})^T \end{aligned} \quad \text{(A59)}$$

■

Note that the variables inside information matrices do not have to be ordered in any particular way, and that the provided above proof is correct for any ordering whatsoever.

2 Appendix B: Sum of Information Gains

Consider action a with *increment* $I(a) = \{F_{new}, X_{new}\}$. Further, consider specific partitioning of a into sub-actions $a = \{a'_1, \ldots, a'_k\}$ where each sub-action a'_i has *increment* $I_i(a'_i) = \{F_{i,new}, X_{i,new}\}$. The factor sets $F_{i,new}$ are disjoint, as also are the new variable sets $X_{i,new}$. Also, for proper action partitioning we will have $\cup_{i=1}^{k} F_{i,new} = F_{new}$ and $\cup_{i=1}^{k} X_{i,new} = X_{new}$.

Next, we will prove that information gain (IG) of a is equal to sum of IGs of sub-actions $\{a'_i\}_{i=1}^{k}$ in `unfocused` scenario. Similar proof can be shown also for `focused` BSP.

The `unfocused` IG of action a by definition is:

$$J_{IG}(a) = \mathcal{H}(b[X_-]) - \mathcal{H}(b[X_+]) \tag{A60}$$

where $b[X_-]$ is a prior state belief before applying action a, and $b[X_+]$ is a posterior state belief after applying it.

Additionally, denote posterior state belief of each sub-action a'_i as $b_i[X_{i,+}]$. When applying sub-actions consecutively in sequence, belief propagation will have next form:

$$b[X_-] \Longrightarrow b_1[X_{1,+}] \Longrightarrow b_2[X_{2,+}] \Longrightarrow \cdots \Longrightarrow b_{k-1}[X_{k-1,+}] \Longrightarrow b[X_+] \tag{A61}$$

Then, the IG of each sub-action is equal to:

$$\begin{aligned}
J_{IG}(a'_1) &= \mathcal{H}(b[X_-]) - \mathcal{H}(b_1[X_{1,+}]) \\
J_{IG}(a'_2) &= \mathcal{H}(b_1[X_{1,+}]) - \mathcal{H}(b_2[X_{2,+}]) \\
&\cdots \\
J_{IG}(a'_{k-1}) &= \mathcal{H}(b_{k-2}[X_{k-2,+}]) - \mathcal{H}(b_{k-1}[X_{k-1,+}]) \\
J_{IG}(a'_k) &= \mathcal{H}(b_{k-1}[X_{k-1,+}]) - \mathcal{H}(b[X_+])
\end{aligned}$$

and sum of these IGs is equal to:

$$\sum_{i=1}^{k} J_{IG}(a'_i) = \mathcal{H}(b[X_-]) - \mathcal{H}(b[X_+]) = J_{IG}(a) \tag{A62}$$

■

References

[1] D. Kopitkov and V. Indelman, "Efficient belief space planning via factor-graph propagation action tree and incremental covariance update," in *The 18th International Symposium on Robotics Research*, Chile, December 2017, submitted.