Active SLAM for Mobile Robots with Area Coverage and Obstacle Avoidance
Supplementary Material

Yongbo Chen*, Shoudong Huang*, Robert Fitch*

December 30, 2019

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

"Y. Chen, S. Huang and R. Fitch, Active SLAM for Mobile Robots with Area Coverage and Obstacle Avoidance, IEEE/ASME Transactions on Mechatronics, (Supplementary Material), 2019. https://github.com/cyb1212/SM-for-TMECH"

Throughout this document, notations used are the same as those in [1].

1 Appendix A: Preliminaries

In this part, we give some preliminary knowledge about the variational description of eigenvalues and the Fischer's min-max theorem

Theorem 1 *The eigenvalues of a block diagonal matrix are the eigenvalues of all block matrices on the principal diagonal.*

Theorem 2 (Variational description of eigenvalues) *([3], p. 232) Let $\boldsymbol{\mathcal{A}}$ be a real symmetric $n \times n$ matrix with eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Let $\boldsymbol{S} = (\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_n)$ be an orthogonal $n \times n$ matrix which diagonalizes $\boldsymbol{\mathcal{A}}$, so that:*

$$\boldsymbol{S}^T \boldsymbol{\mathcal{A}} \boldsymbol{S} = diag(\lambda_1, \lambda_2, \cdots, \lambda_n) \tag{1}$$

Then, for $k = 1, 2, \cdots, n$, we have:

$$\lambda_k = \min_{\boldsymbol{\mathcal{R}}_{k-1}^T \boldsymbol{x}=0} \frac{\boldsymbol{x}^\top \boldsymbol{\mathcal{A}} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} = \max_{\boldsymbol{\mathcal{T}}_{k+1}^\top \boldsymbol{x}=0} \frac{\boldsymbol{x}^T \boldsymbol{\mathcal{A}} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \tag{2}$$

where,

$$\begin{aligned} \boldsymbol{\mathcal{R}}_k &= (\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_k), \\ \boldsymbol{\mathcal{T}}_k &= (\boldsymbol{s}_k, \boldsymbol{s}_{k+1}, \cdots, \boldsymbol{s}_n) \end{aligned} \tag{3}$$

Theorem 3 (The Fischer's min-max theorem) *([3], p. 233) Let $\boldsymbol{\mathcal{A}}$ be a real symmetric $n \times n$ matrix with eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Let $1 \leq k \leq n$. Then, for every $n \times (k-1)$ matrix $\boldsymbol{\mathcal{B}}$,*

$$\min_{\boldsymbol{\mathcal{B}}^T \boldsymbol{x}=0} \frac{\boldsymbol{x}^T \boldsymbol{\mathcal{A}} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \leq \lambda_k \tag{4}$$

and, for every $n \times (n-k)$ matrix $\boldsymbol{\mathcal{C}}$,

$$\max_{\boldsymbol{\mathcal{C}}^T \boldsymbol{x}=0} \frac{\boldsymbol{x}^T \boldsymbol{\mathcal{A}} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \geq \lambda_k \tag{5}$$

*Yongbo Chen, Shoudong Huang and Robert Fitch are with Faculty of Engineering and Information Technology, University of Technology Sydney, Ultimo, NSW, 2007 Australia, e-mail: Yongbo.Chen@student.uts.edu.au, Shoudong.Huang@uts.edu.au, Robert.Fitch@uts.edu.au.

2 Appendix B: Simplification from (3) to (4) in [1]

The objective function of the uncertainty minimization task is based on the generalized belief at the L-th planning step:

$$\begin{aligned} & f_a(u_{k:k+L-1}) = f_a(gb(\boldsymbol{X}_{k+L})) = -\log\left(\det(\boldsymbol{\mathcal{I}}_{k+L}(\boldsymbol{X}^{opt}_{k+L}))\right) \\ & \boldsymbol{X}^{opt}_{k+L} = \underset{\boldsymbol{X}_{k+L}}{\operatorname{argmin}} \; -\log(p(\boldsymbol{X}_{k+L}|\boldsymbol{Z}_{1:k}, u_{0:k+L-1}, \boldsymbol{Z}_{k+1:k+L}) \end{aligned}, \tag{6}$$

Using Bayes' theorem, we have:

$$\begin{aligned} & p(\boldsymbol{X}_{k+L}|\boldsymbol{Z}_{1:k}, u_{0:k-1}, \boldsymbol{Z}_{k+1:k+L}, u_{k:k+L-1}) \\ = & \frac{p(\boldsymbol{Z}_{k+1:k+L}|\boldsymbol{X}_{k+L}, \boldsymbol{Z}_{1:k}, u_{0:k+L-1})p(\boldsymbol{X}_{k+L}|\boldsymbol{Z}_{1:k}, u_{0:k+L-1})}{p(\boldsymbol{Z}_{k+1:k+L}|\boldsymbol{Z}_{1:k}, u_{0:k+L-1})}. \end{aligned} \tag{7}$$

We now make the assumption that the prior $p(\boldsymbol{Z}_{k+1:k+L}|\boldsymbol{Z}_{1:k}, u_{0:k+L-1})$ is uninformative and is set to 1. This assumption is fairly standard in inference, see [4]. The above equation (8) can thus be rewritten as:

$$\begin{aligned} & p(\boldsymbol{X}_{k+L}|\boldsymbol{Z}_{1:k}, u_{0:k-1}, \boldsymbol{Z}_{k+1:k+L}, u_{k:k+L-1}) \\ \propto & p(\boldsymbol{Z}_{k+1:k+L}|\boldsymbol{X}_{k+L}, \boldsymbol{Z}_{1:k}, u_{0:k+L-1})p(\boldsymbol{X}_{k+L}|\boldsymbol{Z}_{1:k}, u_{0:k+L-1}). \end{aligned} \tag{8}$$

The future measurement $\boldsymbol{Z}_{k+1:k+L}$ is a probabilistic event. If the j-th feature is outside the sensor range from the actual position of the robot, it will not be observed. We know nothing about the unmapped features, so, in the MPC predicted framework, we only consider the measurements from the features already mapped by the SLAM method. Even for the mapped features, the exact values of the future measurement $\boldsymbol{Z}_{k+1:k+L}$ are also unknown because of the uncertainties involved. In order to solve this problem, we assume the measurement $\boldsymbol{Z}_{k+1:k+L}$ is perfect (zero-innovation). In other words, the predicted measurement will not affect the SLAM estimated result but change the FIM. Using the Markov property, we obtain the predicted state vector $\boldsymbol{X}^{opt}_{k+L}$:

$$\boldsymbol{X}^{opt}_{k+L} = \underset{\boldsymbol{X}_{k+L}}{\operatorname{argmin}}\, p(\boldsymbol{X}_{k}|\boldsymbol{Z}_{1:k}, u_{0:k-1})p(\boldsymbol{X}_{k+1:k+L}|\boldsymbol{X}_{k}, u_{k:k+L-1}). \tag{9}$$

Equation (9) includes two parts. The first part is a classical SLAM estimation problem, and the second is a prediction process with zero-innovation probabilistic measurement $\boldsymbol{Z}_{k+1:k+L}$. Assuming that the SLAM result at k-th step is $\boldsymbol{X}^{opt}_{k}$, the predicted future pose $\boldsymbol{X}^{opt}_{k+L}$ will be:

$$\boldsymbol{X}^{opt}_{k+L} = \begin{bmatrix} \boldsymbol{X}^{opt}_{k} \\ \boldsymbol{x}^{opt}_{k+1} \\ \vdots \\ \boldsymbol{x}^{opt}_{k+L} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}^{opt}_{k} \\ f_v(\boldsymbol{x}^{opt}_{k}, u_k) \\ \vdots \\ f_v(\boldsymbol{x}^{opt}_{k+L-1}, u_{k+L-1}) \end{bmatrix}, \tag{10}$$

where $f_v(\star)$ is the motion equation without uncertainty shown in Section II.A [1] and $\boldsymbol{x}^{opt}_{k+i}$ is the predicted pose at step $k+i$.

3 Appendix C: Proof of the lower bound in feature-based SLAM

Proof: Introduce the Jacobian matrix, we have:

$$\begin{aligned} & \boldsymbol{\mathcal{I}}(\boldsymbol{X}) = \boldsymbol{J}(\boldsymbol{X})^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{J}(\boldsymbol{X}) \\ = & \begin{bmatrix} \frac{\partial \boldsymbol{h_p}}{\partial \boldsymbol{p}}^\top & \frac{\partial \boldsymbol{h_\theta}}{\partial \boldsymbol{p}}^\top \\ \frac{\partial \boldsymbol{h_p}}{\partial \boldsymbol{\theta}}^\top & \frac{\partial \boldsymbol{h_\theta}}{\partial \boldsymbol{\theta}}^\top \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{p}}^{-1} \otimes \boldsymbol{I}^{2\times 2} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1} \end{bmatrix} \begin{bmatrix} \frac{\partial \boldsymbol{h_p}}{\partial \boldsymbol{p}} & \frac{\partial \boldsymbol{h_p}}{\partial \boldsymbol{\theta}} \\ \frac{\partial \boldsymbol{h_\theta}}{\partial \boldsymbol{p}} & \frac{\partial \boldsymbol{h_\theta}}{\partial \boldsymbol{\theta}} \end{bmatrix} \end{aligned} \tag{11}$$

For each block of the matrix $\boldsymbol{\mathcal{I}}(\boldsymbol{X})$, we have:

$$
\begin{aligned}
&\frac{\partial \boldsymbol{h_p}}{\partial \boldsymbol{p}}^\top \boldsymbol{\Sigma}_p^{-1} \otimes \boldsymbol{I}^{2\times 2} \frac{\partial \boldsymbol{h_p}}{\partial \boldsymbol{p}} = \left(\boldsymbol{A_g} \boldsymbol{R} \boldsymbol{\Sigma_p}^{-1} \boldsymbol{R}^\top \boldsymbol{A_g}^\top\right) \otimes \boldsymbol{I}^{2\times 2} \\
&\frac{\partial \boldsymbol{h_p}}{\partial \boldsymbol{p}}^\top \boldsymbol{\Sigma_p}^{-1} \otimes \boldsymbol{I}^{2\times 2} \frac{\partial \boldsymbol{h_p}}{\partial \boldsymbol{\theta}} = \left(\boldsymbol{A_g} \otimes \boldsymbol{I}^{2\times 2}\right) \boldsymbol{R} \boldsymbol{\Sigma_p}^{-1} \otimes \boldsymbol{I}^{2\times 2} \boldsymbol{\Gamma} \boldsymbol{R}^\top \triangle_{\boldsymbol{w_p}} \\
&\frac{\partial \boldsymbol{h_p}}{\partial \boldsymbol{\theta}}^\top \boldsymbol{\Sigma_p}^{-1} \otimes \boldsymbol{I}^{2\times 2} \frac{\partial \boldsymbol{h_p}}{\partial \boldsymbol{\theta}} + \frac{\partial \boldsymbol{h_\theta}}{\partial \boldsymbol{\theta}}^\top \boldsymbol{\Sigma_\theta}^{-1} \frac{\partial \boldsymbol{h_\theta}}{\partial \boldsymbol{\theta}} \\
= &\triangle^\top \boldsymbol{R} \boldsymbol{\Gamma}^\top (\boldsymbol{\Sigma_p}^{-1} \otimes \boldsymbol{I}^{2\times 2}) \boldsymbol{\Gamma} \boldsymbol{R}^\top \triangle + \boldsymbol{A_p} \boldsymbol{\Sigma_\theta}^{-1} \boldsymbol{A}_p^\top
\end{aligned}
\tag{12}
$$

where $\triangle \in \mathbb{R}^{2m\times n}$ is a matrix defined as follows. Suppose in the k-th measurement, the i_kth node has observed the j_k-th node. For a 2×1 block in $\triangle$,

$$
\triangle_{2k-1:2k, i_k} = \boldsymbol{p}_{j_k} - \boldsymbol{p}_{i_k}. \tag{13}
$$

The remaining elements in $\triangle$ are all zero.

Based on $\boldsymbol{R}\boldsymbol{\Sigma_p}^{-1}\boldsymbol{R}^\top = \boldsymbol{\Sigma_p}^{-1}$, $\boldsymbol{\Gamma}\boldsymbol{R}^\top = \boldsymbol{R}^\top\boldsymbol{\Gamma}$, $\boldsymbol{\Gamma}^\top\boldsymbol{\Gamma} = \boldsymbol{I}^{2m\times 2m}$ and the FIM is a symmetrical matrix, we can get the FIM $\boldsymbol{\mathcal{I}}(\boldsymbol{X})$.

Based on the D-optimality criterion, we have:

$$
\log(\det \begin{bmatrix} \boldsymbol{L^g_{w_p}} \otimes \boldsymbol{I}^{2\times 2} & \boldsymbol{A^g_{w_p}} \otimes \boldsymbol{I}^{2\times 2} \boldsymbol{\Gamma} \triangle_{\boldsymbol{w_p}} \\ (\boldsymbol{A^g_{w_p}} \otimes \boldsymbol{I}^{2\times 2} \boldsymbol{\Gamma} \triangle_{\boldsymbol{w_p}})^\top & \triangle^\top_{\boldsymbol{w_p}} \triangle_{\boldsymbol{w_p}} + \boldsymbol{L^p_{w\theta}} \end{bmatrix}). \tag{14}
$$

The Schur's Determinant Formula says: if $\widetilde{\boldsymbol{A}}^{-1}$ exists,

$$
\det \begin{bmatrix} \widetilde{\boldsymbol{A}} & \widetilde{\boldsymbol{B}} \\ \widetilde{\boldsymbol{C}} & \widetilde{\boldsymbol{D}} \end{bmatrix} = \det(\widetilde{\boldsymbol{A}}) \det(\widetilde{\boldsymbol{D}} - \widetilde{\boldsymbol{C}} \widetilde{\boldsymbol{A}}^{-1} \widetilde{\boldsymbol{B}}). \tag{15}
$$

Using the Schur's Determinant Formula, we have:

$$
\det(\widetilde{\boldsymbol{A}}) = \det(\boldsymbol{L^g_{w_p}} \otimes \boldsymbol{I}^{2\times 2}) = \det(\boldsymbol{L^g_{w_p}})^2, \tag{16}
$$

$$
\begin{aligned}
&\det\left(\widetilde{\boldsymbol{D}} - \widetilde{\boldsymbol{C}} \widetilde{\boldsymbol{A}}^{-1} \widetilde{\boldsymbol{B}}\right) = \\
&\det(\triangle^\top_{\boldsymbol{w_p}} \triangle_{\boldsymbol{w_p}} + \boldsymbol{L^p_{w\theta}} - \triangle^\top_{\boldsymbol{w_p}} \boldsymbol{\Gamma}^\top (\boldsymbol{A^g_{w_p}}^\top \boldsymbol{L^g_{w_p}}^{-1} \boldsymbol{A^g_{w_p}}) \otimes \boldsymbol{I}^{2\times 2} \boldsymbol{\Gamma} \triangle_{\boldsymbol{w_p}}).
\end{aligned}
\tag{17}
$$

Eq.(14) can be written as:

$$
\begin{aligned}
\text{Eq.(14)} =& 2\log\det(\boldsymbol{L^g_{w_p}}) + \log\det(\triangle^\top_{\boldsymbol{w_p}} \boldsymbol{P^g_{w_p}} \triangle_{\boldsymbol{w_p}} + \boldsymbol{L^p_{w\theta}}) \\
\boldsymbol{P}^g_{w_p} =& \boldsymbol{I}^{2m\times 2m} - \boldsymbol{\Gamma}^\top (\boldsymbol{A^g_{w_p}}^\top \boldsymbol{L^g_{w_p}}^{-1} \boldsymbol{A^g_{w_p}}) \otimes \boldsymbol{I}^{2\times 2} \boldsymbol{\Gamma}.
\end{aligned}
\tag{18}
$$

For any two matrixes $\boldsymbol{N}$ and $\boldsymbol{M}$, meeting $\boldsymbol{M} \succeq \boldsymbol{0}$ and $\boldsymbol{N} \succeq \boldsymbol{0}$, we have:

$$
\det(\boldsymbol{M} + \boldsymbol{N}) \geq \det(\boldsymbol{M}). \tag{19}
$$

Because $\boldsymbol{P}^g_{w_p}$ is an orthogonal projection matrix, so we have:

$$
\log(\det(\boldsymbol{\mathcal{I}}(\boldsymbol{X}))) \geq 2\log(\det(\boldsymbol{L^g_{w_p}})) + \log\left(\det\left(\boldsymbol{L^p_{w\theta}}\right)\right). \tag{20}
$$

So (12) in [1] is proved.

4 Appendix D: D-opt objective function gets closed to its lower bound

Fig. 1 shows that the objective function value is very close to its lower bound (Dataset 1).

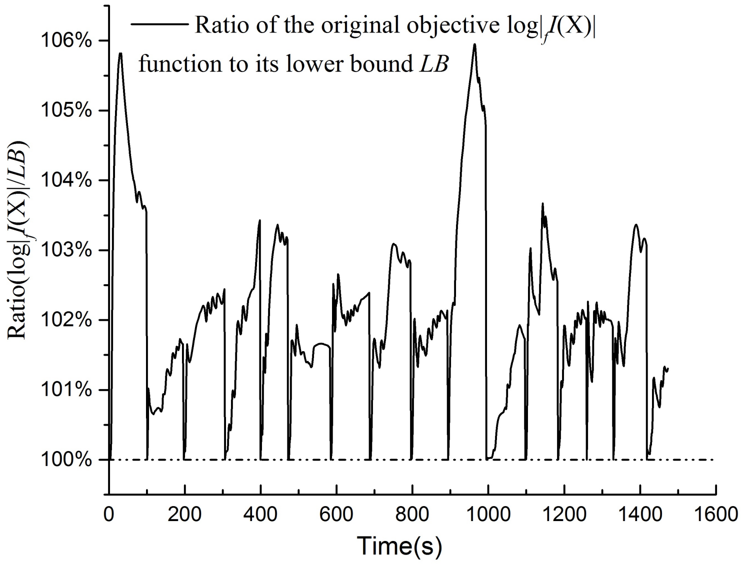

Figure 1: The objective function values of the uncertainty mininization problem are very close to the lower bounds

5 Appendix E: Proof of Conclusion 1

Proof: The coefficient matrix $\boldsymbol{A_Z}$ has a special structure. In fact, it shows the corresponding relationship between the state vectors of the submap and the joined global map. So it is a block matrix with multiple identity matrix and column full rank. At the same time, in every row, there is only one non-zero element, like:

$$\boldsymbol{A_Z} = \begin{bmatrix} & & & \vdots & & & \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ & & & \vdots & & & \end{bmatrix}_{r_1 \times r_2}, \tag{21}$$

where $r_1 \geq r_2$.

So we can get:

$$(\boldsymbol{A}_{\boldsymbol{Z}}^{\top} * \boldsymbol{A_Z})_{i,j} = \begin{cases} zero, & i \neq j, \\ non-zero, & i = j, \end{cases} \tag{22}$$

which is a diagonal matrix.

6 Appendix F: Overview of Linear SLAM [2]

Assuming two submaps 1 and 2 with state vectors $\{\boldsymbol{x^{L1}}, \boldsymbol{x^{L2}}\}$ and their corresponding information matrices $\boldsymbol{\mathcal{I}}(\boldsymbol{x^{L1}})$ and $\boldsymbol{\mathcal{I}}(\boldsymbol{x^{L2}})$, we need to compute the state vectors and the FIM of the joined global map. We assume that the two submaps are dovetailed in that the last pose of submap 1 is equivalent to the first pose of submap 2.

Using the Linear SLAM algorithm, submap 1 is transformed based on its last pose. The new state vector based on the last pose of submap 1 is $\boldsymbol{x^G}$. The relation between $\boldsymbol{x^{L1}}$ and $\boldsymbol{x^G}$ is defined as: $\boldsymbol{x^G} = f_t(\boldsymbol{x^{L1}})$. Then, the FIM of the new transformed map $\boldsymbol{x^G}$ is:

$$\boldsymbol{\mathcal{I}}_{L1} = \boldsymbol{J}_{f_t}(\boldsymbol{x^{L1}})^{\top} \boldsymbol{\mathcal{I}}(\boldsymbol{x^{L1}}) \boldsymbol{J}_{f_t}(\boldsymbol{x^{L1}}), \tag{23}$$

where $\boldsymbol{J}_{f_t}(\star)$ is the Jacobian matrix of the function f_t.

Since the two submaps are now defined in a common coordinate frame, the submap joining problem becomes a linear least squares problem:

$$\min \ f_L(\boldsymbol{X}) = \| \boldsymbol{x^G} - \boldsymbol{X_1} \|^2_{\boldsymbol{\mathcal{I}_{L1}}} + \| \boldsymbol{x^{L2}} - \boldsymbol{X_2} \|^2_{\boldsymbol{\mathcal{I}_{L2}}}, \tag{24}$$

where $\boldsymbol{\mathcal{I}_{L2}} = \boldsymbol{\mathcal{I}}(\boldsymbol{x^{L2}})$, and $\boldsymbol{X}$ is a subset of $[\boldsymbol{X_1}^\top, \boldsymbol{X_2}^\top]^\top$ based on data association. This linear least squares problem can be written in a compact form as:

$$\min \ \|\boldsymbol{A_Z X} - \boldsymbol{Z}\|^2_{\boldsymbol{I_Z}}, \tag{25}$$

where $\boldsymbol{A_Z}$ is a sparse coefficient matrix that gives the data association relationship between $\boldsymbol{X}$ and $[\boldsymbol{X_1}^\top, \boldsymbol{X_2}^\top]$. Matrix $\boldsymbol{\mathcal{I}_Z}$ is the corresponding information matrix given by $\boldsymbol{\mathcal{I}_Z} = \mathrm{diag}\{\boldsymbol{\mathcal{I}}_{L1},\ \boldsymbol{\mathcal{I}}_{L2}\}$. The corresponding global FIM is:

$$\boldsymbol{\mathcal{I}_{all}} = \boldsymbol{A_Z^\top \mathcal{I}_Z A_Z}. \tag{26}$$

7 Appendix G: Proof of Conclusion 2 in [1]

Proof: a. Based on (2), we have:

$$\lambda_i(\boldsymbol{\mathcal{I}_Z}) = \min_{\boldsymbol{\mathcal{R}}_{i-1}^\top \boldsymbol{x}=0} \frac{\boldsymbol{x}^\top \boldsymbol{\mathcal{I}_Z x}}{\boldsymbol{x}^\top \boldsymbol{x}}. \tag{27}$$

When $\boldsymbol{x} = \boldsymbol{A_Z y}$, which is a new constraint for (27), we have:

$$\min_{\boldsymbol{\mathcal{R}}_{k-1}^\top \boldsymbol{x}=0} \frac{\boldsymbol{x}^\top \boldsymbol{\mathcal{I}_Z x}}{\boldsymbol{x}^\top \boldsymbol{x}} \leq \min_{\substack{\boldsymbol{\mathcal{R}}_{k-1}^\top \boldsymbol{x}=0 \\ \boldsymbol{x}=\boldsymbol{A_Z y}}} \frac{\boldsymbol{x}^\top \boldsymbol{\mathcal{I}_Z x}}{\boldsymbol{x}^\top \boldsymbol{x}}. \tag{28}$$

Let $\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{R}}_{k-1}^\top \boldsymbol{A_Z}$, we have:

$$\min_{\substack{\boldsymbol{\mathcal{R}}_{k-1}^\top \boldsymbol{x}=0 \\ \boldsymbol{x}=\boldsymbol{A_Z y}}} \frac{\boldsymbol{x}^\top \boldsymbol{\mathcal{I}_Z x}}{\boldsymbol{x}^\top \boldsymbol{x}} = \min_{\boldsymbol{\mathcal{B} y}=0} \frac{\boldsymbol{y}^\top \boldsymbol{A_Z^\top I_Z A_Z y}}{\boldsymbol{y}^\top \boldsymbol{A_Z^\top A_Z y}}. \tag{29}$$

Because $\boldsymbol{A_Z^\top A_Z}$ is a diagonal matrix, defined as:

$$\boldsymbol{A_Z^\top A_Z} = \mathrm{diag}(\lambda_1(\boldsymbol{A_Z}), \lambda_2(\boldsymbol{A_Z}), \cdots, \lambda_k(\boldsymbol{A_Z})), \tag{30}$$

we can rewrite the denominator $\boldsymbol{y}^\top \boldsymbol{A_Z^\top A_Z y}$ as: $\lambda_1(\boldsymbol{A_Z})y_1^2 + \lambda_2(\boldsymbol{A_Z})y_2^2 + \cdots \lambda_k(\boldsymbol{A_Z})y_k^2$. Then, we have:

$$\boldsymbol{y}^\top \boldsymbol{A_Z^\top A_Z y} = \lambda_1(\boldsymbol{A_Z})y_1^2 + \cdots \lambda_k(\boldsymbol{A_Z})y_k^2 \geq \widehat{\lambda}(\boldsymbol{A_Z})(\boldsymbol{y}^\top \boldsymbol{y}). \tag{31}$$

So, we obtain:

$$\min_{\boldsymbol{\mathcal{B} y}=0} \frac{\boldsymbol{y}^\top \boldsymbol{A_Z^\top \mathcal{I}_Z A_Z y}}{\boldsymbol{y}^\top \boldsymbol{A_Z^\top A_Z y}} \leq \min_{\boldsymbol{\mathcal{B} y}=0} \frac{\boldsymbol{y}^\top \boldsymbol{A_Z^\top \mathcal{I}_Z A_Z y}}{\widehat{\lambda}(\boldsymbol{A_Z})\boldsymbol{y}^\top \boldsymbol{y}}. \tag{32}$$

Based on **The Fischer's min-max theorem** ([3], p. 233) and $\widehat{\lambda}(A_Z) > 0$, we can get:

$$\lambda_i(\boldsymbol{\mathcal{I}_Z}) \leq \min_{\boldsymbol{\mathcal{B}} y=0} \frac{\boldsymbol{y}^\top \boldsymbol{A_Z^\top \mathcal{I}_Z A_Z y}}{\widehat{\lambda}(\boldsymbol{A_Z})\boldsymbol{y}^\top \boldsymbol{y}} \leq \frac{\lambda_i(\boldsymbol{\mathcal{I}_{all}})}{\widehat{\lambda}(\boldsymbol{A_Z})}. \tag{33}$$

$$\Longrightarrow \widehat{\lambda}(\boldsymbol{A_Z})\lambda_i(\boldsymbol{\mathcal{I}_Z}) \leq \lambda_i(\boldsymbol{\mathcal{I}_{all}}). \tag{34}$$

In short, we have proved the first inequality of Conclusion 2.
b. Then, let's prove the other one:
For $n - k + 1 \leq j \leq n$

$$\lambda_j(\boldsymbol{\mathcal{I}_Z}) = \max_{\boldsymbol{\mathcal{T}}_{k+1}^\top \boldsymbol{x}=0} \frac{\boldsymbol{x}^\top \boldsymbol{\mathcal{I}_Z x}}{\boldsymbol{x}^\top \boldsymbol{x}}. \tag{35}$$

When $\boldsymbol{x} = \boldsymbol{A_Z y}$, which is a special solution for (35), we have:

$$\max_{\boldsymbol{\mathcal{T}}_{k+1}^\top \boldsymbol{x}=0} \frac{\boldsymbol{x}^\top \boldsymbol{\mathcal{I}_Z x}}{\boldsymbol{x}^\top \boldsymbol{x}} \geq \max_{\substack{\boldsymbol{\mathcal{T}}_{k+1}^\top \boldsymbol{x}=0 \\ \boldsymbol{x}=\boldsymbol{A_Z y}}} \frac{\boldsymbol{x}^\top \boldsymbol{\mathcal{I}_Z x}}{\boldsymbol{x}^\top \boldsymbol{x}}. \tag{36}$$

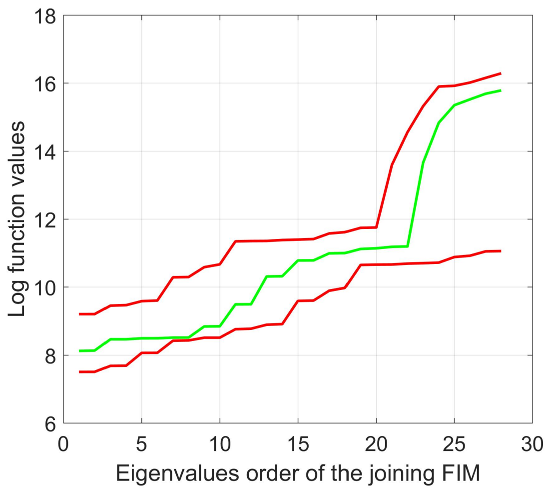

Figure 2: log function of eignvalues and their bounds

Let $\boldsymbol{\mathcal{B}_1} = \mathcal{T}_{k+1}^\top \boldsymbol{A_Z}$, we have:

$$\max_{\substack{\mathcal{T}_{k+1}^\top \boldsymbol{x} = 0 \\ \boldsymbol{x} = \boldsymbol{A_Z y}}} \frac{\boldsymbol{x}^\top \boldsymbol{\mathcal{I}_Z x}}{\boldsymbol{x}^\top \boldsymbol{x}} = \max_{\boldsymbol{\mathcal{B}_1 y}=0} \frac{\boldsymbol{y}^\top \boldsymbol{A_Z^\top \mathcal{I}_Z A_Z y}}{\boldsymbol{y}^\top \boldsymbol{A_Z^\top A_Z y}}. \tag{37}$$

Because $\boldsymbol{A_Z^\top A_Z}$ is a diagonal matrix, we have:

$$\boldsymbol{y}^\top \boldsymbol{A_Z^\top A_Z y} = \lambda_1(\boldsymbol{A_Z}) y_1^2 + \cdots \lambda_k(\boldsymbol{A_Z}) y_k^2 \leq \widetilde{\lambda}(\boldsymbol{A_Z})(\boldsymbol{y}^\top \boldsymbol{y}). \tag{38}$$

So we can get:

$$\max_{\boldsymbol{\mathcal{B}_1 y}=0} \frac{\boldsymbol{y}^\top \boldsymbol{A_Z^\top \mathcal{I}_Z A_Z y}}{\boldsymbol{y}^\top \boldsymbol{A_Z^\top A_Z y}} \geq \max_{\boldsymbol{\mathcal{B}_1 y}=0} \frac{\boldsymbol{y}^\top \boldsymbol{A_Z^\top \mathcal{I}_Z A_Z y}}{\widetilde{\lambda}(\boldsymbol{A_Z})(\boldsymbol{y}^\top \boldsymbol{y})}. \tag{39}$$

Based on **The Fischer's min-max theorem** ([3], p. 233) and $\widetilde{\lambda}(\boldsymbol{A_Z}) > 0$, we can get:

$$\lambda_j(\boldsymbol{\mathcal{I}_Z}) \geq \max_{\mathcal{B}_1 \boldsymbol{y}=0} \frac{\boldsymbol{y}^\top \boldsymbol{A_Z^\top \mathcal{I}_Z A_Z y}}{\widetilde{\lambda}(\boldsymbol{A_Z})(\boldsymbol{y}^\top \boldsymbol{y})} \geq \frac{\lambda_{k-n+j}(\boldsymbol{\mathcal{I}_{all}})}{\widetilde{\lambda}(\boldsymbol{A_Z})}. \tag{40}$$

$$\Longrightarrow \widehat{\lambda}(\boldsymbol{A_Z}) \lambda_j(\boldsymbol{\mathcal{I}_Z}) \geq \lambda_{k-n+j}(\boldsymbol{\mathcal{I}_{all}}). \tag{41}$$

Let's choose $j = n - k + i (1 \leq i \leq k)$:

$$\Longrightarrow \widehat{\lambda}(\boldsymbol{A_Z}) \lambda_{n-k+i}(\boldsymbol{\mathcal{I}_Z}) \geq \lambda_i(\boldsymbol{\mathcal{I}_{all}}). \tag{42}$$

Finally, based on (34) and (42), we have: For $i = 1, 2, \cdots, k$

$$\lambda_i(\boldsymbol{\mathcal{I}_Z}) \widehat{\lambda}(\boldsymbol{A_Z}) \leq \lambda_i(\boldsymbol{\mathcal{I}_{all}}) \leq \lambda_{n-k+i}(\boldsymbol{\mathcal{I}_Z}) \widetilde{\lambda}(\boldsymbol{A_Z}). \tag{43}$$

So Conclusion 2 in [1] is proved.

In order to numerically verify Conclusion 2, we perform a small feature-based Linear SLAM with 6 poses and 5 features based on two small submaps, and then compute the *log* function (monotonous function) of the eigenvalues $\lambda_i(\boldsymbol{\mathcal{I}_{all}})$ of the global joining matrix and their lower $\lambda_i(\boldsymbol{\mathcal{I}_Z})\widehat{\lambda}(\boldsymbol{A_Z})$ and upper bounds $\lambda_{n-k+i}(\boldsymbol{\mathcal{I}_Z})\widetilde{\lambda}(\boldsymbol{A_Z})$, shown in Fig. 2 .

We can see the log function (the green line) of the eigenvalues of joining FIM is bigger than their corresponding lower bound $\lambda_i(\boldsymbol{\mathcal{I}_Z})\widehat{\lambda}(\boldsymbol{A_Z})$ and smaller than their corresponding upper bound $\lambda_{n-k+i}(\boldsymbol{\mathcal{I}_Z})\widetilde{\lambda}(\boldsymbol{A_Z})$ (the red lines).

8 Appendix H: Approximate coordinate transformation for estimating the coordinates of the no-fly zones in every submap

$$
\begin{aligned}
\boldsymbol{x}^{\boldsymbol{o_h}}_{\boldsymbol{N^{sub}_i}} &= \boldsymbol{R}_{N^{sub}_i-1}{}^{\top}(\boldsymbol{x}^{\boldsymbol{o_h}}_{\boldsymbol{N^{sub}_i-1}} - \boldsymbol{x}_{N^{sub}_i-1}), \\
&\cdots \\
\boldsymbol{x}^{\boldsymbol{o_h}}_{\boldsymbol{1}} &= \boldsymbol{R}_1{}^{\top}(\boldsymbol{x}^{\boldsymbol{n}}_{\boldsymbol{h}} - \boldsymbol{x}_1),
\end{aligned}
\tag{44}
$$

where $\boldsymbol{x}^{\boldsymbol{o_h}}_{\boldsymbol{i}}$ and $\boldsymbol{x}^{\boldsymbol{n}}_{\boldsymbol{h}}$ are respectively the coordinate of the h-th no-fly zone in the i-th submap and the global map. $(\boldsymbol{x}_{N^{sub}_i-1}, \boldsymbol{R}_{N^{sub}_i-1})$ are the first pose of the robot in the i-th submap.

References

[1] Y. Chen, S. Huang and R. Fitch, “Active SLAM for Mobile Robots with Area Coverage and Obstacle Avoidance," IEEE/ASME Transactions on Mechatronics, 2018, submitted.

[2] L. Zhao, S. Huang, and G. Dissanayake, “Linear SLAM: A linear solution to the feature-based and pose graph SLAM based on submap joining," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, pp. 24-30, 2013.

[3] J. R. Magnus and H. Neudecker. “Matrix differential calculus with applications in statistics and econometrics." Wiley series in probability and mathematical statistics, 1988.

[4] V. Indelman, L. Carlone, and F. Dellaert, “Planning in the continuous domain: A generalized belief space approach for autonomous navigation in unknown environments,” The International Journal of Robotics Research, vol. 34, no. 7, pp. 849-882, 2015.