

**© 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.**

# Tree-Connectivity: Evaluating the Graphical Structure of SLAM

Kasra Khosoussi, Shoudong Huang and Gamini Dissanayake

**Abstract**—Simultaneous localization and mapping (SLAM) in robotics, and a number of related problems that arise in sensor networks are instances of estimation problems over weighted graphs. This paper studies the relation between the graphical representation of such problems and estimation-theoretic concepts such as the Cramér-Rao lower bound (CRLB) and D-optimality. We prove that the weighted number of spanning trees, as a graph connectivity metric, is closely related to the determinant of CRLB. This metric can be efficiently computed for large graphs by exploiting the sparse structure of underlying estimation problems. Our analysis is validated using experiments with publicly available pose-graph SLAM datasets.

## I. INTRODUCTION

The topology of SLAM [1] and many other estimation problems that arise in the context of sensor networks [2] can naturally be represented by graphs. The graphical representation of such problems, despite its simplicity, can be highly informative. In this paper we extend our previous work [3] and investigate the relation between the graphical representation of SLAM and the performance of sensible estimators.

Consider a pose-graph SLAM problem with a fixed set of robot poses. As we increase the number of pairwise relative measurements between the poses, the uncertainty associated to any sensible estimator will reduce. Increasing the number of measurements is equivalent to adding new edges to the corresponding graph. However, the graphical representation of SLAM is often sparse. Moreover, it is well known that maintaining and exploiting this sparsity is crucial to the efficiency and scalability of solvers [4]. Therefore, there is a delicate trade-off between the computational cost of solving SLAM and reliability of the resulting solution. This trade-off gives rise to a number of open questions such as “how many and which edges are essential to achieve a sufficiently reliable estimate?”. For this reason, it is important to identify and characterize the exact impact of graph structure on the performance of estimators, and pinpoint the key elements in the graph-theoretic and estimation-theoretic facets of SLAM that influence each other. Answering such questions enables us to maintain both sparsity and reliability by distributing the sparse set of available edges such that the graph remains sufficiently connected.

This is a relatively overlooked area. In our previous work [3] we investigated the link between the reduced graph Laplacian matrix and the Fisher information matrix. Here we extend those results and present a proof for our empirical

This work was supported by the Australian Research Council under the Discovery Project DP120102786. The authors are with the Centre for Autonomous Systems (CAS), Faculty of Engineering and Information Technology, University of Technology Sydney, Australia. Kasra.Khosoussi@student.uts.edu.au, {Shoudong.Huang, Gamini.Dissanayake}@uts.edu.au.

observations in [3]. Our approach leads to a better understanding of some of the less-studied intrinsic characteristics of SLAM as an estimation problem over graphs. We provide the metrics required to quantify the quality of SLAM datasets from a graphical perspective. Such tools can ultimately be used to optimize relevant aspects of a SLAM problem in both active and passive scenarios as shown in [3] for graph pruning and active measurement selection.

### A. Contribution

The main contributions of this paper are listed below.

- 1) Proving our conjecture in [3] regarding the impact of the number of spanning trees in planar SLAM (Theorem 4).
- 2) Extending our analysis in [3], [5] to the more realistic case of block-isotropic rotational and translational noise covariance matrices (Theorem 3, Proposition 1, Proposition 2 and Theorem 5).
- 3) Proposing tree-connectivity, a metric for evaluating graph connectivity in the context of SLAM and similar problems involving estimation over networks.
- 4) Extending our numerical results from small graphs in [3] to large-scale datasets.

An early version of this paper [5] was discussed in the workshop on “The Problem of Mobile Sensors”, held in conjunction with the Robotics: Science and Systems 2015. However, [5] relies on a restrictive assumption that all the edges of the graph have equal accuracy. In this paper we extend our theoretical results and analysis from [5] to the more general case where this restriction is removed.

### B. Notation

Bold lower-case and upper-case letters are reserved for real vectors and matrices, respectively. The eigenvalues of symmetric matrix  $\mathbf{M}$  are denoted by  $\lambda_1(\mathbf{M}) \leq \dots \leq \lambda_n(\mathbf{M})$ . Sets are shown by upper-case letters.  $|\mathcal{X}|$  denotes the cardinality of set  $\mathcal{X}$ . Natural logarithm is denoted by  $\log(\cdot)$ .  $\mathbf{1}$ ,  $\mathbf{I}$  and  $\mathbf{0}$  denote the vector of all ones, identity and zero matrix with appropriate sizes, respectively.  $\mathbf{M}_1 \succ \mathbf{M}_2$  means  $\mathbf{M}_1 - \mathbf{M}_2$  is positive-definite. Kronecker product is denoted by  $\otimes$ , and the  $d$ -expansion of matrix  $\mathbf{M}$  refers to  $\mathbf{M}_d \triangleq \mathbf{M} \otimes \mathbf{I}_d$ . Euclidean norm is denoted by  $\|\cdot\|$ . The weighted Euclidean norm of vector  $\mathbf{e}$  with matrix  $\mathbf{W} \succ \mathbf{0}$  is denoted by  $\|\mathbf{e}\|_{\mathbf{W}} \triangleq \sqrt{\mathbf{e}^T \mathbf{W} \mathbf{e}}$ .  $\text{vec}(\mathbf{q}_1, \dots, \mathbf{q}_n)$  is the column vector obtained by stacking  $\mathbf{q}_i$ 's. Moreover,  $\text{diag}(\mathbf{W}_1, \dots, \mathbf{W}_k)$  is the block-diagonal matrix with matrices  $\mathbf{W}_1, \dots, \mathbf{W}_k$  as its blocks. Finally, the set of positive semidefinite matrices in  $\mathbb{R}^{n \times n}$  is denoted by  $\mathbb{S}_+^n$ .

## II. RELATED WORKS

Olson and Kaess [6] propose the average degree of nodes  $\overline{\text{DEG}}$  as a connectivity metric for pose-graphs. They empirically observed that the minimum of the least squares

cost function (i.e., negative log-likelihood)  $f^* \triangleq f(\mathbf{x}^*)$  approaches the cost associated to the true value,  $f^\varsigma \triangleq f(\mathbf{x}^\varsigma)$ , as the average degree of the graph increases. In their view, the ratio  $\gamma \triangleq f^*/f^\varsigma$  reflects the tendency to overfit. Consequently they concluded that having a small average degree leads to overfitting.

Carlone [7] computes a conservative estimate of the basin of attraction of the maximum likelihood estimate under Gauss-Newton. This estimate is related to the smallest eigenvalue of the *reduced* Laplacian matrix of the corresponding graph. Later in [3] we noted that this quantity is closely related to the *algebraic connectivity* of the graph [8].

Our previous work [3] was among the first steps towards understanding the impact of the graphical structure of SLAM on some of the desirable attributes of the underlying estimation problem. In [3] we began where [6] left off; we showed how the expected value of  $\gamma$  can be approximated by a simple function of the average degree of graph:  $1 - 2/\overline{\text{DEG}}$ . We then showed that in specific linear-Gaussian models with isotropic noise (e.g., SLAM with known orientation), the Fisher information matrix (FIM) is proportional to the reduced Laplacian matrix of the corresponding graph. This relation enabled us to establish the missing link between the two facets of such problems. For example, determinant of the reduced Laplacian matrix gives the number of spanning trees, while determinant of the FIM is closely related to the volume of confidence ellipsoids. For 2D SLAM, we demonstrated that the number of spanning trees appears in computing the determinant of the FIM. Our empirical observations suggested that for sufficiently accurate rotational measurements, determinant of the FIM is approximately proportional to  $t(\mathcal{G})^3$ , where  $t(\mathcal{G})$  denotes the number of spanning trees in graph  $\mathcal{G}$ . Finally, we used the number of spanning trees in the graph as a metric to address the active measurement selection and edge pruning problems. We concluded that, even under more general conditions (e.g., correlated noise covariance matrices), the number of spanning trees is an effective measure for evaluating the structural quality of SLAM datasets.

In this paper we extend some of our previous results in [3] and [5] to a more general class of noise covariance matrices (i.e., block-isotropic). We provide a proof to explain our empirical observations regarding the impact of the number of spanning trees. Furthermore, we show how our metric can be used to compare the quality of the graphical structure of datasets with different number of poses.

### III. PRELIMINARIES

In this section we briefly review some basic concepts and results from spectral graph theory and estimation theory that will be used in the following sections.

#### A. Graph Preliminaries

Let  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  be a simple connected graph with  $m = |\mathcal{E}|$  edges whose vertices are labeled by  $\{0, 1, \dots, n\}$ . The degree of vertex  $v \in \mathcal{V}$  is denoted by  $\text{deg}(v)$ .  $\mathbf{A} \in \{-1, 0, 1\}^{n \times m}$  is the *reduced incidence* matrix of  $\mathcal{G}$  after anchoring an arbitrary vertex (e.g., the vertex with the zero label) and assigning arbitrary orientations to its edges (see, e.g., [2]).  $\mathbf{A}$  is obtained by removing the corresponding row from the

incidence matrix of  $\mathcal{G}$ . The *reduced Laplacian* matrix of  $\mathcal{G}$  is defined as  $\mathbf{L} \triangleq \mathbf{A}\mathbf{A}^\top$ . Note that  $\mathbf{L}$  is independent of the chosen edge orientations in  $\mathbf{A}$ .  $\mathbf{L}$  can also be obtained by removing the row and column associated to the anchor from the graph Laplacian matrix. It is well known that  $\mathbf{A}$  is full column rank and, consequently,  $\mathbf{L} \succ \mathbf{0}$ , iff  $\mathcal{G}$  is connected. The  $\ell$ -expanded reduced incidence and reduced Laplacian of  $\mathcal{G}$  are defined as  $\mathbf{A}_\ell \triangleq \mathbf{A} \otimes \mathbf{I}_\ell$  and  $\mathbf{L}_\ell \triangleq \mathbf{L} \otimes \mathbf{I}_\ell = \mathbf{A}_\ell \mathbf{A}_\ell^\top$  for  $\ell \in \mathbb{Z}_{\geq 2}$ , respectively. The reduced incidence and reduced Laplacian matrices are also defined for weighted graphs  $\mathcal{G}^w = (\mathcal{V}, \mathcal{E}, w)$  for some positive weight function  $w : \mathcal{E} \rightarrow \mathbb{R}_{>0}$ . Let us denote the  $i$ th edge with  $e_i \in \mathcal{E}$ . We define  $\mathbf{W} \triangleq \text{diag}(w(e_1), \dots, w(e_m))$ . Then  $\mathbf{A}^w \triangleq \mathbf{A}\mathbf{W}^{\frac{1}{2}}$  and  $\mathbf{L}^w \triangleq \mathbf{A}\mathbf{W}\mathbf{A}^\top = \mathbf{A}^w \mathbf{A}^{w\top}$  are the weighted counterparts of  $\mathbf{A}$  and  $\mathbf{L}$ , respectively.

A spanning tree of  $\mathcal{G}$  is a spanning subgraph of  $\mathcal{G}$  that is also a tree. Let  $t(\mathcal{G})$  denote the number of spanning trees of  $\mathcal{G}$ . Consider an arbitrary tree  $\mathcal{T}$  and the complete graph  $\mathcal{K}$ , both over  $|\mathcal{V}| = n + 1$  vertices. Then from the Cayley's formula [9] we have

$$t(\mathcal{T}) = 1 \leq t(\mathcal{G}) \leq (n + 1)^{n-1} = t(\mathcal{K}). \quad (1)$$

Trivially, if  $\mathcal{G}_1$  is a connected spanning subgraph of  $\mathcal{G}_2$ , then  $t(\mathcal{G}_1) \leq t(\mathcal{G}_2)$ . Thus  $t(\mathcal{G})$  (sometimes referred to as *graph complexity*) can be used to evaluate the graph connectivity. Kirchhoff's matrix-tree theorem provides an expression for  $t(\mathcal{G})$  in the general case [8].

**Theorem 1** (Matrix-Tree Theorem). *The number of spanning trees of  $\mathcal{G}$  is given by  $t(\mathcal{G}) = \det \mathbf{L}$ , in which  $\mathbf{L}$  is the reduced Laplacian matrix of  $\mathcal{G}$  after anchoring an arbitrary vertex.*

Theorem 1 can be generalized to weighted graphs (see, e.g., [10]).

**Theorem 2** (Weighted Matrix-Tree Theorem). *Suppose  $\mathcal{G}^w = (\mathcal{V}, \mathcal{E}, w)$  is a simple weighted graph. Let  $\mathcal{T}_{\mathcal{G}}$  denote the set of spanning trees of graph  $\mathcal{G}^w$ . For any  $\mathcal{T} \in \mathcal{T}_{\mathcal{G}}$ , let  $\mathcal{E}_{\mathcal{T}}$  be its edge set. Then we have*

$$t_w(\mathcal{G}^w) \triangleq \det(\mathbf{L}^w) = \sum_{\mathcal{T} \in \mathcal{T}_{\mathcal{G}}} \text{val}(\mathcal{T}), \quad (2)$$

where  $\text{val}(\mathcal{T}) \triangleq \prod_{e_i \in \mathcal{E}_{\mathcal{T}}} w(e_i)$ .

Note that if  $w(e_i) = 1$  for all  $e_i \in \mathcal{E}$ , then Theorem 2 reduces to Theorem 1 and  $t_w(\mathcal{G}^w) = t(\mathcal{G})$ .

#### B. Estimation Preliminaries

In SLAM and many other estimation problems, our goal is to estimate a set of unknown *parameters*  $\mathbf{x}_0^\varsigma, \mathbf{x}_1^\varsigma, \dots, \mathbf{x}_n^\varsigma$  using a sequence of measurements  $\{\mathbf{z}_{ij}\}_{(i,j) \in \mathcal{E}}$  such that  $\mathbf{z}_{ij}$  only depends on  $\mathbf{x}_i^\varsigma, \mathbf{x}_j^\varsigma$  and a random noise  $\epsilon_{ij}$ .  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  is called the graphical representation of such problem if there is a one-to-one correspondence between the vertex set  $\mathcal{V}$  and  $\{\mathbf{x}_0^\varsigma, \mathbf{x}_1^\varsigma, \dots, \mathbf{x}_n^\varsigma\}$ . Due to the relative nature of measurements in the problems we consider in this work, a standard choice is to choose an arbitrary  $\mathbf{x}_i^\varsigma$  as the *anchor* (e.g.,  $\mathbf{x}_0^\varsigma = \mathbf{0}$ ). The measurements are generated according to  $\mathbf{z} = \mathbf{h}(\mathbf{x}^\varsigma) + \epsilon$ , in which  $\mathbf{x}^\varsigma \triangleq \text{vec}(\mathbf{x}_1^\varsigma, \dots, \mathbf{x}_n^\varsigma) \in \mathcal{X}$ ,  $\mathbf{z} \in \mathcal{Z}$  is the stacked vector of measurements,  $\mathbf{h} : \mathcal{X} \rightarrow \mathcal{Z}$  is called the measurement function and  $\epsilon \sim \mathcal{N}(\mathbf{0}, \Sigma)$  is the

measurement noise. In this paper we are interested in  $\mathbf{x}^*$ , the maximum likelihood estimate (MLE) for  $\mathbf{x}^\zeta$ ,

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \|\mathbf{z} - \mathbf{h}(\mathbf{x})\|_{\Sigma}^2. \quad (3)$$

Under some regularity conditions [11], the covariance matrix of any unbiased estimator of  $\mathbf{x}$ , such as  $\hat{\mathbf{x}}$ , satisfies  $\text{Cov}[\hat{\mathbf{x}}] \succeq \mathcal{I}^{-1}(\mathbf{x}^\zeta)$ , where  $\mathcal{I}(\mathbf{x})$  is the FIM,

$$\mathcal{I}(\mathbf{x}) \triangleq \mathbb{E} \left[ \frac{\partial}{\partial \mathbf{x}} \log p(\mathbf{z}; \mathbf{x}) \frac{\partial^\top}{\partial \mathbf{x}} \log p(\mathbf{z}; \mathbf{x}) \right]. \quad (4)$$

Here the expectation is over  $\mathbf{z}$  and with respect to  $p(\mathbf{z}; \mathbf{x})$ . Recall that if  $\mathbf{M} \succeq \mathbf{N}$ , then  $M_{i,i} \geq N_{i,i}$ . Consequently the diagonal elements of CRLB are lower bounds on the variance of any unbiased estimator for each parameter. Note that FIM depends only on the true value of  $\mathbf{x}$  and  $p(\mathbf{z}; \mathbf{x})$ , and therefore is independent of any particular realization of  $\mathbf{z}$ . An unbiased estimator that achieves CRLB is called an *efficient* estimator. Under some regularity conditions, the maximum likelihood estimator is *asymptotically* unbiased, efficient, and follows a normal distribution with mean  $\mathbf{x}^\zeta$  and covariance  $\mathcal{I}(\mathbf{x}^\zeta)^{-1}$ . It is important to differentiate between the followings: (i)  $\mathcal{I}(\mathbf{x}^\zeta)$  is FIM, and therefore  $\mathcal{I}(\mathbf{x}^\zeta)^{-1}$  gives CRLB, whose trace is a lower bound on the achievable mean squared error (MSE) by any unbiased estimator, (ii)  $\mathcal{I}(\mathbf{x}^*)^{-1}$  is commonly used to approximate the covariance matrix of MLE. Note that unlike FIM, this quantity depends on  $\mathbf{x}^*$  and, consequently, the given realization of  $\mathbf{z}$ . For the measurement model introduced earlier we have

$$\log p(\mathbf{z}; \mathbf{x}) = -\frac{1}{2} \|\mathbf{z} - \mathbf{h}(\mathbf{x})\|_{\Sigma}^2 + \text{const}. \quad (5)$$

Inserting (5) into (4), computing the gradient and the expectation with respect to  $p(\mathbf{z}; \mathbf{x})$  gives  $\mathcal{I}(\mathbf{x}) = \mathbf{J}(\mathbf{x})^\top \Sigma^{-1} \mathbf{J}(\mathbf{x})$ , in which  $\mathbf{J}(\mathbf{x})$  is the Jacobian matrix of the measurement function. Hence for our model,  $\mathcal{I}(\mathbf{x}^\zeta)$  and  $\mathcal{I}(\mathbf{x}^*)$  are obtained by evaluating the Jacobian matrix at  $\mathbf{x}^\zeta$  and  $\mathbf{x}^*$ , respectively.

#### IV. FISHER INFORMATION MATRIX

In this section we derive closed-form expressions for FIM in SLAM and a number of interesting special cases that arise in the context of sensor networks. We will show that in these problems, the FIM is closely related to the graphical representation  $\mathcal{G}$ , mainly through the reduced Laplacian matrix  $\mathbf{L}$ .

##### A. Linear-SN and Compass-SLAM

In [3] we investigated two special classes of linear-Gaussian estimation problems over networks that arise in sensor networks and  $d$ -dimensional ( $d \in \{2, 3\}$ ) SLAM with a compass (i.e., known robot orientation). Linear-SN refers to estimation problems such as time-synchronization [2] with  $\mathbf{h}(\mathbf{x}) = \mathbf{A}_d^\top \mathbf{x}$ . In compass-SLAM, the measurement function is modelled by  $\mathbf{h}(\mathbf{x}) = \mathbf{R}^\top \mathbf{A}_d^\top \mathbf{x}$  in which  $\mathbf{R}$  is a given block-diagonal orthogonal matrix that contains  $d \times d$  rotation matrices. The  $i$ th block of  $\mathbf{R}$  corresponds to the heading of the pose making the  $i$ th observation. Now, under the assumption of additive Gaussian noise  $\epsilon \sim \mathcal{N}(\mathbf{0}, \Sigma)$  with  $\Sigma = \text{diag}(\sigma_1^2 \mathbf{I}_d, \dots, \sigma_m^2 \mathbf{I}_d)$ , the FIM for both problems is given by  $\mathcal{I} = \mathbf{A}_d^\top \Sigma^{-1} \mathbf{A}_d$ . It is easy to verify that  $\mathcal{I}$  in this case is in fact the  $d$ -expansion of the reduced weighted

Laplacian matrix of  $\mathcal{G}$  when each edge  $e_i \in \mathcal{E}$  is weighted according to  $w : e_i \mapsto \sigma_i^{-2}$ .

##### B. 2D SLAM

The state vector in pose-graph SLAM is composed of robot poses captured at discrete time steps along the traversed trajectory. After a permutation, this state vector can be written as  $\mathbf{x} = \text{vec}(\mathbf{p}, \boldsymbol{\theta})$  in which  $\mathbf{p} \in \mathbb{R}^{2n}$  is the stacked vector of robot positions in the plane and  $\boldsymbol{\theta} \in [-\pi, \pi]^n$  is the stacked vector of robot orientations.<sup>1</sup> Each measurement is a noisy 2D rigid body transformation between two robot poses. The measurement function, after computing the correct regularization terms for the rotational component of measurements (see [3], [7], [12]) can be expressed as

$$\mathbf{h}(\mathbf{x}) = \begin{bmatrix} \mathbf{R}^\top \mathbf{A}_2^\top & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^\top \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ \boldsymbol{\theta} \end{bmatrix}, \quad (6)$$

where  $\mathbf{R} \triangleq \text{diag}(\mathbf{R}_{\theta_{k_1}}, \mathbf{R}_{\theta_{k_2}}, \dots, \mathbf{R}_{\theta_{k_m}})$  in which  $k_i$  denotes the label of the origin of the  $i$ th observation, and  $\mathbf{R}_{\theta_{k_i}}$  is the corresponding  $2 \times 2$  rotation matrix.

1) *Isotropic Noise*: The FIM for 2D SLAM (with relative pose measurements) under isotropic rotational and translational noise covariance matrices  $\Sigma = \text{diag}(\sigma_p^2 \mathbf{I}, \sigma_\theta^2 \mathbf{I})$  is given by [5],

$$\mathcal{I}(\mathbf{x}) = \begin{bmatrix} \sigma_p^{-2} \mathbf{L}_2 & \sigma_p^{-2} \mathbf{A}_2 \boldsymbol{\Gamma} \boldsymbol{\Delta} \\ *^\top & \sigma_\theta^{-2} \mathbf{L} + \sigma_p^{-2} \boldsymbol{\Delta}^\top \boldsymbol{\Delta} \end{bmatrix} \quad (7)$$

Here  $\boldsymbol{\Gamma}$  is the following block-diagonal matrix

$$\boldsymbol{\Gamma} \triangleq \mathbf{I}_m \otimes \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad (8)$$

and for each  $e_k = (u_k, v_k) \in \mathcal{E}$ , there is a  $2 \times 1$  block in  $\boldsymbol{\Delta} \in \mathbb{R}^{2m \times n}$  that contains  $(\boldsymbol{\Delta})_{2k-1:2k, u_k} = \mathbf{p}_{v_k} - \mathbf{p}_{u_k}$ . The remaining elements in  $\boldsymbol{\Delta}$  are all zero.

**Remark 1.** The following statements hold regarding  $\mathcal{I}(\mathbf{x})$ .

- $\boldsymbol{\Gamma}^\top \boldsymbol{\Gamma} = \mathbf{I}$ .
- As noted by [7],  $\boldsymbol{\Delta}^\top \boldsymbol{\Delta}$  is a diagonal matrix with an interesting structure.  $(\boldsymbol{\Delta}^\top \boldsymbol{\Delta})_{i,i}$  is equal to the sum of squared distances between the  $i$ th robot pose, and every node observed by it,

$$(\boldsymbol{\Delta}^\top \boldsymbol{\Delta})_{i,i} = \sum_{j \in \mathcal{S}(i)} \|\mathbf{p}_i - \mathbf{p}_j\|^2. \quad (9)$$

Here  $\mathcal{S}(v)$  is the set of nodes observed by  $v \in \mathcal{V}$ .

2) *Block-Isotropic Noise*: A more general noise model emerges when rotational and translational noise covariance matrices are *block-isotropic*. Consider  $\Sigma = \text{diag}(\Sigma_p, \Sigma_\theta)$  in which,

$$\begin{aligned} \Sigma_p &= \text{diag}(\sigma_{p_1}^2 \mathbf{I}_2, \dots, \sigma_{p_m}^2 \mathbf{I}_2), \\ \Sigma_\theta &= \text{diag}(\sigma_{\theta_1}^2, \dots, \sigma_{\theta_m}^2). \end{aligned} \quad (10)$$

<sup>1</sup>In feature-based SLAM problems,  $\mathbf{p}$  also contains the position of landmarks.

Unlike the previous noise model, this model allows measurements to have different “precisions”. In this case the FIM is given by

$$\mathcal{I}^w(\mathbf{x}) = \begin{bmatrix} \mathbf{L}_2^{w_p} & \mathbf{A}_2^{w_p} \mathbf{\Gamma} \mathbf{\Delta}_{w_p} \\ *^\top & \mathbf{L}^{w_\theta} + \mathbf{\Delta}_{w_p}^\top \mathbf{\Delta}_{w_p} \end{bmatrix}, \quad (11)$$

where  $\mathbf{L}_2^{w_p}$  and  $\mathbf{A}_2^{w_p}$  are, respectively, the 2-expansions of the reduced weighted Laplacian  $\mathbf{L}^{w_p}$  and reduced weighted incidence  $\mathbf{A}^{w_p}$  with  $w_p : e_i \mapsto \sigma_{p_i}^{-2}$ ,  $\mathbf{L}^{w_\theta}$  is the reduced weighted Laplacian with  $w_\theta : e_i \mapsto \sigma_{\theta_i}^{-2}$ , and  $\mathbf{\Delta}_{w_p} \triangleq \mathbf{\Sigma}_p^{-\frac{1}{2}} \mathbf{\Delta}$ .

**Remark 2.**  $\mathbf{\Delta}_{w_p}^\top \mathbf{\Delta}_{w_p}$  is a diagonal matrix with the following structure,

$$(\mathbf{\Delta}_{w_p}^\top \mathbf{\Delta}_{w_p})_{i,i} = \sum_{j \in \mathcal{S}(i)} \sigma_{p_{k_{ij}}}^{-2} \|\mathbf{p}_i - \mathbf{p}_j\|^2, \quad (12)$$

where  $k_{ij}$  is the index of the edge  $(i,j)$ .

## V. MAIN RESULTS

The log-determinant of positive-definite covariance matrices can be viewed as a measure of “uncertainty”. For instance, in multivariate Gaussian distributions, the determinant of the covariance matrix is proportional to the square of the hypervolume of confidence ellipsoids. Moreover, from an information-theoretic standpoint, the log-determinant of the covariance matrix of a Gaussian distribution is proportional to its differential entropy up to an additive constant. In the context of optimal experimental design, a design is called D-optimal (determinant-optimal) if it attains the minimum log det of estimation error covariance matrix [13]. It is easy to show that the log det of CRLB is a lower bound for the log det of the covariance matrix of any unbiased estimator. The D-criterion, due to the reasons mentioned above, is a popular choice and has been frequently used as the design criterion in many design problems in robotics, including sensor selection [14] and active SLAM [15]. In this section we define a connectivity metric based on the number of spanning trees, and reveal its impact on the D-criterion.

### A. Criterion: Tree-Connectivity

As we saw earlier,  $t(\mathcal{G})$  is a sensible measure of graph connectivity. It is practically intractable to work directly with  $t(\mathcal{G})$  in large graphs, as it grows too quickly (see [3]). Here we define a closely-related metric based on  $t(\mathcal{G})$  to measure graph connectivity.

**Definition 1.** We formally define the tree-connectivity of graph  $\mathcal{G}$ ,  $\tau(\mathcal{G})$ , as

$$\tau(\mathcal{G}) \triangleq \log t(\mathcal{G}) \stackrel{\text{Theorem 1}}{=} \log \det \mathbf{L}. \quad (13)$$

Similarly, for weighted graph  $\mathcal{G}^w = (\mathcal{V}, \mathcal{E}, w)$  we define the weighted tree-connectivity as

$$\tau_w(\mathcal{G}^w) \triangleq \log t_w(\mathcal{G}^w) \stackrel{\text{Theorem 2}}{=} \log \det \mathbf{L}^w. \quad (14)$$

A key advantage of  $\tau(\mathcal{G})$  over  $t(\mathcal{G})$  is that it can be efficiently computed for large sparse graphs. Algorithm 1 exploits the sparse structure of  $\mathcal{E}$  to compute the tree-connectivity efficiently using the Cholesky decomposition of

**Algorithm 1**  $\tau(\mathcal{G}) \triangleq \log \det(\mathbf{L})$  for sparse symmetric  $\mathbf{L} \succ \mathbf{0}$

---

```

1: function TREE-CONNECTIVITY( $\mathcal{G}$ )
2:   Choose a fill-reducing permutation heuristic  $\mathbf{\Pi}$ .
3:   Compute  $\mathbf{K}$ , the sparse Cholesky factor of  $\mathbf{\Pi} \mathbf{L} \mathbf{\Pi}^\top$ .
4:   return  $\tau(\mathcal{G}) = 2 \sum_i \log \mathbf{K}_{i,i}$ .
5: end function

```

---

$\mathbf{L}$  (or  $\mathbf{L}^w$ ). Here we use the fact that  $\det(\mathbf{K}) = \prod_i \mathbf{K}_{i,i}$  for triangular  $\mathbf{K}$ . For dense graphs, this approach requires cubic time in  $|\mathcal{V}|$ , while in sparse practical scenarios that arise in the context of robotics and sensor networks, Algorithm 1 performs much faster given a sufficiently good fill-reducing permutation heuristic.

### B. Linear-SN and Compass-SLAM

Now we are ready to investigate the link between the (weighted) tree-connectivity and the log-determinant of FIM. The following proposition straightforwardly follows from Section IV-A.

**Theorem 3.** In linear-SN and compass-SLAM, the maximum likelihood estimator is unbiased and efficient. Furthermore,

$$\log \det(\text{Cov}[\mathbf{x}^*]) = -\log \det \mathbf{L}_d^w = -d\tau_w(\mathcal{G}^w) \quad (15)$$

in which  $\mathbf{L}_d^w$  is the  $d$ -expansion of the reduced weighted Laplacian matrix of weighted graph  $\mathcal{G}^w = (\mathcal{V}, \mathcal{E}, w)$  with the weight function  $w(e_i) = \sigma_i^{-2}$  (for  $i = 1, \dots, m$ ).

Theorem 3 ensures that in linear-SN and compass-SLAM and under the specified assumptions, (i) minimizing the differential entropy of  $\mathbf{x}^*$ , (ii) minimizing the volume of uncertainty ellipsoids, or (iii) finding the D-optimal design, are all equivalent to maximizing the (weighted) tree-connectivity of the underlying graph.

### C. 2D SLAM

1) *Isotropic Noise:* In [3] we made an empirical observation regarding the impact of the number of spanning trees on CRLB and the covariance matrix of MLE in 2D SLAM. More recently, we presented a proof in [5] for our earlier observations. Here we review one of our main results. See [5] for the proof and other propositions.

Let  $\mathcal{I}_o(\mathbf{x})$  be the Fisher information matrix of the odometry subgraph, composed of robot poses and odometry edges. Let  $\Phi_{\text{inf}}(\mathcal{G}) \triangleq \log \det \mathcal{I}(\mathbf{x}) - \log \det \mathcal{I}_o(\mathbf{x})$ . Informally speaking,  $\Phi_{\text{inf}}(\mathcal{G})$  can be seen as the “information” gained by having loop-closure edges as compared with the pure dead reckoning scenario. We also define  $\alpha \triangleq \sigma_\theta / \sigma_p$ ,  $\text{dist}_{\text{out,max}}^2 \triangleq \max_i (\mathbf{\Delta}^\top \mathbf{\Delta})_{i,i}$  and  $\psi \triangleq \alpha^2 \text{dist}_{\text{out,max}}^2$ .

**Theorem 4.** For isotropic rotational and translational noise covariance matrices  $\mathbf{\Sigma} = \text{diag}(\sigma_p^2 \mathbf{I}, \sigma_\theta^2 \mathbf{I})$  we have  $\lim_{\psi \rightarrow 0^+} \Phi_{\text{inf}}(\mathcal{G}) = 3\tau(\mathcal{G})$ , in which  $\psi$  and  $\Phi_{\text{inf}}(\mathcal{G})$  are defined above.

Theorem 4 offers an explanation for our empirical observations in [3]. Despite the asymptotic nature of this result, as shown in [5],  $\Phi_{\text{inf}}(\mathcal{G})$  can also be expressed in terms of  $\mathbf{L}$ ,  $\alpha$  and  $\mathbf{\Delta}$  for any value of  $\psi$ . Note that these factors represent topology, sensor quality and geometry, respectively.



2) *Block-Isotropic Noise*: Now we extend Theorem 4 to the more general and realistic case of block-isotropic rotational and translational noise covariance matrices defined in (10). See Appendix I for the proofs.

**Proposition 1.** For the noise model defined in (10) we have

$$\log \det \mathcal{I}^w(\mathbf{x}) = 2\tau_{w_p}(\mathcal{G}^{w_p}) + \log \det(\mathbf{L}^{w_\theta} + \mathbf{E}) \quad (16)$$

where  $\mathbf{E} \triangleq \Delta_{w_p}^\top \mathbf{P}_{w_p}^\perp \Delta_{w_p}$  in which

$$\mathbf{P}_{w_p}^\perp \triangleq \mathbf{I} - \mathbf{P}_{w_p}, \quad (17)$$

$$\mathbf{P}_{w_p} \triangleq \Gamma^\top \mathbf{A}_2^{w_p \top} \mathbf{L}_2^{w_p -1} \mathbf{A}_2^{w_p} \Gamma. \quad (18)$$

Furthermore,  $\mathbf{P}_{w_p}$  and  $\mathbf{P}_{w_p}^\perp$  are orthogonal projections onto range( $\Gamma^\top \mathbf{A}_2^{w_p \top}$ ) and null( $\mathbf{A}_2^{w_p} \Gamma$ ), respectively.

**Proposition 2.** For the noise model defined in (10) we have  $\log \det \mathcal{I}^w(\mathbf{x}) \in [\ell, u]$  where

$$\ell \triangleq 2\tau_{w_p}(\mathcal{G}^{w_p}) + \tau_{w_\theta}(\mathcal{G}^{w_\theta}) \quad (19)$$

$$u \triangleq 2\tau_{w_p}(\mathcal{G}^{w_p}) + \sum_{i=1}^n \log(\lambda_i(\mathbf{L}^{w_\theta}) + \psi^{w_p}), \quad (20)$$

in which  $\psi^{w_p} \triangleq \max_i(\Delta_{w_p}^\top \Delta_{w_p})_{i,i}$ .

**Theorem 5.** For the noise model defined in (10) we have

$$\lim_{\psi^{w_p} \rightarrow 0^+} \log \det \mathcal{I}^w(\mathbf{x}) = 2\tau_{w_p}(\mathcal{G}^{w_p}) + \tau_{w_\theta}(\mathcal{G}^{w_\theta}). \quad (21)$$

#### D. Normalized Tree-Connectivity

Comparing the graphical structure of two SLAM problems based on the tree-connectivity of their graphs is meaningful only if they have the same number of vertices (i.e., same number of robot poses in pose-graphs). A remedy to this limitation is to somehow normalize tree-connectivity.

**Definition 2** (Normalized tree-connectivity). Let  $\mathcal{G}_n$  be a graph with  $n$  vertices. We define the normalized tree-connectivity of graph  $\mathcal{G}_n$ , denoted by  $\bar{\tau}(\mathcal{G}_n)$ , as

$$\bar{\tau}(\mathcal{G}_n) \triangleq \frac{\tau(\mathcal{G}_n)}{\tau(\mathcal{K}_n)} \stackrel{(1)}{=} \frac{\tau(\mathcal{G}_n)}{(n-2)\log(n)}. \quad (22)$$

where  $\mathcal{K}_n$  is the complete graph with  $n$  vertices.

Therefore, to any simple connected graph  $\mathcal{G}$ ,  $\bar{\tau}(\mathcal{G}) \in [0,1]$  assigns a score that reflects the tree-connectivity of  $\mathcal{G}$  relative to the tree-connectivity of the complete graph with the same number of vertices. Theorem 4 ensures that under the conditions stated earlier we have,

$$\lim_{\psi \rightarrow 0^+} \frac{\Phi_{\text{inf}}(\mathcal{G})}{\Phi_{\text{inf}}(\mathcal{K})} = \bar{\tau}(\mathcal{G}). \quad (23)$$

Hence the normalized tree-connectivity can be seen as the ratio of the information gained relative to dead reckoning, between the realized graph  $\mathcal{G}$  and the complete graph  $\mathcal{K}$  (i.e., the graph over the same vertex set that contains all possible loop closures). Note that normalizing the weighted tree-connectivity  $\tau_w(\mathcal{G}^w)$  requires making (arbitrary) assumptions about the weights of the new edges in the corresponding complete graph.

## VI. NUMERICAL EXAMPLES

In this section we present numerical results using publicly available datasets with three objectives. First, we validate the results presented in Section V, in particular Theorem 4 and Theorem 5. Our second objective is to see how small  $\psi$  and  $\psi^{w_p}$  should be in practice for  $\log \det \mathcal{I}(\mathbf{x})$  and  $\log \det \mathcal{I}^w(\mathbf{x})$  to be sufficiently close to their limit values. Finally we evaluate and compare several existing SLAM datasets according to our proposed metric. We first need to define a measure of closeness to the limit values. To validate Theorem 4, we use the relative error (RE) as defined below,

$$\text{RE} \triangleq \left| \frac{\log \det \mathcal{I}(\mathbf{x}) - \mathcal{L}}{\log \det \mathcal{I}(\mathbf{x})} \right|, \quad (24)$$

in which,

$$\mathcal{L} \triangleq \lim_{\psi \rightarrow 0^+} \log \det \mathcal{I}(\mathbf{x}) = 3\tau(\mathcal{G}) + \log \det \mathcal{I}_o(\mathbf{x}). \quad (25)$$

Similarly, to validate Theorem 5 we use

$$\text{RE}^w \triangleq \left| \frac{\log \det \mathcal{I}^w(\mathbf{x}) - \mathcal{L}^w}{\log \det \mathcal{I}^w(\mathbf{x})} \right|, \quad (26)$$

where  $\mathcal{L}^w \triangleq 2\tau_{w_p}(\mathcal{G}^{w_p}) + \tau_{w_\theta}(\mathcal{G}^{w_\theta})$ . The datasets used in this section are all publicly available 2D pose-graph SLAM datasets. Those with arbitrary covariance matrices are modified to be compatible with our assumption about the structure of noise covariance matrix (e.g., isotropic or block-isotropic). Manhattan [16] is the only dataset for which we had a ground truth. Therefore for other datasets, RE is evaluated at the solution obtained by Gauss-Newton initiated from the spanning tree initial guess [17], i.e., (supposedly)  $\mathbf{x}^*$ .

#### A. Results

Figure 1 shows how RE evolves with respect to scaling  $\psi$ . Scaling  $\psi$  can be done by scaling either  $\alpha^2$  or  $\mathbf{p}^5$ . Figure 1 is obtained by scaling  $\psi$  according to  $\beta\psi_{\text{orig}}$  in which  $\psi_{\text{orig}} = 29$  is the original value of  $\psi$  in the Manhattan dataset (see Table I). It is clear that as  $\psi \rightarrow 0^+$ , the log-determinant of FIM converges to  $\mathcal{L}$  as predicted by Theorem 4.

We performed a similar experiment using the Manhattan dataset to validate Theorem 5. To make the original isotropic noise of Manhattan compatible with the (more general case of) block-isotropic noise (10), we add random perturbations to the original noise variances. In Figure 2 we scale  $\psi^{w_p}$  according to  $\beta\psi_0^{w_p}$  ( $\psi_0^{w_p} \approx 6.21 \times 10^4$ ). The FIM is evaluated at the ground truth. It is evident that for relatively small  $\psi^{w_p}$ ,  $\mathcal{L}^w$  provides a good estimate for the log-determinant of FIM as promised by Theorem 5.

The normalized tree-connectivity for several publicly available datasets is shown in Table I. The entries in Table I are sorted according to their normalized tree-connectivity metric  $\bar{\tau}(\mathcal{G})$ . These results together with Figure 1 and Figure 2 indicate that, in practice, RE is relatively small even when  $\psi$  is far from being negligible. In such cases, the log-determinant of FIM (and, CRLB) is entirely characterized by the tree-connectivity of graph. The large RE in the case of Lincoln dataset (highlighted in red) is partially due to the fact that Gauss-Newton has failed to converge to the true  $\mathbf{x}^*$ . The average degree of graph  $\overline{\text{DEG}}$  [3], [6] is not

TABLE I: A summary of results for publicly available 2D pose-graph datasets, sorted according to  $\bar{\tau}(\mathcal{G})$ .

Dataset	$\bar{\tau}(\mathcal{G})$	DEG	$\alpha^2$	$\text{dist}_{\text{out,max}}^2$	$\psi$	RE (%)
M10K	0.2241	12.8622	0.0400	38.9251	1.5570	0.07
Intel	0.1329	3.8918	0.1000	13.9018	1.3902	0.06
City10K	0.1230	4.1374	0.2500	48.9235	12.2309	0.51
Lincoln	0.1155	3.9040	1.0000	50.6824	50.6824	58.00
Manhattan	0.0950	3.1160	1.0000	29.0000	29.0000	1.00
RingCity	0.0585	2.7624	3.0462	2.5370	7.7281	1.08
Freiburg	0.0421	2.4611	0.1000	4.1568	0.4157	0.04
CSAIL	0.0263	2.2411	0.2000	35.7974	7.1595	0.12

sophisticated enough to differentiate between different graph structures with the same number of edges and vertices. This case is depicted in Table I where the highlighted entries in blue show an inconsistency between the rankings based on the normalized tree-connectivity and average degree.

Using Algorithm 1 we can compute  $\tau(\mathcal{G})$  efficiently in large sparse graphs. Figure 3 shows  $\log \det \mathcal{I}(\mathbf{x}^*)$  as a function of  $\tau(\mathcal{G})$  for more than 44,000 random spanning subgraphs of the Intel dataset. Every subgraph has the same vertex set as the original dataset and contains all the odometry edges. Additionally, each subgraph contains a random subset of loop-closure edges of the original dataset. For each possible number of loop-closures, we generated 50 random spanning subgraphs. The predicted value is  $\mathcal{L}$ . Figure 3 indicates that in this case, tree-connectivity almost entirely characterizes the performance of the maximum likelihood estimator and CRLB in terms of log-determinant of the covariance matrices as predicted by our theoretical results.

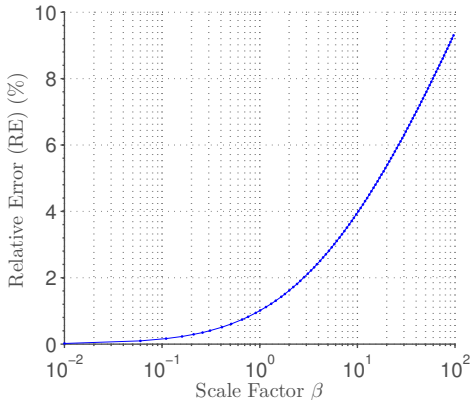


Fig. 1: RE evaluated at  $\mathbf{x}^s$  as a function of scale parameter  $\beta$  for Manhattan. Here  $\psi = \beta\psi_{\text{orig}}$  in which  $\psi_{\text{orig}} = 29$  is the value of  $\psi$  in Manhattan dataset (see Table I). This can be done by scaling either  $\alpha^2$  or  $\mathbf{p}^s$ . Note the logarithmic scale of the horizontal axis.

## VII. CONCLUSION

The graphical structure of SLAM can provide a compact, but rich representation of the underlying estimation problem. As a result, valuable information can be efficiently extracted from the graphical representation of SLAM even *before* solving the underlying estimation problem. This makes decision making (e.g., in active SLAM) based on the graph robust to common convergence issues (e.g., local minima).

This paper is an extension to our previous works [3], [5], where we provided insights into the importance of the

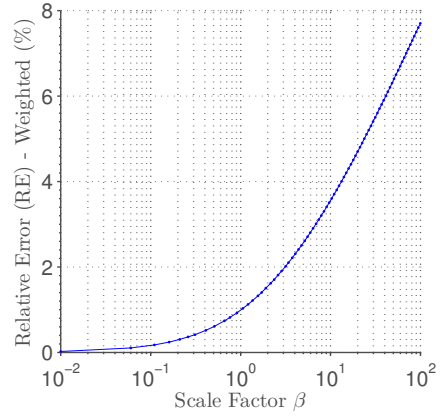


Fig. 2:  $\text{RE}^w$  evaluated at  $\mathbf{x}^s$  as a function of scale parameter  $\beta$  for Manhattan. Here  $\psi^{w_p} = \beta\psi_0^{w_p}$  in which  $\psi_0^{w_p} \approx 6.21 \times 10^4$ . In this experiment different edges have different noise variances. Note the logarithmic scale of the horizontal axis.

graphical representation of SLAM and similar problems. We first presented a proof for our empirical observations in [3], and extended our analysis to more realistic cases in which edges are evaluated based on both their topological importance and information content. We also proposed tree-connectivity as a metric to quantify the connectivity of estimation graphs. It was demonstrated how our metric is related to the log-determinant of (i) covariance matrix of the maximum likelihood estimator and (ii) Cramér-Rao bound for the highest achievable “accuracy” by unbiased estimators. Computing this metric requires performing a (sparse) Cholesky factorization on the reduced Laplacian matrix. Numerical results using publicly available datasets confirmed the validity of our analysis. It was observed that in many practical scenarios,  $\log \det \mathcal{I}(\mathbf{x})$  and  $\log \det \mathcal{I}^w(\mathbf{x})$  are relatively close to their limit values given by Theorem 4 and Theorem 5, respectively.

Our results can be straightforwardly generalized to 2D feature-based SLAM problems. Extensive empirical observations suggest that our analysis can be also extended to 3D SLAM with  $\text{SE}(3)$  relative-pose measurements. We will consider this extension in our future work.

## APPENDIX I PROOFS

We need the following lemma before stating the proofs. A proof for this Lemma 1 is provided in [5].

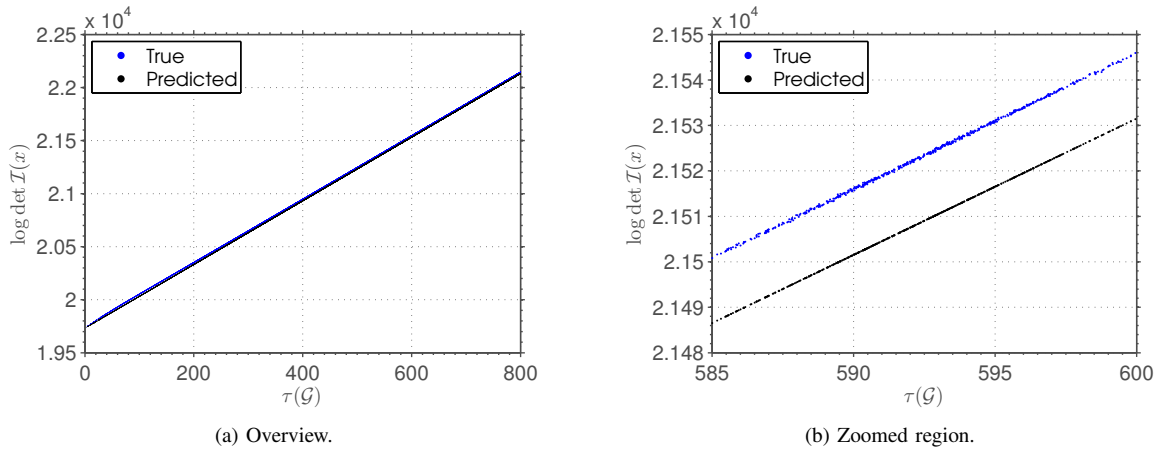


Fig. 3:  $\log \det \mathcal{I}(\mathbf{x}^*)$  as a function of  $\tau(\mathcal{G})$  for over 44,000 randomly generated spanning subgraphs of Intel dataset. Here  $\log \det \mathcal{I}(\mathbf{x})$  is evaluated at the maximum likelihood estimate of the original dataset. The prediction is based on the limit value provided by Theorem 4.

**Lemma 1.** For any two (symmetric)  $\mathbf{N}, \mathbf{M} \in \mathbb{S}_+^n$  we have

$$\det(\mathbf{M} + \mathbf{N}) \geq \det(\mathbf{M}). \quad (27)$$

*Proof of Proposition 1.* Applying the Schur's determinant formula [18] on the top-left block of (11) and the fact that  $\log \det \mathbf{L}_2^{w_p} = 2 \log \det \mathbf{L}^{w_p} = 2\tau_{w_p}(\mathcal{G}^{w_p})$  yields (16). Also note that  $\mathbf{P}_{w_p} = \mathbf{Q}\mathbf{Q}^\dagger$  in which  $\mathbf{Q}^\dagger$  is the Moore-Penrose pseudoinverse of  $\mathbf{Q} \triangleq \mathbf{\Gamma}^\top \mathbf{A}_2^{w_p \top}$ .  $\square$

*Proof of Proposition 2.* First note that  $\mathbf{P}_{w_p} \succeq \mathbf{0}$  and  $\mathbf{P}_{w_p}^\perp \succeq \mathbf{0}$  since their spectrum consists of zeros and ones. For the lower bound we start from (16) and apply Lemma 1,

$$\log \det \mathcal{I}^w(\mathbf{x}) = 2\tau_{w_p}(\mathcal{G}^{w_p}) + \log \det(\mathbf{L}^{w_\theta} + \mathbf{E}) \quad (28)$$

$$\geq 2\tau_{w_p}(\mathcal{G}^{w_p}) + \log \det \mathbf{L}^{w_\theta} \quad (29)$$

$$= 2\tau_{w_p}(\mathcal{G}^{w_p}) + \tau_{w_\theta}(\mathcal{G}^{w_\theta}) \quad (30)$$

where  $\mathbf{E} \triangleq \mathbf{\Delta}_{w_p}^\top \mathbf{P}_{w_p}^\perp \mathbf{\Delta}_{w_p}$ . The upper bound also results from Lemma 1 as shown below.

$$\begin{aligned} \log \det \mathcal{I}^w(\mathbf{x}) &= 2\tau_{w_p}(\mathcal{G}^{w_p}) + \log \det(\mathbf{L}^{w_\theta} + \mathbf{E}) \\ &\leq 2\tau_{w_p}(\mathcal{G}^{w_p}) + \log \det(\mathbf{L}^{w_\theta} + \mathbf{\Delta}_{w_p}^\top \mathbf{\Delta}_{w_p}) \\ &\leq 2\tau_{w_p}(\mathcal{G}^{w_p}) + \log \det(\mathbf{L}^{w_\theta} + \psi^{w_p} \mathbf{I}) \\ &= 2\tau_{w_p}(\mathcal{G}^{w_p}) + \sum_{i=1}^n \log(\lambda_i(\mathbf{L}^{w_\theta}) + \psi^{w_p}). \end{aligned}$$

The second and third lines above follow from applying Lemma 1 on

$$\log \det(\underbrace{\mathbf{L}^{w_\theta} + \mathbf{\Delta}_{w_p}^\top \mathbf{P}_{w_p}^\perp \mathbf{\Delta}_{w_p}}_{> \mathbf{0}} + \underbrace{\mathbf{\Delta}_{w_p}^\top \mathbf{P}_{w_p} \mathbf{\Delta}_{w_p}}_{\geq \mathbf{0}}) \quad (31)$$

and

$$\log \det(\underbrace{\mathbf{L}^{w_\theta} + \mathbf{\Delta}_{w_p}^\top \mathbf{\Delta}_{w_p}}_{> \mathbf{0}} + \underbrace{\psi^{w_p} \mathbf{I} - \mathbf{\Delta}_{w_p}^\top \mathbf{\Delta}_{w_p}}_{\geq \mathbf{0}}), \quad (32)$$

respectively (see Remark 2).  $\square$

*Proof of Theorem 5.* This result directly follows from Proposition 2 and the squeeze theorem, since for  $\psi^{w_p} \rightarrow 0^+$ ,

the upper bound in Proposition 2 approaches the lower bound.  $\square$

## REFERENCES

- [1] S. Thrun and M. Montemerlo, "The graph SLAM algorithm with applications to large-scale mapping of urban structures," *The International Journal of Robotics Research*, vol. 25, no. 5-6, p. 403, 2006.
- [2] P. Barooah and J. P. Hespanha, "Estimation on graphs from relative measurements," *Control Systems, IEEE*, vol. 27, no. 4, pp. 57–74, 2007.
- [3] K. Khosoussi, S. Huang, and G. Dissanayake, "Novel insights into the impact of graph structure on SLAM," in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2014, 2014, pp. 2707–2714.
- [4] F. Dellaert and M. Kaess, "Square Root SAM: Simultaneous localization and mapping via square root information smoothing," *Intl. J. of Robotics Research, IJRR*, vol. 25, no. 12, pp. 1181–1204, Dec. 2006.
- [5] K. Khosoussi, S. Huang, and G. Dissanayake, "Good, bad and ugly graphs for SLAM," in *RSS Workshop on The Problem of Mobile Sensors*, 2015.
- [6] E. Olson and M. Kaess, "Evaluating the performance of map optimization algorithms," in *RSS Workshop on Good Experimental Methodology in Robotics*, 2009, p. 40.
- [7] L. Carlone, "Convergence analysis of pose graph optimization via Gauss-Newton methods," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, May 2013.
- [8] C. Godsil and G. Royle, *Algebraic graph theory*, ser. Graduate Texts in Mathematics Series. Springer London, Limited, 2001.
- [9] A. Cayley, *The collected mathematical papers of Arthur Cayley*. The University Press, 1894, vol. 7.
- [10] M. Mesbahi and M. Egerstedt, *Graph theoretic methods in multiagent networks*. Princeton University Press, 2010.
- [11] H. W. Sorenson, *Parameter estimation: principles and problems*, ser. Control and systems theory. M. Dekker, 1980.
- [12] L. Carlone, R. Aragues, J. A. Castellanos, and B. Bona, "A fast and accurate approximation for planar pose graph optimization," *The International Journal of Robotics Research*, vol. 33, pp. 965 – 987, 06/2014 2014.
- [13] F. Pukelsheim, *Optimal design of experiments*. SIAM, 1993, vol. 50.
- [14] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *Signal Processing, IEEE Transactions on*, vol. 57, no. 2, pp. 451–462, 2009.
- [15] T. A. Vidal-Calleja, A. Sanfeliu, and J. Andrade-Cetto, "Action selection for single-camera SLAM," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 40, no. 6, pp. 1567–1581, 2010.
- [16] E. Olson, "Robust and efficient robotic mapping," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, June 2008.
- [17] K. Konolige, G. Grisetti, R. Kummerle, W. Burgard, B. Limketkai, and R. Vincent, "Efficient sparse pose adjustment for 2D mapping," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE, 2010, pp. 22–29.
- [18] C. D. Meyer, *Matrix analysis and applied linear algebra*. SIAM, 2000.