Good, Bad and Ugly Graphs for SLAM

Kasra Khosoussi, Shoudong Huang, Gamini Dissanayake
Centre for Autonomous Systems, University of Technology Sydney
{kasra.khosoussi, shoudong.huang, gamini.dissanayake}@uts.edu.au

***Abstract*—This paper studies the relation between the graphical structure of simultaneous localization and mapping (SLAM) and the D-criterion. In particular, we define a new connectivity metric for graphs based on the number of spanning trees, and prove that our metric is closely related to the log-determinant of (i) the covariance matrix of the maximum likelihood estimator, and (ii) the Cramér-Rao bound on the highest achievable "accuracy" by unbiased estimators. We also propose an efficient algorithm to compute our metric in large graphs. Our theoretical results are validated using publicly available datasets.**

I. Introduction

SLAM can be naturally represented by a graph whose vertex set represents the set of robot poses (and landmarks) and pairwise measurements are encoded in edges between the corresponding vertices [14]. Other instances of such estimation problems over graphs can be found in the context of sensor networks [1]. Such graphs encapsulate some of the key attributes of the underlying estimation problems. In this paper we extend our previous work [7] and investigate the impact of graph structure in SLAM datasets on the performance of sensible estimators.

Consider a pose-graph SLAM problem with a given 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 due to various reasons such as limited sensing range. It is well known that maintaining and exploiting this sparsity is crucial to the scalability of solvers [4]. Therefore there is a delicate trade-off between the computational cost of solving SLAM and the reliability of the solution. This trade-off gives rise to a number of overlooked questions such as "how many and which edges are needed to have a sufficiently reliable estimate?". For this reason, it is important to identify and characterize the exact impact of graph structure on the performance of sensible 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 [7] 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 some of our empirical observations in [7]. 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 (see [7]).

A. Contribution

The main contributions of this paper are listed below.

1) Proposing $\bar{\tau}(\mathcal{G})$, a metric for evaluating graph connectivity in the context of SLAM and similar problems involving estimation over networks.
2) Proposing an efficient algorithm for computing $\bar{\tau}(\mathcal{G})$ in large SLAM problems by exploiting the sparsity of the edge set (Algorithm 1).
3) Proving our conjecture in [7] regarding the impact of the number of spanning trees in planar SLAM (Propositions 1, 2 and 3).
4) Extending some of our previous results in [7] to the case in which different measurements can have different (but isotropic for relative position and orientation) noise covariance matrices (Proposition 1).
5) Extending our numerical results from small graphs in [7] to large-scale datasets.

B. Outline

In Section II we discuss some of the closely related works. Section III formulates the problem and reviews preliminary concepts from estimation theory and spectral graph theory. In Section IV we compute the Fisher information matrix for SLAM and two other related problems. In Section V we define our connectivity metric and present our main theoretical results regarding the impact of our metric on the D-optimality of a SLAM problem. Numerical results on both synthetic and real datasets are reported in Section VI. Finally, Section VII concludes the paper. The proofs are given in the Appendix.

C. Notation

Throughout this paper bold lower-case and upper-case letters are reserved for (real) vectors and matrices, respectively. The eigenvalues of $\mathbf{M}$ are denoted by $\lambda_1(\mathbf{M}) \leq \cdots \leq \lambda_n(\mathbf{M})$. 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. $\mathbf{S}_1 \succ \mathbf{S}_2$ means $\mathbf{S}_1 - \mathbf{S}_2$ is positive-definite. Kronecker product is denoted by $\otimes$. Euclidean norm is denoted by $\| \cdot \|$. The weighted Euclidean norm of vector $\mathbf{e}$ with matrix $\mathbf{W} \succ \mathbf{0}$

is denoted by $\|\mathbf{e}\|_{\mathbf{W}} \triangleq \sqrt{\mathbf{e}^\top \mathbf{W}\, \mathbf{e}}$. Let $\mathrm{vec}(\mathbf{q}_1,\ldots,\mathbf{q}_n)$ denote the column vector obtained by stacking $\mathbf{q}_i$'s. Moreover $\mathrm{diag}(\mathbf{W}_1,\ldots,\mathbf{W}_k)$ is the block-diagonal matrix with matrices $\mathbf{W}_1,\ldots,\mathbf{W}_k$ as blocks on its main diagonal. Finally $\mathbb{S}^n_+$ and $\mathbb{S}^n_{++}$ denote the positive semidefinite cone and positive definite cone in $\mathbb{R}^{n\times n}$, respectively.

II. Related Works

In [11], Olson and Kaess propose the average degree of nodes as a metric to be used in evaluating connectivity in pose-graphs. They empirically observed that the minimum of cost function $f^\star \triangleq f(\mathbf{x}^\star)$ approaches the cost associated to the ground truth, $f^\varsigma \triangleq f(\mathbf{x}^\varsigma)$, as the average degree of the graph increases. In their view, the ratio $\gamma \triangleq f^\star/f^\varsigma$ reflects the tendency to overfit. Consequently they concluded that having a small average degree leads to overfitting.

Carlone et al. compute a conservative estimate of the basin of attraction of the maximum likelihood estimate under Gauss-Newton [2]. They show that their estimate is related to the smallest eigenvalue of the *reduced* Laplacian matrix of the corresponding graph. Later in [7] we noted that this quantity is closely related to the *algebraic connectivity* of the graph [5].

Our previous work [7] was the first step towards understanding the impact of the graphical structure of SLAM on some of the desirable attributes of the underlying estimation problem. In [7] we began where [11] left off; we showed how the expected value of γ can be approximated by a function of the average degree of graph. It is worth noting that this quantity is closely related to the trace of the reduced graph Laplacian. As shown in [7], our approximation is consistent with the empirical results reported in [11]. We then showed that in particular linear-Gaussian models with isotropic noise (e.g., SLAM with known orientation), the Fisher information matrix 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 we noted that the determinant of the reduced Laplacian matrix gives the number of spanning trees, while determinant of the Fisher information matrix is closely related to the volume of confidence ellipsoids. Furthermore, for such problems we showed that the largest eigenvalue of covariance matrix (CRLB) is bounded from below by $\kappa\lambda_2^{-1}(\mathcal{G})$ in which $\lambda_2(\mathcal{G})$ is the algebraic connectivity of the graph and κ is a constant related to noise variances. For planar SLAM we showed how the number of spanning trees of the graph appears in computing the determinant of the Fisher information matrix. Our empirical observations suggested that for sufficiently accurate rotational measurements, determinant of the Fisher information matrix is approximately proportional to $t(\mathcal{G})^3$, where $t(\mathcal{G})$ denotes the number of spanning trees in graph $\mathcal{G}$. Our observations were based on small portions of both synthetic and real datasets. Finally we used the number 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 of evaluating the structural quality of SLAM datasets.

In this paper we extend some of our previous results in [7] to the case of non-isotropic noise covariance matrices. We also provide a theoretical proof that explains our empirical observations regarding the impact of the number of spanning trees. Furthermore, we propose an efficient algorithm to compute our proposed connectivity metric in large datasets. Finally 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 and Problem Formulation

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

A. Graph-theoretic Preliminaries

Consider a *simple* directed graph $\mathcal{G} = (\mathcal{V},\mathcal{E})$ with vertex set $\mathcal{V} = \{v_0,v_1,\ldots,v_n\}$ and edge set $\mathcal{E} \subseteq \mathcal{V}\times\mathcal{V}$ such that $|\mathcal{E}| = m$. Let $\tilde{\mathcal{G}} = (\mathcal{V},\tilde{\mathcal{E}})$ be the corresponding undirected graph obtained by ignoring the orientation of edges in $\mathcal{G}$. Here we assume $\mathcal{G}$ is weakly connected, or equivalently, $\tilde{\mathcal{G}}$ is connected. The degree of vertex $v_i \in \mathcal{V}$ in $\tilde{\mathcal{G}}$ is denoted by $\deg(v_i)$. The *reduced* incidence matrix of $\mathcal{G}$ after *anchoring* v_0 is denoted by $\mathbf{A} \in \{-1,0,1\}^{n\times m}$. $\mathbf{A}$ is obtained by removing the corresponding row from the incidence matrix of $\mathcal{G}$. Suppose $e_k \triangleq (v_{i_k},v_{j_k})$ denotes the k'th edge. Then for every $e_k \in \mathcal{E}$, $\mathbf{A}_{j_k,k} = 1$, and $\mathbf{A}_{i_k,k} = -1$ if $i_k \neq 0$ and $j_k \neq 0$. The remaining entries of $\mathbf{A}$ are all zero. The *reduced* Laplacian matrix of $\mathcal{G}$ is defined as $\mathbf{L} \triangleq \mathbf{A}\mathbf{A}^\top$.[1] Note that $\mathbf{L}$ can also be obtained by removing the row and column associated to the anchor from the graph Laplacian matrix $\breve{\mathbf{L}}$. It is easy to show that $\mathbf{L}$ can be written as $\mathbf{L} = \mathbf{D} - \mathbf{W}$ in which $\mathbf{D} \triangleq \mathrm{diag}\,(\deg(v_1),\ldots,\deg(v_n))$ and $\mathbf{W}$ is the adjacency matrix of $\tilde{\mathcal{G}}$ after removing the row and column corresponding to the anchor and ignoring the orientation of graph. It is well known that $\mathbf{A}$ is full column rank if and only if $\mathcal{G}$ is weakly connected, and therefore $\mathbf{L} \succ \mathbf{0}$ if and only if $\tilde{\mathcal{G}}$ is connected. The 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.

In connected graphs, the Laplacian matrix $\breve{\mathbf{L}} \succeq 0$ has a zero eigenvalue with eigenvector $\mathbf{1}$. The second smallest eigenvalue of $\breve{\mathbf{L}}$, $\lambda_2(\mathcal{G}) \triangleq \lambda_2(\breve{\mathbf{L}}) > 0$, is known as the *algebraic connectivity* of graph [5]. Note that if $\mathcal{G}_1$ is a connected spanning subgraph of $\mathcal{G}_2$, we have $\lambda_2(\mathcal{G}_1) \leq \lambda_2(\mathcal{G}_2)$. Algebraic connectivity frequently appears as a key factor in many applications in multi-agent systems, often influencing the convergence time in dynamical systems evolving according to $\dot{\mathbf{q}} = -\breve{\mathbf{L}}\,\mathbf{q}$. See [9] for a survey.

A spanning tree of $\tilde{\mathcal{G}}$ is a spanning subgraph of $\tilde{\mathcal{G}}$ that is also a tree. Let $t(\mathcal{G})$ denote the number of spanning trees of $\tilde{\mathcal{G}}$. Consider a tree $\mathcal{T}$ and the complete graph $\mathcal{K}$, both over $\mathcal{V}$

[1] Note that the (reduced) Laplacian defined here is independent of the orientation of edges.

with $|\mathcal{V}| = n+1$ vertices. Then we have[2]

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

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 $t(\mathcal{G})$ can be used to evaluate the graph connectivity. The following theorem provides an expression for $t(\mathcal{G})$ in the general case [5].

Theorem 1: (Kirchhoff's Matrix-Tree Theorem) The number of spanning trees of graph $\mathcal{G}$ 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.

B. Estimation-theoretic Preliminaries

In SLAM and many other estimation problems, our goal is to estimate a set of deterministic unknown *parameters* $\mathbf{x}_0^\varsigma, \mathbf{x}_1^\varsigma, \ldots, \mathbf{x}_n^\varsigma$ using a given 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 $\boldsymbol{\epsilon}_{ij}$. $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is called the graphical representation of such problem if there is a one-to-one correspondence between the vertex set $\mathcal{V}$ and $\{\mathbf{x}_0^\varsigma, \mathbf{x}_1^\varsigma, \ldots, \mathbf{x}_n^\varsigma\}$. 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 $\mathbf{z} = \mathbf{h}(\mathbf{x}^\varsigma) + \boldsymbol{\epsilon}$, in which $\mathbf{x}^\varsigma \triangleq \operatorname{vec}(\mathbf{x}_1^\varsigma, \ldots, \mathbf{x}_n^\varsigma) \in \mathcal{X}$, $\mathbf{z} \in \mathcal{Z}$ is the stacked vector of measurements, $\mathbf{h} : \mathcal{X} \to \mathcal{Z}$ is called the measurement function and $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ is the measurement noise. In this paper we are interested in $\mathbf{x}^\star$, the maximum likelihood estimate (MLE) for $\mathbf{x}^\varsigma$,

$$\mathbf{x}^\star = \arg\min_{\mathbf{x}} \; f(\mathbf{x}) = \arg\min_{\mathbf{x}} \|\mathbf{z} - \mathbf{h}(\mathbf{x})\|^2_{\boldsymbol{\Sigma}^{-1}}. \tag{2}$$

Here we briefly review the Cramér-Rao lower bound (CRLB). Due to the random nature of noise, observations $\mathbf{z}$ and therefore any (reasonable) estimator of $\mathbf{x}^\varsigma$ such as $\hat{\mathbf{x}}$ are random variables. $\hat{\mathbf{x}}$ is called an unbiased estimator if $\mathbb{E}[\hat{\mathbf{x}}] = \mathbf{x}^\varsigma$.[3] Note that the estimation error covariance matrix, $\mathbb{C}\mathrm{ov}\,[\mathbf{x}^\varsigma - \hat{\mathbf{x}}]$, is identical to the covariance of $\hat{\mathbf{x}}$, $\mathbb{C}\mathrm{ov}\,[\hat{\mathbf{x}}]$.

Theorem 2: (CRLB) Under some regularity conditions [13], the covariance matrix of any unbiased estimator of $\mathbf{x}$, such as $\hat{\mathbf{x}}$, satisfies $\mathbb{C}\mathrm{ov}\,[\hat{\mathbf{x}}] \succeq \mathcal{I}^{-1}(\mathbf{x}^\varsigma)$,[4] where $\mathcal{I}(\mathbf{x})$ is the Fisher information matrix (FIM),

$$\mathcal{I}(\mathbf{x}) \triangleq \mathbb{E}\Big[\frac{\partial}{\partial \mathbf{x}} \log p(\mathbf{z};\mathbf{x}) \; \frac{\partial^\top}{\partial \mathbf{x}} \log p(\mathbf{z};\mathbf{x})\Big]. \tag{3}$$

Here the expectation is over $\mathbf{z}$ and with respect to $p(\mathbf{z};\mathbf{x})$. 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, MLE $\mathbf{x}^\star$ is *asymptotically* unbiased, efficient and follows a normal distribution with mean $\mathbf{x}^\varsigma$ and covariance $\mathcal{I}(\mathbf{x}^\varsigma)^{-1}$. It is important to differentiate between the followings.

(i) $\mathcal{I}(\mathbf{x})$ evaluated at $\mathbf{x} = \mathbf{x}^\varsigma$ is FIM, and therefore $\mathcal{I}(\mathbf{x})^{-1}$ evaluated at $\mathbf{x} = \mathbf{x}^\varsigma$ gives CRLB, whose trace is a lower bound on the achievable mean squared error (MSE) by any unbiased estimator.

(ii) The inverse of $\mathcal{I}(\mathbf{x})$ evaluated at $\mathbf{x} = \mathbf{x}^\star$ is commonly used to approximate the covariance matrix of MLE. Note that unlike FIM, this quantity depends on $\mathbf{x}^\star$ and, consequently, the given realization of $\mathbf{z}$.

For the measurement model introduced earlier we have

$$\log p(\mathbf{z};\mathbf{x}) = -\frac{1}{2}\|\mathbf{z} - \mathbf{h}(\mathbf{x})\|^2_{\boldsymbol{\Sigma}^{-1}} + \text{const.} \tag{4}$$

Plugging (4) into (3), computing the gradient vector and the expectation with respect to $p(\mathbf{z};\mathbf{x})$ gives $\mathcal{I}(\mathbf{x}) = \mathbf{J}^\top \boldsymbol{\Sigma}^{-1} \mathbf{J}$, in which $\mathbf{J}$ is the Jacobian matrix of the measurement function. Thus for our model $\mathcal{I}(\mathbf{x}^\varsigma)$ and $\mathcal{I}(\mathbf{x}^\star)$ are obtained by evaluating the Jacobian matrix at $\mathbf{x}^\varsigma$ and $\mathbf{x}^\star$, respectively.

IV. Fisher Information Matrix

In this section we derive closed-form expressions for $\mathcal{I}(\mathbf{x})$ in SLAM and a number of interesting special cases that arise in the context of sensor networks. As will become evident shortly, in such problems, the Fisher information matrix is closely related to the graphical representation $\mathcal{G}$, mainly through the reduced Laplacian matrix $\mathbf{L}$.

A. SLAM

The state vector in planar 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 $\mathbf{x} = \operatorname{vec}(\mathbf{p}, \boldsymbol{\theta})$ 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.[5] Each measurement is a noisy 2D rigid body transformation between two robot poses,

$$\mathbf{z}_{ij} = \begin{bmatrix} \mathbf{R}(\theta_i^\varsigma)^\top (\mathbf{p}_j^\varsigma - \mathbf{p}_i^\varsigma) \\ \operatorname{wrap}(\theta_j^\varsigma - \theta_i^\varsigma) \end{bmatrix} + \boldsymbol{\epsilon}_{\boldsymbol{ij}}, \tag{5}$$

in which $\mathbf{p}_i^\varsigma$ denotes the position of the ith robot pose, $\mathbf{R}(\theta_i^\varsigma) \in \mathrm{SO}(2)$ is the rotation matrix corresponding to θ_i^ς, and $\operatorname{wrap} : \mathbb{R} \to [-\pi, \pi)$ is the function that maps its argument to the equivalent angle in $[-\pi, \pi)$. Therefore the measurement vector can be written as $\mathbf{z} = \operatorname{vec}(\mathbf{z}_p, \mathbf{z}_\theta)$ in which $\mathbf{z}_p$ is composed of translational components and $\mathbf{z}_\theta$ is the stacked vector of the rotational components of measurements. As in [2], in order to replace $\operatorname{wrap}(\theta_j^\varsigma - \theta_i^\varsigma)$ with $\theta_j^\varsigma - \theta_i^\varsigma$ we need to assume that correct regularization terms have been computed for the rotational component of measurements [3]. Then the measurement function can be expressed as below.

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

Here $\mathbf{R} \triangleq \operatorname{diag}(\mathbf{R}(\theta_{k_1}), \mathbf{R}(\theta_{k_2}), \ldots, \mathbf{R}(\theta_{k_m}))$ in which k_i denotes the index of the tail of the ith directed edge in $\mathcal{E}$.

[2]The expression for computing the number of spanning trees in complete graphs, $t(\mathcal{K}_\ell) = \ell^{\ell-2}$, is known as Cayley's formula.

[3]Expectation is over $\mathbf{z}$ and with respect to $p(\mathbf{z};\mathbf{x})$.

[4]If $A - B$ is positive-semidefinite, then $A_{i,i} \geq B_{i,i}$. In other words, the diagonal elements of CRLB are lower bounds on the variance of any unbiased estimator for each parameter.

[5]In feature-based problems, $\mathbf{p}$ also contains the position of landmarks.

Assumption 1: [Noise Covariance Matrix of Planar SLAM] In this paper we assume the noise covariance matrix $\mathbf{\Sigma}$ of planar SLAM has the following structure,

$$\mathbf{\Sigma} = \mathrm{diag}(\sigma_p^2\mathbf{I},\sigma_\theta^2\mathbf{I}).$$

.

The Jacobian of the measurement function and the Fisher information matrix are derived in [7].

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

in which $\mathbf{\Gamma}$ is the following block-diagonal matrix

$$\mathbf{\Gamma} \triangleq \mathbf{I}_m \otimes \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \tag{8}$$

and for each $e_k = (i_k,j_k) \in \mathcal{E}$, there is a 2-by-1 block in $\mathbf{\Delta} \in \mathbb{R}^{2m\times n}$ that contains

$$\big(\mathbf{\Delta}\big)_{2k-1:2k,i_k} = \mathbf{p}_{j_k} - \mathbf{p}_{i_k}. \tag{9}$$

The remaining elements in $\mathbf{\Delta}$ are all zero.

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

- $\mathbf{\Gamma}$ is an orthogonal matrix, i.e., $\mathbf{\Gamma}^\top\mathbf{\Gamma} = \mathbf{I}$.
- As noted by [2], $\mathbf{\Delta}^\top\mathbf{\Delta}$ is a diagonal matrix with an interesting structure. $(\mathbf{\Delta}^\top\mathbf{\Delta})_{i,i}$ is equal to the sum of squared distances between the ith robot pose, and all of the nodes observed by it,

$$(\mathbf{\Delta}^\top\mathbf{\Delta})_{i,i} = \sum_{j\in\{j|(i,j)\in\mathcal{E}\}} \|\mathbf{p}_i - \mathbf{p}_j\|^2. \tag{10}$$

B. Linear-SN and Compass-SLAM

As we did in [7], we define the following linear-Gaussian classes of problems that arise in sensor networks (linear-SN), as well as 2D and 3D SLAM when the robot orientation is known (compass-SLAM).

Definition 1: (**Linear-SN**) The term "linear-SN" refers to the class of estimation problems, often arise in the context of sensor networks, in which each measurement has the following form,

$$\mathbf{z}_{ij} = \mathbf{x}_j^\varsigma - \mathbf{x}_i^\varsigma + \boldsymbol{\epsilon_{ij}}. \tag{11}$$

Therefore the measurement function can be written as $\mathbf{h}(\mathbf{x}) = \mathbf{A}_d^\top\mathbf{x}$ in which d is the dimension of $\mathbf{x}_i$. The time-synchronization problem in sensor networks belongs to this class [1].

The Fisher information matrix in linear-SN is given by

$$\mathcal{I} = \mathbf{A}_d^\top\mathbf{\Sigma}^{-1}\mathbf{A}_d. \tag{12}$$

In the special case of $\mathbf{\Sigma} = \sigma^2\mathbf{I}$, we have $\mathcal{I} = \sigma^{-2}\mathbf{L}_d$. A more general case arise when measurements have different accuracies, i.e., $\mathbf{\Sigma} = \mathrm{diag}(\sigma_1^2\mathbf{I}_d,\ldots,\sigma_m^2\mathbf{I}_d)$. In this case, $\mathcal{I} = \mathbf{A}_d^\top\mathbf{\Sigma}^{-1}\mathbf{A}_d$ is equal to the reduced weighted Laplacian matrix of $\mathcal{G}$ when each edge $e_i \in \mathcal{E}$ is weighted according to $w(e_i) = \sigma_i^{-2}$.

Definition 2: (**Compass-SLAM**) The term "compass-SLAM" refers to the class of SLAM problems (2D and 3D pose-graph and feature-based) in which the robot orientation is known. The measurement function associated to such problems with $\mathbf{p}_i \in \mathbb{R}^d$ ($d \in \{2,3\}$) can be written as $\mathbf{h}(\mathbf{x}) = \mathbf{R}^\top\mathbf{A}_d^\top\mathbf{p}$, in which $\mathbf{R}$ is a given orthogonal matrix. It is straightforward to show that the Fisher information matrix for compass-SLAM with $\mathbf{\Sigma} = \sigma^2\mathbf{I}$ is given by $\mathcal{I} = \sigma^{-2}\mathbf{L}_d$.

V. Main Results

The log-determinant of positive-definite covariance matrices $\log\det : \mathbb{S}_+^n \to \mathbb{R}$, 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 [12]. The D-criterion, due to the reasons mentioned above, is a popular choice and has been frequently used in many design problems in robotics including sensor placement [6] and active SLAM [15]. Finally, it is easy to show that the $\log\det$ of CRLB is a lower bound for the $\log\det$ of covariance matrix of any unbiased estimator (see Theorem 2 and Lemma 1).

A. Criterion: Tree-Connectivity

As we saw earlier, $t(\mathcal{G})$ is a sound measure of graph connectivity. Here we define a metric based on $t(\mathcal{G})$ to measure graph connectivity.

Definition 3: We define the tree-connectivity of graph $\mathcal{G}$, $\tau(\mathcal{G})$, as follows.

$$\tau(\mathcal{G}) \triangleq \log t(\mathcal{G}) = \log\det\mathbf{L}\,. \tag{13}$$

Here $\mathbf{L}$ is the reduced Laplacian matrix of $\mathcal{G}$. Note that $\tau(\mathcal{G})$ is a well-defined metric for graph connectivity since $t(\mathcal{G}_1) \le t(\mathcal{G}_2) \Leftrightarrow \tau(\mathcal{G}_1) \le \tau(\mathcal{G}_2)$.

It is practically infeasible to work directly with the determinant in large SLAM problems and graphs as it often becomes too large (or too small, in the case of $\mathbf{L}^{-1}$ and covariance matrices) for finite-precision arithmetic [7]. 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 $\mathcal{E}$ in order to compute $\tau(\mathcal{G})$ efficiently using the Cholesky decomposition of $\mathbf{L}$. 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 $\mathcal{O}(|V|^3)$ time, while in practical problems that arise in the context of robotics and sensor networks, Algorithm 1 performs much faster given a sufficiently good fill-reducing permutation $\mathbf{\Pi}$.

B. Linear-SN and Compass-SLAM

Now we are ready to investigate the link between $\tau(\mathcal{G})$ and $\log\det\mathcal{I}(\mathbf{x})$. The following proposition straightforwardly follows from Section IV-B.

Proposition 1: In linear-SN and compass-SLAM, the maximum likelihood estimator is unbiased and efficient. Furthermore, if the noise is isotropic, we have

$$\log\det(\mathbb{C}\mathrm{ov}[\mathbf{x}^\star]) = -\log\det\mathcal{I} = -d\tau(\mathcal{G}) + \eta \tag{14}$$

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

```
function tree−connectivity(G)
    Choose a fill-reducing permutation Π.
    Compute K, the sparse Cholesky factor of ΠLΠ^⊤.
    return τ(G) = 2 Σ_i log K_{i,i}.
end function
```

in which $\eta = (|\mathcal{V}|-1)d\log(\sigma^2)$. In linear-SN problems with $\mathbf{\Sigma} = \mathrm{diag}(\sigma_1^2\mathbf{I}_d, \ldots, \sigma_m^2\mathbf{I}_d)$ we have

$$\log\det(\mathbb{C}\mathrm{ov}[\mathbf{x}^\star]) = -\tau_w(\mathcal{G}^w) \triangleq -\log\det\mathbf{L}^w \tag{15}$$

in which $\mathbf{L}^w$ denotes the reduced weighted Laplacian matrix of weighted graph $\mathcal{G}^w = (\mathcal{V},\mathcal{E},w)$ with weight function $w : \mathcal{E} \to \mathbb{R}_{>0}$ defined as $w(e_i) = \sigma_i^{-2}$ and $\tau_w(\mathcal{G}^w)$ is the tree-connectivity of $\mathcal{G}^w$.

Proposition 1 ensures that in linear-SN and compass-SLAM and under the specified assumptions, minimizing the differential entropy of $\mathbf{x}^\star$, minimizing the volume of uncertainty ellipsoids or finding the D-optimal design are all equivalent to maximizing the tree-connectivity of the network. This proposition provides a solid foundation for a framework in which our proposed metric $\tau(\mathcal{G})$ is used to quantify the quality of network topology in such problems.[6]

C. Planar SLAM

In [7] we made a number of empirical observations regarding the relation between $\det\mathcal{I}(\mathbf{x})$ in planar SLAM and the number of spanning trees in the graph. The following propositions provide a formal explanation for those observations. The proofs are provided in Appendix A.

Proposition 2: (Lower and Upper Bound) Suppose $\mathcal{I}_\circ(\mathbf{x})$ is the Fisher information matrix of the odometry subgraph, composed of robot poses and odometry measurements $\mathcal{E}_\circ \subseteq \mathcal{E}$. Let us define $\alpha \triangleq \sigma_\theta/\sigma_p$, $\Phi_{\inf}(\mathcal{G}) \triangleq \log\det\mathcal{I}(\mathbf{x}) - \log\det\mathcal{I}_\circ(\mathbf{x})$ and

$$\mathrm{dist}^2_{\text{out,max}} \triangleq \max_i\ (\mathbf{\Delta}^\top\mathbf{\Delta})_{i,i}, \tag{16}$$

where $\mathbf{\Delta}$ was introduced in Remark 1 (see also Lemma 2 in [2]). Then under the assumptions stated earlier, we have

$$3\tau(\mathcal{G}) \le \Phi_{\inf}(\mathcal{G}) \le 2\tau(\mathcal{G}) + \sum_{i=1}^{n}\log(\lambda_i(\mathbf{L}) + \alpha^2\mathrm{dist}^2_{\text{out,max}}). \tag{17}$$

Informally speaking, $\Phi_{\inf}(\mathcal{G})$ can be seen as the "information" gained by having loop-closure edges as compared with the dead reckoning scenario. Proposition 2 ensures that this gain is bounded by functions of our proposed metric $\tau(\mathcal{G})$ and the distances between the poses along the edges.

Proposition 3: Let us define $\psi \triangleq \alpha^2\mathrm{dist}^2_{\text{out,max}}$. The following statement holds.

$$\lim_{\psi\to 0^+}\Phi_{\inf}(\mathcal{G}) = 3\tau(\mathcal{G}), \tag{18}$$

[6]For other possible metrics see [7].

or, equivalently, as it was empirically observed in [7],

$$\lim_{\psi\to 0^+}\log\det\mathcal{I}(\mathbf{x}) = 3\tau(\mathcal{G}) + \log\det\mathcal{I}_\circ(\mathbf{x}). \tag{19}$$

Proposition 3 provides an explanation for our empirical observations in [7]. It is of utmost importance to note that while Proposition 2 provides lower and upper bounds for $\Phi_{\inf}(\mathcal{G})$, its main purpose is to be used in proving Proposition 3 (see Appendix A). In fact $\Phi_{\inf}(\mathcal{G})$ can be expressed exactly in terms of $\mathbf{L}$ (topology), α (sensor quality) and $\mathbf{\Delta}$ (geometry) (see the proof of Proposition 2 in Appendix A).

D. Normalized tree-connectivity

Comparing the graphical structure of two SLAM problems based on our proposed tree-connectivity metric is meaningful only if the two graphs have the same number of vertices. A remedy to this limitation is to use the normalized tree-connectivity defined below.

Definition 4: (Normalized tree-connectivity) We define the *normalized* tree-connectivity of graph $\mathcal{G}$, denoted by $\bar\tau(\mathcal{G})$, as

$$\bar\tau(\mathcal{G}) \triangleq \frac{\tau(\mathcal{G})}{\tau(\mathcal{K})} \overset{(1)}{=} \frac{\tau(\mathcal{G})}{(n-1)\log(n+1)}. \tag{20}$$

Note that $\tau(\mathcal{K})$ is the maximum achievable tree-connectivity among graphs with the same number of vertices as $\mathcal{G}$.

Therefore for any (simple connected) graph $\mathcal{G}$, $\bar\tau(\mathcal{G}) \in [0,1]$ assigns a score that describes the tree-connectivity of $\mathcal{G}$ relative to the tree-connectivity of the complete graph with the same number of vertices. Proposition 3 ensures that under the stated conditions we have,

$$\lim_{\psi\to 0^+}\frac{\Phi_{\inf}(\mathcal{G})}{\Phi_{\inf}(\mathcal{K})} = \bar\tau(\mathcal{G}). \tag{21}$$

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 that contains all possible loop closures). In the following section we use $\bar\tau(\mathcal{G})$ to compare the quality of the graphical structure of some of the publicly available SLAM datasets.

VI. Numerical Results

In this section we present numerical results using publicly available datasets with two objectives. First, we validate our theoretical results presented in Section V and in particular Proposition 3. Our second objective is to see how small ψ should be in practice for $\log\det\mathcal{I}(\mathbf{x})$ to be sufficiently close to its limit when $\psi \to 0^+$. To do so, we need to define a measure of closeness to the limit value of $\log\det\mathcal{I}(\mathbf{x})$. Here we use the relative error (RE) as defined below,

$$\mathrm{RE} \triangleq \Big|\frac{\log\det\mathcal{I}(\mathbf{x}) - \mathcal{L}}{\log\det\mathcal{I}(\mathbf{x})}\Big|, \tag{22}$$

in which,

$$\mathcal{L} \triangleq \lim_{\psi\to 0^+}\log\det\mathcal{I}(\mathbf{x}) = 3\tau(\mathcal{G}) + \log\det\mathcal{I}_\circ(\mathbf{x}). \tag{23}$$

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

Dataset	$\bar{\tau}(\mathcal{G})$	$\deg_{\text{avg}}$	α^2	$\text{dist}^2_{\text{out,max}}$	ψ	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

The datasets used in this section are all publicly available 2D pose-graph SLAM datasets.[7] For datasets with non-isotropic covariance matrices, the original covariance matrices are modified to satisfy Assumption 1. Manhattan 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 [8], i.e., (supposedly) $\mathbf{x}^\star$.

A summary of results is shown in Table I. The entries in Table I are sorted according to their normalized tree-connectivity metric $\bar{\tau}(\mathcal{G})$. Our results indicate that in practice RE is relatively small even when ψ is far from being negligible. This implies that in such cases, the log-determinant of Fisher information matrix (or covariance matrix) is entirely characterized by the tree-connectivity of graph. The large RE in the case of Lincoln dataset (highlighted in red) is due to the fact that Gauss-Newton has failed to converge to the true $\mathbf{x}^\star$. The average degree of graph [7, 11] is not sufficiently sophisticated 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.

Figure 1 shows how RE evolves with respect to scaling ψ. Scaling ψ can be done by scaling either α^2 or $\mathbf{p}^\varsigma$ which lead to identical results in terms of $\log\det\mathcal{I}(\mathbf{x}^\varsigma)$.[8] Figure 1 is obtained by scaling ψ according to $\beta\psi_{\text{orig}}$ in which $\psi_{\text{orig}} = 29$ is the original value of ψ in the Manhattan dataset (see Table I). This figure clearly shows that as $\psi \to 0^+$, the log-determinant of CRLB converges to $-\mathcal{L}$ as predicted by Proposition 3.

Figure 2 shows $\log\det\mathcal{I}(\mathbf{x}^\star)$ 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 (i.e., $1, \ldots, |\mathcal{E}| - |\mathcal{V}|$), we generated 50 random spanning subgraphs. The predicted value is $\mathcal{L}$. Figure 2 indicates that in this case, tree-connectivity almost entirely characterises the performance of the maximum likelihood estimator and CRLB in terms of log-determinant of the covariance matrices as predicted by our theoretical results. Note that the difference between Figure 2 and Figure 4 in [7] are the sizes of the subgraphs and range of horizontal axis, which indicate the key role of Algorithm 1 in computing $\tau(\mathcal{G})$ efficiently for large graphs.

[7]Note that Proposition 2 can be straightforwardly extended to the case of 2D feature-based SLAM. However in this paper we only provide numerical results for pose-graph problems.

[8]See (38) in the Appendix.

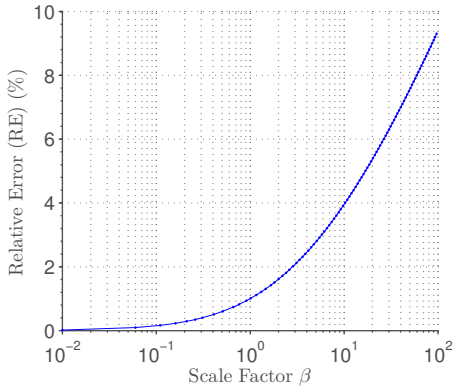


Fig. 1: RE evaluated at $\mathbf{x}^\varsigma$ as a function of scale parameter β for Manhattan. Here $\psi = \beta\psi_{\text{orig}}$ in which $\psi_{\text{orig}} = 29$ is the value of ψ in Manhattan dataset (see Table I). This can be done by scaling either α^2 or $\mathbf{p}^\varsigma$. Note the logarithmic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

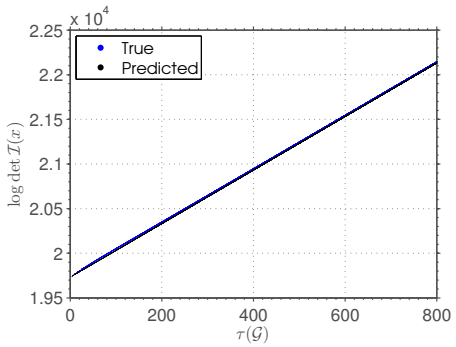
TABLE II: Connectivity of publicly available 3D pose-graph datasets. Entries are sorted according to $\bar{\tau}(\mathcal{G})$.

Dataset	$\bar{\tau}(\mathcal{G})$	$\deg_{\text{avg}}$
Sphere 2500	0.2444	7.8392
Pittsburgh A	0.0014	2.0038
The Hague	0.0012	2.0000
New College	0.0010	2.0037

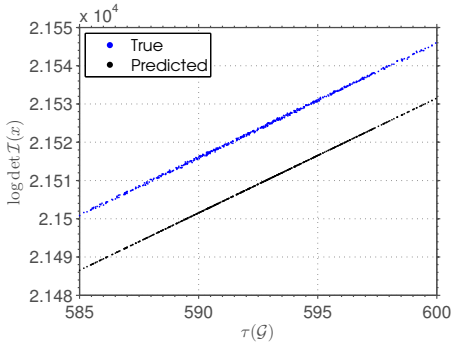
Our extensive empirical observations suggest that $\Phi_{\text{inf}}(\mathcal{G})$ asymptotically behaves as a linear function of $\tau(\mathcal{G})$ in 3D pose-graph SLAM datasets with SE(3) measurements under isotropic noise. Table II shows the tree-connectivity and $\bar{\tau}(\mathcal{G})$ for a number of publicly available 3D pose-graph datasets. We plan to address this extension in our future work.

VII. Conclusion

In this paper we proposed tree-connectivity, a metric to quantify connectivity of estimation graphs. We also proposed an algorithm to efficiently compute our metric for sparse graphs. Furthermore, we proved the relation between our metric and 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.



(a) Overview.



(b) Zoomed region.

Fig. 2: $\log\det\mathcal{I}(\mathbf{x}^\star)$ 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.

Numerical results using publicly available datasets confirmed the validity of our analysis. It was observed that in many practical scenarios, ψ does not need to be very small for $\log\det\mathcal{I}(\mathbf{x})$ to converge to its limit $\mathcal{L}$.

According to [2], graph connectivity, α and $\mathrm{dist}_{\mathrm{out,max}}$ affect the convergence of Gauss-Newton to the optimal solution $\mathbf{x}^\star$. Therefore there is an intriguing overlap between the elements that appear in our analysis and those reported in [2]. This connection requires further investigation.

As mentioned earlier, our results can be straightforwardly generalized to 2D feature-based problems. Our extensive empirical observations suggest that Proposition 3 can be also extended to 3D SLAM with $\mathrm{SE}(3)$ measurements. We will consider this extension in our future work.

Finally, we hope that our analysis will ultimately be used to collect reliable datasets in active scenarios (see [7] for two simplified applications in active measurement selection and edge pruning). The graphical structure of any SLAM problem can provide a compact but rich representation of the underlying estimation problem. Making decisions and planning (e.g., in active SLAM) based on the graph is robust to common convergence issues (e.g., local minima). Our metric can also be used as a benchmark for SLAM datasets. As recommended by Olson and Kaess, we also encourage researchers to compute and report connectivity metrics such as our normalized tree-connectivity in their experimental results as it "conveys significant information about the problem" [11].

Appendix A
Proofs

Before proving the main propositions, we need the following lemmas.

Lemma 1: For any two (symmetric) $\mathbf{N},\mathbf{M}\in\mathbb{S}^n_+$ we have[9]

$$\det(\mathbf{M}+\mathbf{N})\geq\det(\mathbf{M}). \tag{24}$$

Proof: It is trivial to verify the lemma when $\mathbf{M}$ is singular. For $\mathbf{M}\succ\mathbf{0}$, we can decompose $\mathbf{M}$ as $\mathbf{M}=\mathbf{M}^{\frac{1}{2}}\mathbf{M}^{\frac{1}{2}}$ in which $\mathbf{M}^{\frac{1}{2}}\in\mathbb{S}^n_{++}$. Then we have

$$\det(\mathbf{M}+\mathbf{N}) = \det(\mathbf{M})\det(\mathbf{I}+\mathbf{M}^{-\frac{1}{2}}\mathbf{N}\mathbf{M}^{-\frac{1}{2}})) \tag{25}$$

$$= \det(\mathbf{M})\prod_{i=1}^{n}(1+\underbrace{\lambda_i\big(\mathbf{M}^{-\frac{1}{2}}\mathbf{N}\mathbf{M}^{-\frac{1}{2}}\big)}_{\geq 0}) \tag{26}$$

$$\geq \det(\mathbf{M}). \tag{27}$$

■

Lemma 2: The following statements hold about

$$\mathbf{P}\triangleq\mathbf{\Gamma}^\top\mathbf{A}_2^\top\mathbf{L}_2^{-1}\mathbf{A}_2\mathbf{\Gamma}. \tag{28}$$

1) $\mathbf{P}$ is the orthogonal projection onto $\mathrm{range}(\mathbf{\Gamma}^\top\mathbf{A}_2^\top)$.
2) If $\mathcal{G}$ is a tree then $\mathbf{P}=\mathbf{I}$.

Proof: From Remark 1 recall that $\mathbf{\Gamma}$ is orthogonal. Let us define $\mathbf{Q}\triangleq\mathbf{\Gamma}^\top\mathbf{A}_2^\top$. Then it is easy to verify that $\mathbf{P}=\mathbf{Q}\mathbf{Q}^\dagger$ in which

$$\mathbf{Q}^\dagger\triangleq(\mathbf{Q}^\top\mathbf{Q})^{-1}\mathbf{Q}^\top$$

is the Moore-Penrose pseudoinverse of $\mathbf{Q}$. This proves the first statement. The second statement follows from the fact that for trees, $\mathbf{A}_2$ and consequently $\mathbf{Q}$ are full rank square matrices and therefore $\mathrm{range}(\mathbf{Q})=\mathbb{R}^n$, i.e., the orthogonal projection onto $\mathrm{range}(\mathbf{Q})$ is the identity map. ■

Lemma 3: Under the stated assumptions we have

$$\Phi_{\mathrm{inf}}(\mathcal{G})\triangleq\log\det\mathcal{I}(\mathbf{x})-\log\det\mathcal{I}_\circ(\mathbf{x}) \tag{29}$$

$$= 2\tau(\mathcal{G})+\log\det(\mathbf{L}+\alpha^2\mathbf{\Delta}^\top\mathbf{P}^\perp\mathbf{\Delta}) \tag{30}$$

in which $\mathbf{P}^\perp\triangleq\mathbf{I}-\mathbf{P}$ is the orthogonal complement of $\mathbf{P}$ projecting onto $\mathrm{null}(\mathbf{A}_2\mathbf{\Gamma})$.

Proof: We begin by applying Schur's determinant formula [10] on (7).

$$\log\det\mathcal{I}(\mathbf{x})=\log\det(\sigma_p^{-2}\mathbf{L}_2)+\log\det(\boldsymbol{\mathcal{S}}) \tag{31}$$

in which $\boldsymbol{\mathcal{S}}$ is the Schur complement of the top-left block in (7),

$$\boldsymbol{\mathcal{S}}\triangleq\sigma_\theta^{-2}\mathbf{L}+\sigma_p^{-2}\mathbf{\Delta}^\top\mathbf{\Delta}-\sigma_p^{-2}\mathbf{\Delta}^\top\mathbf{\Gamma}^\top\mathbf{A}_2^\top\mathbf{L}_2^{-1}\mathbf{A}_2\mathbf{\Gamma}\mathbf{\Delta} \tag{32}$$

$$=\sigma_\theta^{-2}\mathbf{L}+\sigma_p^{-2}\mathbf{\Delta}^\top\mathbf{P}^\perp\mathbf{\Delta}. \tag{33}$$

[9] As it will become clear shortly, although stronger results do exist, this lemma is sufficient for our proof.

Now note that $\det(\mathbf{L}_2) = t(\mathcal{G})^2$. Therefore (31) can be written as

$$\log\det\mathcal{I}(\mathbf{x}) = 2\tau(\mathcal{G}) + \log\det(\sigma_\theta^2\boldsymbol{\mathcal{S}}) + \log(\sigma_p^{-4n}\sigma_\theta^{-2n}). \tag{34}$$

Following the same steps for $\log\det\mathcal{I}_\circ(\mathbf{x})$ of the odometry subgraph $\mathcal{G}_\circ = (\mathcal{V},\mathcal{E}_\circ)$, using Lemma 2 and noting the fact that $\tau(\mathcal{G}_\circ) = 0$ results in

$$\log\det\mathcal{I}_\circ(\mathbf{x}) = \log\det(\sigma_\theta^2\boldsymbol{\mathcal{S}}_\circ) + \log(\sigma_p^{-4n}\sigma_\theta^{-2n}) \tag{35}$$

$$= \underbrace{\log\det(\mathbf{L}_\circ)}_{\tau(\mathcal{G}_\circ)=0} + \log(\sigma_p^{-4n}\sigma_\theta^{-2n}) \tag{36}$$

$$= \log(\sigma_p^{-4n}\sigma_\theta^{-2n}). \tag{37}$$

Plugging (34) and (37) into the definition of $\Phi_{\inf}(\mathcal{G})$ concludes the proof. ∎

Proof of Proposition 2: Note that both $\mathbf{P}$ and $\mathbf{P}^\perp$ are positive semidefinite.[10] For the lower bound note that from Lemma 3 we have

$$\Phi_{\inf}(\mathcal{G}) = 2\tau(\mathcal{G}) + \log\det(\mathbf{L} + \underbrace{\alpha^2\boldsymbol{\Delta}^\top\mathbf{P}^\perp\boldsymbol{\Delta}}_{\succeq\mathbf{0}}). \tag{38}$$

Recall that $\mathbf{L} \succ \mathbf{0}$ and therefore applying Lemma 1 results in

$$\Phi_{\inf}(\mathcal{G}) \geq 2\tau(\mathcal{G}) + \log\det(\mathbf{L}) = 3\tau(\mathcal{G}). \tag{39}$$

Similarly, the upper bound results from Remark 1 and applying Lemma 1 twice as shown below.

$$\Phi_{\inf}(\mathcal{G}) = 2\tau(\mathcal{G}) + \log\det(\mathbf{L} + \alpha^2\boldsymbol{\Delta}^\top\mathbf{P}^\perp\boldsymbol{\Delta}) \tag{40}$$

$$\leq 2\tau(\mathcal{G}) + \log\det(\underbrace{\mathbf{L} + \alpha^2\boldsymbol{\Delta}^\top\mathbf{P}^\perp\boldsymbol{\Delta}}_{\succ\mathbf{0}} + \underbrace{\alpha^2\boldsymbol{\Delta}^\top\mathbf{P}\boldsymbol{\Delta}}_{\succeq\mathbf{0}}) \tag{41}$$

$$= 2\tau(\mathcal{G}) + \log\det(\mathbf{L} + \alpha^2\boldsymbol{\Delta}^\top\boldsymbol{\Delta}) \tag{42}$$

$$\leq 2\tau(\mathcal{G}) + \tag{43}$$

$$\log\det(\underbrace{\mathbf{L} + \alpha^2\boldsymbol{\Delta}^\top\boldsymbol{\Delta}}_{\succ\mathbf{0}} + \underbrace{\alpha^2\mathrm{dist}^2_{\text{out,max}}\mathbf{I} - \alpha^2\boldsymbol{\Delta}^\top\boldsymbol{\Delta}}_{\succeq\mathbf{0}}) \tag{44}$$

$$= 2\tau(\mathcal{G}) + \log\det(\mathbf{L} + \alpha^2\mathrm{dist}^2_{\text{out,max}}\mathbf{I}) \tag{45}$$

$$= 2\tau(\mathcal{G}) + \log(\prod_{i=1}^{n}(\lambda_i(\mathbf{L}) + \alpha^2\mathrm{dist}^2_{\text{out,max}})) \tag{46}$$

$$= 2\tau(\mathcal{G}) + \sum_{i=1}^{n}\log(\lambda_i(\mathbf{L}) + \alpha^2\mathrm{dist}^2_{\text{out,max}}). \tag{47}$$

∎

Proof of Proposition 3: Proposition 3 follows directly from Proposition 2 and the squeeze theorem in calculus. To see this note that as $\psi \to 0^+$, the upper bound in Proposition 2 approaches the lower bound since

$$\sum_i \log(\lambda_i(\mathbf{L})) = \log(\prod_i \lambda_i(\mathbf{L})) = \tau(\mathcal{G}). \tag{48}$$

∎

[10] The spectrum of $\mathbf{P}$ and $\mathbf{P}^\perp$ consists of only zeros and ones.

REFERENCES

[1] Prabir Barooah and Joao P Hespanha. Estimation on graphs from relative measurements. *Control Systems, IEEE*, 27(4):57–74, 2007.

[2] Luca 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.

[3] Luca Carlone, Rosario Aragues, José A Castellanos, and Basilio Bona. A fast and accurate approximation for planar pose graph optimization. *The International Journal of Robotics Research*, 33:965 – 987, 06/2014 2014. ISSN 1741-3176.

[4] Frank Dellaert and Michael Kaess. Square Root SAM: Simultaneous localization and mapping via square root information smoothing. *Intl. J. of Robotics Research, IJRR*, 25(12):1181–1204, December 2006.

[5] Chris Godsil and Gordon Royle. *Algebraic graph theory*. Graduate Texts in Mathematics Series. Springer London, Limited, 2001. ISBN 9780387952413.

[6] Siddharth Joshi and Stephen Boyd. Sensor selection via convex optimization. *Signal Processing, IEEE Transactions on*, 57(2):451–462, 2009.

[7] Kasra Khosoussi, Shoudong Huang, and Gamini 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*, pages 2707–2714, 2014.

[8] Kurt Konolige, Giorgio Grisetti, Rainer Kummerle, Wolfram Burgard, Benson Limketkai, and R. Vincent. Efficient sparse pose adjustment for 2D mapping. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 22–29. IEEE, 2010.

[9] Mehran Mesbahi and Magnus Egerstedt. *Graph theoretic methods in multiagent networks*. Princeton University Press, 2010.

[10] Carl D Meyer. *Matrix analysis and applied linear algebra*. SIAM, 2000.

[11] Edwin Olson and Michael Kaess. Evaluating the performance of map optimization algorithms. In *RSS Workshop on Good Experimental Methodology in Robotics*, page 40, 2009.

[12] Friedrich Pukelsheim. *Optimal design of experiments*, volume 50. SIAM, 1993.

[13] Harold Wayne Sorenson. *Parameter estimation: principles and problems*. Control and systems theory. M. Dekker, 1980. ISBN 9780824769871.

[14] Sebastian Thrun and Michael Montemerlo. The graph SLAM algorithm with applications to large-scale mapping of urban structures. *The International Journal of Robotics Research*, 25(5-6):403, 2006.

[15] Teresa A Vidal-Calleja, Alberto Sanfeliu, and Juan Andrade-Cetto. Action selection for single-camera SLAM. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 40(6):1567–1581, 2010.