FT-BSP: Focused Topological Belief Space Planning

Moshe Shienman[1], Andrej Kitanov[2], and Vadim Indelman[2]

***Abstract*—At its core, decision making under uncertainty can be regarded as sorting candidate actions according to a certain objective. While finding the optimal solution directly is computationally expensive, other approaches that produce the same ordering of candidate actions, will result in the same selection. With this motivation in mind, we present a computationally efficient approach for the focused belief space planning (BSP) problem, where reducing the uncertainty of only a predefined subset of variables is of interest. Our approach uses topological signatures, defined over the topology induced from factor graph representations of posterior beliefs, to rank candidate actions. In particular, we present two such signatures in the context of information theoretic focused decision making problems. We derive error bounds, with respect to the optimal solution, and prove that one of these signatures converges to the true optimal solution. We also derive a second set of bounds, which is more conservative, but is only a function of topological aspects and can be used online. We introduce the Von Neumann graph entropy for the focused case, which is based on weighted node degrees, and show that it supports incremental update. We then analyze our approach under two different settings, measurement selection and active focused 2D pose SLAM.**

***Index Terms*—Motion and Path Planning, View Planning for SLAM, Reactive and Sensor-Based Planning.**

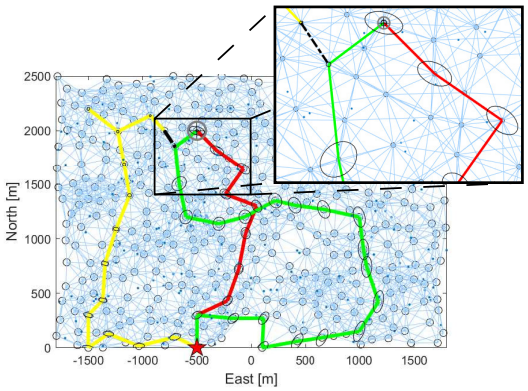


Fig. 1: Chosen candidate paths, in active 2D pose SLAM scenario, with respect to different objectives. The red path considers reducing uncertainty with respect to the entire trajectory while the green path considers reducing uncertainty only over the final pose. The yellow path represents the locations that the robot had previously visited and acts as the initial belief. Stating position is denoted with a red star and the goal is denoted with a circle. Black ellipses denote the marginal covariance at each step and black dashed lines denote loop-closures. Notice how the marginal covariance shrinks in the green path, right before it reaches the goal, due to a big loop closure.

I. Introduction

REASONING about uncertainty, as part of the decision making problem is at the core of robotics and artificial intelligence. In a partially observable setting, the robot has to maintain a probability distribution function (belief) over appropriate random variables that correspond to the robot's and environment state. Furthermore, the robot has to autonomously determine its next actions by solving the corresponding belief space planning (BSP) problem, while reasoning about future belief evolution.

For instance, a robot must often autonomously construct a map of the environment and determine which action to take given a certain objective. Moreover, in different tasks, specific landmarks within the map or certain states of the robot, might be more important than others.

These types of problems are commonly referred to as active *focused* inference, in which the objective function prioritizes or only considers a predefined subset of *focused* variables. e.g. a robot navigating in an uncertain environment must reach a goal position with maximum accuracy; a reconstruction task in which the robot is required to map a specific scene in the environment with high accuracy; collision avoidance, where localizing obstacles is of the greatest importance. As shown in [23] and Fig. 1 the optimal solution in the *focused* case can be significantly different from the one in the *unfocused* case, i.e. considering all variables.

To solve these *focused* problems for Gaussian distributions, exact state-of-the-art algorithms calculate the marginal posterior covariance (information) matrix for each candidate action. While the set of *focused* variables can be small, these calculations require a computationally expensive Schur complement operation. Moreover, many problems in robotics require a high n-dimensional state space model to represent robot poses, obstacles and landmarks. In terms of the D-criterion for example, which is common in information theoretic problems, the general complexity of calculating the determinant of the marginal information matrix is $O(n^3)$.

Solving an alternative problem in a different domain, with respect to the same set of candidate actions, can be used to reduce the computational complexity. Such an alternative should discriminate between candidate actions, be less expensive to compute and would ideally yield a solution which is consistent with the optimal solution of the original problem.

Efficient inference algorithms, operating on different probabilistic graphical models, are commonly used to improve the computational complexity. Modern approaches even suggest to use the topological signature, induced from the connectivity of

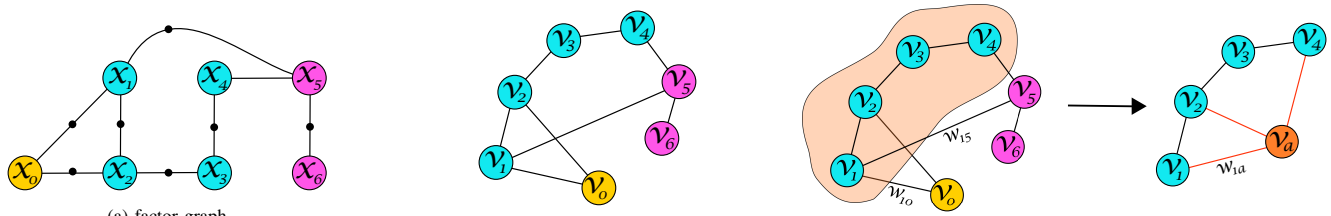
Manuscript received: October, 15, 2020; Revised January, 17, 2021; Accepted February, 20, 2021.

This paper was recommended for publication by Editor Nancy Amato upon evaluation of the Associate Editor and Reviewers' comments. This work was partially supported by the Israel Science Foundation.

[1]Moshe Shienman is with the Technion Autonomous Systems Program (TASP), Technion - Israel Institute of Technology, Haifa 32000, Israel, smoshe@campus.technion.ac.il

[2]Andrej Kitanov and Vadim Indelman are with the Department of Aerospace Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel. {andrej, vadim.indelman}@technion.ac.il

Digital Object Identifier (DOI): see top of this page.



(a) factor graph (b) topological graph (c) *unfocused* augmented topological graph

Fig. 2: Graph representations of a posterior belief. (a) factor graph *FG* with *unfocused* variables in blue, *focused* variables in purple and *anchor* variable in yellow; (b) the corresponding topological graph $\mathcal{G}$; (c) the *unfocused* topological augmented graph $\mathcal{G}^{U,A}$. The blob encapsulating the *unfocused* nodes represents the reduced Laplacians in (26) and the edges in orange are new edges weighted according to (27). Specifically $w_{1a} = w_{10} + w_{15}$.

variables (see e.g. Fig 2b) within these graphical representations, as a further improvement in the context of information theoretic decision making under uncertainty.

However, to the best of our knowledge, considering topological aspects with respect to a *focused* set of variables, within information theoretic BSP is a novel concept.

In this work we study the relation between the graphical structure of SLAM and an information theoretic objective function in the active *focused* case. We present a novel Focused Topological Belief Space Planning (FT-BSP) concept that uses topological signatures, associated with posterior probabilities, to guide the search for an optimal action.

Our contributions in this paper are as follows: (a) We introduce a novel approach, FT-BSP that, addresses the *focused* information theoretic BSP problem via topological aspects; (b) Within FT-BSP we derive two topological signatures to approximate the *focused* cost function; (c) We prove asymptotic convergence and develop bounds for one of the signatures; (d) We provide empirical results that show that the topological signatures are highly correlated with the *focused* information theoretic objective function considering measurement selection and *focused* BSP, and are significantly faster to calculate.

II. Related Work

In the last several decades, information theoretic decision making approaches have been investigated in the context of different problems. These include active perception, e.g. [1], sensor deployment and measurement selection, e.g. [20] and active SLAM, e.g. [9], [18], [28].

As initiations of a Partially Observed Markov Decision Process (POMDP) [21], finding the optimal solution to these problems is known to be computationally intractable [10]. As such, different approaches have been developed to approximate the optimal solution while reducing the computational complexity e.g. [9], [13] and [28].

While numerous graphical models based inference and learning algorithms were proposed, e.g. [11] and [12], the structural properties of such models and their use in decision making under uncertainty was only recently studied. In [14], [15] the authors show that under some conditions, the D-optimality criterion is strongly correlated with the graph topology of SLAM, described by the weighted number of spanning trees. The authors of [17] introduced a topological BSP concept, based on topological signatures, to approximate the solution to BSP in multi-robot active 2D SLAM. They showed, that both the number of spanning trees and the Von Neumann graph entropy signatures [24], [27] have a strong correlation with the information theoretic cost. Moreover, they proposed to use an approximation of the Von Neumann graph entropy, as presented in [7], which only depends on the graph node degrees, to further reduce the computational complexity. In [16], they provide performance guarantees for their method. In [4] the authors suggested to use graph topologies, including weighted number of spanning trees and weighted node tree (T-optimality metric) to reduce computational complexity, in order to find the best trajectory for loop-closures in active 3D pose-graph SLAM. While all of these works utilize topological aspects, none of them considers the *focused* case.

Active *focused* inference approaches aim to reduce the uncertainty only over a predefined subset of variables. In [23] the authors determine MI between *unfocused* and *focused* variable using a message-passing algorithm. In the context of resource-constrained systems, where some data will eventually have to be discarded, the authors of [25] offered a two stage approach where they calculated the posterior covariance matrix for each measurement to obtain the marginal over a *focused* set of variables. The rAMDL approach, presented in [18], [19] efficiently evaluates the information theoretic cost for each candidate action in the *focused* and *unfocused* cases. It avoids a computationally expensive Schur complement operation for each candidate action by performing a one-time calculation of the marginal covariance, associated with the variables involved in candidate actions. While all of these works utilize different graphical models representations to reduce the computational complexity, considering the *focused* case, none of them incorporate any topological aspects in inference nor in planning. Furthermore, one of the main advantages of a topological BSP approach, is that it does not have to propagate the belief nor recover marginals during planning. Yet, under some conditions and given tight bounds, it can find the optimal solution.

III. Notation and Problem Formulation

In this section we provide the theoretical background for *focused* Belief Space Planning and briefly review the factor graph model [22] from which we induce the topological signature associated with a posterior belief.

A. Focused Belief Space Planning

Consider a robot operating in a partially known environment, aiming to autonomously decide its future actions based on information accumulated thus far and a user defined

objective function J. In this work, for simplicity, we assume a 2D pose SLAM framework with relative pose measurements.

Let $p \in \mathbb{R}^2$ and $\theta \in [-\pi, \pi)$ denote the robot position and orientation, respectively. Let x_k denote the robot's state at time instant k where the state vector is defined as $x_k = \begin{bmatrix} p_k^T & \theta_k \end{bmatrix}^T$. The joint state, up to and including time k, is defined as $X_k = \{x_0, x_1, ..., x_k\}$. A *focused* subset of states is denoted by $X_k^F \subseteq X_k$ while the remaining *unfocused* states are denoted by $X_k^U = X_k / X_k^F$.

Let $z_{0:k}$ and $u_{0:k-1}$ denote, respectively, all observations and controls up to time k. The motion and observation models are given by

$$x_{k+1} = f(x_k, u_k) + w_k \quad , \quad z_{jk} = h(x_j, x_k) + v_{jk} \tag{1}$$

where $w_k \sim \mathcal{N}(0, \Sigma_w)$ and $v_{jk} \sim \mathcal{N}\left(0, \Sigma_{v_{jk}}\right)$ are the process and measurement noise terms. The posterior probability density function (pdf) over the joint state, denoted as the belief, is given by

$$b[X_k] \doteq \mathbb{P}(X_k | z_{0:k}, u_{0:k-1}) = \mathbb{P}(X_k | H_k), \tag{2}$$

where $H_k \doteq \{z_{0:k}, u_{0:k-1}\}$ represent history at time k. Given a candidate action sequence $u_{k:k+L-1}$ and future observations $z_{k+1:k+L}$, the future joint belief is given by

$$b[X_{k+L}] \doteq \mathbb{P}(X_{k+L} | z_{0:k+L}, u_{0:k+L-1}). \tag{3}$$

The future joint belief can be expressed in terms of $b[X_k]$ and the corresponding motion and observation models

$$b[X_{k+L}] \propto \mathbb{P}(X_k | H_k) \prod_{l=k+1}^{k+L} \mathbb{P}(x_l | x_{l-1}, u_{l-1}) \mathbb{P}(Z_l | X_l), \tag{4}$$

where the measurement likelihood term $\mathbb{P}(Z_l | X_l) = \prod_{m=1}^{n_l} \mathbb{P}(z_{l,i_m} | x_l, x_{i_m})$ represents all n_l observations acquired at time l between involved variables x_l and x_{i_m}, $\{i_m\} \subseteq \{0, 1, \ldots, l-1\}$.

Given a user defined objective function J, a belief $b[X_k]$ and a set of candidate actions $\mathcal{U}_k$, the goal of BSP is to find the optimal action given by

$$\mathcal{U}^* = \underset{\mathcal{U}}{\operatorname{argmin}} J(\mathcal{U}). \tag{5}$$

For the *unfocused* case, a general objective function in BSP can be written as

$$J(\mathcal{U}) = \underset{Z_{k+1:k+L}}{\mathbb{E}} \left[\sum_{l=0}^{L-1} c_l(b[X_{k+l}], u_{k+l}) + c_L(b[X_{k+L}]) \right], \tag{6}$$

where c_l represents a cost function for each look-ahead step, c_L represents the cost function of the terminal belief (at the end of the planning horizon), and the expectation is taken with respect to future observations. While each cost function can include a number of different terms such as distance to goal, energy spent and information measures of future beliefs, in this work, we only consider the information theoretic term of a terminal belief (to be defined) at time step $k+L$. We note that different cost terms c_l can be treated in a similar manner. Specifically, given an appropriate posterior belief, we aim to minimize the differential entropy $\mathcal{H}$.

Evaluating the objective function (6) involves inference over the appropriate posterior beliefs. Within a pose SLAM framework with relative pose measurements, the Maximum Likelihood (ML) estimation is obtained by the optimal state $X_k^\star$ that maximizes the belief (2). By fixing an arbitrary pose as an *anchor*, e.g. x_0, and considering the rest as unknown, $X_k^\star$ is obtained by minimizing the sum of weighted squared errors between the predicted and measured relative poses

$$X_k^\star = \underset{X_k}{\arg\min} ||\Delta_k - h(X_k)||^2_{\Sigma^{-1}}, \tag{7}$$

where the measurement model

$$\Delta_k = h(X_k) + \epsilon, \ \epsilon \sim \mathcal{N}(0, \Sigma), \tag{8}$$

represents a vector of m stacked relative pose measurements $z_{ij}^r \in SE(2), r = 1, 2, ..., m$ generated according to the motion and observation models (1). In this work, we assume the noise covariance matrices, for both the rotational and translational measurements, have a block-isotropic structure, i.e. $\Sigma = \text{diag}(\Sigma_p, \Sigma_\theta)$ where $\Sigma_p = \text{diag}\left(\sigma_{p_1}^2 \mathrm{I}_2, ..., \sigma_{p_m}^2 \mathrm{I}_2\right)$ and $\Sigma_\theta = \text{diag}\left(\sigma_{\theta_1}^2, ..., \sigma_{\theta_m}^2\right)$.

According to [29] the information matrix $\Lambda(X_k)$, referred from here on simply as Λ_k, of the ML estimation is given by $\Lambda_k = M^T \Sigma^{-1} M$, where $M = \partial h / \partial X_k$ is a measurement Jacobian. Evaluating Λ_k at the true value of X_k is known as the Fisher Information Matrix (FIM) while evaluating Λ_k at $X_k^\star$ is used to approximate the FIM.

Given a multivariate Gaussian posterior belief and taking the ML observations assumption [28], the differential entropy, considering only the terminal joint belief, for the *unfocused* case is given by (log denotes the natural logarithm in this paper)

$$J_{\mathcal{H}}(\mathcal{U}) = \mathcal{H}(b[X_{k+L}]) = \frac{n}{2} \log(2\pi e) - \frac{1}{2} \log|\Lambda_{k+L}|, \tag{9}$$

where n is the dimension of the joint state X_{k+L} and Λ_{k+L} is the estimated FIM associated with the joint posterior belief $b[X_{k+L}]$. As we are only interested in a *focused* set of variables X_{k+L}^F, the differential entropy, considering only the terminal marginal belief, over the *focused* set, is given by

$$J_{\mathcal{H}}^F(\mathcal{U}) = \mathcal{H}\left(b\left[X_{k+L}^F\right]\right) = \frac{n^F}{2} \log(2\pi e) - \frac{1}{2} \log\left|\Lambda_{k+L}^{M,F}\right|, \tag{10}$$

where n^F is the dimension of X_{k+L}^F and $\Lambda_{k+L}^{M,F}$ is the marginal FIM of the marginal posterior belief $b\left[X_{k+L}^F\right]$.

B. Belief Topology

The joint belief at time k can be represented by a factor graph [22]. A factor graph (*FG*) is a probabilistic graphical model that represents a factorization of a probability density function in terms of process and measurement models. It is a bipartite graph whose nodes consist of factors $\mathcal{F}$ and variables $\mathcal{X}$. The variables $\mathcal{X}$ represent the random variables in the estimation problem while the factors represent probabilistic information on those variables. The FG edges encode connectivity according to the variables involved in each factor. In this work, we only consider pairwise factors without self-loops.

A topological representation of a joint belief is defined as a topological graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ associated to a posterior FG where each node $v_i \in \mathcal{V}$ corresponds to a variable node

$x_i \in \mathcal{X}$, as illustrated in Fig 2. Specifically, if x_0 is fixed as an *anchor* pose then v_0 is the corresponding *anchor* node. We also note that $\mathcal{V}^F$ and $\mathcal{V}^U$ correspond to X^F and X^U, respectively. The edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is defined as

$$e_{ij} \doteq (v_i, v_j) \in \mathcal{E} \iff f_{ij} \doteq (x_i, x_j) \in \mathcal{F}, \tag{11}$$

and the weight function $w : \mathcal{E} \to \mathbb{R}_{>0}$ assigns each edge with a weight w_{ij} derived from the appropriate term in the associated noise model of the corresponding factor.

Let W be a diagonal matrix of size $|\mathcal{V}| \times |\mathcal{V}|$ defined as

$$W_{ii} = \sum_j w_{ij}, \tag{12}$$

i.e. each element on the main diagonal W_{ii} is equal to the sum of weights of all edges connected to node v_i. The weighted Laplacian matrix $\mathcal{L}_w$ associated to a topological graph $\mathcal{G}$ is defined in [6] as

$$\mathcal{L}_w(i,j) = \begin{cases} W_{ii} & \text{if } i = j, \\ -w_{ij} & \text{if } v_i \text{ and } v_j \text{ are adjacent,} \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

The weighted normalized Laplacian matrix $\hat{\mathcal{L}}_w$ associated to a topological graph $\mathcal{G}$ is defined in [5] as

$$\hat{\mathcal{L}}_w(i,j) = \begin{cases} 1 & \text{if } i = j, \\ \frac{-w_{ij}}{\sqrt{W_{ii} \cdot W_{jj}}} & \text{if } v_i \text{ and } v_j \text{ are adjacent,} \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

The reduced weighted Laplacian matrix L_w and the reduced weighted normalized Laplacian matrix $\hat{L}_w$ are retrieved by removing the row and column that correspond to the *anchor* node from $\mathcal{L}_w$ and $\hat{\mathcal{L}}_w$, respectively.

We also denote the reduced incidence matrix of $\mathcal{G}$ by A. It is obtained by removing the row that corresponds to the *anchor* node from the incidence matrix of $\mathcal{G}$.

C. Weighted Tree Connectivity

Given a weighted topological graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ the authors of [14] introduce a weighted version to Kirchoff's matrix tree theorem [2], where the value of each spanning tree $T_\mathcal{G}$ is given by

$$\mathbb{V}_w(T_\mathcal{G}) = \prod_{e \in \mathcal{E}(T_\mathcal{G})} w(e), \tag{15}$$

and the weighted number of spanning trees is defined as

$$t_w(\mathcal{G}) = \sum_{t \in T_\mathcal{G}} \mathbb{V}_w(t). \tag{16}$$

The Weighted Tree Connectivity (WTC) of $\mathcal{G}$ is defined as

$$\tau_w(\mathcal{G}) = \log t_w(\mathcal{G}). \tag{17}$$

Considering the 2D pose SLAM framework addressed in this work, the FIM of a posterior belief is given by (see eq. (18) in [14])

$$\Lambda = \begin{bmatrix} L_{w_p} \otimes I_2 & (A_{w_p} \otimes I_2)\Gamma\Delta_{w_p} \\ *^T & L_{w_\theta} + \Delta_{w_p}^T \Delta_{w_p} \end{bmatrix}, \tag{18}$$

where * denotes the top-right block; I_2 is the identity matrix of size 2×2 and $\otimes$ denotes the Kronecker product; L_{w_p}, L_{w_θ} are the reduced weighted Laplacian matrices of $\mathcal{G}$ when edges are weighted according to $w_p : e_i \to \sigma_{p_i}^{-2}, w_\theta : e_j \to \sigma_{\theta_j}^{-2}$ respectively; $A_{w_p} = A \cdot \text{diag}(\sigma_{p_1}, ..., \sigma_{p_m})$ is the reduced weighted incidence matrix when edges are weighted by w_p; Γ is defined as

$$\Gamma = I_{|\mathcal{E}|} \otimes \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \tag{19}$$

and $\Delta_{w_p}^T \Delta_{w_p} = D_{w_p}$ is a diagonal matrix where

$$D_{w_p}(i,i) = \sum_{j \in \mathcal{N}_{\text{out}(i)}} w_p(i,j) \, ||p_i - p_j||^2, \tag{20}$$

i.e. $D_{w_p}(i,i)$ is equal to the weighted sum of squared distances between the i'th robot pose and every node observed by it.

Under Theorem 3 in [14] the authors also provide lower and upper bounds on the actual D-criterion

$$\tau_w(\mathcal{G}) \leq \log|\Lambda| \leq \tau_w(\mathcal{G}) + n \cdot \log(1 + \delta/\lambda_1), \tag{21}$$

where $\tau_w(\mathcal{G}) = 2\tau_{w_p}(\mathcal{G}) + \tau_{w_\theta}(\mathcal{G})$; $\delta = ||\Delta_{w_p}^T \Delta_{w_p}||_\infty$ and $\lambda_1 = \lambda_{\min}(L_{w_\theta})$. Moreover, they show that under some conditions, these bounds become asymptotically tight and the D-criterion is characterized solely by the weighted tree connectivity of $\mathcal{G}$

$$\lim_{\delta/\lambda_1 \to 0^+} \log|\Lambda| = 2\tau_{w_p}(\mathcal{G}) + \tau_{w_\theta}(\mathcal{G}). \tag{22}$$

From hereon, we denote $\tau_w(\mathcal{G}) \doteq \tau_w$.

IV. Approach

In *focused* topological BSP, we aim to rank candidate actions by redefining the original *focused* problem (10) in a topological space, where it can be solved more efficiently. Given a belief $b[X_k]$ and a set of candidate actions $\mathcal{U}_k$, we look for a topological representation $\mathcal{G}(u_k)$ for each $u_k \in \mathcal{U}_k$ and a signature $S : \mathcal{G}(u_k) \to \mathbb{R}$, that would ideally yield a solution which is consistent with the optimal solution of the original *focused* problem (10) such that $\hat{\mathcal{U}} = \mathcal{U}^*$ where $\hat{\mathcal{U}} = \min_{\mathcal{U}} S(\mathcal{G}(\mathcal{U}))$.

In this work, we derive topological signatures inspired by the works of [14] and [17] on measurement selection and active SLAM in the *unfocused* case, respectively.

A. Manipulating $J_\mathcal{H}^F$

In practice, at each time step, we maintain a factor graph that represents the joint posterior belief from which we can induce the topological graph. Specifically, at time step $k+L$ we have access to FG_{k+L}. As we are interested in a topological representation which corresponds to the marginal posterior belief over the *focused* variables X_{k+L}^F, we would first need to marginalize out the *unfocused* variables X_{k+L}^U. Although such algorithms exist, e.g. the sum-product algorithm [22], as the set of *focused* variables is often small with respect to the entire estimation problem, e.g. when we are interested in reducing the entropy only over the robot's last pose, calculating the

marginal posterior information matrix involves an expensive Schur complement operation.

As the joint posterior information matrix Λ_{k+L} is positive-definite and symmetric, we can partition it such that

$$\Lambda_{k+L} = \begin{bmatrix} \Lambda^F_{k+L} & \Lambda^{F,U}_{k+L} \\ (\Lambda^{F,U}_{k+L})^T & \Lambda^U_{k+L} \end{bmatrix}, \tag{23}$$

where $\Lambda^F_{k+L} \in \mathbb{R}^{n^F \times n^F}$ and $\Lambda^U_{k+L} \in \mathbb{R}^{n^U \times n^U}$ are constructed from Λ_{k+L} by retrieving only the rows and columns related to X^F_{k+L} and X^U_{k+L} respectively. Notice that Λ^U_{k+L} is the conditional posterior information matrix of X^U_{k+L}, conditioned on the rest of the variables X^F_{k+L}. The remaining blocks, $\Lambda^{F,U}_{k+L}$, contain the mixed information between *focused* and *unfocused* variables.

While the marginal posterior information matrix $\Lambda^{M,F}_{k+L}$ can be calculated using the Schur complement, to calculate the *focused* objective function (10) we only need to evaluate $\left|\Lambda^{M,F}_{k+L}\right|$. Using theorem 2.1 from [26] where

$$\left|\Lambda_{k+L}\right| = \left|\Lambda^{M,F}_{k+L}\right| \cdot \left|\Lambda^U_{k+L}\right| \Rightarrow \left|\Lambda^{M,F}_{k+L}\right| = \left|\Lambda_{k+L}\right| / \left|\Lambda^U_{k+L}\right|, \tag{24}$$

and substituting (24) into (10) we rewrite $J^F_{\mathcal{H}}$ as

$$J^F_{\mathcal{H}}(\mathcal{U}) = \frac{n^F}{2}\log(2\pi e) - \frac{1}{2}\log\left|\Lambda_{k+L}\right| + \frac{1}{2}\log\left|\Lambda^U_{k+L}\right|. \tag{25}$$

Using (21), we can approximate $\log|\Lambda_{k+L}|$. To derive an approximation for (25) that is characterized solely by structural properties, we now aim to approximate $\log\left|\Lambda^U_{k+L}\right|$ using topological aspects.

From here on, we drop the time indices and refer to a general posterior belief $b[X]$ with the associated FIM Λ.

B. The unfocused augmented graph $\mathcal{G}^{A,U}$

Our goal is to find a topological graph, and a corresponding signature that would approximate $\log\left|\Lambda^U\right|$.

According to (23), as Λ^U is a principal submatrix of Λ, it is not hard to see that by taking only the rows and columns that correspond to X^U, the *unfocused* block Λ^U is given by

$$\Lambda^U = \begin{bmatrix} L^U_{w_p} \otimes I_2 & (A^U_{w_p} \otimes I_2)\Gamma\Delta^U_{w_p} \\ *^T & L^U_{w_\theta} + (\Delta^U_{w_p})^T\Delta^U_{w_p} \end{bmatrix}. \tag{26}$$

We observe that both $L^U_{w_p}$ and $L^U_{w_\theta}$ represent reduced weighted Laplacian matrices, that contain all *unfocused* variables and weighted edges connected to those variables (see e.g. Fig 2c). As such, to construct a topological graph, we can add a *virtual* node v_a, to create the augmented weighted topological graph $\mathcal{G}^{U,A}$. Note that the *anchor* node and all nodes that represent *focused* variables do not belong to $\mathcal{G}^{U,A}$.

This augmented weighted topological graph is formally defined as $\mathcal{G}^{U,A}\left(\mathcal{V}^U \bigcup v_a, \mathcal{E}^{U,A}, w^{U,A}\right)$ where each node $v_i \in \mathcal{V}^U$ corresponds to a variable $x_i \in X^U$; the edge set $\mathcal{E}^{U,A}$ is defined as a union of three groups $\{e^{U,A}_{ij} | v_i, v_j \in \mathcal{V}^U \wedge e_{ij} \in \mathcal{E}\} \cup \{e^{U,A}_{ia} | v_i \in \mathcal{V}^U, v_j \in \mathcal{V}^F \wedge e_{ij} \in \mathcal{E}\} \cup \{e^{U,A}_{ia} | v_i \in \mathcal{V}^U \wedge e_{i0} \in \mathcal{E}\}$ where e_{i0} denotes an edge connected to the *anchor* node; each edge $e^{U,A}_{ia} \in \mathcal{E}^{U,A}$ connected to the *virtual* node v_a is weighted by

$$w^{U,A}_{ia} = \sum_j w_{ij} \text{ such that } v_i \in \mathcal{V}^U \text{ and } v_j \in \mathcal{V}^F, \tag{27}$$

while the remaining edges are weighted according to $w^{U,A}_{ij} = w_{ij}$. Note that in (27), edges to the *virtual* node are weighted such that we construct a Laplacian matrix, i.e. where the sum of each row and each column is exactly zero. We can now use the weighted version of Kirchhoff's theorem (17) as a topological signature for this graph. Specifically, we define

$$\tau^{U,A}_w = 2\tau_{w_p}(\mathcal{G}^{U,A}) + \tau_{w_\theta}(\mathcal{G}^{U,A}). \tag{28}$$

C. Weighted Tree Connectivity Difference Signature

We approximate the *focused* objective function (25), using the Weighted Tree Connectivity Difference (WTCD) between $\mathcal{G}$ and the augmented graph induced from it $\mathcal{G}^{U,A}$. This topological signature is defined as

$$S_{WTCD} = \frac{n^F}{2}\log(2\pi e) - \frac{1}{2}\left[\tau_w - \tau^{U,A}_w\right]. \tag{29}$$

Using this signature we can get an approximated solution $\hat{\mathcal{U}}$ whose properties we analyze from here on.

Theorem 1. *The WTC of $\mathcal{G}^{U,A}$ asymptotically bounds $\log\left|\Lambda^U\right|$.*

Proof. Following a similar proof to theorem 3 in [14] we get

$$\tau^{U,A}_w \le \log|\Lambda^U| \le \tau^{U,A}_w + n^U \cdot \log(1 + \delta^U/\lambda^U_1), \tag{30}$$

where $\delta^U = ||\left(\Delta^U_{w_p}\right)^T \Delta^U_{w_p}||_\infty$ and $\lambda^U_1 = \lambda_{\min}(L^U_{w_\theta})$. It is easy to see that

$$\log\left|\Lambda^U\right| \underset{\delta^U/\lambda^U_1 \to 0^+}{=} \tau^{U,A}_w. \tag{31}$$

■

Theorem 2. *Let $\epsilon(J^F_{\mathcal{H}}) \doteq J^F_{\mathcal{H}} - S_{WTCD}$ be the approximation error, then it is bounded by*

$$-\frac{n}{2}\log\left(1 + \frac{\delta}{\lambda_1}\right) \le \epsilon(J^F_{\mathcal{H}}) \le \frac{n^U}{2}\log\left(1 + \frac{\delta^U}{\lambda^U_1}\right). \tag{32}$$

Proof. Using inequalities (21), (30) and eq. (25) we get

$$\begin{aligned} \mathcal{UB}[J^F_{\mathcal{H}}] &= \frac{n^F}{2}\log(2\pi e) - \frac{1}{2}\tau_w + \\ &+ \frac{1}{2}\left[\tau^{U,A}_w + n^U \cdot \log\left(1 + \frac{\delta^U}{\lambda^U_1}\right)\right] \\ &= S_{WTCD} + \frac{n^U}{2}\log\left(1 + \frac{\delta^U}{\lambda^U_1}\right). \end{aligned} \tag{33}$$

Similarly, the lower bound is given by

$$\begin{aligned} \mathcal{LB}[J^F_{\mathcal{H}}] &= \frac{n^F}{2}\log(2\pi e) - \frac{1}{2}\left[\tau_w + n \cdot \log\left(1 + \frac{\delta}{\lambda_1}\right)\right] + \\ &+ \frac{1}{2}\tau^{U,A}_w = S_{WTCD} - \frac{n}{2}\log\left(1 + \frac{\delta}{\lambda_1}\right). \end{aligned} \tag{34}$$

■

Lemma 1.

$$\delta/\lambda_1 \to 0^+ \Rightarrow \delta^U/\lambda^U_1 \to 0^+.$$

Proof. Since the FIM is Hermitian and since Λ^U is a principal submatrix of Λ, according to Cauchy's interlacing theorem [8], the eigenvalues must satisfy

$$\lambda_1 \leq \lambda_1^U \leq \lambda_2 \leq .... \leq \lambda_n \leq \lambda_n^U \leq \lambda_{n+1}. \tag{35}$$

In addition, as n^U is a finite number, i.e. represents the size of a finite graph, and as $w^{U,A}$ is defined over finite non negative weights, $\exists \alpha < \infty$ such that $\delta^U < \alpha\delta$. Combined with (35), we get

$$\delta/\lambda_1 \to 0^+ \Rightarrow \alpha\delta/\lambda_1 \to 0^+ \Rightarrow \delta^U/\lambda_1^U \to 0^+. \tag{36}$$

■

Lemma 1 implies that when $\delta/\lambda_1 \to 0^+$, the approximation error in (32) approaches zero.

While S_{WTCD} asymptotically converges to the *focused* objective function (25), evaluating this signature still requires calculating the determinant of the associated Laplacian matrices. Moreover, if we would like to use these bounds to eliminate candidate actions, we need to perform eigenvalue decomposition to retrieve λ_1, which cannot be done *online*.

Theorem 3. *The approximation error $\epsilon(J_{\mathcal{H}}^F)$ can also be bounded using topological aspects only, where*

$$\frac{1}{2}\left(\tau_{w_\theta} - \sum_{i=1}^{n}\log\left[W_\theta(i,i)+\delta\right]\right) \leq \epsilon\left(J_{\mathcal{H}}^F\right) ... \tag{37}$$

$$\leq \frac{1}{2}\left(\sum_{i=1}^{n}\log\left[W_\theta^{U,A}(i,i)+\delta^{U,A}\right] - \tau_{w_\theta}^{U,A}\right). \tag{38}$$

Proof. We follow [16] and derive the Hadamard bounds for the *focused* case. Following the proof to Theorem 3 in [14] we know that

$$\log|\Lambda| \leq 2\tau_{w_p} + \log|L_{w_\theta} + \delta I|. \tag{39}$$

We denote by W_θ the diagonal matrix defined in (12) based on weights w_θ defined in (18). Since $L_{w_\theta} + \delta I$ is a positive-definite matrix, applying Hadamard inequality we get

$$\log|L_{w_\theta} + \delta I| \leq \sum_{i=1}^{n}\log\left[W_\theta(i,i)+\delta\right]. \tag{40}$$

Replacing (40) into (39) we get

$$\log|\Lambda| \leq 2\tau_{w_p} + \sum_{i=1}^{n}\log\left[W_\theta(i,i)+\delta\right]. \tag{41}$$

Similarly, for $\log|\Lambda^U|$ we get

$$\log|\Lambda^U| \leq 2\tau_{w_p}^{U,A} + \sum_{i=1}^{n}\log\left[W_\theta^{U,A}(i,i)+\delta^{U,A}\right]. \tag{42}$$

Replacing (41) and (42) into the definition of S_{WTCD} we get new bounds for $\epsilon\left(J_{\mathcal{H}}^F\right)$. ■

While these bounds are somewhat more conservative, they are functions of topological aspects only. We avoid the eigenvalue decomposition and can use them *online*.

Algorithm 1: Focused Topological BSP

Input: Set of factor graphs corresponding to candidate actions $\mathcal{U}$ and a graph signature S
Output: approximate solution to FT-BSP: $\hat{\mathcal{U}}$
foreach *candidate action* $\mathcal{U}$ **do**
  infer topological graph $\mathcal{G}$ (section III. B)
  construct *unfocused* augmented topological graph $\mathcal{G}^{U,A}$ (section IV. B)
  evaluate S using $\mathcal{G}$ and $\mathcal{G}^{U,A}$ (eq. 29 or eq. 49)
end
return $\hat{\mathcal{U}} = \min_{\mathcal{U}} S(\mathcal{U})$

D. Von Neumann Difference Signature

Inspired by [17], we also propose a second topological signature to be evaluated *online*. The Von Neumann entropy $\hat{\mathcal{H}}_{VN}$ of a topological graph $\mathcal{G}$ was introduced in [27] as the Shannon entropy associated with the eigenvalues of its normalized Laplacian

$$\mathcal{H}_{VN}\left(\hat{\mathcal{L}}_w\right) = -\sum_{i=1}^{n}\frac{\hat{\lambda}_i}{2}\log\frac{\hat{\lambda}_i}{2}. \tag{43}$$

Using the quadratic approximation, as proposed in [7], we rewrite (43) for the weighted case as

$$\mathcal{H}_{VN} \approx \tilde{\mathcal{H}}_{VN} = \frac{n\log 2}{2} - \frac{1}{2}\left(\mathrm{Tr}\left[\hat{\mathcal{L}}_w^2\right] - n\right). \tag{44}$$

Following a similar derivation as for equation (8) in [7], the trace of the square of the weighted normalized Laplacian is

$$\mathrm{Tr}\left[\hat{\mathcal{L}}_w^2\right] = n + 2\sum_{e_{ij}\in\mathcal{E}}\frac{w_{ij}}{W_{ii}\cdot W_{jj}}. \tag{45}$$

Substituting (45) into (44) we get an expression for the approximated Von Neumann graph entropy for the weighted case

$$\tilde{\mathcal{H}}_{VN}\left(\hat{\mathcal{L}}_w\right) = \frac{n\log 2}{2} - \sum_{e_{ij}\in\mathcal{E}}\frac{w_{ij}}{W_{ii}\cdot W_{jj}}. \tag{46}$$

We denote

$$h_w = 2\tilde{\mathcal{H}}_{VN}\left(\hat{\mathcal{L}}_{w_p}\right) + \tilde{\mathcal{H}}_{VN}\left(\hat{\mathcal{L}}_{w_\theta}\right) \tag{47}$$

$$h_w^{U,A} = 2\tilde{\mathcal{H}}_{VN}\left(\hat{\mathcal{L}}_{w_p}^{U,A}\right) + \tilde{\mathcal{H}}_{VN}\left(\hat{\mathcal{L}}_{w_\theta}^{U,A}\right), \tag{48}$$

and formally define the Von Neumann Difference (VND) topological signature as

$$S_{VND} = \frac{n^F}{2}\log\left(2\pi e\right) - \frac{1}{2}\left[h_w - h_w^{U,A}\right]. \tag{49}$$

Calculating (46) is dependent on the diagonal matrix W and generally has a quadratic complexity in the number of nodes $O(n^2)$. However, in the context of BSP, as the dimensionality of n grows with time, the information matrix and the topological representation become sparse. Evaluating (46), in this case, only depends on a small number of non-zero elements, i.e. the number of edges $|\mathcal{E}|$.

Incremental aspects: Calculating (49) requires evaluating $\tilde{\mathcal{H}}_{VN}$ for the weighted Laplacian matrices of both $\mathcal{G}$ and $\mathcal{G}^{U,A}$. However, we can re-use calculations rather than evaluating each from scratch. The approximated Von Neumann entropies (44) for a posterior belief, at time $k+L$, are given by

$$\tilde{\mathcal{H}}_{VN}\left(\hat{\mathcal{L}}_w\right) = \frac{(k+L+1)\,\log 2}{2} - \sum_{e_{ij}\in\mathcal{E}} \frac{w_{ij}}{W_{ii}\cdot W_{jj}} \tag{50}$$

$$\tilde{\mathcal{H}}_{VN}\left(\hat{\mathcal{L}}_w^{U,A}\right) = \frac{\left(k+L+2-\left|n^F\right|\right)\,\log 2}{2} - \sum_{e_{ij}\in\mathcal{E}^{U,A}} \frac{w_{ij}^{U,A}}{W_{ii}^{U,A}\cdot W_{jj}^{U,A}} \tag{51}$$

Subtracting (51) from (50) and rearranging the result we get

$$\tilde{\mathcal{H}}_{VN}\left(\hat{\mathcal{L}}_w^{U,A}\right) = \tilde{\mathcal{H}}_{VN}\left(\hat{\mathcal{L}}_w\right) + \frac{\left(1-\left|n^F\right|\right)\,\log 2}{2} + \Delta, \tag{52}$$

where

$$\Delta = \sum_{e_{ij}\in\mathcal{E}} \frac{w_{ij}}{W_{ii}\cdot W_{jj}} - \sum_{e_{ij}\in\mathcal{E}^{U,A}} \frac{w_{ij}^{U,A}}{W_{ii}^{U,A}\cdot W_{jj}^{U,A}}. \tag{53}$$

According to (27), all edges that are not connected to the *virtual* node in the augmented graph, share the same weights in both graphs. We denote an edge that does not connect to an *unfocused* node by $e^{\neg U}$ and by e^V an edge which is connected to the *virtual* node. Reducing all shared terms between the two sums in (53), we rewrite Δ as

$$\Delta = \sum_{e_{ij}^{\neg U}\in\mathcal{E}} \frac{w_{ij}}{W_{ii}\cdot W_{jj}} - \sum_{e_{ij}^{V}\in\mathcal{E}^{U,A}} \frac{w_{ij}^{U,A}}{W_{ii}^{U,A}\cdot W_{jj}^{U,A}}. \tag{54}$$

Notice that evaluating Δ is dependent on the number of *focused* variables and their connectivity. The smaller they are, with respect to the entire problem, the more we gain in terms of computational costs.

Following similar steps, we can also derive a recursive update rule for calculating the VND signature for the posterior graph $S_{VND}\left(\mathcal{G}_{k+L}\right)$ from the VND signature of the prior $S_{VND}\left(\mathcal{G}_k\right)$ for each candidate action. As such, calculating this signature is computationally very efficient.

signature	measurement selection	active SLAM
S_{WTCD}	18.88	1.21
S_{VND}	**12.02**	**0.14**
$J_{\mathcal{H}}^F$	146.24	6.34

TABLE I: Average time in milliseconds for calculating both signatures and the *focused* objective function for each candidate action. We note that while the $WTCD$ signature is calculated using a highly optimized Matlab code with a C back-end and that $J_{\mathcal{H}}^F$ is calculated in C++, we calculate the VND signature purely in Matlab without code optimization.

V. Results

We evaluate our approach considering a measurement selection problem and an active 2D pose SLAM simulation, to empirically study the two topological signatures.

In our first experiment, we use the 2D pose SLAM Intel dataset [3], which contains $n = 1228$ robot poses and 278 loop-closure observations. We modified the original covariance matrix to have a block-isotropic structure. In this measurement selection problem, the goal is to find the most informative subset of observations, with respect to a random *focused* set of variables of size $n^F < n/2$. All sub-graphs share the same vertex set and contain all odometry edges. In total, we generate $n_L = 278$ such sub-graphs where, for each sub-graph we randomly choose a subset of loop closure edges. Given the *focused* set of variables, we evaluate the original *focused* objective function and the two topological signatures. This experiment was performed several times for different *focused* sets in different sizes. In Fig. 3 we see that both signatures are highly correlated with the *focused* objective function (25) given a specific *focused* set. In Table I we report the average run time statistics for all runs showing the improvement in computational cost.

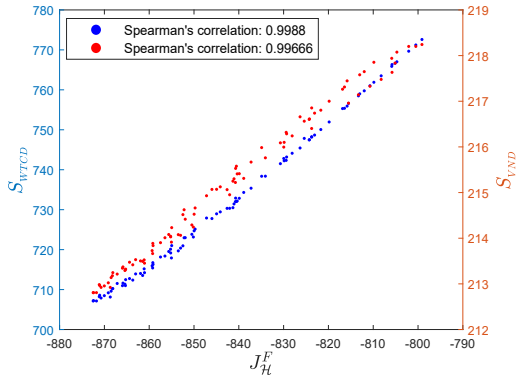


Fig. 3: Topological signatures vs marginal entropy for different loop-closure measurements given a random *focused* set of poses in the Intel dataset.

In our second experiment, we evaluate our approach in an active 2D pose SLAM simulation. In this scenario, the robot's objective is to reach a predefined goal with maximum accuracy, while navigating in an unknown environment. Using a probabilistic roadmap (PRM) [13] we first discretize the environment. We assume that the robot previously visited some areas within the map and that all planning sessions start right after. We then randomly generate a set of candidate paths, over the roadmap, all ending at the predefined goal (see e.g. Fig 4). As such, the *focused* set of variables is defined as the last robot pose in each candidate path. i.e. the objective is to reduce the uncertainty over that position. Given all candidate paths, we evaluate the original *focused* objective function and the two topological signatures. As seen in Fig. 5 both signatures are highly correlated with the *focused* objective function (25). The figure also shows that in this setting, the bounds developed in Theorem 2 are sufficiently tight to allow action elimination. However, the Hadamard bounds developed in Theorem 3, are not informative in this case as they are not tight enough to allow action elimination. In general, Hadamard bounds are more informative in diagonally dominant matrices. In the specific case presented δ was relatively small.

VI. Conclusions

In this work, we introduced a novel concept that considers topological aspects for decision making under uncertainty, with respect to a *focused* set of variables. In the context of information theoretic problems, we developed two such topological signatures and empirically showed that there is a very high correlation between both signatures and the *focused*

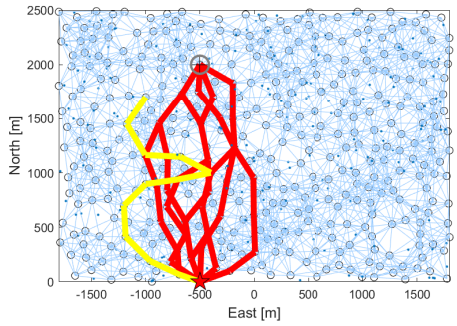


Fig. 4: different candidate paths generated on top of a PRM in a single planning session. Stating position is denoted with a red star and the goal is denoted with a circle. The yellow path represents the locations that the robot had previously visited and acts as the initial belief.

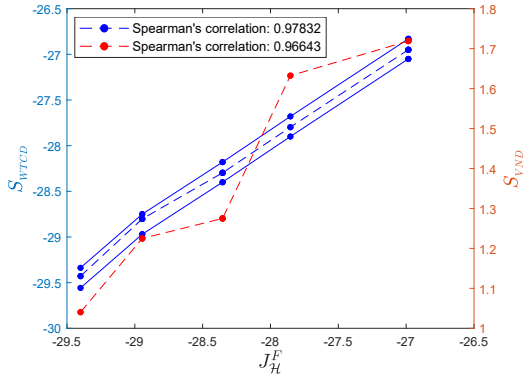


Fig. 5: Topological signatures in dashed lines vs $J^F_{\mathcal{H}}$ in a *focused* BSP problem. The solid lines represent the bounds developed in Theorem 2.

information theoretic objective function in two different problems. We developed bounds for the approximation error of the WTCD signature and showed that under some conditions it converges to the optimal solution. As such, we can use this signature and bounds to discriminate between candidate actions. Furthermore, as calculating these bounds requires eigenvalue decomposition, we also introduced the Hadamard bounds, that are based solely on topological aspects and can be calculated *online*. We then derived the VND signature, which extends the work of [17], to handle the *focused* case and weighted Laplacian matrices. Not only that calculating this signature is more computationally efficient, we also showed that it can be done in incremental fashion.

Finally, as can be seen in Table I, we found that calculating both signatures is an order of magnitude faster then calculating the objective function in the measurement selection experiment. In the active SLAM experiment, the advantage of using the VND signature is even more prominent.

REFERENCES

[1] Ruzena Bajcsy. Active perception. *Proceedings of the IEEE*, 76(8):966–1005, 1988.

[2] Norman Biggs, Norman Linstead Biggs, and Biggs Norman. *Algebraic graph theory*, volume 67. Cambridge university press, 1993.

[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(7):965–987, 2014.

[4] Yongbo Chen, Shoudong Huang, Robert Fitch, Liang Zhao, Huan Yu, and Di Yang. On-line 3d active pose-graph slam based on key poses using graph topology and sub-maps. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 169–175. IEEE, 2019.

[5] Fan Chung and Ross M Richardson. Weighted laplacians and the sigma function of a graph. *Contemporary Mathematics*, 415:93–108, 2006.

[6] Fan RK Chung and Fan Chung Graham. *Spectral graph theory*. Number 92. American Mathematical Soc., 1997.

[7] Lin Han, Francisco Escolano, Edwin R Hancock, and Richard C Wilson. Graph characterizations from von neumann entropy. *Pattern Recognition Letters*, 33(15):1958–1967, 2012.

[8] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.

[9] Vadim Indelman, Luca Carlone, and Frank Dellaert. Planning in the continuous domain: A generalized belief space approach for autonomous navigation in unknown environments. *The International Journal of Robotics Research*, 34(7):849–882, 2015.

[10] Leslie Pack Kaelbling, Michael L Littman, and Anthony R Cassandra. Planning and acting in partially observable stochastic domains. *Artificial intelligence*, 101(1-2):99–134, 1998.

[11] Michael Kaess, Viorela Ila, Richard Roberts, and Frank Dellaert. The bayes tree: An algorithmic foundation for probabilistic robot mapping. In *Algorithmic Foundations of Robotics IX*, pages 157–173. Springer, 2010.

[12] Michael Kaess, Ananth Ranganathan, and Frank Dellaert. isam: Incremental smoothing and mapping. *IEEE Transactions on Robotics*, 24(6):1365–1378, 2008.

[13] Lydia E Kavraki, Petr Svestka, J-C Latombe, and Mark H Overmars. Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *IEEE transactions on Robotics and Automation*, 12(4):566–580, 1996.

[14] Kasra Khosoussi, Matthew Giamou, Gaurav S Sukhatme, Shoudong Huang, Gamini Dissanayake, and Jonathan P How. Reliable graphs for slam. *The International Journal of Robotics Research*, 38(2-3):260–298, 2019.

[15] Kasra Khosoussi, Shoudong Huang, and Gamini Dissanayake. Good, bad and ugly graphs for slam. In *RSS Workshop on the problem of mobile sensors*, 2015.

[16] A. Kitanov and V. Indelman. Topological information-theoretic belief space planning with optimality guarantees. *arXiv preprint arXiv:1903.00927*, 2019.

[17] Andrej Kitanov and Vadim Indelman. Topological multi-robot belief space planning in unknown environments. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1–7. IEEE, 2018.

[18] Dmitry Kopitkov and Vadim Indelman. No belief propagation required: Belief space planning in high-dimensional state spaces via factor graphs, the matrix determinant lemma, and re-use of calculation. *The International Journal of Robotics Research*, 36(10):1088–1130, 2017.

[19] Dmitry Kopitkov and Vadim Indelman. General-purpose incremental covariance update and efficient belief space planning via a factor-graph propagation action tree. *The International Journal of Robotics Research*, 38(14):1644–1673, 2019.

[20] Andreas Krause, Ajit Singh, and Carlos Guestrin. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *Journal of Machine Learning Research*, 9(Feb):235–284, 2008.

[21] Vikram Krishnamurthy. *Partially observed Markov decision processes*. Cambridge University Press, 2016.

[22] Frank R Kschischang, Brendan J Frey, and H-A Loeliger. Factor graphs and the sum-product algorithm. *IEEE Transactions on information theory*, 47(2):498–519, 2001.

[23] Daniel S Levine and Jonathan P How. Sensor selection in high-dimensional gaussian trees with nuisances. In *Advances in Neural Information Processing Systems*, pages 2211–2219, 2013.

[24] Abbe Mowshowitz and Matthias Dehmer. Entropy and the complexity of graphs revisited. *Entropy*, 14(3):559–570, 2012.

[25] Beipeng Mu, Ali-akbar Agha-mohammadi, Liam Paull, Mathew Graham, Jonathan How, and John Leonard. Two-stage focused inference for resource-constrained collision-free navigation. 2015.

[26] Diane Valerie Ouellette. Schur complements and statistics. *Linear Algebra and its Applications*, 36:187–295, 1981.

[27] Filippo Passerini and Simone Severini. Quantifying complexity in networks: the von neumann entropy. *International Journal of Agent Technologies and Systems (IJATS)*, 1(4):58–67, 2009.

[28] Robert Platt Jr, Russ Tedrake, Leslie Kaelbling, and Tomas Lozano-Perez. Belief space planning assuming maximum likelihood observations. 2010.

[29] Harold W Sorenson. Parameter estimation, volume 9 of control and system theory, 1980.