2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

Scalable Networked Feature Selection with Randomized Algorithm for Robot Navigation

Vivek Pandey*, Arash Amini*, Guangyi Liu, Ufuk Topcu, Qiyu Sun, Kostas Daniilidis, and Nader Motee

***Abstract*— We address the problem of sparse selection of visual features for localizing a team of robots navigating in an unknown environment, where robots can exchange relative position measurements with neighbors. We select a set of the most informative features by anticipating their importance in robots localization by simulating trajectories of robots over a prediction horizon. Through theoretical proofs, we establish a crucial connection between graph Laplacian and the importance of features. We leverage a scalable randomized algorithm for sparse sums of positive semidefinite matrices to efficiently select a set of the most informative features.**

I. Introduction

Navigation is a critical component of robotics, enabling robots to perform various tasks autonomously. For successful navigation, robots rely on continuous position estimation or localization. To obtain a good estimate of their position, robots rely on algorithms like Kalman Filter [12].

A key challenge in robot navigation is achieving accurate localization with provable performance bounds while minimizing computational costs due to onboard power limitations. Extensive research has addressed this challenge, focusing on efficient methods for robot localization [9, 4]. These methods typically involve selecting informative features based on the robot's current position and projected trajectory. In [14, 13], the authors utilize computationally efficient sparse extended information filters for simultaneous localization and mapping (SLAM), avoiding the computationally expensive matrix inversion step. In [2], the authors leverage a linear motion and vision model to track features along the robot's predicted trajectory. They then select the most informative features using the greedy method. In [3], the authors utilize the stochastic greedy algorithm proposed in [5] for feature selection for LiDAR SLAM. In [6], the authors propose a randomized algorithm that offers probabilistic performance guarantees while improving the algorithmic complexity of the feature selection problem. Feature selection methods have traditionally targeted single agents, but networked systems like consensus networks [7] demand techniques suited for interconnected systems.

This work was supported partly by ONR N00014-23-1-2779.

∗ Vivek Pandey and Arash Amini contributed equally to this work.

V. Pandey, G. Liu and N. Motee are with the Department of Mechanical Engineering and Mechanics, Lehigh University. {vkp219,gliu,motee}@lehigh.edu.

A. Amini and U. Topcu are with the Department of Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin. {a.amini,utopcu}@utexas.edu.

Q. Sun is with the Department of Mathematics, University of Central Florida. qiyu.sun@ucf.edu.

K. Daniilidis is with the Department of Computer and Information Sciences, University of Pennsylvania. kostas@cis.upenn.edu.

Building upon the sparse feature selection [6], we address the localization problem for a team of robots navigating in an unknown environment. The robots observe relative position measurements with their neighbors and are assumed to follow a predefined trajectory. Each robot utilizes a camera to track selected features, aiming to enhance collective position estimation. We propose a randomized algorithm for feature selection in a multi-agent scenario where robots exchange relative measurement over a communication graph. To quantify the importance of features, we simulate the trajectories of robots over a fixed prediction horizon and obtain information matrices corresponding to each feature. We employ recent findings on tail bounds for sums of random matrices [15] to significantly enhance the probabilistic guarantees of our proposed algorithm. We further analyze the crucial role of network connectivity in both feature selection problem and performance guarantees.

Our contribution: We address the problem of sparse feature selection for localizing a team of agents, where they exchange relative measurements leading to a graphical network. Compared to [6], we significantly improve the probabilistic bound of the randomized feature selection algorithm. By fusing the inter-agent measurement information, we establish crucial connection between graph Laplacian and leverage score of features. The proofs of all theoretical results are provided in the Appendix of [8].

II. Mathematical Notation

We denote a vector and matrix valued variable by lowercase and uppercase letters respectively. We reserve the notation $S^n_{++}(S^n_+)$ to denote the cone of symmetric positive (semi) definite $n \times n$ matrices. The Kronecker product of two matrices X_1, X_2 is denoted by $X_1 \otimes X_2$. For every $X_1, X_2 \in S^n_{++}$, we write $X_2 \preceq X_1$ if and only if $X_1 - X_2 \in S^n_+$. For a matrix X, we denote its transpose by X^T and its trace by $\mathrm{Tr}(X)$. We denote the special orthogonal group in $\mathbb{R}^3$ by $SO(3)$. The cardinality of a set A is given by $|A|$. We reserve the symbol I_n for Identity matrix of size n. A Gaussian random vector z with mean μ and covariance matrix Σ is denoted by $z \sim \mathcal{N}(\mu, \Sigma)$. The set of nonnegative integers and reals are denoted by $\mathbb{Z}_+$ and $\mathbb{R}_+$ respectively.

Graph Theory: A weighted graph is defined by $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \omega)$, where $\mathcal{V}$ is the set of nodes, $\mathcal{E}$ is the set of edges, and $\omega : \mathcal{V} \times \mathcal{V} \to \mathbb{R}_+$ is the weight function that assigns a nonnegative number to every link. Two nodes are directly connected if and only if $(i, j) \in \mathcal{E}$.

The Incidence matrix of a directed graph $\mathcal{G}$ is a $|\mathcal{V}| \times |\mathcal{E}|$ matrix C such that $C_{ij} \in \{-1, 0, 1\}$. The Laplacian matrix of $\mathcal{G}$ is $L = CWC^T$ where W is a $|\mathcal{E}| \times |\mathcal{E}|$ diagonal matrix containing the edge weights [10].

Definition 1. *A map $\rho : S_+^n \to \mathbb{R}$ is called monotonically decreasing if for every $X_2 \preceq X_1 \implies \rho(X_2) \geq \rho(X_1)$.*

III. Problem Statement

We consider a multi-agent navigation scenario where agents localize themselves using camera and inter-agent measurement communication. Due to large number of visual features extracted from camera, selecting a subset of informative features becomes necessary to address computational complexity. We address the problem of sparse feature selection to estimate positions of a team of N robots navigating in an unknown environment, where robots exchange inter-agent measurements through a network. Using a randomized algorithm, we select a set of the most informative features by anticipating their importance over a fixed time horizon. We further analyze how graph Laplacian affects performance measures and relative importance of features.

Let $\boldsymbol{x}_\tau \in \mathbb{R}^{3N}$ be the state vector of N robots at time $\tau \in \mathbb{Z}_+$. For $M \in \mathbb{Z}_+$, consider the vector of time horizon given by $[t : t+M] = [t, t+1, \cdots, t+M]$. The state vector of N robots for all $\tau \in [t : t+M]$ can be written as

$$\mathbf{x}_\star = \left[\boldsymbol{x}_t^T, \boldsymbol{x}_{t+1}^T, \cdots, \boldsymbol{x}_{t+M}^T\right]^T \in \mathbb{R}^{3N(M+1)}. \tag{1}$$

State estimation often requires the Kalman Filter, which yields the mean vector $\bar{\boldsymbol{\mu}}_\star \in \mathbb{R}^{3N(M+1)}$, and the covariance matrix $\bar{\boldsymbol{\Sigma}}_\star \in S_+^{3N(M+1)}$ of the state $\mathbf{x}_\star$. To avoid matrix inversion step in Kalman Filter, we utilize information filter proposed in [14], which updates information vector and information matrix instead of the mean vector and the covariance matrix.

For the mean vector $\bar{\boldsymbol{\mu}}_\star$, and the covariance matrix $\bar{\boldsymbol{\Sigma}}_\star$, the corresponding information vector $\bar{\mathbf{b}}_\star^T \in \mathbb{R}^{3N(M+1)}$, and information matrix $\bar{\mathbf{H}}_\star \in S_+^{3N(M+1)}$, can be written as

$$\bar{\mathbf{b}}_\star = \bar{\boldsymbol{\mu}}_\star \bar{\boldsymbol{\Sigma}}_\star^{-1}, \qquad \bar{\mathbf{H}}_\star = \bar{\boldsymbol{\Sigma}}_\star^{-1}. \tag{2}$$

Information filters offer computational advantages because the information vector and matrix can be updated linearly by adding the contribution from new observations.

Let Θ_t be the set of all features available at time t. Furthermore, let the vector $\mathbf{b}_\star^f$ and the matrix $\mathbf{H}_\star^f$ be the contribution of new features to the information filter, then the information vector and information matrix are updated as

$$\mathbf{b}_\star(\Theta_t) = \bar{\mathbf{b}}_\star + \sum_{f \in \Theta_t} \mathbf{b}_\star^f, \qquad \mathbf{H}_\star(\Theta_t) = \bar{\mathbf{H}}_\star + \sum_{f \in \Theta_t} \mathbf{H}_\star^f. \tag{3}$$

Due to limitations in on-board computational resources, it is often desirable to select a smaller subset of features Φ_t from Θ_t. This subset should prioritize features that contribute significantly to the overall information matrix, maximizing the information gain while minimizing the computational burden.

Our primary aim in this work is to select informative features that collectively lead to enhanced position estimates for all robots within the team. To achieve this, we model the team as a network, fusing relative position measurements with their individual visual observations within a comprehensive robot motion model. The subsequent section delves further into the details of each of these models.

IV. Model for Robot Motion, Relative Measurements and Vision System

This section begins by introducing the robot motion model, which lays the foundation for incorporating measurement models, used to refine the position estimates based on sensor data.

A. Model for Robots Motion

We consider a team of N robots such that the dynamics of the i'th robot at time τ is governed by

$$x_\tau^i = A x_{\tau-1}^i + B u_{\tau-1}^i + \delta_\tau^i, \tag{4}$$

for all $\tau \in [t : t+M]$, where x_τ^i and $u_{\tau-1}^i$ are the position vector and the control input respectively. $\delta_\tau^i \sim \mathcal{N}(0, \Lambda_\tau)$ such that $\mathbb{E}[\delta_{\tau_1}^i \delta_{\tau_2}^i] = 0$ for all $\tau_1 \neq \tau_2$.

Each robot moves along a pre-defined trajectory such that the control input for the i'th robot is given by

$$u_{\tau-1}^i = \boldsymbol{h}\left(u_{\tau-1}^{i,\text{ref}}, x_{\tau-1}^i\right). \tag{5}$$

Since the state $x_{\tau-1}^i$ is unknown, we design the control law using $\mu_{\tau-1}^i$ as

$$u_{\tau-1}^i = \boldsymbol{h}\left(u_{\tau-1}^{i,\text{ref}}, \mu_{\tau-1}^i\right). \tag{6}$$

The overall dynamics of N robots can be written as

$$\boldsymbol{x}_\tau = \boldsymbol{A}\boldsymbol{x}_{\tau-1} + \boldsymbol{B}\boldsymbol{u}_{\tau-1} + \boldsymbol{E}\boldsymbol{\delta}_\tau, \tag{7}$$

where $\boldsymbol{A} = I_N \otimes A, \boldsymbol{B} = I_N \otimes B$, and $E = I_N \otimes I_3$.

Let $\bar{\boldsymbol{\mu}}_t = \mathbb{E}[\boldsymbol{x}_\tau]$ and $\bar{\boldsymbol{\Sigma}}_t = \mathbb{E}[(\boldsymbol{x}_\tau - \bar{\boldsymbol{\mu}}_t)(\boldsymbol{x}_\tau - \bar{\boldsymbol{\mu}}_t)^T]$, then the mean and covariance of $\mathbf{x}_\star$ can be expressed as

$$\bar{\boldsymbol{\mu}}_\star = \begin{bmatrix} \bar{\boldsymbol{\mu}}_t & \bar{\boldsymbol{\mu}}_{t+1} & \cdots & \bar{\boldsymbol{\mu}}_{t+M} \end{bmatrix} \tag{8}$$

$$\bar{\boldsymbol{\Sigma}}_\star = \begin{bmatrix} \bar{\boldsymbol{\Sigma}}_t & \bar{\boldsymbol{\Sigma}}_{t,t+1} & \cdots & \bar{\boldsymbol{\Sigma}}_{t,t+M} \\ \bar{\boldsymbol{\Sigma}}_{t,t+1}^T & \bar{\boldsymbol{\Sigma}}_{t+1} & \cdots & \bar{\boldsymbol{\Sigma}}_{t+1,t+M} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\boldsymbol{\Sigma}}_{t,t+M}^T & \bar{\boldsymbol{\Sigma}}_{t+1,t+M}^T & \cdots & \bar{\boldsymbol{\Sigma}}_{t+M} \end{bmatrix} \tag{9}$$

where $\bar{\boldsymbol{\Sigma}}_\tau = \boldsymbol{A}\bar{\boldsymbol{\Sigma}}_{\tau-1}\boldsymbol{A}^T + I_N \otimes \Lambda_\tau$ for all $\tau \in [t : t+M]$ and

$$\bar{\boldsymbol{\Sigma}}_{\tau_1,\tau_2} = \boldsymbol{A}^{(\tau_2-\tau_1)}\bar{\boldsymbol{\Sigma}}_{\tau_1}$$

for all $\tau_1, \tau_2 \in [t : t+M]$ with $\tau_1 < \tau_2$.

The information matrix of $\mathbf{x}_\star$ given by

$$\bar{\mathbf{H}}_\star = \bar{\boldsymbol{\Sigma}}_\star^{-1}, \tag{10}$$

is a measure of information content of $\mathbf{x}_\star$, when the system dynamics is predicted over the time horizon M.

B. Information from Relative Measurements

To improve estimate of $\mathbf{x}_\star$ from predicted dynamics, we fuse relative measurements obtained through connected network $\mathcal{G}_\tau$.

The relative measurement vector $\xi_\tau^{i,j}$ at time τ between two neighboring agents i and j can be expressed as

$$\xi_\tau^{i,j} = x_\tau^i - x_\tau^j + \zeta_\tau^{i,j}, \tag{11}$$

where $\zeta_\tau^{i,j} \sim \mathcal{N}(\mathbf{0}, \mathrm{P}_\tau)$.

The relative measurement vector for the network $\mathcal{G}_\tau$ can be written as

$$\boldsymbol{\xi}_\tau = (C_\tau \otimes I_3)^T \boldsymbol{x}_\tau + \boldsymbol{\zeta}_\tau, \tag{12}$$

where C_τ is the incidence matrix of $\mathcal{G}_\tau$ and $\boldsymbol{\zeta}_\tau \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_\tau)$. By choosing the weight matrix W of the network such that $\mathbf{P}_\tau^{-1} = W_\tau \otimes I_3$, the information matrix for relative measurement can be written as

$$\hat{\mathbf{H}}_\tau = L_\tau \otimes I_3, \tag{13}$$

where $L_\tau = C_\tau W_\tau C_\tau^T$ is the Laplacian matrix of $\mathcal{G}_\tau$.

The $M-$step relative measurement model is given by

$$\boldsymbol{\xi}_\star = \mathcal{C}_\star \mathbf{x}_\star + \boldsymbol{\zeta}_\star, \tag{14}$$

where $\mathcal{C}_\star = \text{blkdiag}\left[(C_t \otimes I_3)^T, \ldots, (C_{t+M} \otimes I_3)^T\right]$ is block diagonal augmentation of matrices $(C_\tau \otimes I_3)^T$ for all $\tau \in [t : t+M]$. The information matrix of $\mathbf{x}_\star$ obtained using relative measurements can be written using (14) as

$$\hat{\mathbf{H}}_\star = \mathcal{L}_\star, \tag{15}$$

where $\mathcal{L}_\star = \text{blkdiag}\left[(L_t \otimes I_3)^T, \ldots, (L_{t+M} \otimes I_3)^T\right]$ is block diagonal augmentation of matrices $(L_\tau \otimes I_3)^T$ for all $\tau \in [t : t+M]$. The information matrix $\hat{\mathbf{H}}_\star$ quantifies the amount of information of $\mathbf{x}_\star$ over the time horizon M due to the network $\mathcal{G}_\tau$. The information matrix of $\mathbf{x}_\star$ using the relative measurements can be updated according to

$$\tilde{\mathbf{H}}_\star = \bar{\mathbf{H}}_\star + \hat{\mathbf{H}}_\star. \tag{16}$$

The information matrix $\tilde{\mathbf{H}}_\star$ quantifies information of $\mathbf{x}_\star$ available through system dynamics (7) and relative measurement model (14). We discuss the vision model next.

C. Model for Vision System

Let us denote the position vector of the feature $f \in \Theta_t$ by $y_f \in \mathbb{R}^3$, the position vector of the i'th camera with respect to i'th robot by x_c^i, and the unit vector (with respect to i'th camera) corresponding to pixel measurement of feature f at time τ by $(u_\tau^{i,f})^T \in \mathbb{R}^3$. The rotation matrices describing the orientation of the i'th camera and robot are $R_c^i \in SO(3)$ and $R_\tau^i \in SO(3)$ respectively. For a given vector $(u_\tau^{i,f})^T$, its cross product with another vector v can be written as the product of a skew-symmetric matrix $U_\tau^{i,f}$ and vector v as

$$(u_\tau^{i,f})^T \times v = U_\tau^{i,f} v. \tag{17}$$

Then the noisy vision model for the i'th robot proposed in [2] is given by

$$U_\tau^{i,f}(R_\tau^i R_c^i)^T \left(y_f - (x_\tau^i + R_\tau^i x_c^i)\right) = \eta_\tau^{i,f}. \tag{18}$$

The model in (18) can be written equivalently as

$$z_\tau^{i,f} = U_\tau^{i,f}(R_\tau^i R_c^i)^T \left(x_\tau^i - y_f\right) + \eta_\tau^{i,f}, \tag{19}$$

where $\eta_{\tau,T}^{i,f}$ is the measurement noise such that $\eta_{\tau,T}^{i,f} \sim \mathcal{N}(0, \sigma_i^2 I_3)$ and $z_\tau^{i,f} = (U_\tau^{i,f})^T (R_c^i)^T x_c^i$.

The vision model for the team of robots can be written as

$$z_\tau^f = F_\tau^f \mathbf{x}_\star + E_\tau^f y_f + \eta_\tau^f, \tag{20}$$

for appropriate matrices F_τ^f and E_τ^f.

We assume that robots are capable of running $M-$step forward simulations to determine the number of frames n_f in which the feature f is visible. The $M-$step vision model for a feature f can be written as

$$\mathbf{z}_\star^f = \mathbf{F}_\star^f \mathbf{x}_\star + \mathbf{E}_\star^f y_f + \boldsymbol{\eta}_\star^f. \tag{21}$$

Let us denote the covariance matrix of $\boldsymbol{\eta}_\star^f$ by

$$\mathbb{E}\left[\boldsymbol{\eta}_\star^f \left(\boldsymbol{\eta}_\star^f\right)^T\right] = \sigma_i^2 I_{3n_f} \in S_+^{3n_f},$$

then the information matrix of $\mathbf{x}_\star$ and y_f is given by

$$\Omega_\star^f = \sigma_i^{-2} \begin{bmatrix} \left(\mathbf{F}_\star^f\right)^T \mathbf{F}_\star^f & \left(\mathbf{F}_\star^f\right)^T \mathbf{E}_\star^f \\ \left(\mathbf{E}_\star^f\right)^T \mathbf{F}_\star^f & \left(\mathbf{E}_\star^f\right)^T \mathbf{E}_\star^f \end{bmatrix}. \tag{22}$$

TABLE I: Performance Measures

Performance Measures	Matrix Operator Form
Variance of the Error $\rho_v(\mathbf{H}_\star(\Phi_t))$	$\text{Tr}\left(\mathbf{H}_\star(\Phi_t)^{-1}\right)$
Differential Entropy $\rho_e(\mathbf{H}_\star(\Phi_t))$	$-\log\left(\det\left(\mathbf{H}_\star(\Phi_t)\right)\right)$
Spectral Variance $\rho_\lambda(\mathbf{H}_\star(\Phi_t))$	$\lambda_{\min}\left(\mathbf{H}_\star(\Phi_t)^{-1}\right)$

By taking the Schur complement of the block $\left(\mathbf{E}_\star^f\right)^T \mathbf{E}_\star^f$, the information matrix

$$\mathbf{H}_\star^f = \sigma_i^{-2}\left(\left(\mathbf{F}_\star^f\right)^T \mathbf{F}_\star^f - \left(\mathbf{F}_\star^f\right)^T \mathbf{E}_\star^f \left(\left(\mathbf{E}_\star^f\right)^T \mathbf{E}_\star^f\right)^{-1} \left(\mathbf{E}_\star^f\right)^T \mathbf{F}_\star^f\right), \tag{23}$$

quantifies information content in feature f for estimation of $\mathbf{x}_\star$.

For any new observations by tracking a feature f, the information matrix of $\mathbf{x}_\star$ is updated as

$$\mathbf{H}_\star(\{f\}) = \tilde{\mathbf{H}}_\star + \mathbf{H}_\star^f. \tag{24}$$

For a set Φ_t of selected features, the information matrix of $\mathbf{x}_\star$ is updated by adding the information matrices of individual features (24) according to the following:

$$\mathbf{H}_\star(\Phi_t) = \tilde{\mathbf{H}}_\star + \textstyle\sum_{f \in \Phi_t} \mathbf{H}_\star^f. \tag{25}$$

After accounting for all visible features in the set Φ_t the covariance matrix of $\mathbf{x}_\star$ can be obtained as

$$\boldsymbol{\Sigma}_\star(\Phi_t) = \mathbf{H}_\star(\Phi_t)^{-1}. \tag{26}$$

To select a feature, it must be triangulated and the information matrix $\left(\mathbf{E}_\star^f\right)^T \mathbf{E}_\star^f$ corresponding to its position vector y_f must be invertible.

To identify the most informative features, we define suitable metrics that quantify information content or estimated accuracy. We discuss some of these metrics next.

D. Performance Measures

We quantify the informativeness of a feature subset Φ_t, using established performance measures [6, 1], which are monotonically increasing map of the covariance matrix. For accurate multi-agent localization, lower values of these performance measures indicate a better estimate of the robots' positions. Specific examples of these performance measures are listed in Table I.

As robots share relative measurement information to localize themselves, the graph Laplacian L_τ, which describes their communication network at time τ, plays a crucial role in determining the information matrix of their positions. In the next result, we show the monotonicity of the performance measures as a function of the graph Laplacian.

Theorem 1. *The performance measures stated in Table I are monotonically decreasing as connectivity of the communication graph $\mathcal{G}_\tau$ is increased, i.e., if $\mathcal{L}_{t:t+M} \preceq \mathcal{L}'_{t:t+M}$, then $\rho_\square(\mathbf{H}'_\star(\Phi_t)) \leq \rho_\square(\mathbf{H}_\star(\Phi_t))$ in which $\square \in \{v, e, \lambda\}$, $\mathbf{H}'_\star(\Phi_t)$ and $\mathbf{H}_\star(\Phi_t)$ are information matrices corresponding to $\mathcal{L}'_{t:t+M}$ and $\mathcal{L}_{t:t+M}$ respectively.*

Theorem (1) establishes that improved network connectivity among agents leads to better estimation quality for the robots' position vectors. This indicates that if the robots

share relative information over a strongly connected communication graph, the uncertainty in estimate of their positions will decrease. This finding is crucial for designing effective communication graphs for the robot team. By enabling them to share relative information efficiently (strongly connected graphs), we can significantly reduce their position estimation uncertainty.

V. Randomized Sampling Algorithm

This section presents a randomized algorithm for efficiently sampling a subset of candidate information matrices from the entire set. Originally developed for graph sparsification [11], this class of algorithms has found broader applications in diverse fields, including network observability [1]. Unlike the greedy algorithm, which can be computationally expensive, the randomized algorithm achieves similar results with lower computational complexity [6]. To quantify the level of uncertainty in our estimates, we analyze the algorithm's performance by establishing probabilistic bounds for estimation accuracy.

We define the maximal information matrix as the sum of information matrices of all features $f \in \Theta_t$ such that

$$\mathbf{H}_\star(\Theta_t) = \tilde{\mathbf{H}}_\star + \textstyle\sum_{f\in\Theta_t} \mathbf{H}_\star^f. \tag{27}$$

The maximal matrix specified in (27) can be decomposed into a sum of $|\Theta_t|$ individual matrices as

$$\tilde{\mathbf{H}}_\star^f := \frac{1}{|\Theta_t|}\tilde{\mathbf{H}}_\star + \mathbf{H}_\star^f, \tag{28}$$

for every $f \in \Theta_t$. Each matrices $\tilde{\mathbf{H}}_\star^f$ captures the information content of a feature f observed by the robot team. We utilize (27) and (28) to quantify leverage score of features.

A. Sampling using Leverage Score

While leverage score-based sampling has previously been applied for information matrix selection [6], our approach focuses on achieving tighter probabilistic guarantees. This is formally established in Theorem (2). Algorithm 1 outlines the steps involved in sampling a feature $f \in \Theta_t$ based on the leverage score of its information matrix.

Definition 2. *For a feature $f \in \Theta_t$, the leverage score $r_f : S_+^n \to \mathbb{R}_+$ is defined as*

$$r_f := \mathrm{Tr}\left(\mathbf{H}_\star(\Theta_t)^{-1}\tilde{\mathbf{H}}_\star^f\right). \tag{29}$$

The leverage score for a feature f quantifies its relative importance in state estimation. This can be formalized by introducing a discrete probability measure over the set of all features, where features with higher leverage scores correspond to a higher probability of being selected during sampling.

Remark 1. *The probability mass function over the set of all available features, denoted by $\pi : \Theta_t \to [0,1]$ can be defined as*

$$\pi_f := \frac{r_f}{n}, \tag{30}$$

where $n = 3N(M+1)$.

B. Performance Bound

To evaluate performance of our algorithm, we compare the information captured by the selected features, which is quantified by $\mathbf{H}_\star(\Phi_t)$ to the information available from all possible features quantified by $\mathbf{H}_\star(\Theta_t)$. Theorem 2 establishes a bound on the information content of our algorithm.

Algorithm 1: Randomized Sampling Algorithm

input: initial information matrix $\tilde{\mathbf{H}}_\star = \bar{\mathbf{H}}_\star + \hat{\mathbf{H}}_\star$,
set of all features Θ_t,
and number of feature samples q
output: set of selected features Φ_t, information matrix $\mathbf{H}_\star$
initialize: $\Phi_t = \emptyset$, $\mathbf{H}_\star = \tilde{\mathbf{H}}_\star$
for k = 1 to q **do**
  sample a feature from Φ_t using distribution $\pi \to f$
  select the information matrix

$$\mathbf{H} \leftarrow \mathbf{H}_\star^f$$

  if $f \notin \Phi_t$, **then**
    add f to Φ_t
    update the information matrix:

$$\mathbf{H}_\star \leftarrow \mathbf{H} + \mathbf{H}_\star$$

  end if
end for

For the exposition of our next result, we define

$$\kappa \in \{(0, 3/4) \mid \log(n/\kappa) = O(\log n)\}. \tag{31}$$

Theorem 2. *For a given $\epsilon \in (0,1)$ and $\delta \in [\kappa, \frac{3}{4})$, suppose that the number of selected features $q = 2n(\log(n/\delta))/\epsilon^2 = O(n\log n/\epsilon^2) < |\Theta_t|$, then the information matrix of the set of features Φ_t satisfies*

$$\mathbf{H}_\star(\Phi_t) \succeq \tfrac{(1-\epsilon)}{4\bar{\chi}}\mathbf{H}_\star(\Theta_t), \tag{32}$$

with probability greater than $\frac{3}{4} - \delta$ for a number $\bar{\chi}$.

The cone inequality in Theorem 2 helps us to quantify bounds on the accuracy of the estimated positions of robots using performance measures listed in Table I.

Building upon Theorem 1, which established the link between network connectivity and performance measures, we now explore another key implication of the network structure on the feature selection problem. Theorem 3 focuses on the intricate relationship between the leverage score of a feature f and the specific properties of the graphical network $\mathcal{G}_\tau$.

Theorem 3. *For a given point feature $f \in \Theta_t$, the leverage score, as defined in* (29)*, is dependent on the underlying graph Laplacian of the communication network. It is monotonically decreasing when*

$$\mathbf{H}_\star^f \succeq \tfrac{1}{|\Theta_t|}\textstyle\sum_{f\in\Theta_t}\mathbf{H}_\star^f, \tag{33}$$

meaning that if $\mathcal{L}_{t:t+M} \preceq \mathcal{L}'_{t:t+M}$, then $r'_f \le r_f$, where r'_f and r_f represent the leverage scores corresponding to the network information matrices $\mathcal{L}'_{t:t+M}$, and $\mathcal{L}_{t:t+M}$ respectively. In contrast, the leverage score is monotonically increasing under the condition

$$\mathbf{H}_\star^f \preceq \tfrac{1}{|\Theta_t|}\textstyle\sum_{f\in\Theta_t}\mathbf{H}_\star^f, \tag{34}$$

which implies that if $\mathcal{L}_{t:t+M} \preceq \mathcal{L}'_{t:t+M}$, then $r_f \le r'_f$.

Theorem 3 states that as the connectivity of the underlying communication graph of the network $\mathcal{G}_\tau$ increases, the leverage score of all features tend towards the mean

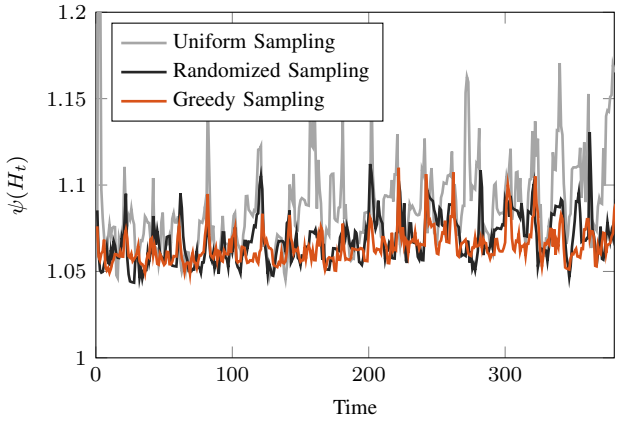

Fig. 1: Ratio of spectral norm of estimated covariance to spectral norm of covariance when states are measured (37).

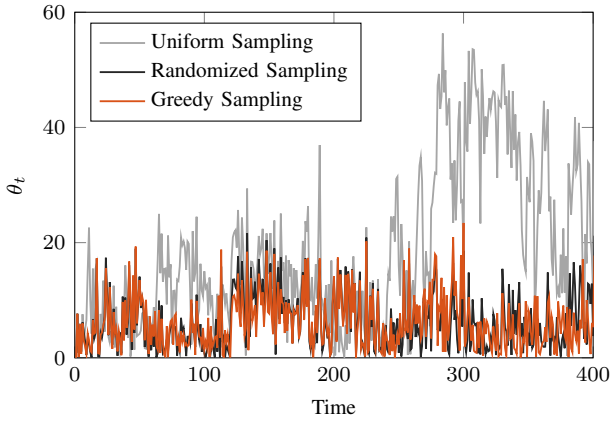

Fig. 2: Multi-agent localization error for different feature selection algorithms.

leverage score. In a strongly connected network of agents, all features become increasingly similar in their informativeness (leverage score). This allows for uniform random sampling of features to achieve good estimation quality. Consequently, the computational complexity of the algorithm is significantly reduced. Calculating the leverage score for features is no longer necessary, as random sampling becomes equally effective.

We evaluate the performance of the algorithm 1 next.

VI. Case Study

We conduct simulation studies to evaluate the performance of the randomized algorithm 1 against greedy and uniform sampling algorithms. The details of robot models and environment are presented in [8].

A. Simulation Setup: Robots and Environment

We consider a team of 10 robots that share relative measurements with each other over a communication graph $\mathcal{G}_\tau$. The weights of the edges in $\mathcal{G}_\tau$, which represent the communication strength between robots decay exponentially with the increasing distance between robots i and j as follows:

$$\omega_{ij}(\tau) = \alpha e^{-\beta \|x_\tau^i - x_\tau^j\|}, \tag{35}$$

where $\omega_{ij}(\tau)$ represents the weight of the communication edge between robots i and j for all $i, j \in \{1, \dots, N\}$, α and β are constants that control the decay rate. This weight change is linked to the covariance matrix of the relative measurement model, implying that the uncertainty in robot localization grows as the distance between communicating robots increases. By analyzing these weight changes and their impact on the estimation covariance matrix, we can gain valuable insights into how the network topology influences the overall accuracy and performance of robot localization algorithms.

B. Feature Selection Algorithms and Comparison Metrics

We compare the proposed Randomized sampling algorithm with Uniform randomized sampling and the Greedy sampling algorithm as the baseline algorithms. In randomized sampling, we select features based on a probability distribution (30), outlined in Algorithm 1. We assign a uniform probability to all features for uniform randomized sampling. For the greedy algorithm, we iteratively select a single feature at a time that leads to the most improvement in a chosen performance measure.

We employ the mean square estimation error denoted by

$$\theta_t = \|\boldsymbol{x}_t - \boldsymbol{\mu}_t\|_2, \tag{36}$$

to measure how close the state estimation is to the true trajectory. For uncertainty estimation, we employ the ratio of the spectral norm of estimated covariance to the spectral norm of covariance when states are measured. This ratio is given by

$$\psi(H_t) = \frac{\rho_\lambda(H_t)}{\lambda_{\min}(I_N \otimes \Lambda_t)}, \tag{37}$$

a lower ratio suggests that feature selection algorithm has effectively reduced the uncertainty in robots' positions.

C. Simulation Results

Figure 2 depicts the mean squared estimation error for all three feature selection algorithms. Figure 1 illustrates the reduction in localization uncertainty, quantified by the norm of estimated covariance. Our proposed algorithm achieves comparable performance to the greedy algorithm. Notably, both approaches significantly outperform uniform sampling.

While our randomized algorithm achieves slightly lower accuracy compared to the greedy approach as evident in Figures 2 and 1, it offers significant advantages in terms of computational complexity. These results underscore the critical role of feature selection algorithms in achieving accurate multi-agent localization.

We now examine the role of network connectivity in feature selection algorithms. Figure 3 illustrates that as the connections between robots strengthen, quantified by a decrease in the parameter β (35), the covariance noticeably diminishes. This observation aligns perfectly with our theoretical findings in Theorem 1, which states that stronger network connectivity leads to reduced estimation error. The peaks observed in Figure 3 suggest that selected features are informative only within a specific prediction horizon. Figure 4 shows a histogram of feature leverage scores for varying communication strength β (35). As network connectivity strengthens (lower β), the distribution of feature leverage scores becomes more uniform. This is reflected in

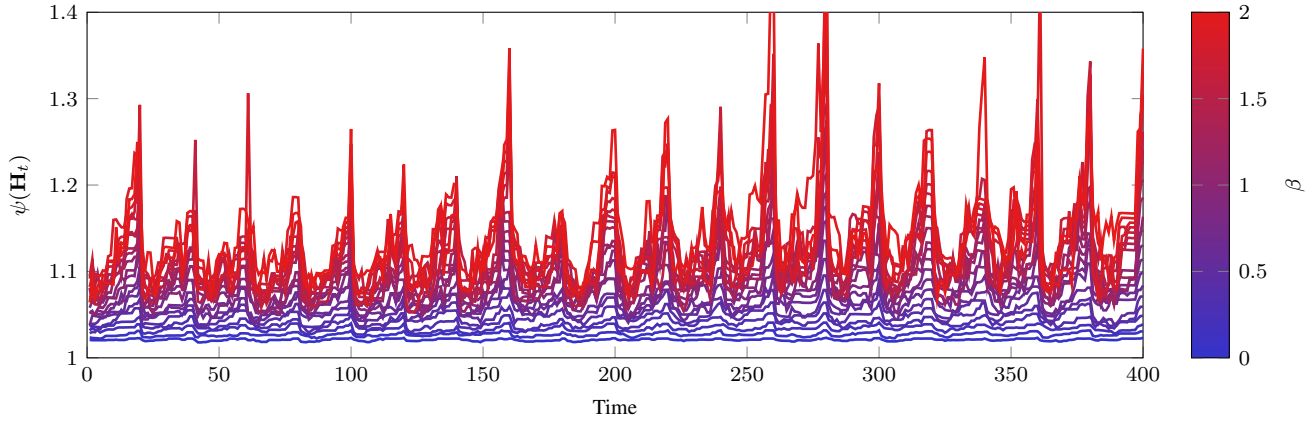

Fig. 3: Impact of network connectivity on estimation covariance.

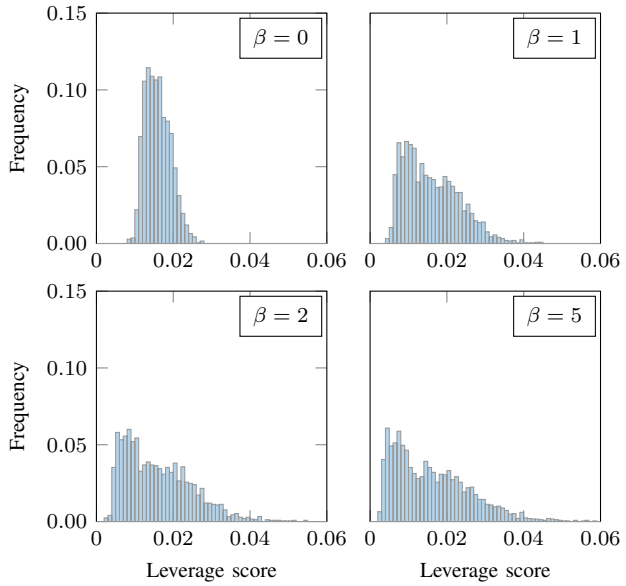

Fig. 4: Histogram of leverage scores of features for varying communication strength β and $\alpha = 1$, where α and β are as defined in (35).

the steeper histogram for $\beta = 0$. This validates our result in Theorem 3, which states that leverage scores of features become more uniform with stronger network connectivity.

VII. Conclusion

We propose a randomized algorithm for sparse visual feature selection, enabling efficient multi-agent localization. Compared to uniform random sampling, our approach achieves significantly better performance by prioritizing informative features. While exhibiting comparable accuracy to the greedy algorithm, our randomized algorithm offers potential advantages in terms of improving algorithmic complexity. We also investigate the impact of network structure on multi-agent localization performance as higher connectivity translates to lower uncertainty in localization. These findings suggest the promise of randomized algorithms for achieving efficient and accurate multi-agent localization with sparse feature sets.

References

[1] A. Amini, H. K. Mousavi, Q. Sun, and N. Motee. "Space Time Sampling for Network Observability". In: *IEEE Transactions on Control of Network Systems* 10.3 (2023), pp. 1159–1171.

[2] L. Carlone and S. Karaman. "Attention and Anticipation in Fast Visual-Inertial Navigation". In: *IEEE Transactions on Robotics* 35.1 (2019).

[3] J. Jiao, Y. Zhu, H. Ye, H. Huang, P. Yun, L. Jiang, L. Wang, and M. Liu. "Greedy-Based Feature Selection for Efficient LiDAR SLAM". In: *IEEE International Conference on Robotics and Automation*. IEEE. 2022, pp. 5222–5228.

[4] R. Lerner, E. Rivlin, and I. Shimshoni. "Landmark Selection for Task-oriented navigation". In: *IEEE Transactions on Robotics* 23.3 (2007), pp. 494–505.

[5] B. Mirzasoleiman, A. Badanidiyuru, A. Karbasi, J. Vondrak, and A. Krause. "Lazier Than Lazy Greedy". In: *Twenty-Ninth AAAI Conference on Artificial Inteliigence*. 2015, pp. 182–1818.

[6] H. K. Mousavi and N. Motee. "Estimation with Fast Feature Selection in Robot Visual Navigation". In: *IEEE Robotics and Automation Letters* 5.2 (2020), pp. 3572–3579.

[7] H. K. Mousavi and N. Motee. "Explicit Characterization of Performance of a Class of Networked Linear Control Systems". In: *IEEE Transactions on Control of Network Systems* 7.4 (2020), pp. 1688–1699.

[8] V. Pandey, A. Amini, G. Liu, U. Topcu, Q. Sun, K. Daniilidis, and N. Motee. "Scalable Networked Feature Selection with Randomized Algorithm for Robot Navigation". In: *arXiv preprint arXiv:2403.12279* (2024).

[9] P. Sala, R. Sim, A. Shokoufandeh, and S. Dickinson. "Landmark Selection for Vision-based navigation". In: *IEEE Transactions on Robotics* 22.2 (2006), pp. 334–349.

[10] M. Siami, S. Bolouki, B. Bamieh, and N. Motee. "Centrality Measures in Linear Consensus Networks with Structured Network Uncertainties". In: *IEEE Transactions on Control of Network Systems* 5.3 (2018), pp. 924–934.

[11] D. A. Spielman and N. Srivastava. "Graph Sparsification by Effective Resistances". In: *SIAM Journal of Computing* 40.6 (2011), pp. 1913–1926.

[12] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*. MIT Press, 2005.

[13] S. A. Thrun and Y. Liu. "Multi-robot SLAM with Sparse Extended Information Filters". In: *Proceedings of the th International Symposium of Robotics Research (ISRR'03)*. 2003.

[14] S. A. Thrun, Y. Liu, D. Koller, A. Y. Ng, Z. Ghahramani, and H. Durrant-Whyte. "Simultaneous Localization and Mapping with Sparse Extended Information Filters". In: *The International Journal of Robotics Research* 23.7–8 (2004), pp. 693–716.

[15] J. A. Tropp. "User-Friendly Tail Bounds for Sums of Random Matrices". In: *Foundations of Computational Mathematics* 12.4 (2012), pp. 389–434.