IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

Globally Optimal Data-Association-Free Landmark-Based Localization Using Semidefinite Relaxations

Vassili Korotkine, Mitchell Cohen, and James Richard Forbes[1]

***Abstract*—This paper proposes a semidefinite relaxation for landmark-based localization with unknown data associations in planar environments. The proposed method simultaneously solves for the optimal robot states and data associations in a globally optimal fashion. Relative position measurements to a fixed set of known landmarks are used, but the data association is unknown in that the robot does not know which landmark each measurement is generated from. The relaxation is shown to be tight in a majority of cases for moderate noise levels. The proposed algorithm is compared to local Gauss-Newton baselines initialized at the dead-reckoned trajectory, and is shown to significantly improve convergence to the problem's global optimum in simulation and experiment. Accompanying software and supplementary material can be found at https://github.com/decargroup/certifiable_uda_loc.**

***Index Terms*—Sensor Fusion, Localization, Optimization and Optimal Control, Probabilistic Inference, SLAM**

I. Introduction and Related Work

ESTIMATING the state of a robot from noisy and incomplete sensor data is a central task associated with autonomy. In the landmark-based localization task, the robot infers its position and orientation from measurements from landmarks with known positions. State estimation methods for localization can be split into filtering methods and batch optimization methods [1]. Filtering methods include Monte Carlo based localization [2], where probabilistic sampling is used to represent the posterior belief on the robot state. This work focuses on batch optimization methods, where all data over a given time window is used to estimate all the robot states within the window. An example is [3], where a joint probability data association algorithm is used to perform localization based on a high-definition map.

An important component of landmark-based localization is determining the *data associations* between measurements and landmarks, where a correspondence must be drawn between the measurements received and the landmark they are assumed to be received from. The data associations can either be solved for before, or simultaneously with, the robot states. Solving for the data associations before the robot states can be done with a front-end sensor dependent system, such as by tracking specific camera pixel coordinates between frames [4], or it can be done in the optimization back-end by using the provided data. In practice, the need to explicitly estimate data associations in the backend can arise from landmark detections that do not provide rich information, as is common in semantic object-level SLAM [5], or failures due to incorrect data association in the frontend [6]. Algorithms such as joint compatibility branch and bound [7], combined constraint data association [8], and the CLEAR alignment and rectification algorithm [9] fall into this category. In all of the above algorithms [7–9], the data associations are solved for *before* solving for the robot states with the data associations fixed, which can result in a suboptimal solution if the initial data associations are incorrect.

Another approach consists of solving for the data associations *simultaneously* with the robot states, a problem that is NP-hard and combinatorial in nature [6]. When applied to this problem formulation, the Max-Mixture approach of [10] iterates through optimization steps selecting the most likely data association at each step. The data-association-free landmark-based simultaneous localization and mapping (SLAM) algorithm of [6] simultaneously solves for robot

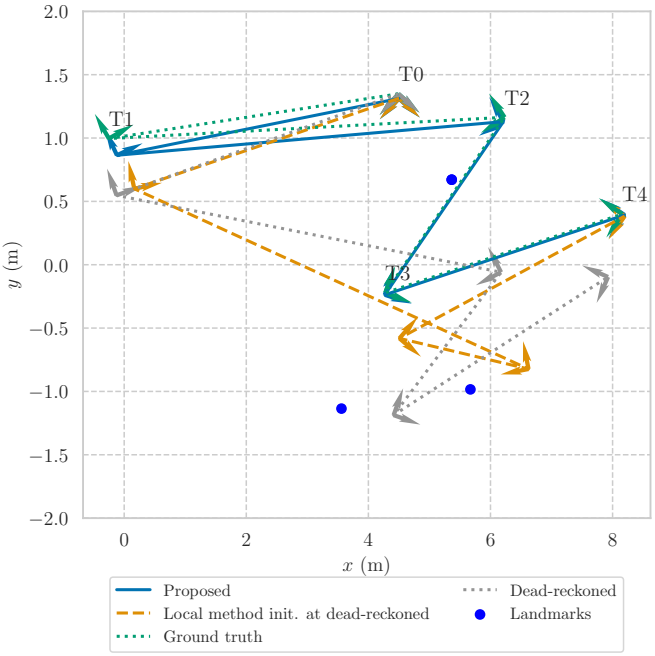

Fig. 1. Algorithm comparison on a subsequence of the Lost in the Woods dataset. The proposed SDP method converges to a solution with the correct data associations, while the local method initialized using dead-reckoning is unable to find the globally optimal data associations, leading to an incorrect robot trajectory estimate. The poses are labeled T_i, where i corresponds to the time index. Adjacent poses are 20 seconds apart.

Manuscript received: April 10, 2025; Revised July 3, 2025; Accepted July 27, 2025.

This paper was recommended for publication by Editor Sven Behnke upon evaluation of the Associate Editor and Reviewers' comments.

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Alliance Grant in collaboration with Denso Corporation.

[1]The authors are with the Department of Mechanical Engineering, McGill University, Montreal, QC H3A 0C3, Canada (e-mails: vassili.korotkine@mail.mcgill.ca, mitchell.cohen3@mail.mcgill.ca, james.richard.forbes@mcgill.ca).

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

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

states, landmark positions, and the data associations in an inner loop using the Max-Mixture algorithm, while solving for the number of landmarks in an outer solver loop. Probabilistic data association in a Max-Mixture framework is also presented in [5] for the application of semantic SLAM. DC-SAM [11] is a general purpose method for problems involving continuous and discrete variables that involves alternation of solving for the continuous variables and then solving for the discrete variables.

The algorithms described above that simultaneously solve for the data associations and robot state are inherently *local* in nature, as they require an initial guess that is refined to obtain the optimal solution. The quality of the solution is impacted by the initial guess. Furthermore, global optimality of the algorithm is not guaranteed. The incorrect minimum may be reached without any indication that it is the local minimum.

The present work focuses on globally optimal solutions obtained from semidefinite relaxations of the problem. Globally optimal methods based on semidefinite relaxations are applicable to problems that may be formulated as polynomial optimization problems, which are written as quadratically constrained quadratic programs (QCQPs) using a suitable change of variables. The QCQP admits a semidefinite program (SDP) relaxation, termed Shor's relaxation [12, Chap. 3]. The SDP relaxation may be solved to global optimality, and if the result is feasible for the non-relaxed problem, it is guaranteed to be the global optimum of the non-relaxed problem. SDP relaxations have been used to treat outlier rejection [13, 14], where the algorithm must decide if a given measurement is an outlier or not, a choice that is encoded using discrete boolean states that are solved for alongside the continuous robot states. *The contribution of this paper may thus be stated as a relaxation-based globally optimal method, in an optimization cost sense, for the data-association-free planar landmark-based localization problem where the number of landmarks is known.* To the best of the authors' knowledge, this is the first such method that scales in polynomial time. An example result is presented in Fig. 1, where for an example with five poses and three landmarks, the proposed method obtains the globally optimal solution with the correct data associations, unlike the local method initialized using dead-reckoning. The method is applicable to ground vehicle localization, in situations where relative pose measurements as well as measurements to known landmarks are available. The implementation is open-sourced at https://anonymous.4open.science/r/certifiable_uda_loc-9BCF/README.md. Supplementary material at the same URL contains further details on the cost and constraint matrix setup.

The rest of this paper is organized as follows. Sec. II describes the problem setup of landmark-based localization with unknown data associations. Sec. III describes the proposed SDP relaxation. Sec. IV presents results on simulated and experimental data, while Sec. V concludes and provides directions for future work.

TABLE I
NOTATION TABLE

		Indices
i	$\triangleq$	Timestep, sometimes generic index
N	$\triangleq$	Number of timesteps
j	$\triangleq$	Landmark index, sometimes generic index
n_ℓ	$\triangleq$	Number of landmarks
k	$\triangleq$	Index of measurement received at given timestep, sometimes generic index
K	$\triangleq$	Number measurements received at given timestep
n_θ	$\triangleq$	Number of discrete problem variables
n_x	$\triangleq$	Number of columns in optimization variable $\mathbf{X}$
		Quantities
$\mathbf{1}$	$\triangleq$	Identity matrix with unspecified dimension
$\mathbf{1}^{n\times n}$	$\triangleq$	Identity matrix of dimension $n \times n$
$\mathbf{T}_i \in SE(3)$	$\triangleq$	Robot pose at time index i
$\mathcal{X} \in SE(3)^N$	$\triangleq$	Tuple of all robot poses
$\theta \in \{0,1\}$	$\triangleq$	Single discrete problem variable
$\boldsymbol{\theta} \in \mathbb{R}^{n_\theta}$	$\triangleq$	All discrete problem variables stacked into column matrix
$\boldsymbol{\ell}_j \in \mathbb{R}^2$	$\triangleq$	j'th landmark position
$\mathbf{H} \in \mathbb{R}^{2\times 2}$	$\triangleq$	Homogenization variable
$\mathbf{X} \in \mathbb{R}^{2\times n_x}$	$\triangleq$	Matrix variable with variables of interest concatenated
$\boldsymbol{\Xi} \in \mathbb{R}^{2\times(2+3N)}$	$\triangleq$	Concatenation of continuous problem variables
$\mathbf{Z} \in \mathbb{R}^{n_x\times n_x}$	$\triangleq$	Optimization matrix variable in Shor's relaxation
		Operators
$\otimes$	$\triangleq$	Kronecker product
$\langle \mathbf{A}, \mathbf{B} \rangle$	$\triangleq$	Matrix inner product. For $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$, $\langle \mathbf{A}, \mathbf{B} \rangle = \mathrm{tr}(\mathbf{A}\mathbf{B}^\mathsf{T})$
$\mathcal{N}(\mathbf{x}; \bar{\mathbf{x}}, \mathbf{P})$	$\triangleq$	A multivariate Gaussian distribution in $\mathbf{x}$, with mean $\bar{\mathbf{x}}$ and covariance $\mathbf{P}$

II. Problem Setup

For convenience, commonly used notation is summarized in Tab. I. The state of interest $\mathcal{X}$ is given by $\mathcal{X} = \{\mathbf{T}_1, \ldots, \mathbf{T}_i, \ldots, \mathbf{T}_N\} \in SE(2)^N$ where the timestep index is denoted i and N is the number of timesteps. The robot poses are denoted $\mathbf{T}_i = (\mathbf{C}_i, \mathbf{r}_i) \in SE(2)$. $\mathbf{C}_i \in SO(2)$ denotes the direction cosine matrix at timestep i describing the change of basis from the robot frame to the inertial frame, and $\mathbf{r}_i$ denotes the robot position at timestep i resolved in the inertial frame. At the i'th timestep, the robot receives K_i relative landmark position measurements $\mathbf{y}_{k_i}$, to landmarks with known positions, $\boldsymbol{\ell}_j$ for $j = 1, \ldots, n_\ell$. The number of measurements received at the i'th timestep K_i is typically less than the number of landmarks n_ℓ. The generative isotropic noise model for the relative landmark position measurements is given by

$$\mathbf{y}_{k_i} = \mathbf{C}_i^\mathsf{T}(\boldsymbol{\ell}_j - \mathbf{r}_i) + \mathbf{v}, \quad \mathbf{v} \sim \mathcal{N}\left(\mathbf{v}; \mathbf{0}, \mathbf{1}\sigma^2_{\text{lndmrk}}\right), \quad (1)$$

where the covariance $\mathbf{1}\sigma^2_{\text{lndmrk}}$ is a multiple of identity scaled by variance $\mathbf{1}\sigma^2_{\text{lndmrk}}$, since isotropic noise is considered. At timestep i, the data association variables are encoded as $\theta_{ik_ij} \in \{0,1\}$. A value of $\theta_{ik_ij} = 1$ encodes that, at timestep i, the k_i'th measurement received at that timestep corresponds to landmark $\boldsymbol{\ell}_j$. The data association variables are further constrained by

$$\sum_{j=1}^{n_\ell} \theta_{ikj} = 1, \quad \forall i, k, \quad (2)$$

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

meaning that every measurement comes from a single landmark. All the $n_\theta = \sum_{i=1}^{N}\sum_{k_i=1}^{K_i} n_\ell$ data association variables may be stacked into a column vector $\boldsymbol{\theta}$. The problem formulation may thus be stated as

$$\underset{\mathcal{X},\boldsymbol{\theta}}{\text{minimize}} \quad J_{\text{odom}}(\mathcal{X}) + J_{\text{prior}}(\mathcal{X}) + J_{\text{uda, lndmrk}}(\mathcal{X},\boldsymbol{\theta}), \tag{Localization}$$

with $\mathcal{X} \in SE(2)^N$ and $\boldsymbol{\theta} \in \mathbb{R}^{n_\theta}$. The first term is the odometric cost [15] given by

$$J_{\text{odom}}(\mathcal{X}) = \sum_{i}^{N} \kappa_{\text{odom}} \left\|\mathbf{C}_i - \mathbf{C}_i\Delta\mathbf{C}_i\right\|_{\text{F}}^2 + \frac{1}{\sigma_{r,\text{odom}}^2}\left\|\mathbf{r}_{i+1} - \mathbf{r}_i - \mathbf{C}_k\Delta\mathbf{r}_i\right\|_2^2, \tag{3}$$

where $\Delta\mathbf{C}_i$ is the relative direction cosine measurement and $\Delta\mathbf{r}_i$ is the relative position measurement. The prior cost inserts a prior on the first pose and is given by

$$J_{\text{prior}}(\mathcal{X}) = \kappa_{\text{prior}}\left\|\mathbf{C}_1 - \check{\mathbf{C}}_1\right\|_{\text{F}}^2 + \frac{1}{\sigma_{r,\text{prior}}^2}\left\|\mathbf{r}_1 - \check{\mathbf{r}}_1\right\|_2^2, \tag{4}$$

where $\check{\mathbf{C}}_1$ and $\check{\mathbf{r}}_1$ are the prior orientation and position, respectively. The log-likelihood residual arising from (1) can be written as [15],

$$r(\mathbf{C}_i,\mathbf{r}_i;\boldsymbol{\ell}_j,\mathbf{y}) = \frac{1}{\sigma_{\text{lndmrk}}^2}\left\|(\boldsymbol{\ell}_j - \mathbf{r}_i) - \mathbf{C}_i\mathbf{y}\right\|_2^2, \tag{5}$$

Such that the unknown data association landmark measurement cost $J_{\text{uda, lndmrk}}(\mathcal{X},\boldsymbol{\theta})$ becomes

$$J_{\text{uda, lndmrk}}(\mathcal{X},\boldsymbol{\theta}) = \sum_{i=1}^{N}\sum_{k_i=1}^{K_i}\sum_{j=1}^{n_\ell}\theta_{ikj}\, r(\mathbf{C}_i,\mathbf{r}_i;\boldsymbol{\ell}_j,\mathbf{y}_{k_i}). \tag{6}$$

The outer loop is over the time indices i. The middle loop is over the k'th received measurement at timestep i, $\mathbf{y}_{k_i}$. The inner loop is over the possible data associations. The SDP formulation used for treating unknown data associations is very similar to the certifiable outlier rejection case [13, 14], However, the present work uses the assumption that a measurement came from one of a set of landmarks, as opposed to [13, 14] where the measurements are assumed to follow a given probability distribution. The outlier problem solutions [13, 14] add cost terms to the optimization, while the data-association-free assumption of the present work leads to additional SDP constraints arising from the sum constraint (2). Outlier rejection using SDPs has been considered by encoding the outlier status of measurements as boolean variables that are solved for alongside the continuous states of interest [13, 14]. Cost terms that are dependent on the boolean variables that depend on the assumed outlier distribution are included in the optimization. Furthermore, the number of landmarks is assumed known in this problem setup. Thus, the problem setup essentially corresponds to the inner loop of the local algorithm proposed in [6]. A nuance of the proposed approach is in the assumption that a measurement comes from one of a set of landmarks. This assumption does not prevent many measurements being assigned to a single landmark in a "many-to-one" fashion, leaving open the possibility of using a null hypothesis landmark in future work, similarly to [16].

III. Proposed Approach

The proposed approach consists of a semidefinite relaxation in the spirit of the outlier rejection approach in [13].

A. Mathematical Formulation

A homogenization variable $\mathbf{H} \in \{-\mathbf{1}_{2\times2}, \mathbf{1}_{2\times2}\}$, which can be enforced quadratically, is introduced to be able to write cost and constraint terms that are linear in $\mathcal{X}$ and $\boldsymbol{\theta}$. The impact of homogenization is discussed in [17]. To fix ideas, especially related to cost functions and constraints, $\mathbf{H} = \mathbf{1}_{2\times2}$ can be assumed. The direction cosine matrix variables may be concatenated to yield $\mathbf{C} = \begin{bmatrix}\mathbf{C}_1 & \cdots & \mathbf{C}_i & \cdots & \mathbf{C}_N\end{bmatrix}$, while the position variables may be concatenated as $\mathbf{r} = \begin{bmatrix}\mathbf{r}_1 & \cdots & \mathbf{r}_i & \cdots & \mathbf{r}_N\end{bmatrix}$. The continuous problem variable may be written

$$\boldsymbol{\Xi} = \begin{bmatrix}\mathbf{H} & \mathbf{C} & \mathbf{r}\end{bmatrix} \in \mathbb{R}^{2\times(2+3N)}. \tag{7}$$

The columns of $\boldsymbol{\Xi}$ are denoted as $\boldsymbol{\xi}_k$ such that $\boldsymbol{\Xi} = \begin{bmatrix}\boldsymbol{\xi}_1 & \cdots & \boldsymbol{\xi}_k & \cdots & \boldsymbol{\xi}_K\end{bmatrix}$. The odometric cost (3) may be written as $\left\langle \boldsymbol{\Xi}^{\mathsf{T}}\boldsymbol{\Xi}, \mathbf{Q}_{\text{odom}}\right\rangle$ where the matrix inner product $\langle\cdot,\cdot\rangle$ is defined for matrices $\mathbf{A},\mathbf{B} \in \mathbb{R}^{m\times n}$ as $\langle\mathbf{A},\mathbf{B}\rangle = \text{tr}(\mathbf{A}\mathbf{B}^{\mathsf{T}}) = \sum_{i=1}^{M}\sum_{j=1}^{N} a_{ij}b_{ij}$, and $\mathbf{Q}_{\text{odom}}$ is a data matrix that depends on the measurements $\Delta\mathbf{C}_i, \Delta\mathbf{r}_i$. Similarly, the prior cost (4) can be written as $\left\langle \boldsymbol{\Xi}^{\mathsf{T}}\boldsymbol{\Xi}, \mathbf{Q}_{\text{prior}}\right\rangle$ with $\mathbf{Q}_{\text{prior}}$ a data matrix that depends on the prior values. The *known* data association loss term corresponding to a given timestep and landmark measurement (5) can also be written as $\left\langle \boldsymbol{\Xi}^{\mathsf{T}}\boldsymbol{\Xi}, \mathbf{Q}_{\text{lndmrk},i,k,j}\right\rangle$ where $\mathbf{Q}_{\text{lndmrk},i,k,j}$ depends on the j'th landmark position $\boldsymbol{\ell}_j$ as well as on the measurement received, $\mathbf{y}_{k_i}$.

To write the SDP formulation for the unknown data association, let the optimization variable $\mathbf{X}$ be defined as

$$\mathbf{X} = \begin{bmatrix}\mathbf{H} & \boldsymbol{\theta}^{\mathsf{T}}\otimes\mathbf{1}^{2\times2} & \boldsymbol{\theta}^{\mathsf{T}}\otimes\boldsymbol{\Xi} & \boldsymbol{\Xi}\end{bmatrix}, \tag{8}$$

where $\mathbf{X}$ has $n_x = (2 + 2n_\theta + 3N(n_\theta+1))$ columns such that $\mathbf{X} \in \mathbb{R}^{2\times n_x}$. The unknown data association loss (6) may then be written as $\left\langle \mathbf{X}^{\mathsf{T}}\mathbf{X}, \mathbf{Q}_{\text{uda}}\right\rangle$, which contains the individual known data association cost matrices $\mathbf{Q}_{\text{lndmrk},i,k,j}$ to achieve equality between $\left\langle \mathbf{X}^{\mathsf{T}}\mathbf{X}, \mathbf{Q}_{\text{uda}}\right\rangle$ and (6). The overall problem cost matrix may be written $\mathbf{Q} = \mathbf{Q}_{\text{uda}} + \mathbf{Q}_{\text{odom}} + \mathbf{Q}_{\text{prior}}$, and Problem (Localization) may be formulated as

$$\begin{aligned}\underset{\mathbf{X}}{\text{minimize}} \quad & \sum_{i=1}^{n_\theta}\left\langle \mathbf{X}^{\mathsf{T}}\mathbf{X}, \mathbf{Q}\right\rangle \\ \text{subject to} \quad & \mathbf{X} \in \{-\mathbf{1},\mathbf{1}\}\times\{\mathbf{0},\mathbf{1}\}^{n_\theta}\times \\ & \qquad \left(SO(2)^N \times (\mathbb{R}^2)^N\right)^{(1+n_\theta)}.\end{aligned} \tag{Reform. Prob}$$

The convention of [13] for overloading the $\times$ operator is used for defining blockwise matrix membership constraints. Relaxing the $\mathbf{C}_i \in SO(2)$ orthonormality constraint to the orthogonal constraint $\mathbf{C}_i \in O(2)$ yields the constraint $\mathbf{X} \in \{-\mathbf{1},\mathbf{1}\}\times\{\mathbf{0},\mathbf{1}\}^{n_\theta}\times\left(O(2)^N\times(\mathbb{R}^2)^N\right)^{(1+n_\theta)}$, which can be written fully in terms of a homogenization constraint $\left\langle \mathbf{A}_{\text{hom}}, \mathbf{X}^{\mathsf{T}}\mathbf{X}\right\rangle = 1$ as well as quadratic constraints of the form $\left\langle \mathbf{A}_{i,\text{init}}, \mathbf{X}^{\mathsf{T}}\mathbf{X}\right\rangle = 0, \quad i = 1,\ldots,n_{\text{constr, init}}$. These constraints, with constraint matrices $\mathbf{A}_{i,\text{init}}$, are termed the initial

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

constraints, as additional redundant constraints have to then be added to tighten the relaxation. The orthogonality relaxation can thus be written

$$\begin{aligned} \underset{\mathbf{X}}{\text{minimize}} \quad & \left\langle \mathbf{X}^\mathsf{T}\mathbf{X}, \mathbf{Q} \right\rangle && \text{(Orth. Relax.)} \\ \text{subject to} \quad & \left\langle \mathbf{A}_{\text{hom}}, \mathbf{X}^\mathsf{T}\mathbf{X} \right\rangle = 1 \\ & \left\langle \mathbf{A}_{i,\text{init}}, \mathbf{X}^\mathsf{T}\mathbf{X} \right\rangle = 0, \quad i = 1, \ldots, n_{\text{constr, init}}. \end{aligned}$$

Introducing a matrix variable $\mathbf{Z}$ that corresponds to $\mathbf{X}^\mathsf{T}\mathbf{X}$ an equivalent problem to Problem (Orth. Relax.) as

$$\begin{aligned} \underset{\mathbf{Z}}{\text{minimize}} \quad & \langle \mathbf{Z}, \mathbf{Q} \rangle && \text{(Orth. Relax. Z.)} \\ \text{subject to} \quad & \langle \mathbf{A}_{\text{hom}}, \mathbf{Z} \rangle = 1 \\ & \langle \mathbf{A}_{i,\text{init}}, \mathbf{Z} \rangle = 0, \quad i = 1, \ldots, n_{\text{constr, init}} \\ & \operatorname{rank}(\mathbf{Z}) = 2. \end{aligned}$$

Relaxing the $\operatorname{rank}(\mathbf{Z}) = 2$ constraint yields the semidefinite program,

$$\begin{aligned} \underset{\mathbf{Z}}{\text{minimize}} \quad & \langle \mathbf{Q}, \mathbf{Z} \rangle && \text{(Rank Relax.)} \\ \text{subject to} \quad & \langle \mathbf{A}_{\text{hom}}, \mathbf{Z} \rangle = 1 \\ & \langle \mathbf{A}_{i,\text{init}}, \mathbf{Z} \rangle = 0, \quad i = 1, \ldots, n_{\text{constr, init}} \\ & \mathbf{Z} \succcurlyeq 0. \end{aligned}$$

(Rank Relax.) is convex and admits a globally optimal solution in polynomial time. If the solution to (Rank Relax.) has rank two, then it is feasible for (Orth. Relax. Z.), and is thus the global optimum for (Orth. Relax. Z.). The relaxation is then considered *tight*, specifically in the rank sense as the solution to (Rank Relax.) would satisfy the rank two constraint of (Orth. Relax. Z.). The variable $\mathbf{X}$ may then be extracted from $\mathbf{Z}$ and is the global optimum for (Orth. Relax.). In turn, if this $\mathbf{X}$ is also feasible for (Reform. Prob), such that the direction cosine matrix blocks of $\mathbf{X}$ have determinant equal to one, the globally optimal solution to (Reform. Prob), and thus the localization problem (Localization), is obtained. However, tightness is not guaranteed since the explicit rank constraint is dropped from (Orth. Relax. Z.). To encourage tightness, redundant constraints of the form $\langle \mathbf{A}_{\text{red},i}, \mathbf{Z} \rangle = 0$ are added to the relaxation to yield

$$\begin{aligned} \underset{\mathbf{Z}}{\text{minimize}} \quad & \langle \mathbf{Q}, \mathbf{Z} \rangle, && \text{(Tightened Relax.)} \\ \text{subject to} \quad & \langle \mathbf{A}_{\text{hom}}, \mathbf{Z} \rangle = 1 \\ & \langle \mathbf{A}_{i,\text{init}}, \mathbf{Z} \rangle = 0, \quad i = 1, \ldots, n_{\text{constr, init}} \\ & \langle \mathbf{A}_{j,\text{red}}, \mathbf{Z} \rangle = 0, \quad j = 1, \ldots, n_{\text{constr, red}} \\ & \mathbf{Z} \succcurlyeq 0. \end{aligned}$$

B. Redundant Constraints

Finding the redundant constraints is problem-specific and can be difficult. A general purpose method is proposed in [18] that exploits the fact that any constraint matrix $\mathbf{A}$ lies in the nullspace, in a matrix inner product sense, of all feasible points for the SDP. By generating many such feasible points and examining the nullspace of the resulting data matrix, redundant constraints may be obtained. The work of [18] proposes both a method for finding constraints for a given problem instance, `AutoTight`, as well as a method of generalizing them, `AutoTemplate`. The present work used `AutoTight` on small problem instances, which yielded interpretable constraints that were then applied to larger problem instances. These analytical constraints are presented in the hope that they illuminate the problem structure and can be useful to potential analytical future work.

Herein, given a matrix $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_i & \ldots & \mathbf{a}_N \end{bmatrix}$, the integer $i = \text{slice}_{\mathbf{A}}(\mathbf{a}_i)$ is used to denote the column index corresponding to $\mathbf{a}_i$ in $\mathbf{A}$. For instance, denoting the first column of the homogenization variable $\mathbf{H}$ as $\mathbf{h}_1$, then $\mathbf{Z}[\text{slice}_{\mathbf{X}}(\mathbf{h}_1), \text{slice}_{\mathbf{X}}(\mathbf{h}_1)]$ is used to refer to the entry of $\mathbf{Z}$ that has both the row and column index equal to the position of $\mathbf{h}_1$ in the optimization variable $\mathbf{X}$ in (8).

1) Discrete Variable Constraints: Let $\tilde{\boldsymbol{\theta}}^\mathsf{T} = \begin{bmatrix} 1 & \boldsymbol{\theta}^\mathsf{T} \end{bmatrix}$. These constraints arise from algebraic observations of the form $\tilde{\boldsymbol{\theta}}^\mathsf{T}\mathbf{A}_\theta\tilde{\boldsymbol{\theta}} = 0$

2) Continuous Variable Constraints: The continuous variable constraints are of the form

$$\left\langle \mathbf{A}_\Xi, \begin{bmatrix} \mathbf{1} \\ \boldsymbol{\Xi}^\mathsf{T} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \boldsymbol{\Xi} \end{bmatrix} \right\rangle = 0. \tag{9}$$

For the robot navigation problem considered, these constraints correspond to the direction cosine matrix parts being constrained as orthogonal.

3) Combining Discrete and Continuous Variable Constraints: Constraints acting on rows and columns corresponding to combinations of discrete and continuous variables can be formed as follows. First, considering two columns of $\boldsymbol{\Xi}$, denoted $\boldsymbol{\xi}_k, \boldsymbol{\xi}_\ell$, as well as two discrete variables in the problem θ_i, θ_j, a set of redundant constraints comes from the algebraic observation given by

$$\begin{aligned} \boldsymbol{\xi}_k^\mathsf{T}\boldsymbol{\xi}_\ell\tilde{\boldsymbol{\theta}}^\mathsf{T}\mathbf{A}_\theta\tilde{\boldsymbol{\theta}} &= \sum_{i,j} a_{ij}\tilde{\theta}_i\tilde{\theta}_j\boldsymbol{\xi}_k^\mathsf{T}\boldsymbol{\xi}_\ell && (10) \\ &= \sum_{i,j} a_{ij} \left(\tilde{\theta}_i\boldsymbol{\xi}_k^\mathsf{T}\right)\left(\tilde{\theta}_j\boldsymbol{\xi}_\ell\right), && (11) \end{aligned}$$

which corresponds to constraint matrix $\mathbf{A}$ with elements

$$\mathbf{A}[\text{slice}_{\mathbf{X}}(\tilde{\theta}_i\boldsymbol{\xi}_k), \text{slice}_{\mathbf{X}}(\tilde{\theta}_j\boldsymbol{\xi}_\ell)] = \mathbf{A}_{\tilde{\theta}}[\text{slice}_{\tilde{\boldsymbol{\theta}}^\mathsf{T}}(\tilde{\theta}_i), \text{slice}_{\tilde{\boldsymbol{\theta}}^\mathsf{T}}(\tilde{\theta}_j)]. \tag{12}$$

Second, given a constraint of the form

$$g(\boldsymbol{\Xi}) = \left\langle \mathbf{A}_{\boldsymbol{\Xi}}, \begin{bmatrix} \mathbf{H} & \boldsymbol{\Xi} \end{bmatrix}^\mathsf{T} \begin{bmatrix} \mathbf{H} \\ \boldsymbol{\Xi} \end{bmatrix} \right\rangle = \mathbf{0}, \tag{13}$$

it can be premultiplied by θ_i to yield

$$\theta_i g(\boldsymbol{\Xi}) = \left\langle \mathbf{A}_{\boldsymbol{\Xi}}, \begin{bmatrix} \boldsymbol{\Theta}_i & \theta_i\boldsymbol{\Xi} \\ \theta_i\boldsymbol{\Xi}^\mathsf{T} & (\theta_i\boldsymbol{\Xi})^\mathsf{T}(\theta_i\boldsymbol{\Xi}) \end{bmatrix} \right\rangle, \tag{14}$$

where the boolean constraint $\theta_i^2 = \theta_i$ was used. In addition, moment constraints as well as constraints due to the particular matrix structure of (8) have to be added to the relaxation. These constraints are described in the supplementary material.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

C. Sparsity

A sparse basis is used to reduce the computational footprint of the algorithm. The $\boldsymbol{\theta}^{\mathsf{T}} \otimes \boldsymbol{\Xi}$ portion of the optimization variable definition (8) is reduced solely to the terms necessary to express the unknown data association cost (6). Furthermore, redundant constraints that involve subsets of the discrete variables θ_i and continuous variable columns $\boldsymbol{\xi}_k$ are reduced to those where θ_i and $\boldsymbol{\xi}_k$ correspond to the same timestep i. This is found necessary to keep the problem size and computational footprint tractable, while still yielding an acceptably tight relaxation.

D. Extraction of Robot States From Optimization Variable

The extraction of the robot states from the solution to (Tightened Relax.) is done in two steps. First, the optimization variable $\mathbf{X}$ of (Reform. Prob) is extracted from $\mathbf{Z}$. This is done by recognizing that $\mathbf{Z} = \mathbf{X}^{\mathsf{T}}\mathbf{X}$, and thus that the first two rows of $\mathbf{Z}$ are $\mathbf{Z}[:2,:] = \mathbf{H}^{\mathsf{T}}\mathbf{X}$, and thus directly correspond to $\mathbf{X}$. The robot states then correspond to the continuous variable $\boldsymbol{\Xi}$, extracted from the last block matrix of $\mathbf{X}$ in (8). The data association variables are then extracted from the $\boldsymbol{\theta}^{\mathsf{T}} \otimes \mathbf{1}^{2\times 2}$ block of $\mathbf{X}$ in (8).

IV. Simulation and Experiments

The experiments aim to demonstrate that the global minimum of the objective in Problem (Localization) is attained by the proposed approach, while the local method baseline gets trapped in local minima with incorrect data associations. The proposed algorithm is evaluated on simulated examples as well as real-world data. The parameters of interest are given by relative pose measurement noise levels, the relative landmark position measurement noise, the prior on the first pose, as well as the number of poses and landmarks in the setup. The proposed method consists of solving the SDP (Rank Relax.), and is referred to as the "SDP" method in text and figures. The local method consists of using the Max-Mixture method [10], applied to the objective function of Problem (Reform. Prob), with the association variables analytically eliminated [19, Chap. 4]. The data associations are solved for implicitly and recovered from the continuous robot states upon convergence. The solution considered the "true" solution is obtained by initializing the Max-Mixture baseline at the ground-truth robot states. The Gauss-Newton method is used, with a right perturbation Lie group formulation used to handle the orientations. The same loss function is used as for the relaxation (Reform. Prob), with the Frobenius norm handled using the approach outlined in [20]. The baseline algorithm, termed Max-Mix DR, consists of the Max-Mixture baseline initialized at the dead-reckoned state estimates. The proposed SDP relaxation only yields a certifiably optimal result when it is tight and the obtained solution is feasible for the non-relaxed problem, with $\operatorname{rank}(\mathbf{Z}) = 2$. This is measured using the eigenvalue ratio of the second and third-largest eigenvalues of $\mathbf{Z}$. In an ideal scenario, this eigenvalue ratio is infinite, and the matrix rank is exactly two. In practice, a numerical cutoff threshold is used due to the numerical nature of the solver. An eigenvalue ratio threshold of 10^6 is used in this work, following [21]. The experiments were run on an OptiPlex 7000 workstation with an Intel Core i9-12900K processor and 32 GB of memory.

A. Metrics

Position and data association errors are used to quantify algorithm performance. The position error for a given trial is computed as the absolute trajectory error (ATE) quantified using the root-mean-square error (RMSE) [22], $\text{ATE}_{\text{pos}} = \frac{1}{N}\sum_i^N \|\hat{\mathbf{r}}_i - \bar{\mathbf{r}}_i\|_2$. where $\hat{\mathbf{r}}_i$ and $\bar{\mathbf{r}}_i$ are the estimated and true robot positions at timestep i. When computing ATE_{pos} for many trials, such as for trials corresponding to given noise values, the median across the trials is used. The data association error is computed as a binary quantity, corresponding to whether all the data associations are computed correctly in the problem. When computing the data association error for many trials, the fraction of trials where the data associations are computed correctly is used.

B. Simulated Example

1) Setup: Ground truth data for simulated examples was generated by first generating N random poses on $SE(2)$ as $\mathbf{T}_i = \exp(\boldsymbol{\xi}^{\wedge})$, with $\boldsymbol{\xi}^{\mathsf{T}} = \begin{bmatrix}\xi^{\phi} & \boldsymbol{\xi}^{\mathrm{r}^{\mathsf{T}}}\end{bmatrix}$, with ξ^{ϕ} generated using the uniform distribution $\xi^{\phi} \sim \mathcal{U}(0, 2\pi)$ and $\boldsymbol{\xi}^{\mathrm{r}}$ generated from the Gaussian distribution $\boldsymbol{\xi}^{\mathrm{r}} \sim \mathcal{N}(\boldsymbol{\xi}^{\mathrm{r}}; \mathbf{0}, \mathbf{1})$. The exponential map $\exp(\boldsymbol{\xi}^{\wedge})$ is associated with the $SE(2)$ matrix Lie group [23]. Landmarks are generated on a grid with $\boldsymbol{\ell}_j \sim \mathcal{U}([0,10] \times [0,10])$. Relative pose and landmark measurements are generated from the ground truth and corrupted by noise to be used in the estimator. The same approach for parametrizing relative pose measurement noise is used as in [15]. Baseline noise levels are set separately for the orientation and position components, then scaled by a multiplier $m_{\text{noise, rel}}$. The multiplier is used as a measure of the noise. Relative orientation noise is parametrized by $1/\kappa$ of (3). The value of $1/\kappa$ increases with higher noise. The relative pose position component has covariance $\sigma_r^2\mathbf{1}$. The base values of $1/\kappa$ and σ_r^2 are set to $0.01\frac{1}{\text{rad.}^2}$ and 0.745 m^2. The parameters used for the Monte Carlo simulation runs are given by $(n_{\text{poses}}, n_{\text{lndmrk}}, m_{\text{noise}}, \sigma_p^2) \in (3,5) \times (2,3) \times (0.1, 1, 10, 20, 30, 40, 50, 60) \times (0.5, 1, 2, 3, 4, 5)$, with 10 Monte Carlo trials generated for each parameter configuration, for a total of 1920 trials. The relative landmark position measurement is directly parametrized by the scalar variance multiplied by identity covariance, $\sigma_p^2\mathbf{1}$. Priors on the first pose were set to the ground truth with $\kappa_{\text{prior, rot}} = 100$ and $1/\sigma^2_{\text{prior, pos}} = 0.01$. To evaluate optimality of the proposed approach, when state errors are computed, unless explicitly stated otherwise, they are computed *with respect to a local method result initialized at the ground truth*, since the ground truth may not correspond to the global objective minimum.

2) Discussion: The proposed method is tight for reasonable noise levels, as demonstrated in Figure 2. The method starts breaking down to around 60% tightness rate for a landmark covariance of 4 m^2 and a relative pose noise multiplier of 40. The expected behavior is for tightness fraction to go

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

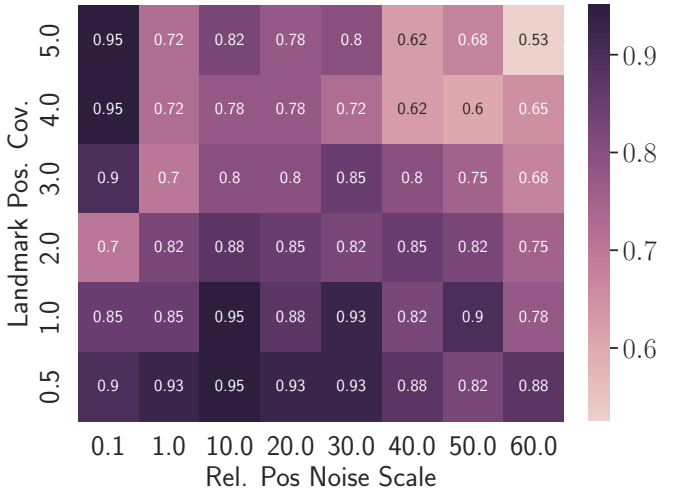

Fig. 2. Fraction of tight cases in simulated example, evaluated using a threshold of $\lambda_2/\lambda_3 \geq 10^6$. The breakdown occurs at a relative pose measurement noise scale of around 30 and a landmark position noise scale of 3.

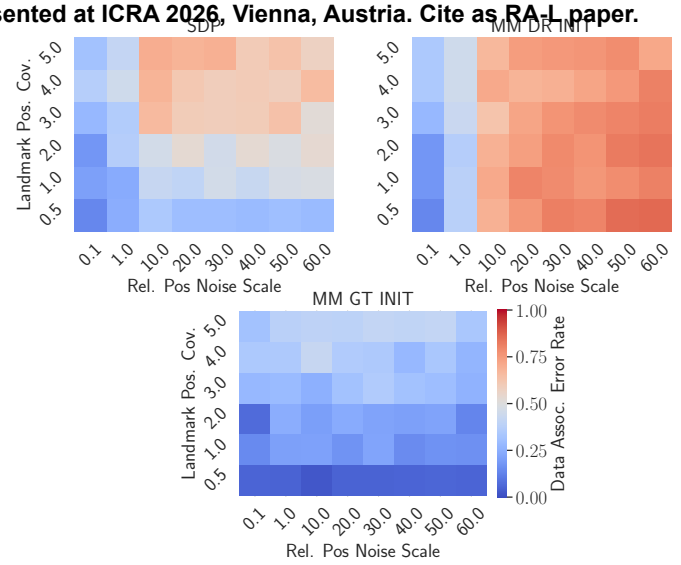

Fig. 3. Evaluation of data association error rate for tight cases in simulated example. For each noise level combination, the corresponding heatmap value denotes the fraction of trials in which at least one data association is incorrect.

down with noise. The trend is not perfect, as particularly for Rel. Pos Noise Scale=1, the tightness fraction decreases to 0.7 before going up again and resuming the expected trend, which is attributed to experimental spread. However, the general trend is as expected, with tightness rates going down with increasing landmark measurement noise as well as increasing relative pose measurement noise. The data association correctness for tight cases is presented in Figure 3, where the heatmaps plot the fraction of cases where at least one computed data association did not match the original data associations versus the relative pose and relative landmark position measurement noise levels. For the tight cases, it is expected that the SDP method attains the correct data associations, while the Max-Mix DR performs worse, especially with increasing noise levels. Furthermore, it is expected that the Max-Mix GT attains the correct data associations in all cases. While it is observed that the Max-Mix DR does perform worse than SDP, neither SDP nor Max-Mix GT attain the ground truth data associations in all cases. This is because the minimum of the objective function with *noisy* measurements of Problem (Localization) does not necessarily have the same data associations as the ground truth *noiseless* setup. Similarly, the SDP solution does not always match the Max-Mix GT solution because the ground truth initialization might not attain the global minimum of the problem. This was supported by finding that, in tight cases, the cost attained by SDP was always lower than that of Max-Mix GT. Furthermore, the improved data association obtained by SDP compared to Max-Mix DR is shown to improve position error in Figure 4. The drawback of the proposed implementation is scalability. A runtime comparison is shown in Table II, where the proposed method has a median runtime of 32.81 seconds for the case of 5 poses and 3 landmarks. Generic SDP solvers such as `MOSEK` do not scale well past medium-size problems, a threshold that is reached quickly since the problem size scales linearly with the amount of discrete variables added.

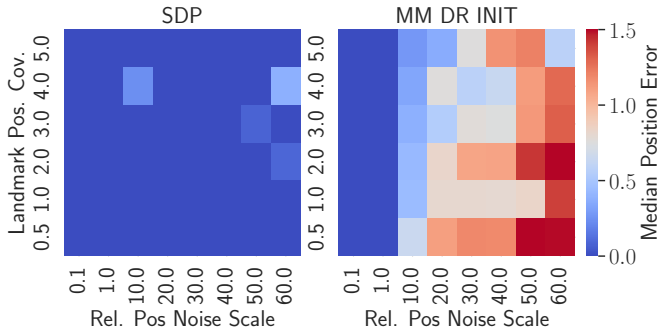

Fig. 4. Evaluation of position error for tight cases in simulated example. Each entry in the heatmap corresponds to the median position error for trials for the corresponding noise values.

C. Lost in the Woods Dataset

1) Setup: The Lost in the Woods dataset [24] is used to validate the proposed approach on real data. It consists of a wheeled robot driving in a forest of plastic tubes that are treated as landmarks. The robot receives wheel odometry measurements that provide forward and angular velocity measurements. A laser rangefinder provides range-bearing measurements to the landmarks, whose positions are provided in the dataset. The wheel odometry measurements are integrated to provide relative pose measurements using the body frame velocity process model [25], while the range-bearing measurements are converted to relative position measurements. The orientation component of the relative pose cost (3) arises as a log-likelihood of an isotropic Langevin noise distribution on $SO(2)$ [26]. Given a set of pairs of true relative orientations and corresponding measured relative orientations computed from odometry, the parameter κ in (3) is computed by reverse-engineering the isotropic Langevin distribution sampling procedure in [26]. Priors on the first pose were set to the ground truth with $\kappa_{\text{prior, rot}} = 100$ and $\sigma^2_{\text{prior, pos}} = 0.01$. The relative pose position components and the relative landmark

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

TABLE II
MEDIAN SOLVER TIME FOR SIMULATION EXPERIMENT (SECONDS)

Landmarks	3 Poses		5 Poses	
	SDP	MM DR INIT	SDP	MM DR INIT
2	0.64	**0.03**	2.19	**0.04**
3	6.14	**0.03**	32.81	**0.06**

position measurement noise covariances are also computed using the measurements compared to the ground-truth data. The dataset is used to extract subsequences, parametrized by the number of poses n_{poses}, the number of landmarks used n_{lndmrk}, as well as the spacing between poses Δt, measured in seconds. The higher the spacing between the poses, the more error accumulates for the relative pose measurements obtained by integrating the odometry. The subsequences extracted for each dataset parameter setting $(n_{\text{poses}}, n_{\text{lndmrk}}, \Delta t)$ are non-overlapping. While the dataset provides 17 landmarks, only a subset is used for each sequence. For a set amount of landmarks, n_{lndmrk}, the landmarks used are chosen that are visible at the largest amount of the selected subsequence timesteps. The Monte Carlo trials are run for parameter configurations given by $(n_{\text{poses}}, n_{\text{lndmrk}}, \Delta t) \in (3, 5) \times (2, 3) \times (20, 40, 60)$. The overall dataset is 20 minutes long. For each parameter set $(n_{\text{poses}}, n_{\text{lndmrk}}, \Delta t)$, the first subsequence starts at $t = 0$, with each following subsequence starting at the end of the previous one. For each combination of $n_{\text{poses}}, \Delta t$, the number of subsequences is chosen to be the maximum allowed for by the dataset length.

2) Discussion: The parameters used for the x and y axes are different for the experiment section compared to the simulated examples. Since the landmark measurement noise is fixed, the varied parameters consist of n_{poses} and Δt. The Δt essentially corresponds to the odometry measurement noise, while n_{poses} corresponds to increasing the problem size and trajectory length. Thus, increasing either Δt or n_{poses} is expected to worsen the quality of the dead-reckoned initialization. Furthermore, increasing Δt is expected to degrade the tightness of the relaxation. The fraction of tight cases for each parameter configuration is presented in Fig. 5. While the method is tight in the majority of cases, it is hard to extract a clear pattern of worsening tightness with noise in this setup. This can be ascribed to the fewer amount of subsequences used in this setup compared to the simulation. In Fig. 2, each square corresponds to 40 Monte Carlo trials, while in Fig. 5 the number of subsequences used ranges from 21 for $\Delta t = 20, n_{\text{poses}} = 3$ to only four for the $\Delta t = 60, n_{\text{poses}} = 5$ case. The data association results are presented in Fig. 6, where both tight and nontight results are plotted together. This is unlike the simulated case where only the tight results were plotted. In this experiment, both tight and nontight solutions were shown to yield good results. In both tight and nontight cases, SDP matches Max-Mix GT in all cases, while Max-Mix DR has a very high failure rate for $n_{\text{poses}} = 5$. This is expected, since with a higher number of poses, the quality of the dead-reckoned initialization worsens leading to degradation of performance of the local method that is dependent on the initialization. The position error in Table III follows the same

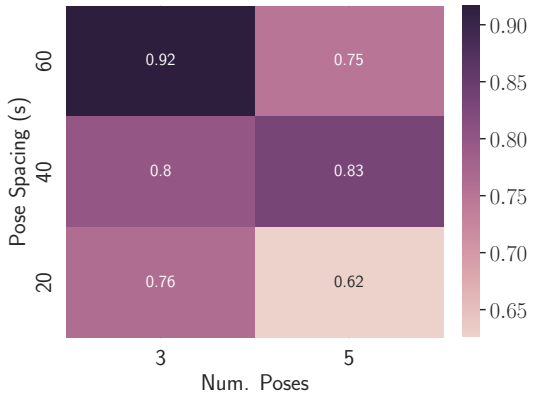

Fig. 5. Fraction of tight cases in Lost in the Woods, evaluated using a threshold of $\lambda_2/\lambda_3 \geq 10^6$. While the method is tight in the majority of cases, there is no clear relationship demonstrated between an increase in pose spacing, essentially corresponding to odometry noise, and the tightness of the relaxation.

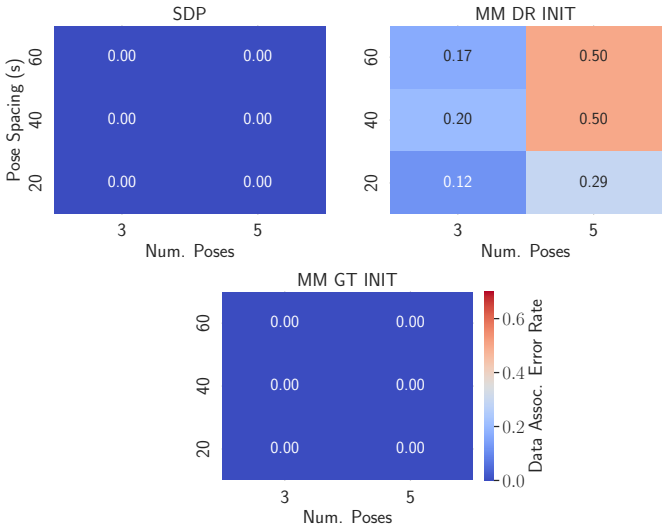

Fig. 6. Evaluation of data association correctness for the Lost in the Woods example. This includes all results, both tight and nontight.

trend as the data association plot in Fig. 6, with position errors increasing for $n_{\text{poses}} = 5$, showing that the incorrect data association of the local method also leads to a degradation of the estimated robot state. The position error is computed with respect to the trajectory computed by the Max-Mix GT method. A scalar position error is computed for each trajectory, and the median for each parameter configuration is shown in the table.

TABLE III
MEDIAN POSITION ERROR FOR LOST IN THE WOODS DATASET

Pose Spacing (s)	3 Poses		5 Poses	
	SDP	MM DR INIT	SDP	MM DR INIT
20	1.8×10^{-5}	$\mathbf{2.4 \times 10^{-11}}$	5.3×10^{-5}	$\mathbf{3.2 \times 10^{-11}}$
40	7.8×10^{-6}	$\mathbf{2.0 \times 10^{-11}}$	$\mathbf{5.5 \times 10^{-5}}$	1.2×10^{-1}
60	8.9×10^{-6}	$\mathbf{7.5 \times 10^{-11}}$	$\mathbf{1.8 \times 10^{-5}}$	1.7×10^{-1}

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

V. Conclusion

This paper proposes a relaxation-based globally optimal method for robot localization. The relaxation is tight in the majority of cases in simulation, and the majority of cases in real-world experiment. While offering global optimality in cases where tightness is attained, the scalability of the algorithm is limited by the difficulty in solving large-scale SDPs. While the scaling of semidefinite program solvers is generally still prohibitive, advances have been made for specific problem instances, particularly for SDPs where the solution is of low rank [14, 26–28], and future SDP solvers may allow scaling of the proposed method beyond the small problem instances analyzed in this paper. Future work also includes development of certification schemes based on the relaxation. The use of a null hypothesis to account for misidentified landmarks is also a promising direction of future research, where, similarly to [16], outlier measurements are assigned to a dummy landmark. The null hypothesis would make use of the possibility of "many-to-one" assignment of measurements to landmarks. Furthermore, while this paper specifically considers the setting of robot localization, the overall approach and redundant constraints used are applicable to other settings with unknown data association, investigating which is another topic of future work.

References

[1] T. D. Barfoot, *State Estimation for Robotics*. Cambridge, UK: Cambridge University Press, 2024.

[2] F. Dellaert, D. Fox, W. Burgard, and S. Thrun, "Monte carlo localization for mobile robots," in *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No.99CH36288C)*, vol. 2, 1999, 1322–1328 vol.2.

[3] Z. Gu, S. Cheng, C. Wang, R. Wang, and Y. Zhao, "Robust visual localization system with hd map based on joint probabilistic data association," *IEEE Robotics and Automation Letters*, vol. 9, no. 11, pp. 9415–9422, 2024.

[4] J. Shi and Tomasi, "Good features to track," in *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1994, pp. 593–600.

[5] K. J. Doherty, D. P. Baxter, E. Schneeweiss, and J. J. Leonard, "Probabilistic data association via mixture models for robust semantic SLAM," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 1098–1104.

[6] Y. Zhang, O. A. Severinsen, J. J. Leonard, L. Carlone, and K. Khosoussi, "Data-association-free landmark-based slam," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 8349–8355.

[7] J. Neira and J. Tardos, "Data association in stochastic mapping using the joint compatibility test," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 6, pp. 890–897, 2001.

[8] T. Bailey, "Mobile robot localisation and mapping in extensive outdoor environments," PhD thesis, University of Sydney, 2002.

[9] K. Fathian, K. Khosoussi, Y. Tian, P. Lusk, and J. P. How, "Clear: A consistent lifting, embedding, and alignment rectification algorithm for multiview data association," *IEEE Transactions on Robotics*, vol. 36, no. 6, pp. 1686–1703, 2020.

[10] E. Olson and P. Agarwal, "Inference on networks of mixtures for robust robot mapping," *The International Journal of Robotics Research*, vol. 32, no. 7, pp. 826–840, Jun. 2013.

[11] K. J. Doherty, Z. Lu, K. Singh, and J. J. Leonard, "Discrete-continuous smoothing and mapping," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 12 395–12 402, 2022.

[12] H. Yang, *Semidefinite optimization and relaxation*, Working Draft Edition. 2024. [Online]. Available: https://hankyang.seas.harvard.edu/Semidefinite/.

[13] P.-Y. Lajoie, S. Hu, G. Beltrame, and L. Carlone, "Modeling perceptual aliasing in slam via discrete–continuous graphical models," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1232–1239, 2019.

[14] H. Yang and L. Carlone, "Certifiably optimal outlier-robust geometric perception: Semidefinite relaxations and scalable global optimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 3, pp. 2816–2834, 2023.

[15] C. Holmes and T. D. Barfoot, "An efficient global optimality certificate for landmark-based slam," *IEEE Robotics and Automation Letters*, vol. 8, no. 3, pp. 1539–1546, 2023.

[16] J. Pöschmann, T. Pfeifer, and P. Protzel, "Factor graph based 3d multi-object tracking in point clouds," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 10 343–10 350.

[17] D. Cifuentes, S. Agarwal, P. A. Parrilo, *et al.*, "On the local stability of semidefinite relaxations," *Mathematical Programming*, vol. 193, pp. 629–663, 2022.

[18] F. Dümbgen, C. Holmes, B. Agro, and T. Barfoot, "Toward globally optimal state estimation using automatically tightened semidefinite relaxations," *IEEE Transactions on Robotics*, vol. 40, pp. 4338–4358, 2024.

[19] K. J. Doherty, *Lifelong, Learning-Augmented Robot Navigation*. Massachusetts Institute of Technology, 2023.

[20] F. Dellaert, D. M. Rosen, J. Wu, R. Mahony, and L. Carlone, "Shonan rotation averaging: Global optimality by surfing $SO(p)^n$," in *Computer Vision – ECCV 2020*, Springer International Publishing, 2020, pp. 292–308.

[21] C. Holmes, F. Dümbgen, and T. Barfoot, "On semidefinite relaxations for matrix-weighted state-estimation problems in robotics," *IEEE Transactions on Robotics*, vol. 40, pp. 4805–4824, 2024.

[22] Z. Zhang and D. Scaramuzza, "A tutorial on quantitative trajectory evaluation for visual(-inertial) odometry," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 7244–7251.

[23] J. Solà, J. Deray, and D. Atchuthan, "A micro Lie theory for state estimation in robotics," pp. 1–17, 2021. arXiv: 1812.01537 [cs.RO]. [Online]. Available: http://arxiv.org/abs/1812.01537.

[24] T. D. Barfoot, "AER1513 course notes, assignments, and data sets," *University of Toronto, Institute for Aerospace Studies*, 2011.

[25] C. C. Cossette, A. Walsh, and J. R. Forbes, "The Complex-Step Derivative Approximation on Matrix Lie Groups," *IEEE Robotics and Automation Letters (RA-L)*, vol. 5, no. 2, pp. 906–913, 2020.

[26] D. Rosen, L. Carlone, A. Bandeira, and J. Leonard, "SE-Sync: A certifiably correct algorithm for synchronization over the special euclidean group," *The International Journal of Robotics Research*, vol. 38, no. 2-3, pp. 95–125, 2019.

[27] F. Dümbgen, C. Holmes, and T. D. Barfoot, *Exploiting chordal sparsity for fast global optimality with application to localization*, 2025. arXiv: 2406.02365 [cs.RO]. [Online]. Available: https://arxiv.org/abs/2406.02365.

[28] A. Papalia, A. Fishberg, B. W. O'Neill, J. P. How, D. M. Rosen, and J. J. Leonard, "Certifiably correct range-aided slam," *IEEE Transactions on Robotics*, vol. 40, pp. 4265–4283, 2024.