

MM-LINS: a Multi-Map LiDAR-Inertial System for Over-Degenerate Environments

Yongxin Ma*, Jie Xu*, Shenghai Yuan†, Tian Zhi, Wenlu Yu, Jun Zhou†, and Lihua Xie, *Fellow, IEEE*

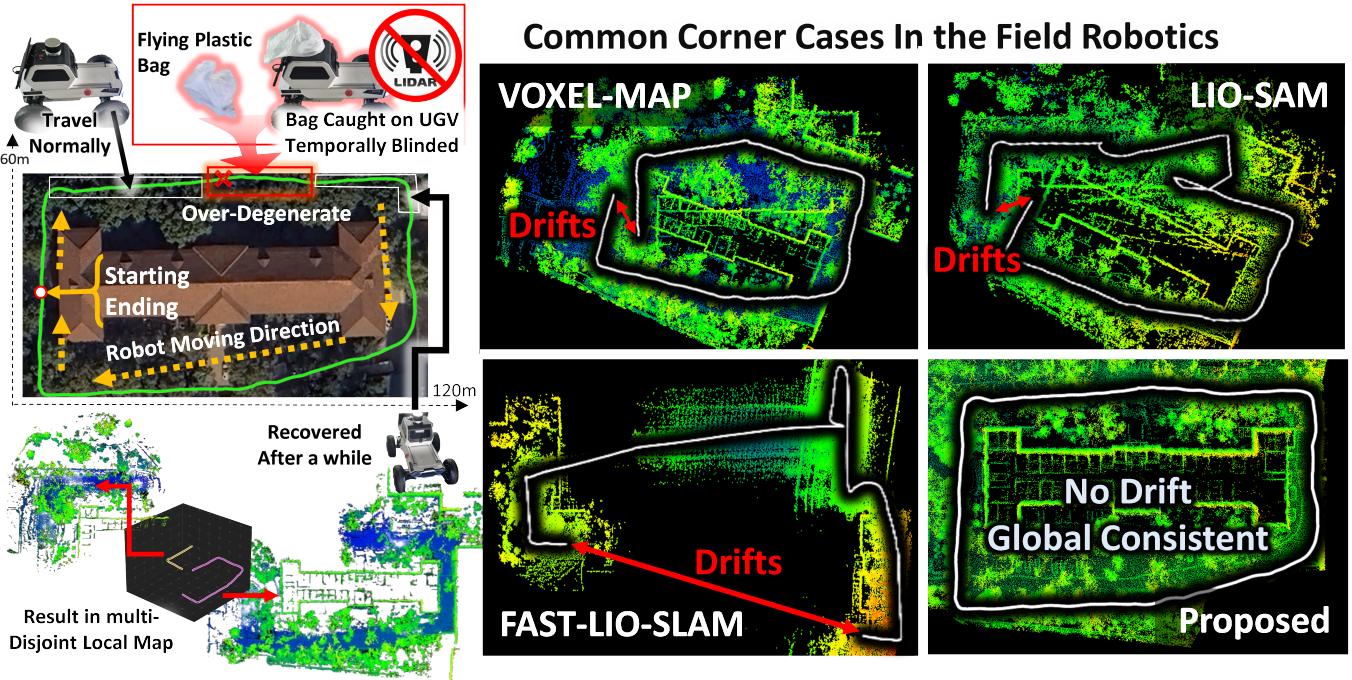


Fig. 1: Left Column: Satellite map and trajectory of the robot during the real-world campus dataset. LiDAR obstruction caused by floating garbage bags is illustrated through real images. Demonstration implementation of our algorithm. Right Column: Point cloud maps and trajectories created using SOTA algorithms and ours.

Abstract—SLAM plays a crucial role in automation tasks, such as warehouse logistics, healthcare robotics, and restaurant delivery. These scenes come with various challenges, including navigating around crowds of people, dealing with flying plastic bags that can temporarily blind sensors, and addressing reduced LiDAR density caused by cooking smoke. Such scenarios can result in over-degeneracy, causing the map to drift. To address this issue, this paper presents a multi-map LiDAR-inertial system (MM-LINS) for the first time. The front-end employs an iterated

error state Kalman filter for state estimation and introduces a reliable evaluation strategy for degeneracy detection. If over-degeneracy is detected, the active map will be stored into sleeping maps. Subsequently, the system continuously attempts to construct new maps using a dynamic initialization method to ensure successful initialization upon leaving the over-degeneracy. Regarding the back-end, the Scan Context descriptor is utilized to detect inter-map similarity. Upon successful recognition of a sleeping map that shares a common region with the active map, the overlapping trajectory region is utilized to constrain the positional transformation near the edge of the prior map. In response to this, a constraint-enhanced map fusion strategy is proposed to achieve high-precision positional and mapping results. Experiments have been conducted separately on both public datasets that exhibited over-degenerate conditions and in real-world environments. These tests demonstrated the effectiveness of MM-LINS in over-degeneracy environment. Our codes are open-sourced on Github¹.

Index Terms—Multi-map, over-degenerate, pose graph optimization, simultaneous location and mapping.

I. INTRODUCTION

Simultaneous localization and mapping (SLAM) are the fundamental components for enabling applications like Mars

* denotes equal contribution, † denotes corresponding author.

This work is supported in part by the Shandong Key R&D Program of China under Grant 2021SFJC0202. (Corresponding author: Jun Zhou and Shenghai Yuan.)

Yongxin Ma and Jun Zhou are with school of Mechanical Engineering, Shandong University, Jinan 250061, China and Key Laboratory of High Efficiency and Clean Mechanical Manufacture, Ministry of Education, Jinan 250061, China (e-mail: yxma@mail.sdu.edu.cn; zhoujun@sdu.edu.cn)

Jie Xu and Wenlu Yu are with State Key Laboratory of Robotics and Systems, Harbin Institute of Technology, Harbin 150001, China (e-mail: n23087161@e.ntu.edu.sg; yuwenlu@stu.hit.edu.cn).

Tian Zhi is with the College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, Jiangsu, China (e-mail: zt1117@nuaa.edu.cn).

Jie Xu, Shenghai Yuan, and Lihua Xie are with the School of Electrical and Electronic Engineering, Nanyang Technological University, 639798, Singapore (e-mail: shyuan@ntu.edu.sg; elhxie@ntu.edu.sg).

¹ <https://github.com/lian-yue0515/MM-LINS>

rover exploration, warehouse AMR, healthcare robotics, restaurant delivery robots, construction mapping, etc. [1]–[3]. Compared to other visual [4]–[11] and radar-based methods [12], LiDAR SLAM [13], [14] can perceive 3D environmental data accurately and in a dense manner [15], making it a reliable tool for various applications [16]–[18].

Existing LiDAR-based SLAM [19], [20] struggle in dynamic and cluttered environments [21], [22], such as navigating through crowds of people [23], getting blocked by things like flying plastic bags, or dealing with smoke/fog. These scenarios frequently lead to significant drift issues, which we classify as over-degenerated cases. And there is no simple solution for over-degeneracy.

Adding proprioceptive sensors like IMUs can partially mitigate the drift issue [24]–[28], but prolonged perception interruptions can still affect the overall trajectory and map quality. By converting the map representation using different voxel sizes [29], [30], the system can partially address the issue of degeneracy. However, this is not a permanent solution, and the overall map may still become drifted over time.

We hypothesize that in severe cases of over-degeneracy, certain map regions may become irreversibly drifted. This requires reliance on multiple disjoint sub-maps of varying quality to construct a larger map and enable localization. However, this approach poses several challenges, including (1) effectively identifying degeneracy, (2) managing disjoint multi-map sections for determining when and where to use them, and (3) conducting global pose graph optimization with degenerated sections.

This paper presents MM-LINS, a system that focuses on utilizing a multi-map strategy to mitigate the impact of map drift resulting from self-localization failures in over-degenerate areas. It addresses challenges by continuously updating an active map and archiving it into sleeping maps during over-degeneracy. Meanwhile, while in over-degeneracy, the system continuously attempts to construct new maps using a dynamic initialization method to ensure a high-quality initialization of the new map upon leaving these areas. By using Scan Context [31] descriptors, similar keyframes are detected and corresponding maps are reactivated. These maps are then merged for pose graph optimization, resulting in a globally accurate trajectory and map. Experimental results show MM-LINS outperforms traditional algorithms in over-degenerate scenarios (Fig. 1).

The main contributions of this paper are as follows:

- A multi-map LiDAR-inertial SLAM system is proposed for the first time, which mitigates the problem of error accumulation or system failure in LiDAR SLAM systems when faced with over-degenerate scenarios.
- In front-end, we propose a reliable degeneracy detection strategy, coupled with a dynamic initialization method for LiDAR SLAM. Regarding back-end, a constraint-enhanced map fusion strategy is proposed, aimed at enhancing the accuracy of map merging.
- The system performance is evaluated both in degenerated public datasets and in a real-world environment. The results indicate that MM-LINS can withstand various over-degenerate scenarios, demonstrating higher robust-

ness and accuracy than current state-of-the-art (SOTA) methods under over-degenerate scenarios.

II. RELATED WORKS

Our work intersects the following related fields of degeneracy detection, multi-robot and multi-session SLAM system.

A. Degeneracy Detection

In this paper, perceiving degenerate phenomena is crucial. Zhang et al. [32] proposed a method for detecting degeneracy in optimization problems. They described the degeneracy detection problem as a state estimation issue in least squares optimization, formulating it as a multicriteria problem. In the filter system, Zhen et al. [33] performed degeneracy judgment based on the Singular Value Decomposition (SVD) of the constraint matrix formed by the normal vectors of the LiDAR beams projected onto the environment plane. X-ICP [34] utilizes the elements of the Jacobian from the derivation of the matrix for a point-to-plane ICP cost function for degeneracy detection, simplifying the formulation and enabling more practical deployment in various environments. Different from the above methods, we focus on the characteristics of the covariance matrix of the system to improve the efficiency of degeneracy detection.

B. Multi-Robot and Multi-Session SLAM

Multi-robot and multi-session SLAM has garnered considerable attention in contemporary research endeavors [35]–[39]. Our work pays additional attention to achieving merging between different maps. Dubé et al. [40] proposed a robust method for inter-robot closed-loop selection for map fusion aimed at solving relative positional transformations among multiple robots. This method leverages the consistency of the closed-loop and data similarity to prevent mismatches. Huang et al. [41] first implemented the Scan Context descriptor in a multi-robot LiDAR SLAM to address the multi-robot problem, thereby enhancing the data transfer efficiency for public area detection. Do et al. [42] suggested an inter-robot closed-loop selection method to augment the robustness of map fusion. The two-by-two consistency strategy of the internal data introduced in the original closed-loop detection aids in averting false positives in public area detection. Kim et al. [43] put forth the LT-mapper, with its key module, the LT-SLAM, designed to handle multi-session SLAM for long-term operations. The core concept involves solving pose transformations with the Iterative Closest Point (ICP) algorithm and further optimizing through the distance to the loop closure factor. Inspired by the above, global descriptors between different maps are the key to inter-map similarity detection. This paper uses the Scancontext descriptor, which is currently recognized as having better robustness. In contrast to the above approaches, this paper introduces a constraint-enhanced strategy to improve the accuracy of map merging.

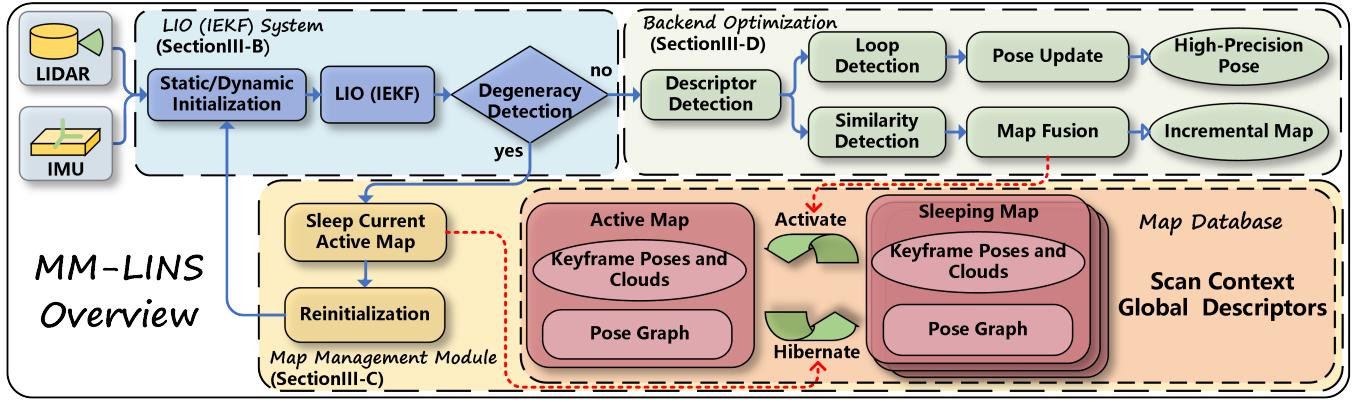


Fig. 2: MM-LINS overview.

III. METHODOLOGY

A. Framework Overview

In this paper, we propose a robust multi-map LiDAR-inertial system framework, dubbed MM-LINS. The proposed framework, shown in Fig. 2, is designed to operate effectively in environments subject to over-degeneracy. We employ an error state-iterated Kalman filter for pose estimation and propose a new degeneracy detection method to identify over-degenerate scenes. Upon detection of such a scene, the active map is stored into sleeping maps and is reactivated when similar regions are encountered in the future. The system continuously attempts to generate new maps using a dynamic initialization method. Concurrently, the system's back-end leverages feature descriptors based on global geometric information for inter-map similarity detection, enabling the identification of identical regions. When similarities are identified, the corresponding sleeping maps are integrated into the currently active map using an innovative constraint-enhanced map fusion strategy, consequently forming a highly accurate global map upon completion of the system run.

B. Degeneracy-Aware LIO System

The FAST-LIO2 [25] is widely recognized as the SOTA LIO system. We have leveraged it as our foundational framework and implemented enhancements to align with the specific demands rising from the challenges. We have innovatively designed and integrated a specialized module for over-degeneracy detection, enabling the system to proactively identify and respond to instances of over-degeneracy. Additionally, a dynamic initialization method has been added to improve performance in new map construction.

1) Over-Degeneracy Detection: During robot operation, LIO can produce significant errors attributed to LiDAR degeneracy resulting from the lack of geometrically informative structures. This implies that the robot's own 6-DOF are inadequately constrained, particularly concerning the over-degenerate phenomenon described in this paper. This not only leads to the challenge of generating large errors but also poses the risk of system drift or even failure, precisely the issue we aim to proactively mitigate. Hence, The issue we prioritize is on perceiving degenerate phenomena.

For this system, the covariance matrix from [25] contains information about the strength of the constraints for the robot state, which can be divided into submatrices:

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{rr} & \mathbf{P}_{rt} \\ \mathbf{P}_{tr} & \mathbf{P}_{tt} \end{bmatrix}_{6 \times 6}, \quad (1)$$

where r represents the rotation, t denotes the translation, and \mathbf{P}_{rr} exclusively contains information related to the rotation variables, \mathbf{P}_{tt} exclusively contains information related to the translation variables.

It can be explicitly divided into constraints on translations and rotations. Since the scales and types of rotation and translation are different. If analyzed for the covariance matrix as a whole, this would lead to coupling of rotation and translation, resulting in complications with the parameter settings for the degeneracy threshold. Therefore, feature decomposition is performed separately for rotation and translation:

$$\mathbf{P}_{rr} = \mathbf{V}_r \Sigma_r \mathbf{V}_r^\top, \quad \mathbf{P}_{tt} = \mathbf{V}_t \Sigma_t \mathbf{V}_t^\top, \quad (2)$$

where \mathbf{V}_r and \mathbf{V}_t are the eigenvectors in matrix, Σ_r and Σ_t are diagonal matrices with the eigenvalues of \mathbf{P}_{rr} and \mathbf{P}_{tt} as the diagonal entries, respectively.

The eigenvalues in \mathbf{V}_r and \mathbf{V}_t provide a direct measure of the strength of the systematic observation with respect to the rotation and translation constraints. Due to the varying magnitudes, separate thresholds are set for their degenerate assessments. As a result, we can monitor the maximum eigenvalues $\lambda^{(i)r}_{max}$ and $\lambda^{(i)t}_{max}$ corresponding to the eigenvalue matrices $\mathbf{V}^{(i)r}$ and $\mathbf{V}^{(i)t}$ for each frame to evaluate the degeneracy or not of the system.

To achieve this, the major degeneracy threshold ξ_x^r and ξ_x^t , the minor degeneracy threshold ξ_ϕ^r and ξ_ϕ^t , and the persistence threshold N_k of the system are determined empirically. It is worth noting that the relevant threshold settings are closely related to the sensor type. Therefore, we determined the thresholds through experiments on datasets from different sensor types to ensure degeneracy detection performance. These thresholds reflect the occurrence of over-degenerate phenomena and also act as criteria for map hibernation.

Whether the over-degeneracy occur is represented in (3).

$$I^{(i)} = \begin{cases} 1, & \text{if } \lambda_{max}^{(i)r} > \xi_{\kappa}^r, \lambda_{max}^{(i)t} > \xi_{\kappa}^t \text{ or } \Gamma_{\lambda} > N_{\kappa} \\ 0, & \text{otherwise} \end{cases}, \quad (3)$$

where $I^{(i)}$ is the flag for the occurrence of the over-degenerate phenomena in i_{th} frame, Γ_{λ} represents the count of consecutive frames with degeneracy falling between the major and minor degeneracy thresholds ($\xi_{\varphi}^r < \lambda_{max}^{(i)r} < \xi_{\kappa}^r$ or $\xi_{\varphi}^t < \lambda_{max}^{(i)t} < \xi_{\kappa}^t$).

2) Initialization Method: The objective of the initialization module is to estimate pivotal parameters at the onset of the SLAM system: velocity \mathbf{v} , gravitational acceleration \mathbf{g} , IMU bias \mathbf{b}_a , \mathbf{b}_{ω} .

Initially, this system's LIO employs a static initialization method like FAST-LIO2. However, during motion, when our system has experienced over-degeneracy and requires re-initialization for SLAM, the static method fails, making a robust dynamic initialization necessary.

Extensive research has been conducted in the field of vision for the dynamic initialization of SLAM systems [44]–[46], primarily due to the lack of scale information in monocular systems, necessitating system motion to recover scale. Since LiDAR sensors can directly sense scale information and are robustly initialized in a stationary state, there has been limited focus on their dynamic initialization, which is crucial for our system. Unlike methods that require the estimation of extrinsic calibration of LiDAR and an IMU [47], as well as scale information, our approach does not necessitate sufficient excitation in all directions. Inspired by VINS-MONO [46], our method omits the scale initialization step and directly obtains the LiDAR point clouds with scale, estimating the pose by ICP. Finally, constructs a linear system using the pose and IMU pre-integrations to solve for the initial parameters.

i) Data Collection: The gyroscope bias of the parameters to be solved is determined through rotation constraints, while the velocity and gravitational acceleration are resolved via velocity and translation constraints. These constraints are established using at least four frames of data to guarantee solvability of the system. In order to achieve real-time initialization, only the poses computed from four frames of the LiDAR point clouds (i.e., $\mathbf{P}_{L_i}^L(q_{L_i}^L, t_{L_i}^L)$, where $q_{L_i}^L$ and $t_{L_i}^L$ represent the i th frame rotation and translation components relative to the first LiDAR frame in the pose, respectively.) and IMU pre-integrations (i.e., $\alpha_{I_{i+1}}^{I_i}, \beta_{I_{i+1}}^{I_i}, \gamma_{I_{i+1}}^{I_i}$) are collected for subsequent solving.

ii) \mathbf{b}_{ω} Initialization: Obtain the rotations $q_{L_{i+1}}^L$ and $q_{L_i}^L$ for two consecutive frames and the relative rotation constraints $\hat{\gamma}_{I_{i+1}}^{I_i}$ from IMU pre-integrations. Construct the following equations to solve:

$$\min_{\delta\mathbf{b}_{\omega}} \sum_{i=0}^3 \left\| \mathbf{q}_{I_{i+1}}^L - 1 \otimes \mathbf{q}_{I_i}^L \otimes \hat{\gamma}_{I_{i+1}}^{I_i} \right\|^2 \quad (4)$$

$$\hat{\gamma}_{I_{i+1}}^{I_i} \approx \hat{\gamma}_{I_{i+1}}^{I_i} \otimes \left[\begin{array}{c} 1 \\ \frac{1}{2}\mathbf{J}_{\mathbf{b}_{\omega}}^{\gamma} \delta\mathbf{b}_{\omega} \end{array} \right],$$

where $\mathbf{J}_{\mathbf{b}_{\omega}}^{\gamma}$ denotes the Jacobi matrix of the relative change in rotation with respect to \mathbf{b}_{ω} , additional note $q_{L_i}^L$ is the same as $q_{L_i}^L$. The solved $\delta\mathbf{b}_{\omega}$ is the initial \mathbf{b}_{ω} .

iii) \mathbf{v} and \mathbf{g} Initialization: Re-propagate IMU pre-integration using \mathbf{b}_{ω} solved above. A linear measurement model is constructed with relative translation and relative velocity constraints for initial parameter solving. The variables to be solved for this process are as follows:

$$\mathbf{x} = [\mathbf{v}_{I_0}^L, \mathbf{v}_{I_1}^L, \dots, \mathbf{v}_{I_3}^L, \mathbf{g}^L]^T, \quad (5)$$

where $\mathbf{v}_{I_i}^L$ is the velocity in the LiDAR coordinates when the i_{th} frame finishes. \mathbf{g}^L is the gravitational acceleration when completes the first LiDAR frame.

Construct the equations as follows:

$$\begin{aligned} \mathbf{q}_{I_{i+1}}^{I_i} &= \mathbf{q}_L^{I_i} \left(\mathbf{t}_{I_{i+1}}^L - \mathbf{t}_{I_i}^L + \frac{1}{2} \mathbf{g}^L \Delta t^2 - \mathbf{v}_{I_i}^L \Delta t \right) \\ \mathbf{q}_{I_{i+1}}^{I_i} &= \mathbf{q}_L^{I_i} \left(\mathbf{q}_{I_{i+1}}^L \mathbf{v}_{I_{i+1}}^L + \mathbf{g}^L \Delta t - \mathbf{q}_{I_i}^L \mathbf{v}_{I_i}^L \right), \end{aligned} \quad (6)$$

where Δt is the time interval between two frames. Combining the extrinsic calibration of LiDAR and an IMU (i.e., $\mathbf{t}_{I_i}^L = \mathbf{t}_{L_i}^L - q_{I_i}^L t_L^I$, where t_L^I denotes the translational extrinsic calibration.) yields the following linear model:

$$\begin{aligned} \hat{\mathbf{z}}_{I_{i+1}}^{I_i} &= \begin{bmatrix} \hat{\alpha}_{I_{i+1}}^{I_i} - \mathbf{q}_L^{I_i} \left(\mathbf{t}_{L_{i+1}}^L - \mathbf{t}_{L_i}^L - q_{I_{i+1}}^L t_L^I \right) - \mathbf{t}_L^I \\ \hat{\beta}_{I_{i+1}}^{I_i} \end{bmatrix} \\ &= \mathbf{H}_{I_{i+1}}^{I_i} \mathbf{x}_i + \mathbf{n}_{I_{i+1}}^{I_i}, \end{aligned} \quad (7)$$

where

$$\begin{aligned} \mathbf{H}_{I_{i+1}}^{I_i} &= \begin{bmatrix} -\mathbf{I} \Delta t & \mathbf{0} & \frac{1}{2} \mathbf{q}_L^{I_i} \Delta t^2 \\ -\mathbf{I} & \mathbf{q}_L^{I_i} q_{I_{i+1}}^L & \mathbf{q}_L^{I_i} \Delta t \end{bmatrix}, \\ \mathbf{x}_i &= [\mathbf{v}_{I_i}^L, \mathbf{v}_{I_{i+1}}^L, \mathbf{g}^L]^T, \end{aligned} \quad (8)$$

solve the linear least squares problem to obtain \mathbf{x} :

$$\min_{\mathbf{x}_i} \sum_{i=0}^3 \left\| \hat{\mathbf{z}}_{I_{i+1}}^{I_i} - \mathbf{H}_{I_{i+1}}^{I_i} \mathbf{x}_i \right\|^2. \quad (9)$$

Next, we refine the gravitational acceleration, as done similarly to [46], by constraining its magnitude to the corresponding local value. Finally, we rotate all LiDAR frame $(\cdot)^L$ to the IMU frame $(\cdot)^I$.

C. Map Management Module

The map management module primarily handles the tasks of storing, converting and merging maps.

1) Multi-Map Representation: The map database, as shown in Fig. 2, comprises several maps, divided into two categories: an active map that the system updates in real-time and a series of sleeping maps derived from an active map. Each map possesses its own keyframe point clouds, keyframe poses and pose graph. The pose graph is constructed based on the starting node of the active map as an a priori constraint. As the system operates, odometry constraints and loop closure constraints are continuously added to perform pose graph optimization, thereby reducing accumulated errors within the maps. To facilitate information exchange between different maps, the database includes a Scan Context global descriptor, falling into two categories: loop closure detection and similarity detection.

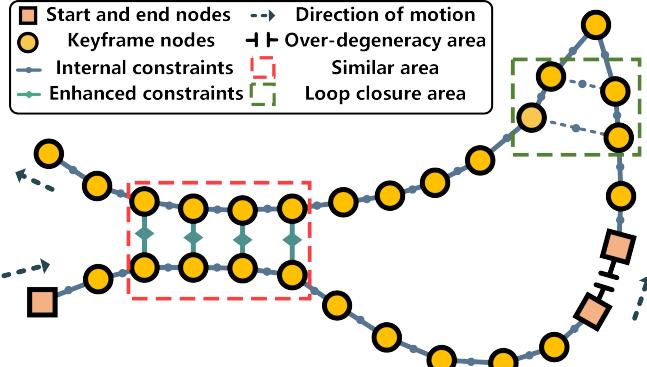


Fig. 3: The conceptual overview of the map merging idea. The internal constraints include both loop closure constraints and odometry constraints, which are respectively denoted by dashed and solid lines. Enhanced constraints comprise those detected by similarity measures.

Loop closure detection occurs in the active map maintained by the system, while similarity detection takes place in the sequence of sleeping maps. Once similarity detection concludes, the identified similar sleeping maps are activated for map fusion operations.

2) *Map Reconstruction Strategy*: When the over-degeneracy indicator is met, we carry out the “hibernate” operation to store active map information in sleeping maps. Subsequently, the system continuously employs the dynamic initialization method to ensure new map construction can be completed promptly upon leaving the over-degenerate area.

3) *Map Fusion*: We employ subscripts a , s and m to distinguish between active, sleeping and merged maps, respectively. First, through similarity detection, we identify keyframe point clouds of the common area, represented by C_a and C_s , between M_a and M_s , and calculate the initial pose transformation, T_a^s . We apply T_a^s to all keyframe poses and point clouds in map M_a to transform them to the M_s coordinate system, and merge M_s and M_a into M_m . Subsequently, we carry out pose graph fusion. During this process, the pose graph of M_a is aligned with the pose graph of M_s in preparation for fusion. However, this transformation renders the a priori constraint in the pose graph of M_a unreliable, prompting us to completely discard the priori constraint in the pose graph of M_a . As a result, the pose graph becomes under-defined. To counteract this issue, we introduce similarity constraints, which in turn ensure the completeness of the pose graph. Lastly, the system carries out pose graph optimization to finalize the optimization of T_a^s , and M_s is used as the subsequent M_a for updates.

D. Constraint-Enhanced Pose Graph Optimization

The back-end primarily focuses on accurate pose estimation and global consistent point cloud mapping. Accordingly, the back-end takes in the odometry poses and point clouds output from the front-end and further optimizes them utilizing the pose graph. Building on this, the back-end framework concentrates on stitching together the multi-maps derived after experiencing over-degeneracy to create a high-precision global

map. The conceptual overview of the map merging idea is presented in Fig. 3.

1) *Descriptor Detection*: This system employs the Scan Context descriptor [31] for loop closure detection and similarity detection. To meet the demands of map merging, we impose stringent screening conditions for similarity detection to avoid misidentification of similar maps. We hypothesize that similarity detection does not usually occur by chance in a single frame, but occurs continuously. Therefore, we augment the original loop closure detection with a time consistency condition for similarity detection.

The time consistency check essentially quantifies the number of consecutive occurrences of similarity detection (as shown in (10)). A similarity detection is only accepted if the similarity consistently occurs over a successive period. This can influence the quality of the subsequent optimization for the map pose transformation.

$$\Gamma_\varsigma > \varepsilon_{th}, \quad (10)$$

where Γ_ς signifies the count of similar Scan Context features occurring consecutively, while ε_{th} represents the threshold of temporal consistency, a value determined based on the experimental environment.

2) *Constraint-Enhanced Fusion Optimization*: The back-end maintains the active map in real-time by building a keyframe pose graph, where a new keyframe is initiated when the distance or angle between adjacent frames exceeds a certain threshold. As the robot moves, optimization is performed by adding odometry constraints and loop closure constraints. In this context, Let $P = \{p_0, \dots, p_t\}, p \in SE(3)$ represent a set of keyframe poses filtered from the odometry information passed from the front-end, encompassing a set of 6-DoF robot poses from frame 0 to frame t . C encompasses all constraints within the map, including both the odometry constraints inferred between keyframe nodes and the loop closure constraints identified through loop detection. For each pair of pose constraints, denoted as $(i, j) \in C$, we define the error between the observed transformation $O_{ij} \in SE(3)$, corresponding to the loop constraint between nodes, and the predicted transformation $\hat{O}_{ij} \in SE(3)$, derived from odometry inference, as e_{ij} :

$$e_{ij}(p_i, p_j, O_{ij}) = O_{ij} \boxminus \hat{O}_{ij}(p_i, p_j), \quad (11)$$

where \boxminus encapsulation operator denotes a mapping from a local neighborhood on $SE(3)$ to its tangent space [48], and $\hat{O}_{ij}(p_i, p_j) = p_i^{-1}p_j$.

This optimization problem can be represented as a nonlinear least squares problem in (12), The system aims to optimize the robot’s pose to minimize the error.

$$P^* = \arg \min_P \sum_{(i,j) \in C} R_{ij}, \quad (12)$$

$$R_{ij} = e_{ij}^T \Omega_{ij} e_{ij}, \quad (13)$$

where R_{ij} signifies the negative log-likelihood function of a constraint between p_i and p_j , Ω_{ij} stands for the covariance.

TABLE I: Comparison of ATE (RMSE) in meter Across Various Degenerate Datasets

Sequences (Range:m)	FAST-LIO2	VOXEL-MAP	LIO-SAM	FAST-LIO-SLAM	FAST-LIO-SAM	MM-LINS	
						w/o	PGO
M2DGR-07-31 (289.7m)	23.65	13.27	1049.781	8.653	5.221	<u>0.346</u>	0.322
M2DGR-08-06 (340.6m)	35.951	41.959	1177.675	25.814	389.838	<u>0.221</u>	0.186
UrbanLoco-04-26 (741.3m)	38.667	26.472	× ^a	76.571	28.292	<u>1.979</u>	1.978
UrbanLoco-03-16 (601.1m)	7.27	36.697	×	10.549	14.109	<u>1.157</u>	1.155
NCLT-01-10 (1139.2m)	0.997	— ^b	—	29.146	1.203	0.785	0.811
UTBM-07-16 (5044.9m)	25.718	25.481	720.591	81.736	17.817	<u>5.118</u>	2.362
UTBM-04-18 (5112.3m)	67.481	95.458	1845.723	96.728	43.345	<u>5.945</u>	5.033

^a “×” denotes the result drifted, and the table below follows the same rules.

^b “—” denotes numerical instability leading to crashing either at Start-up or at degeneracy encounter.

Best results are **bolded**, and second-best results are underlined, and the table below follows the same rules.

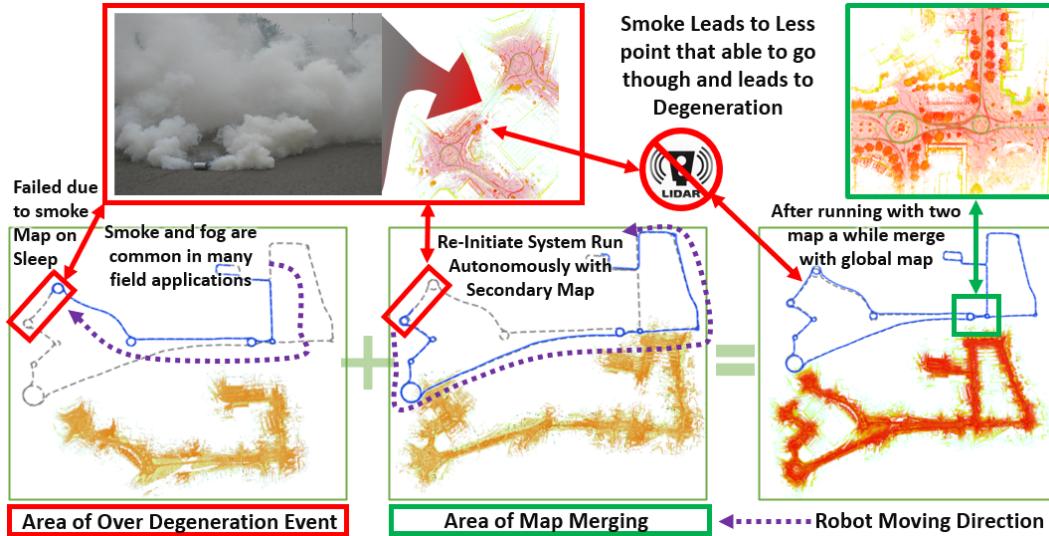


Fig. 4: A visual demonstration of degeneracy event and map merging.

When map fusion is not occurring, the back-end optimization process is shown by the above. During map merging, the primary expectation is to accurately solve the pose transformation between the maps, which directly affects the integrity and consistency of the global map. Therefore, the map optimization objective changes, as described below.

In the context of merging two maps, the corresponding subindexes are defined as follows: $\psi = \{a, s\}$, a , and s represent the active and sleeping maps, respectively. $\forall \alpha \in \psi$, P_α denote the set of poses for map α . $\mathbb{P} = \{P_\alpha \mid \alpha \in \psi\}$ contains all poses of the map being merged. C_α represents all pose constraints within M_a and M_s , and C_a^s denotes the inter-map transformation constraints between M_a and M_s . With these definitions, the variables to be optimized are as follows:

$$\mathbb{X} = \{\mathbb{P}, T_a^s\}. \quad (14)$$

For the optimization problem, the equation is as follows:

$$\mathbb{P}^* = \arg \min_P \left\{ \sum_{(i,j) \in C_\alpha} R_{ij}(\mathbb{X}) + R_a^s(\mathbb{X}) \right\}, \quad (15)$$

where is composed of two components: The first half consists of the negative log-likelihood function of constraints in intra-

map and the second half is the negative log-likelihood function of inter-map transformation.

Due to the inability to provide an accurate constraint relationship between the maps directly, we constrain the transformation relationship between the maps by fixing the overlapping part of the trajectory of the active map and the sleeping map in the common area. If only one similarity constraint occurs during merging, the accuracy of map merging is not high enough. Therefore, a constraint-enhanced strategy is proposed, as shown in Fig. 3, during map merging, we optimize the pose transformation between maps using a certain number of similarity constraints to form enhanced constraints. The objective function is modified into the following form:

$$\mathbb{P}^* = \arg \min_P \left\{ \sum_{(i,j) \in C_\alpha} R_{ij}(\mathbb{X}) + \sum_{C_Y} R_t(\mathbb{X}) \right\}, \quad (16)$$

where $R_t(\mathbb{X})$ denotes the enhanced constraints between maps, composed of pairs of similar constraints between two maps, indicated by $(i,j)_{i \in \alpha, j \in \beta, \alpha \neq \beta} \in C_Y$. It is evident that a higher quantity of these similarity constraints (indicating more pronounced constraint enhancement) leads to more accurate calculations of pose transformations. This quantity is determined by the threshold ε_{th} in the similarity detection.

TABLE II: End-to-End Distance Comparison (meter) of Real-World Proprietary Datasets (Indoor)

Sequence (Range:m)	Fast	Voxel	Fast-LIO	Fast-LIO	MM-LINS	
	LIO2	Map	SLAM	SAM	w/o	PGO
Fac.1(87.2)	29.6	67.26	116.57	×	0.22	0.18
Fac.2(88.1)	27.25	23.01	66.60	0.11	0.15	0.05
Fac.3(96.9)	1.41	8.81	25.49	17.35	0.53	0.31

"Fac." denotes handheld factory dataset with a single event of Over-degeneracy.

IV. EXPERIMENT RESULTS

In this work, we extensively evaluated the efficacy of our proposed MM-LINS system using both public datasets and real-world field-collected proprietary datasets. All experimental procedures were executed on an old laptop equipped with an Intel Core i7-9750H (2.60GHz) CPU.

A. Public Dataset Validation

We conducted thorough experiments on SOTA methods and MM-LINS using publicly available datasets, with a special focus on scenarios characterized by loop closure. These scenarios assist in identifying maps with analogous regions and facilitate map fusion. Given the lack of real-world public datasets for accurate ground truth, over-degeneracy, and loop closure characteristics, we opted to modify part of the public datasets to simulate over-degeneracy. To streamline the modification process, we significantly downsampled the point clouds for a duration prior to the loop closure phase of the dataset, thereby emulating a period of limited observational information and producing an effect analogous to over-degeneracy. To validate the algorithm, we selected seven sequences from four reputable public datasets: M2DGR [49], NCLT [50], utbm [51], and UrbanLoco [52]. The selected sequences encapsulate a diverse range of scenarios, including large-scale urban environments featuring ramps and high-rise buildings, as well as suburban terrain. Each dataset is equipped with different LiDAR and IMU configurations, showcasing the versatility and applicability of the proposed system across various environments and hardware.

In our multi-map strategy, when it encounters an over-degenerate scenario during mapping, it archives the active map into sleeping maps and initiates the construction of a new map. Upon completion of subsequent similarity detection, map fusion becomes feasible. Fig. 4 provides an example of map fusion. It is apparent that our algorithm can more resiliently navigate the challenges posed by over-degenerate scenarios.

To quantitatively analyse the quality of the mapping, we conducted a comparative analysis of the absolute trajectory errors (ATE) between the trajectories generated by the algorithms and the ground truth. To highlight the accuracy and robustness of our approach, we contrasted the results of FAST-LIO [25], VOXEL-MAP [30], LIO-SAM [53], FAST-LIO-SLAM [54], and FAST-LIO-SAM [55] algorithms under degeneracy augmented datasets, as delineated in Table I. The experimental results indicate that the multi-map strategy proposed in this paper generally maintains the error within

a range of 5m or lower. Compared to the SOTA methods, it can more effectively manage over-degenerate scenarios and enhances robustness. Ablation studies were further conducted on constraint-enhanced pose graph optimization methods (labeled as **MM-LINS (w/o PGO)** and **MM-LINS (PGO)** in the table). The findings demonstrated a marginal reduction in algorithmic error following the incorporation of the constraint-enhanced method, thereby affirming the necessity of this method.

B. Real-world Field Testing

It is crucial to ensure that the algorithm works reliably in real-world situations. We tested it out in various environments, both indoors and outdoors, using handheld and vehicle-mounted setups. It is important to highlight that we ensured consistent start and end points for each test, serving as our ground truth measurement. This precaution was particularly necessary because the testing area was densely populated with trees, making RTK/GPS unreliable, Leica Tracking unfeasible and prior map matching inaccurate. By relying on end-to-end distance measurement, we were able to assess the SLAM performance accurately.

1) *Indoor Experiment:* In this environment, data collection was done by a handheld device equipped with a MID-360 LiDAR. The environment mapped was a factory spanning an area of over 1200 m². Data collection involved making a complete circuit within the factory premises, during which over-degeneracy was simulated by obstructing the LiDAR while in motion. The specific results are displayed in Table II. All algorithms, barring the one outlined in this paper, exhibited errors of over 1m, sometimes even higher. This is an unacceptable margin of error for a SLAM system. Conversely, our algorithm was able to constrain the error within 0.4m. Meanwhile, we also show the performance of different algorithms, as shown in Fig. 5.

2) *Outdoor Experiment:* The efficacy of the proposed MM-LINS method was also evaluated outdoors. Data collection was done by a vehicle-mounted device equipped with RS-LIDAR-16 LiDAR and IMU in KY-INS180-A0, as shown in Fig. 7. Fig. 6 illustrates scenarios of moving a complete circuit around a building on the campus with over-degeneracy occurring at different locations. Additionally, it presents the trajectories of different algorithms in their corresponding scenarios. The specific results are displayed in Table III. MM-LINS can handle the case of over-degeneracy at any position as well as a certain number of over-degeneracies, limiting the error to less than 15 centimeters. Other existing algorithms are subject to huge errors unacceptable to the SLAM system when faced with the same situation.

C. Degeneracy Detection and Dynamic Initialization Performance

For degeneracy detection, Zhang et al. [32] proposes a degeneracy detection metric called the degeneracy factor, which is considered SOTA. It directly perform singular value decomposition on the information matrix. We used it as a baseline method for comparison in this work, highlighting the

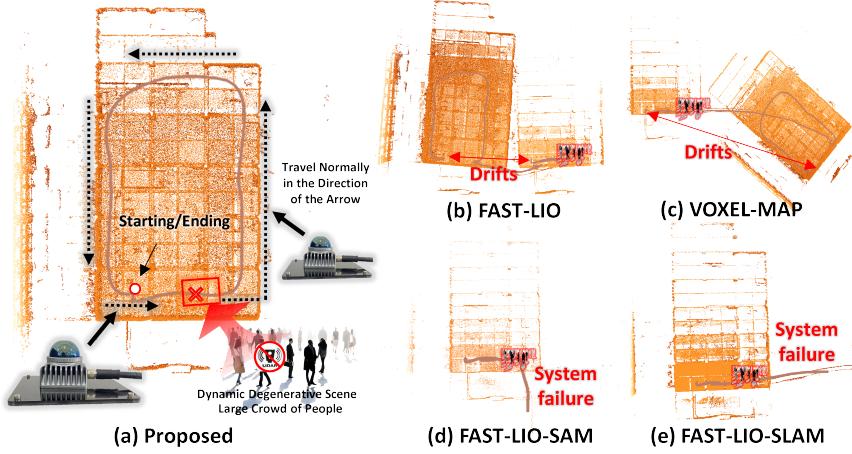


Fig. 5: Performance comparison of different algorithms on the Factory-1 dataset. The subfigures depict the following: (a) Real-world operation of the handheld device and the excellent performance of our algorithm in this case, (b) - (e) Point cloud maps and trajectories of SOTA algorithms in the same situation.

TABLE III: End-to-End Distance Comparison (meter) of Real-World Proprietary Datasets (Outdoor)

Sequences (Range:m)	FAST-LIO2	VOXEL-MAP	LIO-SAM	FAST-LIO-SLAM	FAST-LIO-SAM	MM-LINS		Remark
						w/o	PGO	
Campus-1 (350.6m)	29.705	13.026	19.941	16.610	×	0.126	0.092	
Campus-2 (351.1m)	391.773	×	×	62.072	×	0.603	<u>1.295</u>	Single Event.
Campus-3 (347.6m)	43.995	67.96	×	7.586	8.014	0.137	0.093	
Campus-4 (348.7m)	137.555	146.381	34.152	×	91.339	0.128	0.051	Two Events.
Campus-5 (349.4m)	306.398	×	×	44.195	×	0.159	0.132	Three Events.

All campus series of proprietary datasets are collected via vehicle-mounted perception suits following exactly the same path. The only difference is the physical interruption of sensing caused by wrapping a plastic bag on top of the LiDAR. The number of events denotes the number of physical interferences used to simulate the degeneracy.

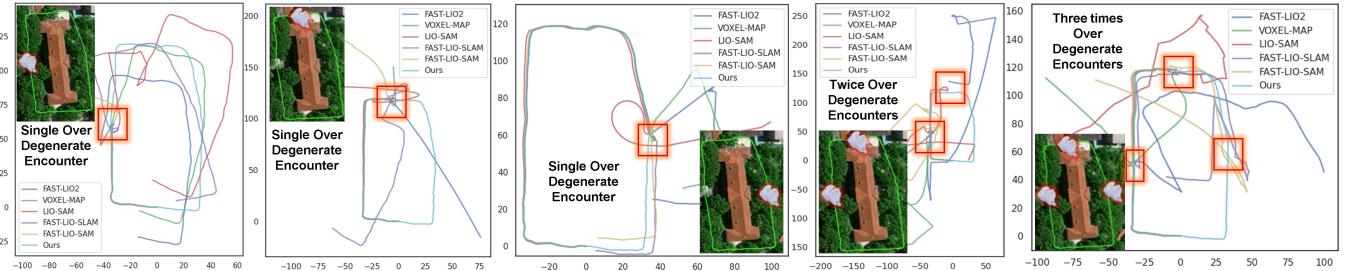


Fig. 6: Robot trajectories in campus scenes highlight LiDAR obstruction by drifting garbage bags on various occasions. Compared with various state-of-the-art methods, MM-LINS consistently performs in challenging conditions.

TABLE IV: The Number of Sub-maps Generated by the System with Different Degeneracy Detection Methods

Sequence	Zhang's	Ours	N.DE	LiDAR type
M2DGR-07-31	9	10	9	
M2DGR-08-06	11	10	9	Velodyne
NCLT-01-10	10	10	9	

"N.DE" denotes the number of degeneracy events. The theoretical count of submaps is one more than the number of degeneracy events.

superior performance of our method. Given that degeneracy detection is difficult to measure through objective metrics. We applied different degeneracy detection methods to our

system with map fusion disabled. The degeneracy detection performance was evaluated by comparing the number of submaps generated by the system and the number of degeneracy scenarios. To ensure the number of degeneracy scenarios is controllable, we modified some public datasets to simulate different levels of degeneracy scenarios, while keeping their quantity within a certain limit. Zhang's method sets the threshold to 100, while our method sets the translation and rotation thresholds to 0.005 and 0.0006, respectively. The results are shown in Table. IV. Zhang's method exhibits mismatches in quantity, whereas ours does not. This reflects that separating rotation and translation components can improve detection performance.

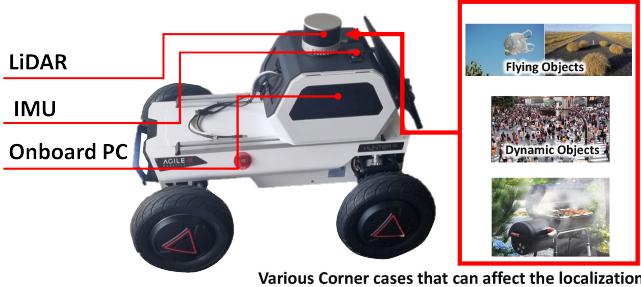


Fig. 7: Vehicle-mounted device used for the experiment.

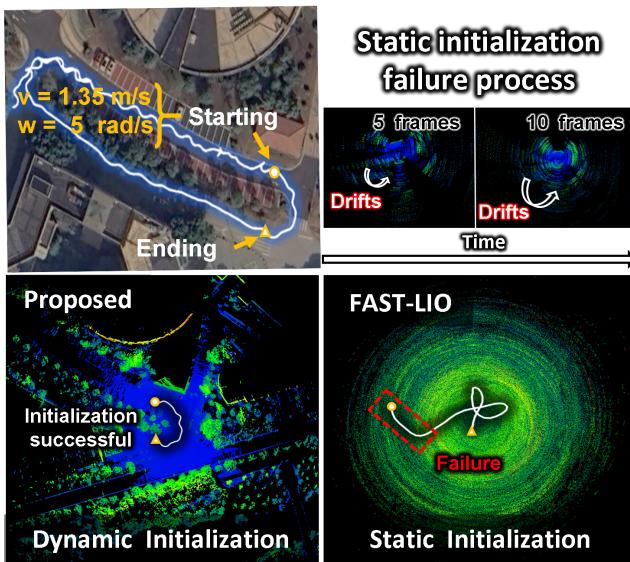


Fig. 8: Dynamic initialization performance demonstration on the M2DGR-08-06 sequence. The process begins playing the sequence at the 220s mark (during the robot's traversal) and lasts for 50 seconds. The static initialization method of FAST-LIO exhibits rotational drift, leading to system failure, whereas our method completes initialization and commences mapping.

For LiDAR-inertial systems, there is a scarcity of research on dynamic initialization. We employed both dynamic initialization and static initialization methods to start map construction as the robot moves. By observing the map quality, we further highlight the importance of dynamic initialization research, as shown in Fig. 8.

D. Computation Efficiency

The time consumed in each module is shown in Table. V. For the NCLT-01-10, UTBM-04-18, and Campus-1 sequences, degeneracy scenarios occurred only once (DI and CE-PGO executed only once), with all modules consuming less than 1.5% of the total time, which indicates a low additional computation cost compared to map construction. Overall, ODD accompanies the system throughout its lifetime but consumes minimal time. DI's time consumption is mainly related to the number of submaps, CE-PGO's time consumption is related to both the scale and the number of submaps. The larger the number and scale of submaps, the more time they consume.

TABLE V: The Percentage of Time Consumed by Each Module (%)

Sequence	ODD	DI	CE-PGO	Total (s)	Remark
NCLT-01-10	0.04	0.26	1.02	780	Single Event.
UTBM-04-18	0.06	0.15	0.95	1050	Single Event.
Campus-1	0.01	0.38	0.51	394	Single Event.
Campus-4	0.02	0.53	1.08	469	Two Events.
Campus-5	0.01	0.92	1.43	491	Three Events.

“ODD” denotes over-degeneracy detection module, “DI” denotes dynamic initialization module, “CE-PGO” denotes constraint-enhanced pose graph optimization module.

V. CONCLUSION

In this paper, we present and validate the multi-map LiDAR-inertial system. During periods of over-degeneracy, this system significantly reduces localization and mapping errors and prevents system failure, thanks to innovative aspects such as the reliable over-degeneracy detection method, robust multi-map strategy and constraint-enhanced map fusion strategy. The system's multi-map management module and constraint-enhanced pose graph optimization back-end are independent of the LIO system, enabling seamless transplantation to any superior LIO framework. However, our work still has some limitations. The accuracy of map fusion highly depends on the precision of point cloud recognition and registration. If there is minimal overlap in point clouds, the accuracy of map fusion will be significantly reduced. In the future, we will consider integrating semantic information to improve the accuracy of point cloud recognition and registration, making the system more robust.

REFERENCES

- [1] Y. Jia, X. Yan, and Y. Xu, “A survey of simultaneous localization and mapping for robot,” in *2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, vol. 1, 2019, pp. 857–861.
- [2] J. Li, S. Yuan, M. Cao, T.-M. Nguyen, K. Cao, and L. Xie, “Hcto: Optimality-aware lidar inertial odometry with hybrid continuous time optimization for compact wearable mapping system,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 211, pp. 228–243, 2024.
- [3] T.-M. Nguyen, S. Yuan, T. H. Nguyen, P. Yin, H. Cao, L. Xie, M. Wozniak, P. Jensfelt, M. Thiel, J. Ziegelnbein, and N. Blunder, “McD: Diverse large-scale multi-campus dataset for robot perception,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 6 2024. [Online]. Available: <https://mcdrviral.github.io/>
- [4] K. Xu, Y. Hao, S. Yuan, C. Wang, and L. Xie, “Airvo: An illumination-robust point-line visual odometry,” in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023, pp. 3429–3436.
- [5] Y. Lyu, M. Cao, S. Yuan, and L. Xie, “Vision-based plane estimation and following for building inspection with autonomous uav,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023.
- [6] S. Chen, K. Liu, C. Wang, S. Yuan, J. Yang, and L. Xie, “Salient sparse visual odometry with pose-only supervision,” *IEEE Robotics and Automation Letters*, 2024.
- [7] T. Deng, Y. Chen, L. Zhang, J. Yang, S. Yuan, D. Wang, and W. Chen, “Compact 3d gaussian splatting for dense visual slam,” *arXiv preprint arXiv:2403.11247*, 2024.
- [8] T. Deng, N. Wang, C. Wang, S. Yuan, J. Wang, D. Wang, and W. Chen, “Incremental joint learning of depth, pose and implicit scene representation on monocular camera in large-scale scenes,” *arXiv preprint arXiv:2404.06050*, 2024.

- [9] M. A. Esfahani, K. Wu, S. Yuan, and H. Wang, "From local understanding to global regression in monocular visual odometry," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 34, no. 01, p. 2055002, 2020.
- [10] ———, "Deepdsair: Deep 6-dof camera relocalization using deblurred semantic-aware image representation for large-scale outdoor environments," *Image and Vision Computing*, vol. 89, pp. 120–130, 2019.
- [11] W. Zhao, H. Sun, X. Zhang, and Y. Xiong, "Visual slam combining lines and structural regularities: Towards robust localization," *IEEE Transactions on Intelligent Vehicles*, pp. 1–18, 2023.
- [12] J. Zhang, H. Zhuge, Y. Liu, G. Peng, Z. Wu, H. Zhang, Q. Lyu, H. Li, C. Zhao, D. Kircali *et al.*, "Ntu4dradlm: 4d radar-centric multi-modal dataset for localization and mapping," in *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2023, pp. 4291–4296.
- [13] X. Ji, S. Yuan, P. Yin, and L. Xie, "Lio-gvm: an accurate, tightly-coupled lidar-inertial odometry with gaussian voxel map," *IEEE Robotics and Automation Letters*, 2024.
- [14] T.-M. Nguyen, D. Duberg, P. Jensfelt, S. Yuan, and L. Xie, "Slicht: Multi-input multi-scale surfel-based lidar-inertial continuous-time odometry and mapping," *IEEE Robotics and Automation Letters*, vol. 8, no. 4, pp. 2102–2109, 2023.
- [15] S. Yuan, H. Wang, and L. Xie, "Survey on localization systems and algorithms for unmanned systems," *Unmanned Systems*, vol. 9, no. 02, pp. 129–163, 2021.
- [16] G. Bresson, Z. Alsayed, L. Yu, and S. Glaser, "Simultaneous localization and mapping: A survey of current trends in autonomous driving," *IEEE Transactions on Intelligent Vehicles*, vol. 2, no. 3, pp. 194–220, 2017.
- [17] M. Ramezani, K. Khosoussi, G. Catt, P. Moghadam, J. L. Williams, P. Borges, F. Pauling, and N. Kottege, "Wildcat: Online continuous-time 3d lidar-inertial slam," *ArXiv*, vol. abs/2205.12595, 2022. [Online]. Available: <https://api.semanticscholar.org/CorpusID:249062855>
- [18] S. Li, G. Li, L. Wang, and Y. Qin, "Slam integrated mobile mapping system in complex urban environments," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 166, pp. 316–332, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0924271620301386>
- [19] T.-M. Nguyen, M. Cao, S. Yuan, Y. Lyu, T. H. Nguyen, and L. Xie, "Liro: Tightly coupled lidar-inertia-ranging odometry," in *2021 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2021, pp. 14484–14490.
- [20] T.-M. Nguyen, S. Yuan, M. Cao, L. Yang, T. H. Nguyen, and L. Xie, "Miliom: Tightly coupled multi-input lidar-inertia odometry and mapping," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5573–5580, 2021.
- [21] X. Tian, Z. Zhu, J. Zhao, G. Tian, and C. Ye, "Dl-slot: Tightly-coupled dynamic lidar slam and 3d object tracking based on collaborative graph optimization," *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 1, pp. 1017–1027, 2024.
- [22] K. Nielsen and G. Hendeby, "Multi-hypothesis slam for non-static environments with reoccurring landmarks," *IEEE Transactions on Intelligent Vehicles*, 2022.
- [23] T. Ji, S. Yuan, and L. Xie, "Robust rgb-d slam in dynamic environments for autonomous vehicles," in *2022 17th International Conference on Control, Automation, Robotics and Vision (ICARCV)*. IEEE, 2022, pp. 665–671.
- [24] J. Zhang and S. Singh, "Loam: Lidar odometry and mapping in real-time," in *Robotics: Science and Systems*, 2014. [Online]. Available: <https://api.semanticscholar.org/CorpusID:18612391>
- [25] W. Xu, Y. Cai, D. He, J. Lin, and F. Zhang, "Fast-llo2: Fast direct lidar inertial odometry," *IEEE Transactions on Robotics*, vol. 38, no. 4, pp. 2053–2073, 2022.
- [26] C. Bai, T. Xiao, Y. Chen, H. Wang, F. Zhang, and X. Gao, "Faster-llo: Lightweight tightly coupled lidar-inertial odometry using parallel sparse incremental voxels," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4861–4868, 2022.
- [27] T.-M. Nguyen, M. Cao, S. Yuan, Y. Lyu, T. H. Nguyen, and L. Xie, "Viral-fusion: A visual-inertial-ranging-lidar sensor fusion approach," *IEEE Transactions on Robotics*, vol. 38, no. 2, pp. 958–977, 2021.
- [28] T.-M. Nguyen, X. Xu, T. Jin, Y. Yang, J. Li, S. Yuan, and L. Xie, "Eigen is all you need: Efficient lidar-inertial continuous-time odometry with internal association," *IEEE Robotics and Automation Letters*, 2024.
- [29] Z. Chen, Y. Xu, S. Yuan, and L. Xie, "ig-llo: An incremental gicp-based tightly-coupled lidar-inertial odometry," *IEEE Robotics and Automation Letters*, vol. 9, no. 2, pp. 1883–1890, 2024.
- [30] C. Yuan, W. Xu, X. Liu, X. Hong, and F. Zhang, "Efficient and probabilistic adaptive voxel mapping for accurate online lidar odometry," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 8518–8525, 2022.
- [31] G. Kim and A. Kim, "Scan context: Egocentric spatial descriptor for place recognition within 3D point cloud map," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Madrid, Oct. 2018.
- [32] J. Zhang, M. Kaess, and S. Singh, "On degeneracy of optimization-based state estimation problems," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 809–816.
- [33] W. Zhen, S. Zeng, and S. Soberer, "Robust localization and localizability estimation with a rotating laser scanner," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 6240–6245.
- [34] T. Tuna, J. Nubert, Y. Nava, S. Khattak, M. H. R. S. Lab, E. Zurich, and A. A.G., "X-icp: Localizability-aware lidar registration for robust localization in extreme environments," *IEEE Transactions on Robotics*, vol. 40, pp. 452–471, 2022. [Online]. Available: <https://api.semanticscholar.org/CorpusID:254069896>
- [35] Y. Xie, Y. Zhang, L. Chen, H. Cheng, W. Tu, D. Cao, and Q. Li, "Rdc-slam: A real-time distributed cooperative slam system based on 3d lidar," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 14 721–14 730, 2022.
- [36] H. Zhou, Z. Yao, Z. Zhang, P. Liu, and M. Lu, "An online multi-robot slam system based on lidar/uwb fusion," *IEEE Sensors Journal*, vol. 22, no. 3, pp. 2530–2542, 2022.
- [37] J. Wu, J. Wang, and C. Wei, "Robust map merging method for collaborative lidar-based slam using gps sensor," *Journal of Physics: Conference Series*, vol. 2402, 2022. [Online]. Available: <https://api.semanticscholar.org/CorpusID:254630391>
- [38] B. Peng, H. Xie, and W. Chen, "ROLL: Long-term robust lidar-based localization with temporary mapping in changing environments," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022, pp. 2841–2847.
- [39] Z. Yu, Z. Qiao, L. Qiu, H. Yin, and S. Shen, "Multi-session, localization-oriented and lightweight lidar mapping using semantic lines and planes," *ArXiv*, vol. abs/2307.07126, 2023. [Online]. Available: <https://api.semanticscholar.org/CorpusID:259924529>
- [40] R. DubÅ©, A. Gawel, H. Sommer, J. Nieto, R. Siegwart, and C. Cadena, "An online multi-robot slam system for 3d lidars," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 1004–1011.
- [41] Y. Huang, T. Shan, F. Chen, and B. Englot, "DiSCo-SLAM: Distributed scan context-enabled multi-robot lidar slam with two-stage global-local graph optimization," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1150–1157, 2022.
- [42] H. Do, S. Hong, and J. Kim, "Robust loop closure method for multi-robot map fusion by integration of consistency and data similarity," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5701–5708, 2020.
- [43] G. Kim and A. Kim, "Lt-mapper: A modular framework for lidar-based lifelong mapping," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE Press, 2022, pp. 7995–8002. [Online]. Available: <https://doi.org/10.1109/ICRA46639.2022.9811916>
- [44] W. Huang and H. Liu, "Online initialization and automatic camera-imu extrinsic calibration for monocular visual-inertial slam," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 5182–5189.
- [45] R. Mur-Artal and J. D. Tardos, "Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras," *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255–1262, 2017.
- [46] T. Qin, P. Li, and S. Shen, "Vins-mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, 2018.
- [47] S. Mishra, G. Pandey, and S. Saripalli, "Target-free extrinsic calibration of a 3d-lidar and an imu," in *2021 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 2021, pp. 1–7.
- [48] W. Xu and F. Zhang, "Fast-llo: A fast, robust lidar-inertial odometry package by tightly-coupled iterated kalman filter," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3317–3324, 2021.
- [49] J. Yin, A. Li, T. Li, W. Yu, and D. Zou, "M2DGR: A multi-sensor and multi-scenario slam dataset for ground robots," *IEEE Robotics and Automation Letters*, pp. 1–1, 2021.
- [50] N. Carlevaris-Bianco, A. K. Ushani, and R. M. Eustice, "University of Michigan North Campus long-term vision and lidar dataset," *International Journal of Robotics Research*, vol. 35, no. 9, pp. 1023–1035, 2015.
- [51] Z. Yan, L. Sun, T. Krajnik, and Y. Ruichek, "EU long-term dataset with multiple sensors for autonomous driving," in *Proceedings of the 2020*

- IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020.*
- [52] W. Wen, Y. Zhou, G. Zhang, S. Fahandezh-Saadi, X. Bai, W. Zhan, M. Tomizuka, and L. Hsu, "Urbanloco: A full sensor suite dataset for mapping and localization in urban scenes," *CoRR*, vol. abs/1912.09513, 2019. [Online]. Available: <http://arxiv.org/abs/1912.09513>
- [53] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and D. Rus, "Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 5135–5142.
- [54] G. Kim, "FAST_LIO_SLAM," 2021, GitHub repository. [Online]. Available: https://github.com/gisbi-kim/FAST_LIO_SLAM
- [55] kahowang, "FAST_LIO_SAM," 2022, GitHub repository. [Online]. Available: https://github.com/kahowang/FAST_LIO_SAM

VI. BIOGRAPHY



Yongxin Ma received the B.E. degree from Jiangsu University of Science and Technology, Zhenjiang, China, in 2022, and he is currently pursuing the Ph.D. degree in Mechanical Engineering with the School of Mechanical Engineering, Shandong University, Jinan, China. His research interests include LiDAR-inertial systems and robot localization.



Jie Xu received the B.E., M.S. degree from the Harbin Institute of Technology in 2018, 2020 respectively. He is currently pursuing a Ph.D. degree with the School of Mechanical and Electrical Engineering, Harbin Institute of Technology. Meanwhile, he is participating in a joint doctoral training program at NTU EEE College.

His research interests include multi-sensor calibration, RGB-D SLAM and multi-modal SLAM.



Shenghai Yuan received the B.E. and Ph.D. degrees in electrical and electronic engineering from Nanyang Technological University, Singapore, in 2019 and 2013, respectively. He is currently a Senior Research Fellow with the Center for Advanced Robotics Technology Innovation (CARTIN) at Nanyang Technological University. His research interests include perception, Sensor fusion for robust navigation, machine learning, and various autonomous systems.



Tian Zhi received his B.S. degree in Mechanical and Electronic Engineering from Suzhou City College, Suzhou, China, in 2017 and his M.S. degree in Aerospace Science and Technology from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2024.

His research interests include robotics navigation, perception, and SLAM.



Wenlu Yu was born in Heilongjiang, China in 2000. He received his B.E. degree from the Harbin Institute of Technology, Harbin, China, in 2022, and now he is currently pursuing an M.S. degree with the State Key Laboratory of Robotics and Systems.

His research interests include SLAM and multi-sensor fusion.



Jun Zhou received the B.S. degree in Mechanical Design and Manufacturing from China University of Mining and Technology, Beijing, China, in 1998. The M.E. degree and the Ph.D. degree in Mechanical Design and Automation from Shandong University, Jinan, China, in 2000 and 2004, respectively.

His current research interests include intelligent control systems, LiDAR SLAM, and robot localization.



Lihua Xie (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of Newcastle, Australia, in 1992. Since 1992, he has been with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, where he is currently a professor and Director, Center for Advanced Robotics Technology Innovation. He served as the Head of the Division of Control and Instrumentation from July 2011 to June 2014 and Co-Director of Delta-NTU Corporate Lab for Cyber-Physical Systems from 2016 to 2021. He held teaching appointments in the Department of Automatic Control, Nanjing University of Science and Technology from 1986 to 1989. His research interests include robust control and estimation, networked control systems, multi-agent networks, localization, and unmanned systems. He is an Editor-in-Chief for Unmanned Systems and an Associate Editor for IEEE Control System Magazine. He has served as Editor of the IET Book Series in Control and Associate Editor of a number of journals, including IEEE Transactions on Automatic Control, Automatica, IEEE Transactions on Control Systems Technology, IEEE Transactions on Network Control Systems, and IEEE Transactions on Circuits and Systems-II. He was an IEEE Distinguished Lecturer (Jan 2012 - Dec 2014). Dr Xie is a Fellow of the Academy of Engineering Singapore, IEEE, IFAC, and CAA.