When to Map? Adaptive Switching Between Localization and SLAM in Multi-Session Systems

Lorenzo Montano-Oliván*, Julio A. Placed, Luis Montano, and María T. Lázaro

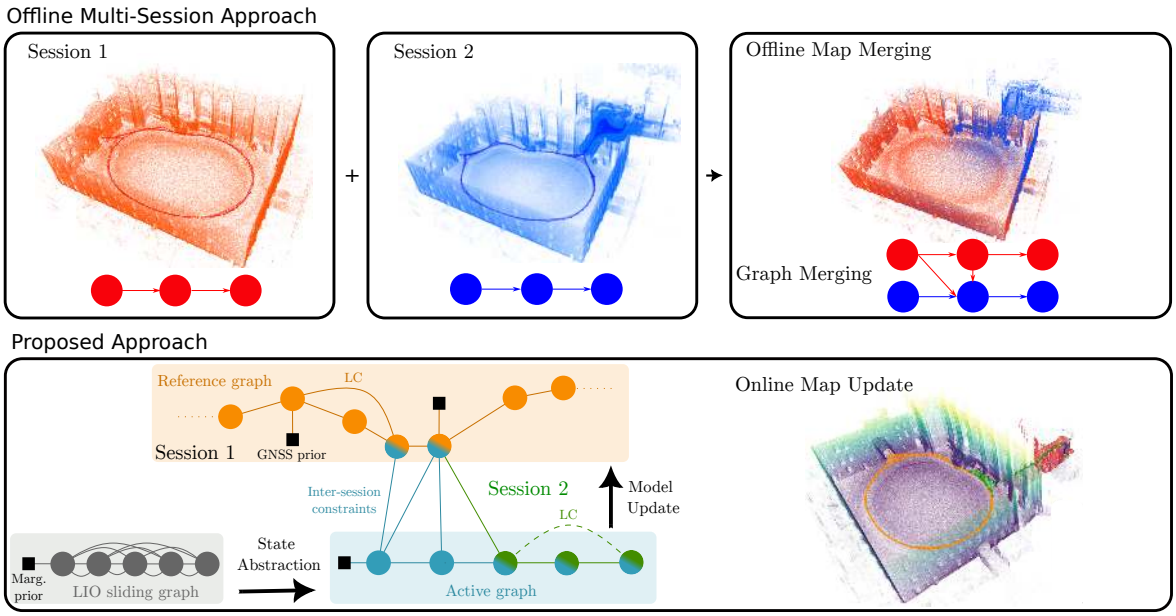

Fig. 1: The top row of the figure shows traditional multi-session approaches, in which separate SLAM sessions are performed and their maps and graphs are combined, typically offline. In contrast, the lower pipeline shows the proposed method, which allows maps to be updated online between sessions. This visualization represents the different graphs used in the multi-session approach: the active graph (blue) is constructed from selected key frames of the odometry graph (gray) along with their condensed measurements, as well as vertices from the reference graph (orange) connected through inter-session correspondences. When triggered by the decision-making module (*e.g.,* in unmapped regions), a subset of the active graph (candidate graph, green) is integrated into the reference model. Finally, loop closing and global pose graph optimization are performed to reconcile measurements across sessions and ensure consistency.

***Abstract*— Autonomous robots often operate repeatedly in previously mapped environments, where running full SLAM in every session can be unnecessary and inefficient. We propose a multi-session framework that instead prioritizes map-based localization, activating mapping only when required. A topology-informed, uncertainty-aware mechanism analyzes the pose-graph structure to detect when the prior map no longer sufficiently supports the current observations, triggering mapping and loop closure selectively. This enables incremental map expansion while maintaining global consistency and reducing redundant computation. Experiments on public datasets and a real-world mine environment show that the approach achieves accuracy comparable to full SLAM while limiting unnecessary mapping.**

This work was partially supported by DGA_FSE T73_23R, the EU project MASTERMINE (HORIZON-CL4-2022-RESILIENCE-01, Grant ID: 101091895), and by project UNDERAIBOT (CPP2022-009792) funded by MICIU/AEI/10.13039/501100011033 and European Union (NextGenerationEU/PRTR).

*Corresponding author.

Lorenzo Montano-Oliván, Julio A. Placed, and María T. Lázaro are with the Instituto Tecnológico de Aragón (ITA), María de Luna 7, Zaragoza, Spain (e-mail: {lmontano, jplaced, mtlazaro}@ita.es).

Luis Montano is with the Instituto de Investigación en Ingeniería de Aragón (I3A), Universidad de Zaragoza, María de Luna 1, Zaragoza, Spain (e-mail: montano@unizar.es).

I. Introduction

Autonomous robotic systems are increasingly required to operate repeatedly in previously visited environments. Applications such as warehouse automation, infrastructure inspection, and long-term deployment in structured indoor spaces demand not only accurate mapping, but also effective reuse of prior knowledge across time.

Simultaneous Localization and Mapping (SLAM) has been the dominant framework for estimating both the trajectory of a robot and a map of the environment [1]. To extend this capability to repeated operation, multi-session SLAM approaches incrementally build, align, and refine maps across sessions (often offline due to computational complexity), enabling consistent long-term representations [2]. In particular, multi-session LiDAR mapping has become increasingly relevant in domains where high-precision maps are required, such as surveying, search and rescue, and autonomous driving [3], [4]. A common strategy consists of executing SLAM independently in each session and subsequently merging the resulting maps, either offline or through incremental alignment and pose-graph optimization [5], [6]. While effective,

these approaches typically involve repeatedly performing full SLAM pipelines, even in regions that have already been properly reconstructed. This can lead to redundant computation, increased system complexity, and challenges related to drift, map consistency, and scalability, particularly when inter-session constraints are sparse.

An alternative perspective is to consider map-based localization as the primary mode of operation once a reliable prior map is available, using SLAM selectively to extend the map when needed. Such a strategy has the potential to reduce unnecessary mapping efforts while maintaining accuracy. However, it raises a key question: *How can we determine when the existing map is sufficient and when new mapping should be triggered?*

In this work, we build on this perspective and propose a principled mechanism to adaptively switch between localization and mapping based on the structure of the underlying Pose-Graph (PG). Rather than assuming that each session requires full SLAM, the system evaluates whether current observations are adequately supported by the prior map. The main contributions of this work are as follows:

- A novel uncertainty-aware decision mechanism based on the connectivity of the PG, enabling automatic switching between localization and mapping modes.
- A multi-session system that integrates this mechanism within state-of-the-art LiDAR-based localization and mapping frameworks.
- Experimental validation on public datasets and a real-world underground mine environment, showing that the proposed approach maintains SLAM-level accuracy while reducing unnecessary mapping operations.

This formulation provides a step toward more efficient multi-session operation by prioritizing localization and invoking mapping only when required, while remaining compatible with existing SLAM pipelines.

II. Related Work

Multi-session mapping capabilities have been extensively studied in the visual SLAM literature. In [6], loop closure factors are employed to jointly optimize multiple trajectories from different sessions. Daoud *et al.* [7] combine multiple maps online by adding these constraints between keyframes with enough similarity. ORB-SLAM Atlas [8] improves the previous by treating the maps as a single entity with a shared descriptor database. In addition, they remove duplicate points, but map merging requires $\approx$ 13x frame rate.

In the domain of LiDAR mapping, where the accuracy and consistency of the maps are crucial, multi-session remains an even more challenging problem. Many works have addressed this problem by creating entirely new maps for each session and combining them either online or offline. LT-Mapper [2] combines multiple maps from independent sessions using ScanContext and ICP verification. The constraints are included as anchor node-based inter-session loop factors to reduce the optimization complexity, and rely on a robust back-end to mitigate inevitable false loop detections. In [9], the optimization complexity across sessions is kept bounded through the use of condensed measurements, which enable the merging of local maps upon detecting inter-session loop closures, while simultaneously removing out-dated nodes during the fusion process.

Automerge [5] demonstrates large-scale (*i.e.,* hundreds of km) map merging. However, it heavily depends on sufficient overlapping between maps and the availability of GNSS. In addition, the matching complexity scales quadratically with the number of session keyframes, requiring minutes to complete. FRAME [10] employs learned descriptors and place recognition to detect overlap between point cloud maps, boosting the map matching process. Despite this is treated as a triggered event, the events are predefined by the user (*e.g.,* when two robots are nearby). LM-Mapping [3] proposes a two-level optimization with local odometry constraints and inter-session factors. Despite this allows to improve local precision and mitigate inter-map inconsistencies, it still requires to rebuild the maps on every session and is relegated to an offline post-processing phase. In MS-Mapping [4], each new LiDAR cloud is registered both sequentially and with chunks of the prior map, incrementally building the graph and a merged map. Once again, the multi-session mapping problem is formulated as a PG optimization, where inter-session constraints are modeled as loop closure factors. This approach is shown to outperform direct map-to-map and scan-to-map registration. LEMON-Mapping [11] presents a scalable multi-robot system that leverages geometric constraints from overlapping point clouds to strengthen the role of inter-session LC. Reliable loop selection and joint inter-session (hierarchical) PG optimization are shown to be crucial for mitigating map divergence across sessions.

In contrast to approaches that repeatedly perform full SLAM across sessions, our method follows a complementary perspective by prioritizing localization in previously mapped regions and invoking mapping selectively when required. Rather than constructing and maintaining complete maps for all sessions, we maintain a local *submap* and extend the representation only when new areas are observed. Similar to [10], mapping can be interpreted as a triggered event; however, in our case, this decision is performed automatically based on the structure of the PG.

The problem of multi-session mapping is closely related to the *kidnapped robot problem* [12], where the goal is to estimate the robot's pose given a prior map and no *a priori* information about its location within it. In this context, **map-based localization** provides an effective solution, although its performance is inherently bounded by the quality and coverage of the prior map. G-Loc [13] proposes a robust PG-based method for LiDAR-inertial localization using prior topological and geometric information. However, its operation in unmapped regions is limited and it does not incorporate a way to expand the prior map. LTA-OM [14] first triggers a localization method and then loads the prior map into the odometry module [15], thus not requiring additional map merging operations nor rebuilding a map when operating in the same area.

In contrast to these approaches, which focus primarily

on localization given a prior model, this work addresses the complementary problem of determining when the prior map is no longer sufficient. We introduce a topology-based decision mechanism that analyzes the connectivity of the PG to decide whether the system should continue in localization mode or activate mapping to incorporate new observations. Furthermore, our approach does not require the prior map to be generated by the same SLAM system, increasing its flexibility in practical deployments.

III. Method

We propose a topology-driven framework for multi-session robotic operations that prioritizes map-based localization. Instead of performing full SLAM in each session, the system primarily operates in localization mode using an existing map. It selectively activates mapping and loop closing to expand the map when necessary. To support this behavior, the system analyzes the joint PG —combining the current session and prior map— to assess whether the robot state is sufficiently constrained. If not, mapping is triggered to incorporate new observations and reinforce connectivity.

The framework integrates three main components: (i) a map-based localization backbone, (ii) a mapping and loop closing module that establishes both intra- and inter-session constraints, and (iii) a topology-aware mechanism that governs the transition between operation modes. See Fig. 1.

A. Multi-session Graph Localization

We adopt G-Loc [13] as the localization backbone, which combines LiDAR-inertial Odometry (LIO) and map matching constraints into a multiple graph optimization framework. The bottom stream in Fig. 1 depicts the different graphs involved in the localization process. First, LIO hyper-edges and scan-to-scan matching are optimized within a sliding-window fashion (gray graph in that figure), producing coarse estimates of the robot poses at IMU rate. A marginalization prior is used to retain information from past states outside the optimization window.

The *active graph* (blue) contains vertices that encode a sparse subset of the LIO states, *i.e.,* LIO-optimized poses of the most recent n key frames, $\mathcal{T} \triangleq \{\mathbf{T}_1, \dots, \mathbf{T}_n \mid \mathbf{T}_i \in SE(3)\}$, together with their associated submaps. Such vertices are connected consecutively via submap-to-submap matching constraints. In addition, this graph includes inter-session constraints (submap-to-map matching) that link $\mathcal{T}$ with nearby vertices of the *reference graph* (orange) and their k-neighbors. The active graph is optimized to obtain a precise localization, keeping the nodes from the reference graph fixed, *i.e.,*

$$\mathcal{T}^\star = \underset{\mathcal{T}}{\operatorname{argmin}} \left(\sum_{(i,j)\in\mathcal{E}_{\text{intra}}} \|\boldsymbol{e}_{i,j}^{\text{intra}}\|^2 + \sum_{(i,r)\in\mathcal{E}_{\text{inter}}} \|\boldsymbol{e}_{i,r}^{\text{inter}}\|^2 \right), \tag{1}$$

where $\|\boldsymbol{e}\|^2 \triangleq (\boldsymbol{e}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{e}) \in \mathbb{R}$ is the quadratic error for a measurement with covariance $\boldsymbol{\Sigma}$, $\boldsymbol{e}_{i-1,i}^{\text{intra}} \in \mathfrak{se}(3)$ are the residuals associated with consecutive submap matching (intra-session constraints), and $\boldsymbol{e}_{i,r}^{\text{inter}} \in \mathfrak{se}(3)$ are those associated with the correspondences between current observations and the reference map (inter-session constraints). The pair set $\mathcal{E}_{\text{intra}} \triangleq \{(i,j) \mid 1 \le i < j \le n, j = i+1\}$ and $\mathcal{E}_{\text{inter}} \triangleq \{(i,r) \mid 1 \le i \le n, r \in \mathcal{I}_{\text{ref}}\}$, with $\mathcal{I}_{\text{ref}}$ the index set of vertices that belong to the reference graph.

Solving (1) enables limited spatial and topological association with previous sessions. G-Loc treats the reference model (PG and point cloud submaps) as a passive input —immutable information used solely for localization. This approach underutilizes the rich information encoded in the PG, including its representation of uncertainty. To address this, we extend [13] by transforming the reference model into an active component of the localization pipeline: when the system traverses an unseen or poorly mapped area, the model is incrementally updated with new observations.

First, all vertices together with their associated submaps that leave the sliding window of key frames ($\mathcal{T}$) are retained. To bound memory usage, temporal and graph-distance limits are imposed, pruning elements that exceed them. When the mapping mode is triggered by the decision-making module, these inactive nodes are selected to form the *candidate graph* (green in Fig. 1), which is then merged into the reference model. As they enter the reference graph, they are automatically incorporated into a LC database. Robust LC search, matching and false-positive rejection (see [16]) occurs in a separate thread, improving the overall connectivity and accuracy. Therefore, under this mapping mode, the optimization problem (1) becomes:

$$\{\mathcal{T}, \mathcal{R}\}^\star = \underset{\{\mathcal{T},\mathcal{R}\}}{\operatorname{argmin}} \left(\sum_{(i,j)\in\mathcal{E}_{\text{intra}}} \|\boldsymbol{e}_{i,j}^{\text{intra}}\|^2 + \sum_{(i,r)\in\mathcal{E}_{\text{inter}}} \|\boldsymbol{e}_{i,r}^{\text{inter}}\|^2 + \sum_{(r,s)\in\mathcal{E}_{\text{LC}}} \|\boldsymbol{e}_{r,s}^{\text{LC}}\|^2 \right), \tag{2}$$

where $\mathcal{R} \triangleq \{\mathbf{T}_r \in SE(3) \mid r \in \mathcal{I}_{\text{ref}}\}$ denotes the set of reference poses, which in mapping mode can grow over time and are jointly optimized together with $\mathcal{T}$. Note that global PGO is now performed, taking into account both the current and reference poses. In addition, $\boldsymbol{e}_{r,s}^{\text{LC}} \in \mathfrak{se}(3)$ are the residuals from LC constraints and $\mathcal{E}_{\text{LC}} \triangleq \{(r,s) \mid r, s \in \mathcal{I}_{\text{ref}}, r \neq s\}$.

B. Switching Mechanism

The decision to localize on a prior map or extend it through mapping is driven by the overall structure of the current and reference PGs. The system dynamically switches between localization and mapping based on the connectivity and uncertainty in the environment. If the graphs are weakly connected or disconnected, this suggests the system is either entering an unvisited area or facing insufficient overlap, both of which require initiating mapping to expand the map and improve localization. Additionally, regions with high uncertainty in the pose-graph are flagged for map reinforcement to ensure robustness and consistency across sessions.

The structural patterns of the PGs reflect the quality and stability of the underlying optimization. Well-distributed, strongly connected loops enhance global consistency, facilitating reliable localization. In contrast, sparse or weakly

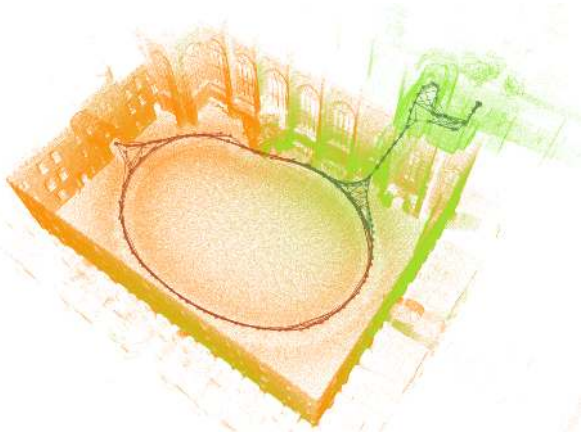

Fig. 2: Multi-session mapping results in NCE dataset Extension [18] (*Quad E*, *Quad H*); each depicted in a different color for visualization.

connected graphs may signal uncertainty or drift, which are often associated with poor optimization and unstable estimates. In such cases, the system detects when the prior map's structure is insufficient to support reliable localization, triggering a switch to mapping mode. This enables the system to incrementally extend the map, reinforce weak connections, and improve global consistency as it encounters new areas or insufficient overlap. In practice, we employ spectral metrics derived from the PG (*e.g.,* Fiedler value, node degree) to assess connectivity and uncertainty.

C. Initial Inter-Session Localization

To initialize localization in a prior map, we employ a place-recognition approach based on ScanContext++ [17]. It retrieves the k-best loop candidates from the prior map, followed by a 3D alignment step to select the most consistent one and obtain an initial pose estimate within the existing map.

IV. Experiments

In order to validate the proposed method, we have conducted two sets of experiments. The first utilizes two widely used datasets to evaluate the performance of our system in a controlled setting, enabling comparison against existing approaches and a detailed analysis. The second experiment involves a real-world deployment in an underground mine, aimed at demonstrating the system's robustness and application in challenging environments. All experiments were performed on a system equipped with an Intel Core Ultra 7 155H processor and an NVIDIA GeForce RTX 4070 GPU.

A shared requirement for all experiments is the availability of prior topological and geometric models of the environment (*i.e.,* a PG and a point cloud). To fulfill this requirement, we use [16], a robust and accurate framework for LiDAR-inertial mapping.

A. Dataset Evaluation

The Newer College (NC) [22] and Newer College Extension (NCE) [18] datasets contain data recorded using a handheld device in diverse environments. In NC, data was recorded using an Ouster OS1-64 LiDAR and its internal Inertial Measurement Unit (IMU) while NCE uses an Ouster OS0-128 and an Alphasense Core Both datasets provide ground-truth trajectories and include sequences with partially overlapping trajectories that have been previously used for multi-session SLAM evaluation [4], [23].

The experiment consisted of gradually processing 3 overlapping sequences from NCE. First, we create a base model with *Quad-e* (orange in Fig. 2). Notably, *Quad-m* sequence completely overlaps with the first, enabling the system to operate in map-based localization mode. Then, we perform incremental updates using the *Quad-h* (green), extending the information provided by the preceding one. Fig. 2 illustrates the joint model after processing both sequences. The new region is shown in a different color, showcasing the contribution of this session, the strong inter-session links, and the fact that overlapping regions are modeled only once.

To assess the performance of the proposed method, we present the Absolute Trajectory Error (ATE) Root Mean Squared Error (RMSE) for each individual sequence, and compare our approach with state-of-the-art LIO/SLAM [16], [19], [20] methods (which build separate maps for each sequence) and map-based localization systems [13], [21] (all of them using *Quad-e* as reference map). Note that, in contrast, our method concatenates the sessions and incrementally leverages the information from previous sequences. The results are reported in Table I.

Overall, the proposed approach produced the best result in every sequence and had the lowest average RMSE, alongside LG-SLAM. Since the mapping module is similar to that of LG-SLAM, both approaches achieved the same level of accuracy in *Quad-e*, where no prior information was available. Similarly, when the sequences fully overlap (*i.e.,* in *Quad-m*), our method exhibits the same behavior as G-Loc. However, in *Quad-h*, where there is a new indoor region not present in the previous sequences, our method outperforms all others. As this new area is small, G-Loc still performs accurately, demonstrating its robustness even in partially unmapped regions. However, map-based localization methods fail or accumulate drift in *Quad-h*, as there is an unvisited region. Notably, our approach achieves accuracy comparable to SLAM methods while substantially reducing graph complexity by alternating between localization and mapping modes, and providing a joint model.

Table II contains a comparison against a multi-session systems in *Parkland* (S1) and *Short* (S2) sequences from [22]. We compare against MS-Mapping [4] and two multi-session localization algorithms (M2F and F2F) reported in [4]. Specifically, M2F localizes the new session point cloud frame-by-frame on the prior map to establish constraints, while F2F localizes frames from the current session against the joint map. The first column reports ATE RMSE after processing each sequence individually, where our approach outperforms the alternatives. The second column shows results when leveraging cross-session information —using *Parkland* while processing *Short*, and vice versa. Results for

Method / Sequence		*Quad-e*	*Quad-m*	*Quad-h*	Avg.
SLAM	LIO-SAM [19]	<u>0.074</u>	**0.067**	0.13	0.090
	NV-LIOM [20]	0.076	0.076	0.18	0.110
	LG-SLAM [16]	**0.066**	<u>0.072</u>	0.090	<u>0.076</u>
Map-based Localization	HDL-Loc [21]	–	0.10	×	×
	G-LOC [13]	–	**0.067**	<u>0.087</u>	0.077
Ours		**0.066**	**0.067**	**0.082**	**0.071**

TABLE I: ATE RMSE (m) in NC and NCE dataset. SLAM methods build separate maps for each sequence. Map-based localization methods employ *Quad-e* as reference map. Our method incrementally concatenates sessions (left to right). Best results are bold and second best are underlined. × indicates an unrecoverable failure.

Method / Seq.	Individual		Combination	
	S1	*S2*	*S1* (given *S2*)	*S2* (given *S1*)
MS-Mapping [4]	<u>0.26</u>	<u>0.45</u>	0.29	**0.40**
Ours	**0.17**	**0.38**	**0.24**	<u>0.47</u>

TABLE II: Comparison of ATE RMSE (m) for multi-session approaches on *Parkland* (S1) and *Short* (S2) sequences.

the combined sequences show comparable results, with lower computational requirements.

Finally, regarding execution time, our approach significantly reduces graph complexity by alternating between localization and mapping modes. For example, the combined graph of S1 and S2 consists of 600 nodes, while the individual S1 and S2 graphs have 770 nodes. This reduction leads to improved computational efficiency, highlighting the benefits of using multi-session information without needing to remap entire sequences. On average, our method processed each keyframe in 50 ms and required 196 ms for loop closure (LC) and pose graph optimization (PGO). In comparison, the SLAM baseline [16] processed each keyframe in 48 ms and took 231 ms for LC and PGO. Notably, in our approach, the LC thread was only active during part of the sequences, further reducing the overhead.

B. Real-World Experiment

The Kemi Mine is the largest underground mine in Findland, with 1 km depth. The experiments covered a distance of 8 km along a series of spiral primary and secondary galleries. Using a prototype vehicle and following diverse trajectories throughout the tunnels, data from the mine was collected using two Robosense M1 solid state LiDARs (front- and rear-facing) and an industrial-grade IMU. Both LiDARs were used to build the prior maps while only the front-facing LiDAR was used for localization to demonstrate the performance with mismatched field-of-view. Although the standard configuration uses both sensors, this experiment demonstrates that reliable localization can still be achieved under these conditions. In addition, Wi-Fi localization was used to provide a rough location estimate and correctly initialize the map-based localization algorithm, thanks to the large amount of access points located throughout the tunnels.

Fig. 3 contains the resulting map for two of the sequences in the mine, which partially overlap. The prior SLAM-generated model is shown in orange, while the extended model from the second session appears in green. Two areas of the prior map are seamlessly expanded into previously unmapped regions. The PG constructed during the second session exhibits strong internal connectivity, as well as robust links to the prior graph —enabled by the intelligent module responsible for orchestrating mapping and LC operations. Notably, one of the new regions (boxed area in Fig. 3) enforces a loop in the prior map. This clearly improves the overall connectivity of the joint PG. Also, thanks to the global optimization performed where uncertainties of both mapping processes are accommodated, the consistency and precision of the point cloud representation are enhanced.

The analysis of the eigenvectors for both graphs reveals that the weakest region lies within the first session in both cases; *i.e.,* the new session did not introduce any weaker edges.

Regarding time consumption, executing SLAM required, on average, 59.9 ms to process each frame (including odometry and mapping). In addition, 149.6 ms were consumed by the LC thread per iteration (including PGO). These values remained similar for both sequences. In contrast, our multi-session approach required 45.3 ms to process each frame. Mapping, LC and global PGO were selectively triggered, requiring 109.9 ms on average but being active just 10% of the time.

V. Conclusions

In this work, we proposed a framework for online multi-session robotic operation that enables continuous map updates across sessions. The method allows for seamless integration of new data into the reference map, avoiding the need for redundant remapping of previously visited areas. By operating incrementally, our approach maintains high localization accuracy while enabling real-time map expansion. Experiments on public datasets and real-world deployments show that our method achieves localization performance comparable to state-of-the-art SLAM systems. Future work will focus on improving real-time change detection in previously mapped areas and refining the integration process to further enhance the overall efficiency and accuracy of multi-session mapping.

Acknowledgement

The authors acknowledge the use of GPT-5.3 for improving readability throughout the manuscript.

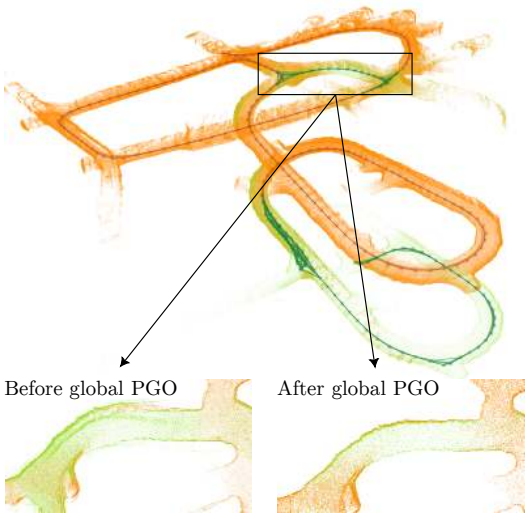

Fig. 3: Multi-session results in the Kemi underground mine. The boxed region shows a loop closure enforced between two areas previously mapped but disconnected. Furthermore, the zoomed-in images further illustrate the impact of global PGO.

REFERENCES

[1] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," *IEEE Trans. on Robotics*, vol. 32, no. 6, pp. 1309–1332, 2017.

[2] G. Kim and A. Kim, "Lt-mapper: A modular framework for LiDAR-based lifelong mapping," in *IEEE Int. Conf. on Robotics & Automation*, 2022, pp. 7995–8002.

[3] C. Pang, Z. Shen, R. Yuan, C. Xu, and Z. Fang, "Lm-mapping: Large-scale and multi-session point cloud consistent mapping," *IEEE Robotics and Automation L.*, 2024.

[4] X. Hu, J. Wu, J. Jiao, B. Jiang, W. Zhang, W. Wang, and P. Tan, "Ms-mapping: an uncertainty-aware large-scale multi-session lidar mapping system," *arXiv preprint arXiv:2408.03723*, 2024.

[5] P. Yin, S. Zhao, H. Lai, R. Ge, J. Zhang, H. Choset, and S. Scherer, "Automerge: A framework for map assembling and smoothing in city-scale environments," *IEEE Trans. on Robotics*, vol. 39, no. 5, pp. 3686–3704, 2023.

[6] B. Kim, M. Kaess, L. Fletcher, J. Leonard, A. Bachrach, N. Roy, and S. Teller, "Multiple relative pose graphs for robust cooperative mapping," in *IEEE Int. Conf. on Robotics & Automation*, 2010, pp. 3185–3192.

[7] H. A. Daoud, A. Q. Md. Sabri, C. K. Loo, and A. M. Mansoor, "SLAMM: Visual monocular SLAM with continuous mapping using multiple maps," *PloS one*, vol. 13, no. 4, p. e0195878, 2018.

[8] R. Elvira, J. D. Tardós, and J. M. Montiel, "ORBSLAM-Atlas: a robust and accurate multi-map system," in *IEEE/RSJ Int. Conf. on Intell. Robots and Sys.*, 2019, pp. 6253–6259.

[9] M. T. Lázaro, R. Capobianco, and G. Grisetti, "Efficient long-term mapping in dynamic environments," in *IEEE/RSJ Int. Conf. on Intell. Robots and Sys.*, 2018, pp. 153–160.

[10] N. Stathoulopoulos, B. Lindqvist, A. Koval, A.-A. Agha-Mohammadi, and G. Nikolakopoulos, "FRAME: A modular framework for autonomous map merging: Advancements in the field," *IEEE Trans. on Field Robotics*, vol. 1, pp. 1–26, 2024.

[11] L. Wang, X. Zhong, Z. Xu, K. Chai, A. Zhao, T. Zhao, C. Jiang, Q. Wang, and F. Gao, "LEMON-Mapping: Loop-enhanced large-scale multi-session point cloud merging and optimization for globally consistent mapping," *arXiv preprint arXiv:2505.10018*, 2025.

[12] J. McDonald, M. Kaess, C. Cadena, J. Neira, and J. J. Leonard, "Real-time 6-DOF multi-session visual SLAM over large-scale environments," *Robotics and Autonomous Sys.*, vol. 61, no. 10, pp. 1144–1158, 2013.

[13] L. Montano-Oliván, J. A. Placed, L. Montano, and M. T. Lázaro, "G-Loc: Tightly-coupled graph localization with prior topo-metric information," *IEEE Robotics and Automation L.*, vol. 9, no. 11, pp. 9167–9174, 2024.

[14] Z. Zou, C. Yuan, W. Xu, H. Li, S. Zhou, K. Xue, and F. Zhang, "LTA-OM: Long-term association LiDAR–IMU odometry and mapping," *J. of Field Robotics*, vol. 41, no. 7, pp. 2455–2474, 2024.

[15] W. Xu, Y. Cai, D. He, J. Lin, and F. Zhang, "Fast-lio2: Fast direct lidar-inertial odometry," *IEEE Trans. on Robotics*, vol. 38, no. 4, pp. 2053–2073, 2022.

[16] L. Montano-Oliván, J. A. Placed, L. Montano, and M. T. Lázaro, "From underground mines to offices: A versatile and robust framework for range-inertial SLAM," in *Iberian Robotics Conf.*, 2024, pp. 1–8.

[17] G. Kim, S. Choi, and A. Kim, "Scan context++: Structural place recognition robust to rotation and lateral variations in urban environments," *IEEE Trans. on Robotics*, vol. 38, no. 3, pp. 1856–1874, 2022.

[18] L. Zhang, M. Camurri, D. Wisth, and M. Fallon, "Multi-camera LiDAR inertial extension to the Newer College dataset," *arXiv preprint arXiv:2112.08854*, 2021.

[19] 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 *IEEE/RSJ Int. Conf. on Intell. Robots and Sys.*, 2020, pp. 5135–5142.

[20] D. Chung and J. Kim, "NV-LIOM: LiDAR-inertial odometry and mapping using normal vectors towards robust SLAM in multifloor environments," *IEEE Robotics and Automation Letters*, 2024.

[21] K. Koide, "Real-time 3D localization using a (velodyne) 3D LiDAR," https://github.com/koide3/hdl_localization, 2018, accessed: 2025.

[22] M. Ramezani, Y. Wang, M. Camurri, D. Wisth, M. Mattamala, and M. Fallon, "The Newer College dataset: Handheld lidar, inertial and vision with ground truth," in *IEEE/RSJ Int. Conf. on Intell. Robots and Sys.*, 2020, pp. 4353–4360.

[23] M. Menezes, M. Muñoz, E. Pignaton de Freitas, S. Cheng, A. de Almeida Neto, P. Ribeiro, and A. Oliveira, "A multisession SLAM approach for RatSLAM," *J. of Intell. & Robotic Sys.*, vol. 108, no. 4, p. 61, 2023.