Collaborative Robot Mapping using Spectral Graph Analysis

Lukas Bernreiter[1], Shehryar Khattak[2], Lionel Ott[1], Roland Siegwart[1], Marco Hutter[2] and Cesar Cadena[1]

Abstract— In this paper, we deal with the problem of creating globally consistent pose graphs in a centralized multi-robot SLAM framework. For each robot to act autonomously, individual onboard pose estimates and maps are maintained, which are then communicated to a central server to build an optimized global map. However, inconsistencies between onboard and server estimates can occur due to onboard odometry drift or failure. Furthermore, robots do not benefit from the collaborative map if the server provides no feedback in a computationally tractable and bandwidth-efficient manner. Motivated by this challenge, this paper proposes a novel collaborative mapping framework to enable accurate global mapping among robots and server. In particular, structural differences between robot and server graphs are exploited at different spatial scales using graph spectral analysis to generate necessary constraints for the individual robot pose graphs. The proposed approach is thoroughly analyzed and validated using several real-world multi-robot field deployments where we show improvements of the onboard system up to 90%.

I. INTRODUCTION

Collaborative multi-robot exploration with heterogeneous platforms and sensors imposes significant challenges on current state-of-the-art mapping and localization approaches. Maintaining a consistent pose estimate across all employed systems is particularly difficult for distributed modules and is mission-critical for the operation of robot teams in applications like disaster response and search and rescue.

With the recent availability of high-bandwidth mobile networks, e.g., 5G networks, centralized and collaborative robotic approaches have received increased attention in the robotics community due to their improved practical feasibility. Typically, mobile robotic systems have limited onboard computational resources for which centralized approaches can assist by offloading computational intensive operations from individual robots to a shared centralized server. A centralized server with more computational capacity can perform expensive operations such as global optimizations, loop closing, exploitation of all available sensor data to improve accuracy and also overcome onboard failure.

Most collaborative mapping approaches focus on building accurate maps on the server and ignore the use of global multi-robot information to provide localization feedback to individual agents to make the onboard and server maps
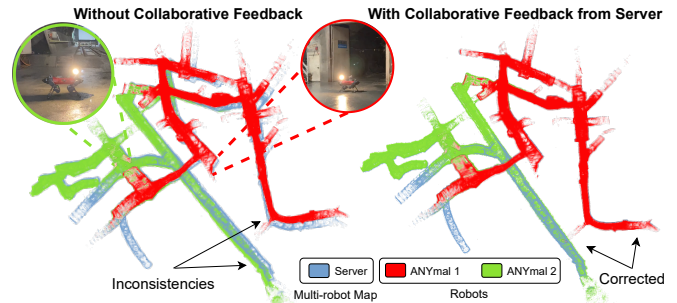
Fig. 1: Overview of a large-scale multi-robot deployment in an underground tunnel system. Structural differences between single-robot maps and a collaborative global map (left) are used to derive necessary constraints for making the onboard and global estimates consistent (right).

consistent w.r.t. each other. Especially in centralized settings without feedback, pose estimation discrepancies may arise between robots during large missions leading to severe drift between robot and server maps resulting in increased optimization time at the server for collaborative mapping. Therefore, it is evident that the provision of additional constraints to improve onboard estimation and collaborative mapping performance can be highly beneficial for large-scale multi-robot missions.

Furthermore, multi-robot missions often deploy a heterogeneous set of robots, e.g., aerial and ground robots, which depending on the type and role of the robot, carry a diverse set of multi-modal sensors onboard and correspondingly utilize different algorithms for robotic operations, e.g., localization and mapping. Consequently, no common layer sharing data to improve pose estimation and mapping accuracy among the employed systems is readily available. Hence, a sensor modality-invariant approach that can incorporate and communicate relevant correction information among robots while maintaining low network bandwidth requirements is essential for large-scale multi-robot field deployments.

Motivated by the discussion above, this paper proposes a novel collaborative multi-robot pose graph approach (cf. Figure 2) which is independent of the underlying robot pose estimation processes and relies only on a sparse abstraction of the estimated poses – a positional graph. Our framework operates in the graph spectral domain of the positional graphs to identify structural anomalies in the individual robot pose graphs using a multi-scale analysis. By examining the structural components of the positional graphs at different scales, our system identifies discrepancies in the local and coarser neighborhoods and adds corresponding constraints to
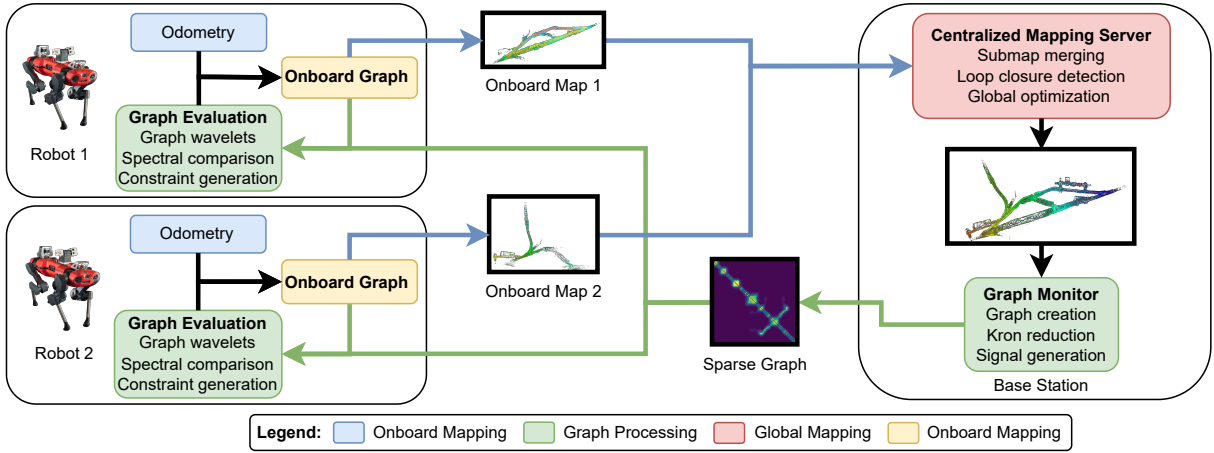
Fig. 2: Our approach considers multiple robots individually exploring an environment and sending incremental mapping data to a centralized server that accumulates and jointly optimizes them. A relaxation of the collaborative multi-robot map is sent back to the robots, where a multi-scale graph spectral analysis is performed to identify discrepancies onboard and server maps and to generate necessary constraints for making them consistent.

improve pose estimation accuracy of individual robots and to make the individual robot and collaborative server maps consistent. Key contributions of this paper are:

- Graph spectral analysis of positional graphs to identify disagreements between onboard and server pose graphs.
- Automatic adaptive inference of multi-scale constraints to achieve a consensus between robot and server maps.
- Comparison against current state-of-the-art approaches on datasets and a thorough quantitative analysis on large-scale multi-robot field deployments are presented to validate the proposed approach.

II. RELATED WORK

In this section, we review the state-of-the-art in collaborative and consistent multi-robot localization and mapping approaches as well as the current applications of graph signal processing.

A. Globally Consistent Multi-Robot Mapping

Collaborative multi-robot approaches can be distinguished into centralized [1]–[3] and distributed solutions [4,5]. Deutsch et al. [1] proposed a vision-based centralized multi-robot SLAM approach where a mapping server performs loop closures and replaces robot pose graphs with corrected graphs. A similar approach was proposed by [2] in which robots send local maps to a mapping server which then returns optimized keyframes and landmarks to each robot to include in their onboard optimizations. To improve the speed of onboard optimizing tasks, CoSLAM [6] proposes to make use of GPU computing, hence requiring a GPU onboard individual robots. Different from vision-only approaches, the work of Ebadi [7] proposes a large-scale collaborative multi-modal SLAM framework. However, their proposed approach does not provide any pose corrections from the centralized server to the individual robots.

In contrast to collaborative approaches, distributed approaches require each robot to run a full onboard SLAM solution [5] and share marginalized information with other robots [4], thus making full information available to each robot but increasing the onboard compute requirements significantly. Independent of the sensing modality used, [8] aims to achieve consistent maps across multiple robots by detecting loop closures between robots and connecting their pose graphs. In this direction, [9,10] aim to robustly select inter-robot loop closure candidates by maintaining pair-wise consistent measurements. More recently [11] proposed a distributed system with distributed loop closure detection.

Most current collaborative multi-robot mapping approaches require the partial or complete substitution of the onboard pose graphs with the server-optimized graph. Conversely, this paper proposes to detect discrepancies between the robot graphs using spectral analysis and a sparse abstraction of the server graph to generate an individual set of constraints for each robot. Hence, the proposed approach achieves high mapping accuracy while maintaining low network and compute requirements.

B. Graph Signal Processing

Spectral graph theory is an active research area and has gained popularity in the past years in the context of robotics. Spectral graph theory approaches have been proposed for robotic mapping [12], planning [13], and more recently, in combination with graph neural networks for various robotic tasks [14,15]. In general, graphs are irregular structures and are capable of modeling large, complex, and distributed problems [16], e.g. [17] proposes an anomaly detection for spatial proximity of graph nodes using spectral graph filtering. Furthermore, graph signal processing aims at applying signal processing techniques on graph structures, thus allowing the use of existing concepts such as the Laplacian operator [18] and multi-scale analysis [19,20]. Similarly, [21] aims to learn a multi-scale structural embedding using graph wavelets by treating the wavelet coefficients as a probability distribution.
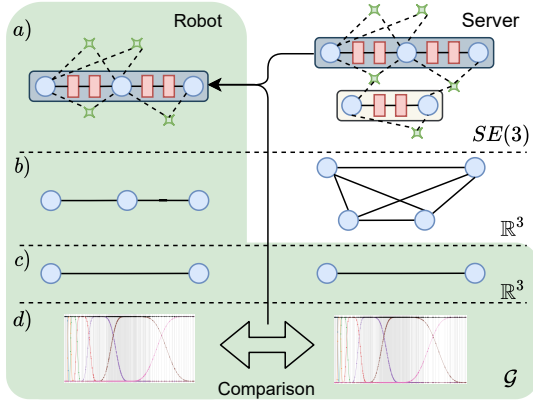
Fig. 3: Different employed graphs: a) Onboard visual-inertial graph, which will be incorporated into the global server map. b) Positional graph proxies including Kron reduction. c) Synchronized graph proxies. d) Multiscale spectral graph analysis. With the results from (d) the system queries poses in $SE(3)$ to generate constraints.

A good introduction and overview of graph signal processing is presented in [22].

Our approach also performs a structural analysis of graph signals to detect discrepancies between the onboard and server graphs. Using localized graph wavelets in the graph domain, our approach directly compares the trajectories at different scales to estimate the severity of the inconsistency of the individual positional graphs.

III. CONSISTENT COLLABORATIVE MULTI-ROBOT MAPPING

This section details the proposed method, for which an overview is presented in Figure 2. Overall, the aim is to identify graphs nodes with high positional drift that will lead to large errors and correct them with only a few constraints. The proposed approach comprises of the following core components: (i) Onboard localization and mapping, (ii) Mapping server at the base station, (iii) Pose graph comparison and correction.

A. Onboard Localization and Mapping

Each robot performs onboard mapping and localization to provide an odometry estimate of its current position. This state estimate is then utilized to build a dense visual-inertial (VI) graph used for global multi-robot mapping on the server. In brief, the odometry estimates are used to triangulate tracked visual features, and throughout the exploration, a factor graph in the form of keyframed [23] submaps is sent to the mapping server, where all submaps are accumulated, merged into a collaborative map, loop closed and globally optimized. Consequently, all the computational-intense multi-robot operations are delegated to the central mapping server while the robots only perform the initial graph building.

B. Centralized Mapping Server

Each submap overlaps with the previous one, allowing the mapping server to readily combine submaps by attaching them, only requiring the handling of potential landmark conflicts. The mapping server holds and continuously operates on a single global multi-robot map. In particular, each iteration includes detection of for intra- and inter-robot loop closures [24] as well as performing a joint optimization on the merged multi-robot map. Once a globally optimized map is available, the mapping server creates, relaxes and broadcasts the global graph. Since for comparison rotational information is not required, the VI graph is relaxed to proxy graphs in $\mathbb{R}^3$ which contains only the positional information.

Global Proxy Graph. The global proxy graph encapsulates the global knowledge of the environment in a compact representation and contains crucial information such as the last known positions of all robots. The global proxy graph is built by defining representative nodes for each incoming submap. The representative nodes can be freely chosen but ought to reflect the robot trajectories to some degree. Moreover, since the global map is continuously optimized, the proxy graph is not immediately built but only a reference to each node is maintained. When an update is triggered after the global optimization, the *graph monitor* (cf. Figure 2) retrieves the latest estimate of the multi-robot pose graph, builds the proxy graph, and sends it to all robots. Although only the proxy graph in $\mathbb{R}^3$ is needed for comparison, auxiliary information is also included for each node in the transmission, such as timestamps and orientations for synchronization and constraint generation purposes.

In general, the proxy graph is a weighted undirected graph $\mathcal{G}_{server} = (\mathcal{E}, \mathcal{V}, w)$ consisting of a set of nodes $\mathcal{V}$, edges $\mathcal{E}$ and weights $w : \mathcal{E} \mapsto \mathbb{R}^+$ denoting how strong two nodes are connected. A radius search ($7\,\text{m}$) is performed around each vertex in the proxy graph and the weight w_D of adjacent nodes is calculated using a squared exponential function where the weight decreases with increasing spatial distance, i.e.

$$w_D(n, m) = \exp\left(-\frac{\|\mathcal{V}_n - \mathcal{V}_m\|_2}{2\sigma^2}\right), \tag{1}$$

with σ being a free parameter. The weighted adjacency matrix $\mathbf{A}$ is then given by $\mathbf{A}_{nm} = w_D(n, m)$.

Consequently, the graph is described by N nodes $n \in \mathbb{R}^3$ and its adjacency matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$ with $\mathbf{A}_{nm} > 0$ if two nodes n and m are connected. Furthermore, the degree matrix $\mathbf{D}$ is a diagonal matrix with entries $\mathbf{D}_{nn} = \sum_m \mathbf{A}_{nm}$ where m are all incident nodes of n. Based on these two matrices $\mathbf{A}$ and $\mathbf{D}$, the graph Laplacian $\mathcal{L}$ can be obtained by

$$\mathcal{L} = \mathbf{D} - \mathbf{A}, \tag{2}$$

which is by definition a symmetric and a positive semidefinite matrix. Hence, it can be decomposed into its eigenvalues and eigenvectors, i.e. $\mathcal{L} = \mathbf{U}\Lambda\mathbf{U}^\top$.

As a final step, $\mathcal{G}_{server}$ is reduced using a Kron reduction [25], when the graph increases to a specific number of nodes, to avoid long computation times. The Kron reduction takes the nodes to keep from the proxy graph as input and reduces them while preserving the spectral properties and the adjacency matrix. The reduced nodes are selected based on

the polarity of the largest eigenvector $u_{\max}$ in $\mathcal{L}$ [26], i.e.

$$\mathcal{V}_{\text{reduced}} = \{n \in \mathcal{V} : u_{\max}(n) \geq 0\}. \qquad (3)$$

Finally, the global proxy graph $\mathcal{G}_{server}$ is broadcasted over the network. Each robot then uses this information to build an onboard graph $\mathcal{G}_{robot}$ using the procedure described above and synchronizes the nodes based on their timestamp (cf. Figure 3c). Next, we perform a graph spectral analysis using $\mathcal{G}_{server}$ and $\mathcal{G}_{robot}$ to identify drifts in the onboard estimation.

C. Spectral Analysis of Graph Signals

The analysis of spectral components of band-limited signals is a well-established and widely used technique in engineering and research. In this paper, we utilize the theory of spectral analysis and signal processing defined on graphs. In contrast to the standard spectral analysis, graphs do not assume any underlying manifold. Therefore, they are well-suited for many robotic applications where graph structures such as pose graphs play a significant role.

The traditional Fourier transform is the inner product of a signal x with a harmonic oscillation i.e.

$$X(\omega) = \langle x(t), \exp(j\omega t)\rangle_t = \int_{-\infty}^{\infty} x(t)\exp(-j\omega t)\, dt, \quad (4)$$

where $X(\omega)$ is the spectrum of $x(t)$ and $\exp(-j\omega t)$ the eigenfunctions of the one-dimensional Laplacian. Analogously, the Graph Fourier transform F of a function f is given by the expansion of f with the eigenfunctions u of the graph Laplacian $\mathcal{L}$ (cf. Eq. (2)), i.e.

$$F(\lambda_l) = \langle f(n), u_l(n)\rangle_n = \sum_n f(n)u_l^*(n), \qquad (5)$$

where λ_l is the l-th non-negative eigenvalue of $\mathcal{L}$ corresponding to u_l and $(\cdot)^*$ denotes a complex conjugation.

The graph Fourier transform of a signal f is then $F = U^\top f$. The eigenvalues Λ are real values, thus can be ordered and correspond to the graph frequencies, allowing a similar intuition as for traditional frequency analysis. Consequently, most of a graph signals energy is preserved in the lower bands of Λ and higher bands correspond to high oscillating frequencies.

Graph Comparison. After the robot has received a global update, a chronological synchronization is performed, yielding a one-to-one mapping of the global graph $\mathcal{G}_{server}$ and onboard graph $\mathcal{G}_{robot}$. Next, the pose information at the nodes of each graph is used to create the functions f and h for server and robot, respectively. In particular, the positional information for each node in the graph is used to compute the relative distance to the origin. Since rotational drifts result in positional structural discrepancies, the inclusion of rotational information is omitted and left for future research.

Generally, wavelets are well-known to be very efficient and flexible for a variety of different tasks in signal processing problems [20]. In traditional wavelet analysis, a signal $x(t)$ is projected onto a scaled (a) and shifted (b) wavelet ψ, i.e.

$$W(a,b) = \langle x(t), \psi_{a,b}(t)\rangle_t = \int_{-\infty}^{\infty} \frac{1}{a}\psi^*\left(\frac{t-b}{a}\right)x(t)\, dt. \qquad (6)$$
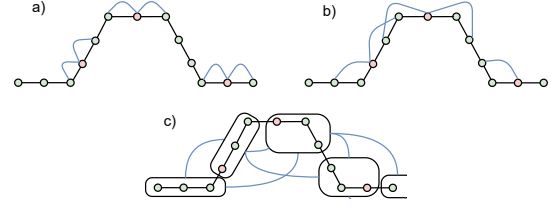

Fig. 4: Illustration of three different relative constraint types. Based on the scale of the structural difference, additional constraints are added to correct (a) adjacent, (b) close neighborhood or (c) submaps.

Using Parseval's theorem, Eq. (6) can also be expressed with the Fourier-transformed signal: $W(a,b) = \langle X(\omega), \Psi_{a,b}(\omega)\rangle_\omega$. Analogously, the graph wavelet transform can be derived using the graph Fourier transform and a wavelet filter kernel on $\mathcal{L}$. For more details, we refer the interested reader to the work of Hammond et. al [19,20].

In this work, the Meyer wavelet using seven scales ($s_{\max} = 7$) is used due to its good localization in the graph and frequency domain. By construction, graph wavelets have the property of being localized on the graph and, therefore, directly relate to its structural properties. The realization of a graph wavelet $\psi_{s,n}$ for a scale s and node n is given by

$$\psi_{s,n} = \mathbf{U}\mathbf{G_s}(\Lambda)\mathbf{U}^\top \delta_n, \qquad (7)$$

where δ_n is a Dirac centered at vertex n and $\mathbf{G_s}$ the wavelet filter bank at scale s. In other words, the filter bank $\mathbf{G_s}$ acts only on the eigenvalues of the graph, i.e. $\mathbf{G_s}(\Lambda) = \text{diag}(g(s\lambda_1), \ldots, g(s\lambda_N))$ and is multiplied with the graph Fourier-transformed Dirac, followed by an inverse transform $\mathbf{U}$. Since $\psi_{s,n}$ lies in the graph domain, we can compute the wavelet coefficients for a graph signal f using $\mathbf{W}_{s,n} = \psi_{s,n}^\top f$.

In summary, the wavelet coefficients up to scale $s_{\max}$ constitute a feature vector that represents multiscale structural information for a node n.

D. Correcting Onboard Estimation

For a server node n with corresponding onboard node n', the scale-wise distance is computed as:

$$d_{n,n'}^s = \|\mathbf{W}_{s,n} - \mathbf{W}_{s,n'}\|_2, \qquad (8)$$

where $\mathbf{W}_{s,n}$ and $\mathbf{W}_{s,n'}$ were computed using $\mathcal{G}_{server}$ and $\mathcal{G}_{robot}$, respectively.

Intuitively, since graph wavelets are localized at a specific node n in the graph [21,27], large scales compress the filter function g leading to a description of the larger neighborhood. In contrast, small scales stretch g and thus yield a description of the closer neighborhood of n. Consequently, three separate cases are distinguished (cf. Figure 4), i.e., a large difference in the lower, mid and higher scales of the coefficients. If $d_{n,n'}^s$ exceeds the threshold for small scales, a relative constraint is added between the direct neighbors of n'. Likewise, for mid-scale differences, a corresponding constraint within a two-node distance of n' is added. In the case of large-scale discrepancies, a constraint between the k-nearest submaps

is added. Since all different types of constraints are relative between nodes, the server and robot graph can be expressed in arbitrary and unknown frames. Additionally, it should be noted that the comparison is based on the proxy graphs, but constraints originate from the optimized poses in $SE(3)$ (cf. Figure 3d).

Furthermore, the existence of already added constraints between the corresponding nodes is checked before the addition of every new constraint to the graph, and constraints are only updated when there is a reasonable difference in translation and rotation. Otherwise, they remain unchanged.

IV. Experiments

We thoroughly evaluate the proposed framework and demonstrate its real-world application using aerial and legged robot datasets. First, we validate our approach and compare its performance to the current state-of-the-art methods using the EuRoC [28] dataset sequences to simulate multi-robot deployments. Next, we demonstrate the real-world performance of our framework during a multi-robot autonomous exploration and mapping mission conducted in an underground tunnel system using ANYmal [29] legged robots. Finally, the localization recovery for an individual agent in case of onboard localization failure is demonstrated during a multi-robot experiment conducted in an environment consisting of indoor and outdoor areas. For all experiments, the root-mean-square of the absolute trajectory error, denoted as RMSE, is used as an evaluation metric. Furthermore, in each experiment, the close neighborhood constraints use a 5-hop distance, and the submap-constraints are added between the four closest submaps. Each robot creates a new submap every $10\,\mathrm{s}$ and sends it to the mapping server.

A. EuRoC Dataset: Validation and Comparison

To validate the proposed approach and compare its performance against the current state-of-the-art collaborative mapping frameworks [2,3,30], Machine Hall (MH) sequences from the EuRoC dataset are used to evaluate single- and multi-robot performance. For each sequence, ROVIO [31] is used to provide monocular visual-inertial odometry for individual aerial agents. First, we evaluate the localization performance by comparing the onboard robot estimates, the collaborative server estimate, and the proposed approach to the provided ground truth; in addition, we also compare with current state-of-the-art approaches with results presented in Table I. It can be noted that despite larger individual onboard error, the proposed framework still attains the lowest collaborative error. Furthermore, correcting the onboard estimation using our multi-scale spectral approach, the lowest single-robot errors are also achieved, demonstrating the proposed approach's effectiveness to correct large onboard estimation errors.

Next, using the experimental setup described in CVI-SLAM [3], it is demonstrated that the proposed approach can
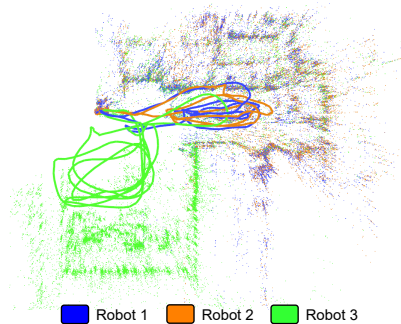

Fig. 5: Collaboratvie VI multi-robot map built by three robots for the EuRoC dataset.

EuRoC Machine Hall - Single and Collaborative

Method / Seq	MH01	MH02	MH03	MH01-03
VINS-mono[1] [32]	0.12 m	0.12 m	0.13 m	0.074 m
ORB-SLAM3[2] [30]	0.062 m	0.037 m	0.046 m	0.037 m
Onboard	0.21 m	0.29 m	0.41 m	**0.025** m
CCM-SLAM[3] [2]	0.061 m	0.081 m	0.048 m	0.077 m
Proposed	**0.029** m	**0.028** m	**0.033** m	**0.025** m

TABLE I: RMSE comparison for the EuRoC dataset. Top part shows the results of single and collaborative approaches while the bottom row shows the individual corrected results.

facilitate accurate pose estimation for individual robots by providing collaborative corrections, as shown in Table II.

EuRoC Machine Hall - Collaborative Corrections

Sequences	CVI-SLAM [3]		Proposed	
	Single	Multi	Single	Multi
MH01 & MH02	0.224 m	0.139 m	0.29 m	**0.030** m
MH02 & MH03	0.295 m	0.256 m	0.41 m	**0.033** m
MH04 & MH05	0.412 m	0.34 m	0.62 m	**0.094** m

TABLE II: Onboard pose RMSE after adding constraints from the centralized server for different dataset combinations.

Finally, to understand the potential for real-world applications, bi-directional network bandwidth utilization between robot and server is analyzed and compared to the network requirements of CVI-SLAM [3]. The results are presented in Figure 6 and show that the robot-to-server network requirements of the proposed approach are comparable to CVI-SLAM. However, the server-to-robot communication requirements are significantly reduced as only a sparse relaxation of the dense server graph is sent to the robots.

B. Large-Scale Multi-Robot Subterranean Exploration

To demonstrate the suitability towards complex real-world applications, the proposed approach was utilized during an autonomous multi-robot exploration [33] and mapping mission conducted at Hagerbach underground facility in Switzerland. Two ANYmal quadrupedal robots were deployed during an hour long mission and autonomously navigated distances of $1.21\,\mathrm{km}$ and $1.1\,\mathrm{km}$ respectively. Each robot is

---

[1]Single robot results from [32]. Collaborative result from [2].

[2]Monocular visual-inertial results from [30].
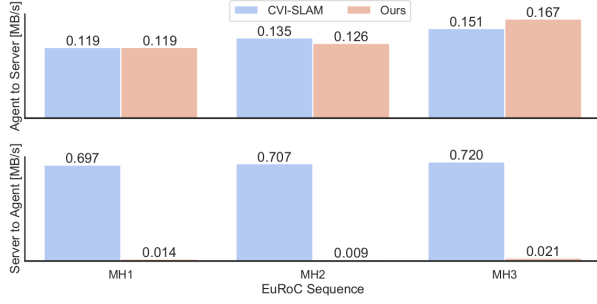
[3]As reported in [2].

Fig. 6: Average network usage for bi-directional communication between robot and server on the EuRoC dataset.

equipped with a Velodyne VLP-16 LiDAR and a Sevensense Alphasense visual-inertial sensor. Onboard robot odometry estimation and mapping is performed by compSLAM [34] and, along with required visual and pointcloud data, sent to the mapping server whenever the robots are within the communication range.

A ground-truth map was created using a Leica RTC360 scanner to evaluate the proposed collaborative mapping framework's performance and quantify the effect of integration of collaborative corrections on individual robot pose accuracy. Ground-truth robot poses were then computed by registering individual robot pointclouds against the ground-truth map following the approach of [35]. The collaborative server and individual robots maps are shown in Figure 1, with quantitative results presented in Table III, demonstrating that the collaborative mapping approach achieves a lower error than either of the individual robot estimates. In addition, benefiting from the feedback constraints, the onboard estimated pose error is significantly reduced. Furthermore, the proposed approach adds the lowest number of additional constraints than the baseline approach of adding a correction to each node of the robot graph or for all nodes corresponding to the proxy graph while achieving comparable accuracy improvement without significantly increasing computational cost. Finally, the robot-server network usage statistics for the mission are presented in Table IV, showing overall low communication bandwidth requirements and demonstrating the feasibility of the proposed approach towards real-world applications. Additionally, by reducing the graph using the Kron reduction, we can further decrease the bandwidth by more than 40%.

Underground Tunnel - Ground Truth Evaluation				
Method	**RMSE**	**Time**	**# Factors**	**Server**
ANYmal 1	1.15 m	68.6 ms	29152	**0.25** m
Baseline	0.51 m	187.2 ms	31736 (+2584)	
Proposed	**0.46** m	71.2 ms	29898 (+746)	
ANYmal 2	0.97 m	53.7 ms	25435	**0.59** m
Baseline	0.72 m	152.6 ms	27815 (+2380)	
Proposed	**0.67** m	59.8 ms	26067 (+632)	

TABLE III: RMSE comparison for the original onboard, server and corrected onboard graph. The numbers in parenthesis denote the improvement and total amount of additional constraints, respectively.

Underground Tunnel - Average Bandwidth				
	Without Reduction		With Reduction	
Robot	**Rob→Srv**	**Srv→Rob**	**Nodes**	**Srv→Rob**
ANYmal 1	0.32 MBps	0.044 MBps	1295→691	0.025 MBps
ANYmal 2	0.30 MBps	0.044 MBps	1191→444	0.025 MBps

TABLE IV: Average bandwidth usage for the multi-robot underground mission. Our approach maintains low bandwidth requirements with the server-to-robot communication being identical as the same global graph is transmitted to all robots.

Indoor/Outdoor Dataset - Ground Truth Evaluation				
Method	**RMSE**	**Time**	**# Factors**	**Server**
ANYmal 1	2.22 m	5.2 ms	3317	**0.21** m
Proposed	**0.30** m	5.9 ms	3339 (+22)	
ANYmal 2	0.25 m	3.4 ms	2239	**0.14** m
Proposed	**0.16** m	4.1 ms	2253 (+14)	

TABLE V: Comparison of the RMSE of the onboard estimation before and after the supplying additional constraints.
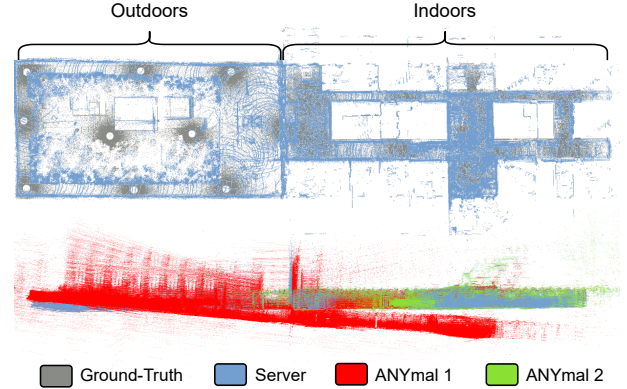

Fig. 7: Mapping results for the indoor/outdoor multi-robot experiment. The top compares the collaborative map to the ground-truth while the bottom row shows a sideview comparison of the robot and server maps. The robot map of ANYmal 1 is misaligned due to onboard localization failure.

C. Localization Recovery

We conducted an experiment in a particularly challenging environment for LiDAR localization due to the absence of surrounding geometric structure to demonstrate the utility of collaborative mapping towards localization recovery for an individual robot in case of an onboard estimation failure. Two ANYmal robots, equipped with the same sensory payload as described in subsection IV-B, were simultaneously deployed in an indoor office environment connected to an outdoor rooftop terrace. The first robot performs a loop indoors while the second robot transitions outdoors through a narrow doorway, navigates a rectangular path, and returns indoors. Due to the absence of surrounding structure outdoors, the onboard localization drifts significantly, skewing the onboard robot map. Nevertheless, the collaborative mapping approach is able to generate a consistent map of the environment,

as shown in Figure 7, due to its loop closure capabilities. Furthermore, the integration of collaborative constraints enables localization recovery for the individual robot leading to a significant reduction in its pose error, as shown in Table V, when compared to ground-truth robot trajectory, generated as described in subsection IV-B.

V. Conclusions

This paper presented a novel framework for creating globally consistent estimates between multiple robots and a centralized mapping server. A graph-based spectral analysis of the robot and server graphs is proposed to identify the underlying structural differences in positional graphs. By adding constraints on drifting nodes efficiently, low additional computation and communication resources are achieved. The presented results using large-scale multi-robot field deployments in challenging environments demonstrate the real-world potential of the proposed approach for both research and industry.

References

[1] I. Deutsch, M. Liu, and R. Siegwart, "A framework for multi-robot pose graph SLAM," *2016 IEEE International Conference on Real-Time Computing and Robotics, RCAR 2016*, pp. 567–572, 2016.

[2] P. Schmuck and M. Chli, "CCM-SLAM: Robust and efficient centralized collaborative monocular simultaneous localization and mapping for robotic teams," *Journal of Field Robotics*, vol. 36, no. 4, pp. 763–781, 6 2019.

[3] M. Karrer, P. Schmuck, and M. Chli, "CVI-SLAM—Collaborative Visual-Inertial SLAM," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 2762–2769, 10 2018.

[4] A. Cunningham, V. Indelman, and F. Dellaert, "DDF-SAM 2.0: Consistent distributed smoothing and mapping," in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 5 2013, pp. 5220–5227.

[5] J. Dong, E. Nelson, V. Indelman, N. Michael, and F. Dellaert, "Distributed real-time cooperative localization and mapping using an uncertainty-aware expectation maximization approach," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, vol. 2015-June, no. June. IEEE, 5 2015, pp. 5807–5814.

[6] D. Zou and P. Tan, "CoSLAM: Collaborative Visual SLAM in Dynamic Environments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 2, pp. 354–366, 2 2013.

[7] K. Ebadi, Y. Chang, M. Palieri, A. Stephens, A. Hatteland, E. Heiden, A. Thakur, N. Funabiki, B. Morrell, S. Wood, L. Carlone, and A. A. Agha-Mohammadi, "LAMP: Large-Scale Autonomous Mapping and Positioning for Exploration of Perceptually-Degraded Subterranean Environments," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 80–86, 3 2020.

[8] B. Kim, M. Kaess, L. Fletcher, J. Leonard, A. Bachrach, N. Roy, and S. Teller, "Multiple relative pose graphs for robust cooperative mapping," in *2010 IEEE International Conference on Robotics and Automation*. IEEE, 5 2010, pp. 3185–3192.

[9] J. G. Mangelson, D. Dominic, R. M. Eustice, and R. Vasudevan, "Pairwise Consistent Measurement Set Maximization for Robust Multi-Robot Map Merging," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 2916–2923, 2018.

[10] J. G. Mangelson, M. Ghaffari, R. Vasudevan, and R. M. Eustice, "Characterizing the Uncertainty of Jointly Distributed Poses in the Lie Algebra," 6 2019.

[11] P. Y. Lajoie, B. Ramtoula, Y. Chang, L. Carlone, and G. Beltrame, "DOOR-SLAM: Distributed, Online, and Outlier Resilient SLAM for Robotic Teams," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1656–1663, 4 2020.

[12] E. Brunskill, T. Kollar, and N. Roy, "Topological mapping using spectral clustering and classification," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 10 2007, pp. 3491–3496.

[13] V. Indelman, "Towards Cooperative Multi-robot Belief Space Planning in Unknown Environments," in *Springer Proceedings in Advanced Robotics*. IEEE, 2018, vol. 2, pp. 441–457.

[14] R. Chandra, T. Guan, S. Panuganti, T. Mittal, U. Bhattacharya, A. Bera, and D. Manocha, "Forecasting Trajectory and Behavior of Road-Agents Using Spectral Clustering in Graph-LSTMs," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4882–4890, 7 2020.

[15] J. Moon and B.-H. Lee, "Object-oriented Semantic Graph Based Natural Question Generation," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 5 2020, pp. 4892–4898.

[16] G. Mateos, S. Segarra, A. G. Marques, and A. Ribeiro, "Connecting the Dots: Identifying Network Structure via Graph Signal Processing," *IEEE Signal Processing Magazine*, vol. 36, no. 3, pp. 16–43, 5 2019.

[17] H. E. Egilmez and A. Ortega, "Spectral anomaly detection using graph-based filtering for wireless sensor networks," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, no. ii. IEEE, 5 2014, pp. 1085–1089.

[18] A. Sandryhaila and J. M. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3042–3054, 2014.

[19] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129–150, 3 2011.

[20] ——, "The Spectral Graph Wavelet Transform: Fundamental Theory and Fast Computation," in *Signals and Communication Technology*, 2019, pp. 141–175.

[21] C. Donnat, M. Zitnik, D. Hallac, and J. Leskovec, "Learning Structural Node Embeddings via Diffusion Wavelets," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York, NY, USA: ACM, 7 2018, pp. 1320–1329.

[22] A. Ortega, P. Frossard, J. Kovacevic, J. M. F. Moura, and P. Vandergheynst, "Graph Signal Processing: Overview, Challenges, and Applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 5 2018.

[23] M. Dymczyk, S. Lynen, M. Bosse, and R. Siegwart, "Keep it brief: Scalable creation of compressed localization maps," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 2015-Decem, no. September. IEEE, 9 2015, pp. 2536–2542.

[24] S. Lynen, T. Sattler, M. Bosse, J. Hesch, M. Pollefeys, and R. Siegwart, "Get Out of My Lab: Large-scale, Real-Time Visual-Inertial Localization," in *Robotics: Science and Systems XI*, vol. 11, no. November. Robotics: Science and Systems Foundation, 7 2015.

[25] F. Dorfler and F. Bullo, "Kron Reduction of Graphs With Applications to Electrical Networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 150–163, 1 2013.

[26] D. I. Shuman, M. J. Faraji, and P. Vandergheynst, "A Multiscale Pyramid Transform for Graph Signals," *IEEE Transactions on Signal Processing*, vol. 64, no. 8, pp. 2119–2134, 2016.

[27] N. Tremblay and P. Borgnat, "Graph wavelets for multiscale community mining," *IEEE Transactions on Signal Processing*, vol. 62, no. 20, pp. 5227–5239, 2014.

[28] M. Burri, J. Nikolic, P. Gohl, T. Schneider, J. Rehder, S. Omari, M. W. Achtelik, and R. Siegwart, "The EuRoC micro aerial vehicle datasets," *The International Journal of Robotics Research*, vol. 35, no. 10, pp. 1157–1163, 9 2016.

[29] M. Hutter, C. Gehring, A. Lauber, F. Gunther, C. D. Bellicoso, V. Tsounis, P. Fankhauser, R. Diethelm, S. Bachmann, M. Blösch, *et al.*, "Anymal-toward legged robots for harsh environments," *Advanced Robotics*, vol. 31, no. 17, pp. 918–931, 2017.

[30] C. Campos, R. Elvira, J. J. G. Rodriguez, J. M. M. Montiel, and J. D. Tardos, "ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual–Inertial, and Multimap SLAM," *IEEE Transactions on Robotics*, pp. 1–17, 2021.

[31] M. Bloesch, M. Burri, S. Omari, M. Hutter, and R. Siegwart, "Iterated extended Kalman filter based visual-inertial odometry using direct photometric feedback," *The International Journal of Robotics Research*, vol. 36, no. 10, pp. 1053–1072, 9 2017.

[32] 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, 8 2018.

[33] T. Dang, M. Tranzatto, S. Khattak, F. Mascarich, K. Alexis, and M. Hutter, "Graph-based subterranean exploration path planning using aerial and legged robots," *Journal of Field Robotics*, 2020.

[34] S. Khattak, H. Nguyen, F. Mascarich, T. Dang, and K. Alexis, "Complementary Multi–Modal Sensor Fusion for Resilient Robot Pose

Estimation in Subterranean Environments," in *2020 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 9 2020, pp. 1024–1029.

[35] 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 *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020.