Cooperative LiDAR Localization and Mapping for V2X Connected Autonomous Vehicles

Bingyi Cao, Claas-Norman Ritter, Khaled Alomari and Daniel Goehring

***Abstract*— Cooperative Simultaneous Localization and Mapping (C-SLAM) is an active research topic in mobile robotics. However, its application in the field of autonomous driving is rare. While the advent of Vehicle-to-Everything (V2X) communication has empowered Connected Autonomous Vehicles (CAV) to exchange data with each other, recent research on CAV cooperation tasks has primarily focused on cooperative perception and global positioning improvement. Techniques for organizing multiple CAV to work together to achieve localization and mapping in unknown environments have not been actively explored. We propose a C-SLAM system for CAVs that employs sparse LiDAR feature representations to enable vehicles to exchange data using standard V2X messages. The system was tested in real environments using two connected vehicles. The results show that the proposed V2X-based C-SLAM system can operate in both centralized and decentralized manners and output accurate pose estimates and global maps, showing promising application possibilities.**

I. INTRODUCTION

Simultaneous Localization and Mapping (SLAM) is an essential technique in robotics. In particular, compared with vision-based methods, laser-based SLAM is widely adopted in areas such as autonomous driving due to its stronger resistance to environmental disturbances and higher measurement accuracy. Multiple robots cooperating with each other can complete localization and mapping tasks faster and more accurately than a single robot, which is particularly advantageous in large-scale environments. The problem to be solved by C-SLAM is to coordinate a group of robots and use all the data to build a consistent global map and, at the same time, localize themselves on the map. It has better robustness than single-robot methods. The failure of a single robot will not affect the completion of the entire task. However, the cost of obtaining these benefits is the need to deal with problems such as communication, coordination, and information fusion.

The V2X technique endows autonomous vehicles with the ability to communicate with other vehicles and smart roadside facilities, making it possible to share data and collaborate among networked nodes. It was developed to enable reliable communication between CAVs while driving at high speeds in urban or highway traffic. Extensive research on the cooperative perception and vehicle global positioning of connected vehicles through V2X communication has recently emerged. Nevertheless, the exploration of V2X-based Cooperative SLAM has remained relatively scarce.

The authors are with the Department of Mathematics and Computer Science, Freie Universität Berlin, Berlin, Germany. This work was supported by the German Federal Ministry for Digital and Transport under grant number 45AVF3001F.

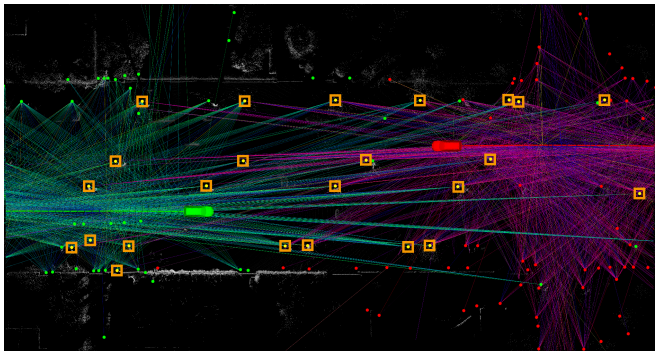

Fig. 1. Demonstration of the proposed V2X-based C-SLAM approach running in a centralized manner. The green car performs localization and mapping while using the received perception data to estimate the state of the red car. The green dots and cyan lines represent the map points of the green car and its measurements of the visible map points at each frame. The red dots and purple lines are the map points and measurements of the red car. The yellow dots highlighted by the orange rectangles represent features shared by both vehicles, which provide constraints between the two systems.

This paper presents a novel framework for cooperative SLAM for Connected Autonomous Vehicles. Our proposed framework offers two distinct operating modes. The first mode employs V2X communication to transmit perception data among a group of vehicles, which enables the creation of a global map and trajectories of all participating vehicles on one of the vehicles. This mode operates in a centralized manner. The second mode enables vehicles to exchange local maps rather than perception data. Each vehicle estimates its poses, creates its local map and merges it with the received maps. This mode operates in a decentralized fashion. The main contributions of our work include the following:

- We investigated the feasibility of implementing V2X-based C-SLAM in both centralized and decentralized manners.
- We developed a framework utilizing sparse representations of map and perception information that can operate efficiently in V2X networks with limited communication capabilities.
- We conducted real-world tests using two V2X connected vehicles and obtained highly accurate localization and mapping results.

These contributions represent a significant step forward in exploring cooperative SLAM for connected vehicles based on V2X communication. The remaining paper is arranged as follows: Section II lists the related research of C-SLAM and the application of V2X in the field of Intelligent Transportation Systems (ITS). Section III details the proposed

method. Sections IV and V present the experimental results and summarise the work.

II. RELATED WORK

Due to its broad application scenarios, cooperative SLAM is an active research topic in mobile robotics. Depending on where the data obtained by the robots are processed, C-SLAM systems can generally be divided into two categories: Centralized methods [1], [2], [3], in which all information is brought together in a powerful server or management robot, and decentralized [4], [5], [6] approach, where all robots communicate with each other, and the computations are performed locally on each node.

Like in single-robot SLAM systems, filtering methods were widely used to solve C-SLAM problems. Early C-SLAM researchers tried to extend various filtering-based single-robot SLAM methods to multi-robot environments. Zhou et al. [7] used the EKF method to fuse robot poses and landmark positions. This method required the robots in the team to identify each other and measure their relative poses in order to merge the global map. Particle filtering methods [8], [9], [10] were also proposed to solve grid-map-based or feature-based C-SLAM problems. In addition, the application of information filtering methods in C-SLAM has been examined [11].

The application of graph-optimization-based SLAM in multi-robot environments has also been extensively studied. Graph optimization searches for the maximum posterior estimate of the system. Cunningham et al. [12] introduced a distributed C-SLAM solution based on a Constrained Factor Graph. The simulated test showed that a consistent feature map could be produced. Dubé et al. [13] adopted iSAM2 [14] and implemented a centralized multi-robot SLAM for 3D LiDAR, having the sub-map association done using SegMatch [15]. Indelman et al. [16] presented a pose graph-based multi-robot localization solution. High robustness for outliers was shown through real data. DCL-SLAM [17] proposed a distributed framework that leverages existing single-robot LiDAR-based SLAM methods to enable collaborative localization of multiple robots. Swarm-SLAM [18] also provided an open-source, decentralized multi-robot SLAM framework capable of supporting different types of sensors, such as LiDAR, stereo, and RGB-D cameras.

Communication is the basis for multiple robots to share data. Processing delays, message losses, and other communication issues are essential to C-SLAM. Extensive studies have focused on the C-SLAM problem in cases of limited communication [19], [20], [21]. In the field of Intelligent Transportation Systems, the application of V2X technology makes intelligent vehicles no longer isolated individuals; instead, they can communicate with the roadside, cloud, and other vehicles, enabling them to complete more complex collaborative tasks while driving normally in urban and non-urban traffic like on highways.

In Europe the ITS-G5 standard [22] is used for V2X communication. ITS-G5 is based on IEEE 802.11p also known as WAVE which is a modification of the IEEE 802.11a wireless standard. It has been specifically designed to tackle communication issues at higher speeds. Lin et al. [23] compared IEEE 802.11a with its modification for V2X showing the advantages of IEEE 802.11p in real-world scenarios and usual urban traffic speeds.

Recent research on V2X-based cooperative tasks has focused more on the area of cooperative perception. Researchers [24], [25] fused LiDAR information from different vehicles to obtain more comprehensive surrounding object detection. Xu et al. [26] integrated information from roadside units to complement the vehicle's environmental understanding. Where2comm [27] explored new methods to improve communication efficiency in cooperative perception. OPV2V [28] provided an open benchmark dataset for V2V-based cooperative perception. Recently, real-world cooperative datasets [29], [30] have been released to advance related research. Unlike the above topics, research about cooperative SLAM using V2X communication is rare in the literature.

III. METHOD

A. V2X Communication

Our work explores a solution that shares LiDAR-detected landmarks, vehicle poses, and local maps among connected vehicles only via the V2X network using Collective Perception Messages (CPM) [31] and Cooperative Awareness Messages (CAM) [32], enabling multiple vehicles to work cooperatively to explore the surrounding environment. All messages are sent with the ITS-G5 default channel bandwidth of 10 Hz using QPSK (Quadrature Phase Shift Keying) modulation and 1/2 FEC (Forward Error Correction) rate, providing theoretically up to 6 Mib/s available data rate, shared between all communicating vehicles. We utilize channel 180 and limit the number of features transmitted in each CPM to 50 to stay below the maximum V2X message size of 1400 bytes [33]. To be able to communicate our perception data and maps, we extended the CPM ASN.1 definition [34] as follows (extensions in bold):

```
OtherSublassType ::= INTEGER {
unknown(0), roadSideUnit(1), pole(2),
wallSegment(3), wallEdge(4)} (0..255)
```

B. System Overview

Fig. 2 shows the structure of the proposed C-SLAM system. We take data from the LiDAR sensor and wheel odometry as input. ${}^{\alpha}P_k$ represents the input of the LiDAR frame k of robot α. ${}^{\alpha}\check{T}_k^{k-1}$ is the prior motion transformation given by wheel odometry. The *Detection* module obtains the features of interest ${}^{\alpha}F_k$ from the input and passes it to the *Odometry and Mapping* module. Integrated with the prior values, the LiDAR *Odometry* estimates the robot's motion transformation ${}^{\alpha}T_k^{k-1}$. The *Mapping* module obtains the robot's pose ${}^{\alpha}T_k^W$ in the map coordinate system through local map optimization. Then the system outputs the optimized vehicle poses at 100 Hz and builds the feature map. The *Cooperation* module exchanges the current local map ${}^{\alpha}M_k$, perception ${}^{\alpha}F_k$ and pose ${}^{\alpha}T_k^W$ with other vehicles through the V2X network via CAMs and CPMs. Information from

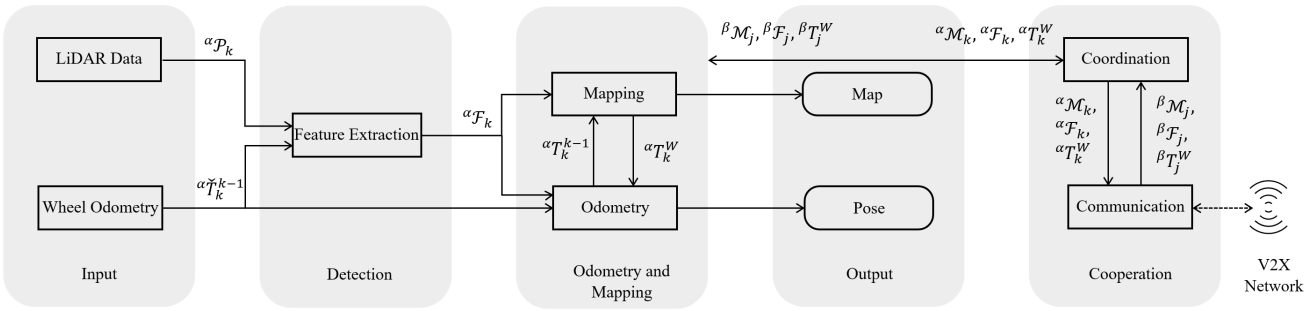

Fig. 2. Schematic representation of the Cooperative SLAM system. A vehicle uses its own sensor information to achieve localization and mapping while exchanging location, perception, and map information with other vehicles through the V2X network.

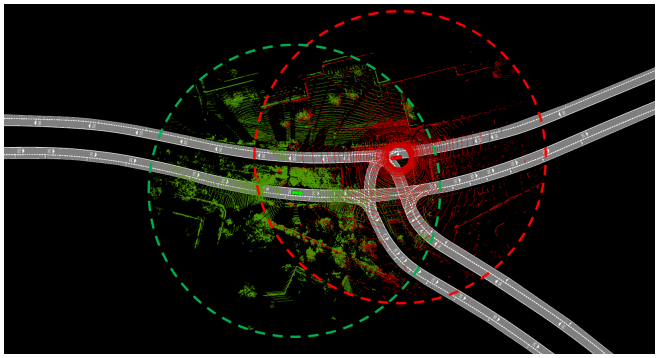

Fig. 3. In the decentralized operation, when two vehicles are within sensor perception range of each other, they start sending local maps to the V2X network.

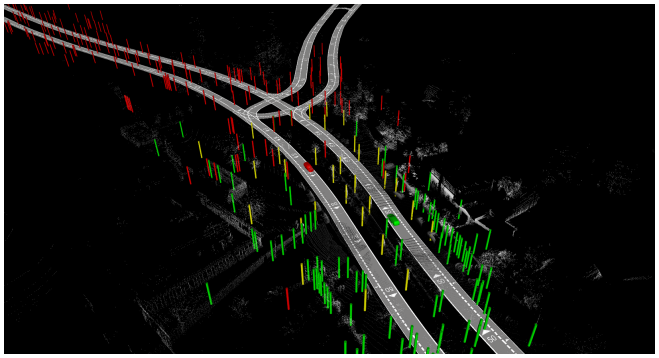

Fig. 4. Successfully merged maps, different colors indicate that they are from different vehicles.

other vehicles is sent to the *Odometry and Mapping* module, which merges with and updates the current map.

C. Cooperation

We have designed two inter-vehicle cooperation strategies to support the proposed C-SLAM system operating in two different modes.

1) Perception Data Exchange: In this setup, the vehicles participating in the C-SLAM task need to keep within communication range. That is, they can *hear* each other. Each vehicle processes its sensor information in real-time, building its trajectory and a local map. Then the features extracted from each LiDAR frame and the current pose are sent over the V2X network as CPMs at 10 Hz, the same rate as the LiDAR measurements. The cooperation module listens to the perception information from other vehicles. With received perception data, the system estimates the driving path and builds a local map for each vehicle it can *hear*.

The map merging algorithm will be triggered when these local maps have overlapping areas. Inter-robot data association can be achieved by matching these two sets of overlapped features. Then the global optimization will be performed to obtain a consistent map. From a data fusion perspective, each vehicle operates in a (redundant) centralized fashion in this setup, computing trajectories and a global map for all vehicles. We can leave one vehicle for centralized data processing, while other vehicles are only responsible for sending data.

2) Local Map Exchange: Each vehicle runs the SLAM system independently in this operation mode. They broadcast their positions over the V2X network at a low frequency of 1 Hz, through CAMs. When two vehicles are within their communication range and can *hear* each other, they calculate and track the distance between them by processing the received messages.

When two vehicles are close enough that their sensor detection ranges overlap, we say they can *see* each other. In this state, the *Cooperation* module triggers the map transfer action. The local map of the ego vehicle is encoded into a series of CPMs and sent over the V2X network at 20 Hz. It simultaneously listens to the network and receives map messages from vehicles that it can *see*, as illustrated in Fig. 3.

When the ego vehicle receives a map from another car, it will try to merge the received map with its local map. As shown in Fig. 4, the ego vehicle (shown in green) receives a map from another vehicle (the second vehicle and its map are shown in red) through the V2X network and merges it into its local map (shown in green), the yellow poles represent the shared features. As a result, the vehicle acquired map data of the unexplored region and can continue driving, achieving self-localization on the expanded map.

In this setup, the vehicles participating in the task operate

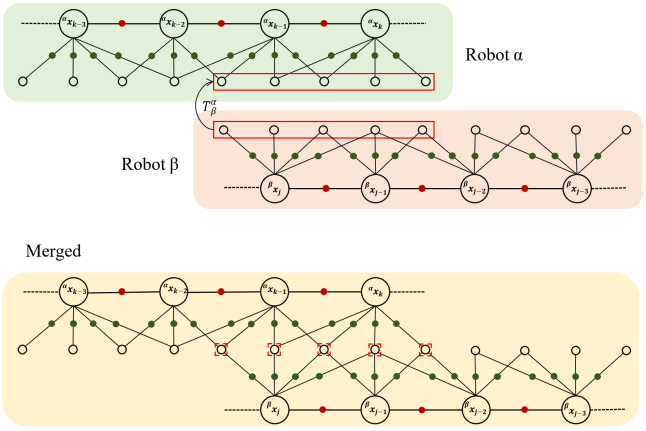

Fig. 5. Factor graph representation of the proposed C-SLAM system. Two sets of features (in red boxes) from different robots are matched to connect the two systems.

in a decentralized fashion. Each vehicle independently runs a SLAM system to compute its state and map and then share the local map with other vehicles through V2X communication.

D. Problem Formulation

The *Feature Detection* module extracts the pole-like structures from the LiDAR range image. In real environments, these are usually tree trunks, poles of street lights, or traffic signs. Such features are common in urban road environments and are therefore frequently adopted in localization and mapping-related research [35], [36]. More importantly, pole-like structures are an essential element in high-definition city maps. Therefore, pole-based SLAM systems have potential applications in the automated updating and expansion of high-definition maps. We adopt the feature detection method described in [37]. The extracted poles are projected onto the ground, and their geometric centers are used as measurement points. Additionally, the wheel odometry is taken to describe the motion of the robot. With these two types of information, we can model the motion and measurement error as follows:

$$\begin{aligned} e_{u,rk} &= f(x_{rk-1}, u_{rk}) - x_{rk} \\ e_{z,rki} &= h(x_{rk}, m_{ri}) - z_{rki} \end{aligned} \tag{1}$$

where f is the motion function and h is the measurement function. x_{rk} represents the pose of the robot r at time k, m_{ri} is feature i in the map of robot r, and z_{rki} represents the measurement of feature i at time k of robot r. Then we can formulate the multi-robot SLAM problem in both centralized and decentralized forms.

1) Centralized C-SLAM: We use $X = \{x_{rk}\}_{r=1,k=1}^{R,K_r}$ to represent the poses of all robots. All involved features are represented by $M = \{m_{ri}\}_{r=1,i=1}^{R,I_r}$. $U = \{u_{rk}\}_{r=1,k=1}^{R,K_r}$ are the input values of all robot odometry data and $Z = \{z_{rki}\}_{r=1,k=1,i=1}^{R,K_r,I_r}$ are all measurement values. Assuming that motion and observation noises are independent at all time frames, we can define the C-SLAM problem as a MAP (Maximum A Posteriori) estimation:

$$\begin{aligned} X^*, M^* &= \underset{X,M}{\operatorname{argmax}} P(X,M|U,Z) = \underset{X,M}{\operatorname{argmax}} P(X|U)P(Z|X,M) \\ &= \underset{X,M}{\operatorname{argmax}} \underbrace{\prod_{r,k} P(x_{rk}|x_{rk-1},u_{rk})}_{odometry} \underbrace{\prod_{r,k,i} P(z_{rki}|x_{rk},m_{ri})}_{measurement} \end{aligned} \tag{2}$$

Note that (2) does not explicitly show how constraints between robots are handled. Fig. 5 shows the C-SLAM problem represented as factor graphs. Large circles represent robot states, and small circles the features on the map. The red and green dots represent the odometry data and the measurements of features, respectively. In the upper part of Fig. 5, two groups of feature points from robot α and robot β are matched through inter-robot data association, as highlighted in red rectangles. Once the data association succeed, the factor graph of robot α and robot β has been merged together as shown in the lower part of the figure. To write the merged factor graph in a mathematical form we can take the negative logarithm of (2) and substitute (1) into it. We then get:

$$\begin{aligned} X^*, M^* &= \underset{X,M}{\operatorname{argmin}}(-\log P(X,M|U,Z)) \\ &= \underset{X,M}{\operatorname{argmin}} \underbrace{\sum_{r=1}^{R}\sum_{k=1}^{K_r} e_{u,rk}^T \Sigma_{u,rk}^{-1} e_{u,rk}}_{odometry} \\ &+ \underbrace{\sum_{r=1}^{R}\sum_{k=1}^{K_r}\sum_{m_i \notin S} e_{z,rki}^T \Sigma_{z,rki}^{-1} e_{z,rki}}_{non-shared\,features} \\ &+ \underbrace{\sum_{r=1}^{R}\sum_{k=1}^{K_r}\sum_{m_i \in S} e_{z,rki}^T \Sigma_{z,rki}^{-1} e_{z,rki}}_{shared\,features} \end{aligned} \tag{3}$$

(3) shows that the error of the whole system can be divided into three parts: the error of the odometry data from all robots, the observation error of the features that can only be observed by a single robot, and the observation error of the features that are shared by multiple robots. S is the set of shared features and Σ^{-1} is the inverse of the covariance of the errors. (3) can be solved using non-linear solvers such as Gauss-Newton and Levenberg-Marquardt methods.

2) Decentralized C-SLAM: In a decentralized system, each robot only estimates its state and builds its local map independently before receiving maps from other robots. In this step, each robot only processes its own odometry and sensor measurements. Therefore, the estimation of the local poses and the local map of robot α can be described as a single-robot SLAM problem:

$$X_\alpha^*, M_\alpha^* = \underset{X_\alpha, M_\alpha}{\operatorname{argmin}} \sum_{k=1}^{K_\alpha} e_{u,k}^T \Sigma_{u,k}^{-1} e_{u,k} + \sum_{k=1}^{K_\alpha}\sum_{i=1}^{I_\alpha} e_{z,ki}^T \Sigma_{z,ki}^{-1} e_{z,ki} \tag{4}$$

After receiving a local map from another robot, the current robot α will try to find matches between the received

Fig. 6. Autonomous vehicles used to test the proposed method. Both vehicles are equipped with 360° LiDAR, Applanix POS LV position and orientation system, and V2X communication devices.

map and its own local map. Map merging will only be performed after a group of feature matches is confirmed. We estimate the transformation between two maps with objective functions:

$$T_\beta^{\alpha *} = \underset{T_\beta^\alpha}{\operatorname{argmin}} \sum_{(m_{\alpha i}, m_{\beta i}) \in S_{\alpha\beta}} \| m_{\alpha i} - T_\beta^\alpha m_{\beta i} \|_2^2 \tag{5}$$

T_β^α is the transformation from the map coordinates of robot α to the map coordinates of robot β and $(m_{\alpha_i}, m_{\beta_i})$ is a pair of matched features in the inter-robot data association set $S_{\alpha\beta}$. After obtaining the optimal estimation of T_β^α, we update the current map α using the following function:

$$M_\alpha^* \leftarrow M_\alpha^* \cup T_\beta^{\alpha *} (M_\beta^* \setminus S_{\alpha\beta}) \tag{6}$$

The new feature points in map β are transformed and added to the current map α. To obtain an extended map, the data association and map merging process needs to be executed for each received map. In case an inter-robot feature match cannot be found, the received map will be discarded. After the map is merged, robot α continues to localize itself on the expanded map, updates it with new sensor data, and waits for the next map transmitted from the V2X network.

E. Inter-Robot Data Association

To associate the local maps of two robots when they overlap, a commonly used approach is to identify the objects' nearest neighbors on the targeted map. However, this method can easily fail due to system drift over time. Therefore, we adopt a more robust point pattern matching method, which has been widely studied in the fields of fingerprint recognition [38], place recognition, and image matching. Specifically, we use the RANSAC-based method described in [37]. Alternative methods, such as the one proposed in [39] and point registration methods like ICP [40], [41] and its variants, are also available. After successful matching, the matched feature pairs are assigned the same ID. In the centralized scenario, the local factor graphs of the two robots

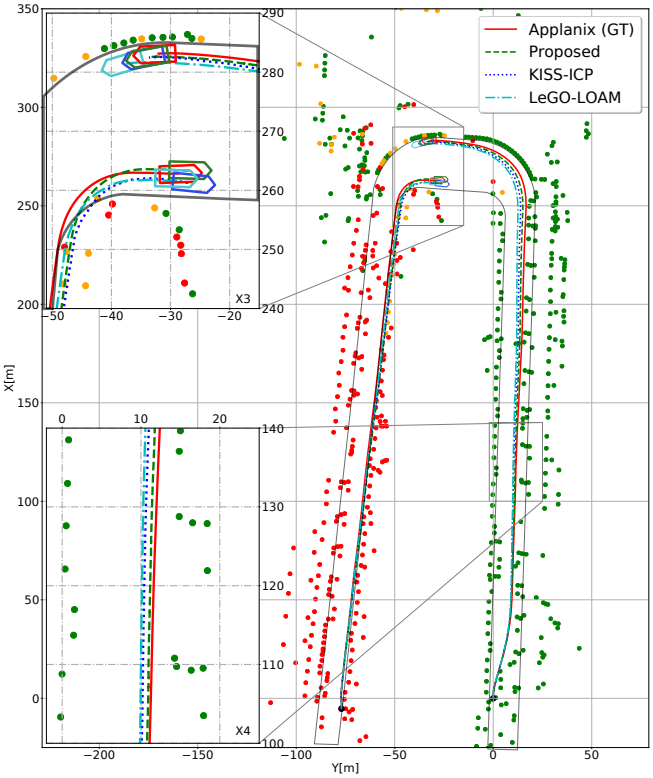

Fig. 7. Localization and mapping results around the Olympic Place in front of Olympiastadion in Berlin. The trajectories generated by the proposed method were compared with GNSS/INS system, KISS-ICP, and LeGO-LOAM methods.

can be connected together, as shown in Fig. 5, and the optimal state of the system can be estimated through non-linear optimization. In the case of distributed execution, an optimal transformation between the two groups of points can be calculated using (5), and the received map is then added to the current robot's local map after transformation.

IV. Experiments

In this section, we present two experiments in the real-world environment. We run the proposed algorithm on two test vehicles. One is equipped with a Velodyne Alpha Prime 128-beam-LiDAR, and the other has a Velodyne HDL-64E S2 64-beam-LiDAR. Both vehicles have an Applanix POS-LV 510 GNSS/INS navigation system with RTK (Real-Time Kinematic) positioning technique and a Cohda MK5 OBU for V2X communication. The algorithm runs in real-time on two laptops equipped with Intel Core i9 CPUs. We verified the feasibility of V2X-based cooperative SLAM and evaluated the accuracy of the merged map. In addition, we compared the localization results with a single-robot laser SLAM method, LeGO-LOAM [42], and a recently proposed ICP method, KISS-ICP [43]. Each vehicle's initial position is estimated by utilizing the on-board GPS.

A. Experiment 1: Centralized Operation

In the first experiment, we performed collaborative mapping of an outdoor parking lot using the two test vehicles

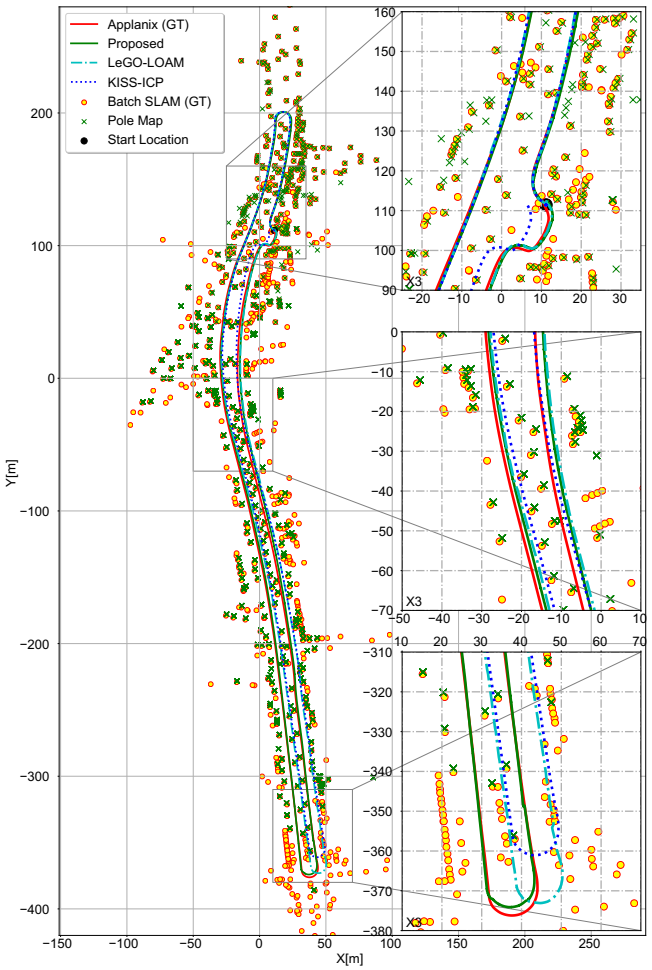

Fig. 8. Comparison of localization and mapping results of the proposed C-SLAM method with ground truth and state-of-the-art methods.

(hereafter referred to as Car 1 and Car 2). Two vehicles drove in parallel from opposite sides of the parking lot, exchanging data simultaneously. For each car, in addition to processing its own sensor information, it had also to process data from the other car: it estimated the poses and feature maps for both cars at the same time. Once the two maps had sufficient overlap, the map merging process was initiated, followed by a global optimization of the merged factor graph to estimate the optimal state of the entire system.

Fig. 7 shows the mapping results, where the green and red dots represent the map points independently observed by Car 1 and Car 2, and the orange dots represent the shared map points. The transformation between the two systems can be calculated through these shared features. We compare the localization results of the two vehicles produced by Car 1 with LeGO-LOAM and KISS-ICP. As shown in the zoomed-in area, the proposed method exhibits similar results to these single-robot methods. The trajectory of Car 1 is closer to the GNSS/INS values compared to the other methods. The quantified pose estimation errors are listed in Table I.

Experiment	Length[m]	Method	APE[m]	RPE[m]
Exp. 1-Part 1	319	Proposed	**0.27**	**0.01**
		KISS-ICP	1.43	0.03
		LeGO-LOAM	2.33	0.06
Exp. 1-Part 2	285	Proposed	**0.38**	0.16
		KISS-ICP	1.64	**0.04**
		LeGO-LOAM	0.88	0.05
Exp. 2	1202	Proposed	**1.47**	**0.03**
		KISS-ICP	6.99	0.04
		LeGO-LOAM	3.32	0.07

TABLE I

APE AND RPE OF THE PROPOSED METHOD, KISS-ICP AND LEGO-LOAM.

B. Experiment 2: Decentralized Operation

In the second experiment, the two test vehicles started outside the communication range from the opposite end of a test road. As they progressed toward each other, both vehicles built local maps and transmitted CAMs at 1 Hz to the V2X network while also listening to incoming messages. Once the two vehicles come within communication range, they will receive these messages and obtain the ID and location information of the other vehicle, thus knowing the distance between them. When the distance was less than the sensor detection range (60 meters in the experiment), they sent their local maps to the V2X network. Upon receipt of the map from the other vehicle at this stage, they activated the map merging algorithm to combine the received map with their local map. After the successful map merge, they continued driving and localized on the merged map.

Fig. 8 compares the pole map produced by our proposed method (represented by green crosses) with the ground-truth map generated by a batch SLAM method (represented by orange circles). The features in the upper part of the figure were generated by Car 1, while the lower part of the map was received through the V2X network, generated by Car 2, and merged with Car 1's local map. Upon closer inspection of the zoomed-in areas, it can be observed that the maps generated by our algorithm mostly match the ground truth, albeit with certain displacements.

We compared the trajectory generated by our proposed algorithm with the results obtained from the Applanix GNSS/INS navigation system, which serves as ground truth, as well as LeGO-LOAM and KISS-ICP methods. As shown in Fig. 8, the trajectory of KISS-ICP failed to return to the starting position, and both KISS-ICP and LeGO-LOAM have clear offsets at the lower region of this test road. In contrast, our proposed method demonstrated better consistency with the ground truth.

Table I presents the Absolute Pose Error (APE) and Relative Pose Error (RPE) per meter for each of the methods evaluated, compared to the Applanix-GNSS data. Experiment 1 consists of two parts. Part 1 is Car 1's trajectory estimated based on its sensor data (right part in Fig. 7), and Part 2 is Car 1's estimation of Car 2 from the received messages (left part in Fig. 7). Our proposed approach achieved smaller APEs across all test roads. As for RPE, our method integrated

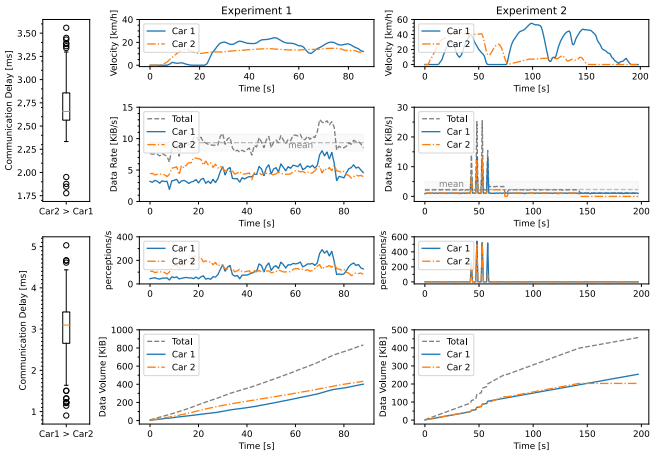

Fig. 9. V2X messages arrive with 1 to 5 ms delay after being send (left box plots). The remaining plots visualize the velocity of both cars (first row), the transmitted data in KiB/s while moving (second row), the perception transfer rate (third row) and the overall transmitted amount of data (last row) in Experiment 1 (middle column) and Experiment 2 (right column),

wheel odometry data and operated at a higher output rate (100 Hz) than the other two methods. This yielded superior relative accuracy in Exp. 1-Part 1 and Exp. 2. However, in Exp. 1-Part 2, where the poses were estimated from the received messages, our method no longer had the high-frequency advantage.

C. Bandwidth and Data Rate

The complete communication data for both experiments was meticulously logged. The box plots in Fig. 9 present the transmission delay between the two vehicles, which ranges from 1–5 ms, with average delays of 2.7 ms and 3.2 ms, respectively. The first row of data plots in Fig. 9 depict the velocity of both cars. In both experiments, the cars did drive normally according to the traffic situation and rules at the parking lot and the urban street. The second row shows the data transfer rate and the third row the corresponding perception transfer rate during the two experiments. In Experiment 1 (middle column), the data was transferred at an average rate of 9.4 KiB/s, with a peak reaching 13.0 KiB/s. Up to 291 perceptions/s have been transferred there. In Experiment 2 (right column), local maps were exchanged at a higher frequency, resulting in several data transmission peaks at around the 50th second, with up to 25.4 KiB/s and up to 542 perception per second. The average data transmission rate in Experiment 2 was 2.3 KiB/s. The last row of line plots in Fig. 9 visualize the total amount of data communicated during the two experiments. Specifically, Experiment 1 and Experiment 2 transferred a total of 833.2 KiB and 457.2 KiB of data, respectively.

V. CONCLUSIONS

This paper proposes a multi-vehicle cooperation framework for simultaneous localization and mapping tasks that utilize V2X messages for data exchange between vehicles. The method employs a pole-based sparse representation of the environment. Therefore, only limited information exchange is required, which can be efficiently achieved using V2X messages. Real-world experiments have been conducted using two V2X-connected test vehicles, showing that the proposed method can operate in centralized and decentralized ways. The results show that our approach can build global feature maps competitive with offline batch SLAM and generate trajectory estimations that outperform single-robot SLAM and recent ICP methods. Future work will focus on exchanging different types of features, scaling up to large scenarios, and testing with more connected vehicles.

REFERENCES

[1] L. Riazuelo, J. Civera, and J. M. Montiel, "C2tam: A cloud framework for cooperative tracking and mapping," *Robotics and Autonomous Systems*, vol. 62, no. 4, pp. 401–413, 2014.

[2] I. Deutsch, M. Liu, and R. Siegwart, "A framework for multi-robot pose graph slam," in *2016 IEEE International Conference on Real-time Computing and Robotics (RCAR)*. IEEE, 2016, pp. 567–572.

[3] P. Schmuck and M. Chli, "Multi-uav collaborative monocular slam," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 3863–3870.

[4] J. Knuth and P. Barooah, "Collaborative localization with heterogeneous inter-robot measurements by riemannian optimization," in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 1534–1539.

[5] A. Cunningham, M. Paluri, and F. Dellaert, "Ddf-sam: Fully distributed slam using constrained factor graphs," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2010, pp. 3025–3030.

[6] R. G. Colares and L. Chaimowicz, "The next frontier: Combining information gain and distance cost for decentralized multi-robot exploration," in *Proceedings of the 31st Annual ACM Symposium on Applied Computing*, 2016, pp. 268–274.

[7] X. S. Zhou and S. I. Roumeliotis, "Multi-robot slam with unknown initial correspondence: The robot rendezvous case," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006, pp. 1785–1792.

[8] S. Thrun, "A probabilistic online mapping algorithm for teams of mobile robots," *International Journal of Robotics Research*, vol. 20, p. 2001, 2001.

[9] A. Howard, "Multi-robot simultaneous localization and mapping using particle filters," *I. J. Robotic Res.*, vol. 25, pp. 1243–1256, 12 2006.

[10] A. Gil, ı. Reinoso, M. Ballesta, and M. Juliá, "Multi-robot visual slam using a rao-blackwellized particle filter," *Robot. Auton. Syst.*, vol. 58, no. 1, p. 68–80, jan 2010. [Online]. Available: https://doi.org/10.1016/j.robot.2009.07.026

[11] F. Conte, A. Cristofaro, A. Renzaglia, and A. Martinelli, *Cooperative Localization and SLAM Based on the Extended Information Filter*. IntechOpen, 01 2011.

[12] A. Cunningham, M. Paluri, and F. Dellaert, "Ddf-sam: Fully distributed slam using constrained factor graphs," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 11 2010, pp. 3025 – 3030.

[13] R. Dubé, 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)*, 09 2017, pp. 1004–1011.

[14] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. Leonard, and F. Dellaert, "isam2: Incremental smoothing and mapping using the bayes tree," *International Journal of Robotic Research - IJRR*, vol. 31, pp. 216–235, 05 2012.

[15] R. Dubé, D. Dugas, E. Stumm, J. Nieto, R. Siegwart, and C. Cadena, "Segmatch: Segment based place recognition in 3d point clouds," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 5266–5272.

[16] V. Indelman, E. Nelson, N. Michael, and F. Dellaert, "Multi-robot pose graph localization and data association from unknown initial relative poses via expectation maximization," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 593–600.

[17] S. Zhong, Y. Qi, Z. Chen, J. Wu, H. Chen, and M. Liu, "Dcl-slam: A distributed collaborative lidar slam framework for a robotic swarm," *arXiv preprint arXiv:2210.11978*, 2022.

[18] P.-Y. Lajoie and G. Beltrame, "Swarm-slam: Sparse decentralized collaborative simultaneous localization and mapping framework for multi-robot systems," *arXiv preprint arXiv:2301.06230*, 2023.

[19] L. Paull, G. Huang, M. Seto, and J. J. Leonard, "Communication-constrained multi-auv cooperative slam," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 509–516.

[20] A. Chapman and S. Sukkarieh, "A protocol for decentralized multi-vehicle mapping with limited communication connectivity," in *2009 IEEE International Conference on Robotics and Automation*, 2009, pp. 357–362.

[21] E. Nettleton, S. Thrun, H. Durrant-Whyte, and S. Sukkarieh, "Decentralised slam with low-bandwidth communication for teams of vehicles," in *Field and Service Robotics: Recent Advances in Reserch and Applications*, vol. 24, 01 2003, pp. 179–188.

[22] "Intelligent transport systems (its); its-g5 access layer specification for intelligent transport systems operating in the 5 ghz frequency band."

[23] W.-Y. Lin, M.-W. Li, K.-C. Lan, and C.-H. Hsu, "A comparison of 802.11a and 802.11p for v-to-i communication: A measurement study," in *Quality, Reliability, Security and Robustness in Heterogeneous Networks*, X. Zhang and D. Qiao, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 559–570.

[24] Q. Chen, X. Ma, S. Tang, J. Guo, Q. Yang, and S. Fu, "F-cooper: Feature based cooperative perception for autonomous vehicle edge computing system using 3d point clouds," in *Proceedings of the 4th ACM/IEEE Symposium on Edge Computing*, 2019, pp. 88–100.

[25] T.-H. Wang, S. Manivasagam, M. Liang, B. Yang, W. Zeng, and R. Urtasun, "V2vnet: Vehicle-to-vehicle communication for joint perception and prediction," in *European Conference on Computer Vision*. Springer, 2020, pp. 605–621.

[26] R. Xu, H. Xiang, Z. Tu, X. Xia, M.-H. Yang, and J. Ma, "V2x-vit: Vehicle-to-everything cooperative perception with vision transformer," *arXiv preprint arXiv:2203.10638*, 2022.

[27] Y. Hu, S. Fang, Z. Lei, Y. Zhong, and S. Chen, "Where2comm: Communication-efficient collaborative perception via spatial confidence maps," *Advances in neural information processing systems*, vol. 35, pp. 4874–4886, 2022.

[28] R. Xu, H. Xiang, X. Xia, X. Han, J. Li, and J. Ma, "Opv2v: An open benchmark dataset and fusion pipeline for perception with vehicle-to-vehicle communication," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 2583–2589.

[29] H. Yu, Y. Luo, M. Shu, Y. Huo, Z. Yang, Y. Shi, Z. Guo, H. Li, X. Hu, J. Yuan *et al.*, "Dair-v2x: A large-scale dataset for vehicle-infrastructure cooperative 3d object detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 21 361–21 370.

[30] R. Xu, X. Xia, J. Li, H. Li, S. Zhang, Z. Tu, Z. Meng, H. Xiang, X. Dong, R. Song *et al.*, "V2v4real: A real-world large-scale dataset for vehicle-to-vehicle cooperative perception," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 13 712–13 722.

[31] "Intelligent transport system (its); vehicular communications; basic set of applications; collective perception service; release 2: Etsi ts 103 324."

[32] "Intelligent transport systems (its); vehicular communications; basic set of applications; part 2: Specification of cooperative awareness basic service: Etsi en 302 637-2."

[33] "Trial-use standard for wireless accesses in vehicular environments (wave);ieee 1609.3 - networking services."

[34] "Intelligent transport systems (its); vehicular communications;basic set of applications;analysis of the collective perception service (cps);release 2: Tr 103 562."

[35] R. Spangenberg, D. Goehring, and R. Rojas, "Pole-based localization for autonomous vehicles in urban scenarios," in *2016 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. IEEE, 2016, pp. 2161–2166.

[36] A. Schaefer, D. Büscher, J. Vertens, L. Luft, and W. Burgard, "Long-term urban vehicle localization using pole landmarks extracted from 3-d lidar scans," in *2019 European Conference on Mobile Robots (ECMR)*. IEEE, 2019, pp. 1–7.

[37] B. Cao, C.-N. Ritter, D. Göhring, and R. Rojas, "Accurate localization of autonomous vehicles based on pattern matching and graph-based optimization in urban environments," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2020, pp. 1–6.

[38] S.-H. Chang, F.-H. Cheng, W.-H. Hsu, and G.-Z. Wu, "Fast algorithm for point pattern matching: invariant to translations, rotations and scale changes," *Pattern recognition*, vol. 30, no. 2, pp. 311–320, 1997.

[39] P. C. Lusk, K. Fathian, and J. P. How, "Clipper: A graph-theoretic framework for robust data association," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 13 828–13 834.

[40] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Sensor fusion IV: control paradigms and data structures*, vol. 1611. Spie, 1992, pp. 586–606.

[41] Y. Chen and G. Medioni, "Object modelling by registration of multiple range images," *Image and vision computing*, vol. 10, no. 3, pp. 145–155, 1992.

[42] T. Shan and B. Englot, "Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 4758–4765.

[43] I. Vizzo, T. Guadagnino, B. Mersch, L. Wiesmann, J. Behley, and C. Stachniss, "Kiss-icp: In defense of point-to-point icp simple, accurate, and robust registration if done the right way," *IEEE Robotics and Automation Letters*, 2023.