SPINS: A structure priors aided inertial navigation system

Yang Lyu1,2 | Thien-Minh Nguyen2 | Liu Liu3 | Muqing Cao2 | Shenghai Yuan2 | Thien Hoang Nguyen2 | Lihua Xie2

1School of Automation, Northwestern Polytechnical University, Xian, China
2School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, Singapore
3College of Engineering and Computer Science, Australian National University, Australian Capital Territory, Canberra, Australia

Correspondence
Yang Lyu, School of Automation, Northwestern Polytechnical University, 1 Dongxiang Road, Xian 710129, China.
Email: lyu.yang0326@gmail.com

Funding information
National Research Foundation Singapore; National Natural Science Foundation of China, Grant/Award Number: 62203358

Abstract
We propose a navigation system combining sensor-aided inertial navigation and prior-map-based localization to improve the stability and accuracy of robot localization in structure-rich environments. Specifically, we adopt point, line, and plane features in the navigation system to enhance the feature richness in low-texture environments and improve the localization reliability. We additionally integrate structure prior information of the environments to constrain the localization drifts and improve the accuracy. The prior information is called structure priors and parameterized as low-dimensional relative distances/angles between different geometric primitives. The localization is formulated as a graph-based optimization problem that contains sliding-window-based variables and factors, including Inertial Measurement Unit, heterogeneous features, and structure priors. A limited number of structure priors are selected based on the information gain to alleviate the computation burden. Finally, the proposed framework is extensively tested on synthetic data, public data sets, and, more importantly, on the real Unmanned Aerial Vehicle flight data obtained from both indoor and outdoor inspection tasks. The results show that the proposed scheme can effectively improve the accuracy and robustness of localization for autonomous robots in civilian applications.

KEYWORDS
localization, nonlinear optimization, SLAM, structure prior, UAV navigation

1 | INTRODUCTION

Autonomous vehicles, especially Unmanned Aerial Vehicles (UAVs), have attracted tremendous research interests in recent years due to their potential in improving efficiency and safety in many military and civilian applications (Cai et al., 2014). One fundamental prerequisite for an autonomous vehicle to successfully execute missions is to accurately and reliably estimate its six-degrees-of-freedom pose (Zhang & Singh, 2015), which requires delicate algorithm development and systematic integration. A detailed review of various localization systems is provided in Yuan et al. (2021). Due to physical and electromagnetic interferences, Global Positioning Systems (GPSs) may not provide persistent and reliable localization information in many complex environments, named as GPS-denied environments, making it a challenging task for a vehicle to carry out missions autonomously.

There are mainly two approaches to handle the localization in a GPS-denied environment. The first approach is the simultaneous localization and mapping (SLAM) framework (Durrant-Whyte & Bailey, 2004), which is to incrementally track the local pose by estimating the relative transformation between two observation frames. The main drawback of a SLAM-based method is that the localization result drifts as time goes on due to the accumulated relative pose estimation errors (Strasdat et al., 2010). One may reason that causes the relative pose estimation error is that the salient point features, on which many localization methods are
dependent, are insufficient in many challenging civilian environments. Line and plane features are considered as promising supplements to point features for robot localization, especially in many civilian environments. Line and plane features are considered as promising supplements to point features for robot localization. First, lines and planes are more structurally salient and therefore can be endowed with more prior information. Second, they are more stable than points with respect to lighting/texture changes. Finally, lines and planes are more ubiquitous in a structure-rich environment, such as most infrastructures in urban cities. Thus, it is a good practice to additionally integrate line and plane features in the localization framework for many civil applications.

Another alternative is to localize the vehicle according to a prior map by matching the local observations to the map directly (Sattler et al., 2011). Although this method has no drift issue, it is not easy to apply this method solely to achieve reliable localization in many challenging environments. First, the prior-map-based localization requires a high-fidelity map that may not be available in many scenarios. Second, even if there is a prior map, the association between local observations and map information is usually sophisticated and time-consuming due to large differences in observation view angle, data resolution, and even data format (Mühlfellner et al., 2016; Sattler et al., 2018).

Reflecting on the pros and cons of the two distinct approaches, it is a good practice to combine them, namely, to lend some prior information of the map to aid the SLAM-based navigation. A loosely coupled framework is proposed in Platinsky et al. (2020), which incorporates global pose factors in the local SLAM optimization. The global pose is obtained by matching images with a global map. Middelberg et al. (2014) use a locally prestored three-dimensional (3D) points cloud map to fix the drift of a local SLAM system in a tightly coupled framework. Similar works are also presented in Mur-Artal and Tardós (2017) and Zuo et al. (2020) based on different sources of maps. Although the combination methods above demonstrate improved localization performances, their frequent and indiscriminate map prior information integration may severely slow down the localization process. Therefore, it is important to determine what information to integrate into the SLAM process to make a balance between localization performance and efficiency.

Inspired by the discussions above, this paper proposes a Structure Priors aided Inertial Navigation System (SPINS) that combines SLAM-based and prior map-based localization methods by lending feature-level structure prior information to restrain the drift of the SLAM-based methods. To deal with the challenges such as feature deficiency, prior information association, and to achieve an efficient combination of the two methods in civilian environments with rich structure information, we make the following contributions:

To summarize, this paper makes the following contributions:

1. First, we extend our previous work (Lyu et al., 2022) and further relieve the feature deficiency problem by integrating 3D point, line, and plane features from various sensor modalities to relieve the feature deficiency in civil environments. The heterogeneous features-based localization is modeled in a sliding-window fashion and solved with a graph optimization method.
2. With the heterogeneous feature used, we integrate more generic and broader range of prior information which we named as structure priors and parameterized as the relative distances/angles between different geometric primitives. The association between observation and map is therefore simplified to 1D level. To ease the burden of integrating too many factors, we develop a structure prior information selection strategy based on the information gain to incorporate only the most effective structure priors for localization.
3. Finally, we test our proposed framework extensively based on synthetic data, public data sets, and real UAV flight data obtained from both indoor and outdoor inspection tasks. The results indicate that the proposed framework can improve the localization robustness even in challenging environments.

The remainder of the paper is organized as follows. In Section 2, the related works are provided. The proposed SPINS framework is formulated in Section 3. The geometric features and structure priors are modeled in Sections 4 and 5, respectively. Experiment validations are provided in Section 6. Section 7 concludes the paper.

2 | RELATED WORKS

SLAM-based localization is considered as one of the most promising approaches for robot localization. By measuring salient features from the environment, the robot can accumulatively estimate its pose in local coordinates, which is preferred in many challenging environments, such as complex indoor or urban cities (Weiss et al., 2011). In this part, we review the recent SLAM results based on different features and structure priors.

Among existing methods reported in the literature, the point feature is most commonly used in the environment perception front-end in SLAM frameworks, especially in the visual-aided navigation frameworks. Some successful demonstrations, such as oriented FAST and rotated BRIEF (ORB)-SLAM (Mur-Artal et al., 2015), and visual-inertial system (VINS)-mono (Qin et al., 2018), utilize 2D point features from vision sensors to estimate the local motion and sparse 3D point clouds of the environment. With the development of more advanced 3D sensing technologies, navigation methods based on 3D sensors, such as stereovision (Lemaire et al., 2007), light detection and ranging (LiDAR; Zhang & Singh, 2014), and RGB-D camera (Sturm et al., 2012), can directly utilize 3D points with metric information in the estimation process, therefore can provide improved localization and reconstruction results. In addition to the most commonly used point features, line and plane features are also adopted as additional features in some mission scenarios, such as indoor servicing (Lu & Song, 2015; Padhy et al., 2019), structure inspection (Hasan et al., 2017; Nguyen, Cao, et al., 2021), and autonomous landing (Chavez et al., 2017), in case that point features are not sufficient.
In the past few years, SLAM methods using heterogeneous features have begun to draw researchers’ attention. The improvement of localization performance has been verified by works with different feature combinations. In vision-based SLAM, line features are considered more effective than point features in a low textural but high structural environment with a proper definition of the state and reprojection error. Extended from the ORB-SLAM, the PL-SLAM (Pumarola et al., 2017) can simultaneously handle both point and line correspondences. The line state is parameterized with its endpoints, and the reprojection error is defined as point-to-plane distances between the projected endpoints of the 3D line and the observed line on the image plane. A tightly coupled Visual Inertial Odometry (VIO) exploiting both point and line features is proposed in He et al. (2018). The line is parameterized as a six-parameter Plücker coordinate (Hodge et al., 1994) for transformation and projection simplicity, and a four-parameter orthonormal representation for optimization compactness. Similar reprojection error to Pumarola et al. (2017) is utilized in He et al. (2018) and Lyu et al. (2022). Comparisons of different line feature parameterization are provided in Yang, Geneva, Eckenhoff et al. (2019) based on the multi-state constraint kalman filter (MSCKF) SLAM framework, which shows that the closest point (CP)-based (Yang & Huang, 2019a) and quaternion-based (Kottas & Roumeliotis, 2013) representations outperform the Plücker representation under noisy measurement conditions. Besides the monocular-based methods above, a stereovision-based VIO using point and line features together is proposed in Zheng et al. (2018) where the measurement model is directly extended from the monocular camera model similar to Pumarola et al. (2017). In the stereovision-based PL-SLAM framework (Gomez-Ojeda et al., 2019), a visual odometry is formulated similarly to Zheng et al. (2018). In addition to that, the key-frame selection and loop closure detection under point and line features setup are also provided.

As 3D sensors such as LiDAR or RGB-D camera become more popular, plane features now can be effectively extracted in a man-made environment. A LiDAR Odometry and Mapping (LOAM) in real-time is proposed in Zhang and Singh (2014), which utilizes plane features to improve the registration accuracy of a point cloud. A LiDAR-inertial SLAM framework based on 3D planes is proposed in Geneva et al. (2018), where the CP representation is utilized for parameterizing a plane. A tightly coupled vision-aided inertial navigation framework combining point and plane features is proposed in Yang, Geneva, Zuo et al. (2019), where a plane is parameterized similar to Geneva et al. (2018). In addition, the point-on-plane constraint is incorporated to improve the VINS performance. Recently, point, line, and plane features are jointly applied in the SLAM frameworks (Aloise et al., 2019; Li et al., 2020; Yang & Huang, 2019b; Zhang et al., 2019) to realize stable localization and mapping in low-texture environments. In Zhang et al. (2019), the line and plane features are tracked simultaneously along a long distance to provide persistent measurements. Also, the relationship between features, such as coplanar points, is implemented to enforce structure awareness. Similar work (Li et al., 2020) utilizes line and plane to improve the feature richness and incorporates more spatial constraints to realize more robust visual-inertial odometry. A pose-landmark graph optimization back-end is proposed in Aloise et al. (2019) based on the three types of features, which are handled in a unified manner in Nardi et al. (2019). A thorough theoretical analysis of implementing point, line, and plane features in VINS is provided in Yang, Geneva, Eckenhoff et al. (2019) where three kinds of features are parameterized as measurements to estimate the local state based on a recursive MSCKF framework. More importantly, the observability analysis of different combinations of features is provided, and the effect of degenerate motion is studied.

Although researchers have begun to introduce heterogeneous geometric features into their SLAM works, integrating prior map information to the localization still draws limited attention. There are mainly two types of prior information that can be obtained from a prior map to aid the SLAM, namely, the global information and the local structure information. By incorporating global pose constraints by matching local observations with a consistent global map, the SLAM drift can be reduced. Inspiring by this, the global information is incorporated in the local SLAM in both loosely coupled manner (Platinsky et al., 2020) and tightly coupled manner (Middelberg et al., 2014).

On the other hand, only limited structure priors information, such as point-on-line, line-on-plane, and point-on-plane constraints, are considered in the SLAM. A visual-inertial navigation system that utilizes both point features and actively selected line features on the image plane is proposed in Lyu et al. (2022), structure prior information based on special point and line relationships are utilized to improve the localization performance. Note that the structure information is ubiquitous in civil environments that are rich of man-made objects. It is a practical wisdom to implement more general structure prior information of the environment to improve the localization and mapping quality rather than to consider the environment as entirely unknown. For instance, in a building inspection environment, the structure information, as elaborate as the computer-aided design models, or as coarse as some common sense such as flat planes, parallel lines, with proper parameterization, can be implemented to aid the localization process (Hasan et al., 2017; Jovančević et al., 2016).

3 | SYSTEM DESCRIPTION

In this section, the SPINS is described from a systematic point of view. The functional blocks of the system are illustrated in Figure 1. The SPINS depends on three types of information to fulfill the task of accurately and reliably localizing an autonomous vehicle in a challenging civilian environment, namely, (1) ego-motion measurements from interoceptive sensors, such as the Inertial Measurement Unit (IMU), at high-frequency feeding streaming, and (2) detected/tracked point, line, and plane features from exteroceptive sensors, and (3) structure priors which are parameterized as pairwise high-fidelity measurements between features (Figure 2).
3.1 Optimization formulation

The aforementioned three types of information within a sliding window are incorporated into a factor graph, which is a bipartite graph that contains variables and factors. Specifically, the variables represent the state of the local vehicle and the heterogeneous features, and factors encode different types of observations and structure priors. The states included in the sliding window at time $t$ are defined as

$$
X_t \triangleq \left[ \begin{array}{c} x_{x_1}^T \end{array} \right]_{m \in \mathcal{X}} \left[ \begin{array}{c} x_{y_1}^T \end{array} \right]_{l \in \mathcal{L}} \left[ \begin{array}{c} x_{z_1}^T \end{array} \right]_{k \in \mathcal{K}},
$$

(1)

where $X_{x_1}, X_{y_1}$, and $X_{z_1}$ contain active IMU states within the sliding window at time instance $t$. $T_t$ denotes the set of IMU measurements at $t$. $P_l, L_l, \Pi_l$ denote the sets of point, line, and plane features that are observed within the sliding window at time $t$, respectively. The IMU state is

$$
x_t \triangleq \left[ \begin{array}{c} q^T \end{array} \right]_{l \in \mathcal{L}} \left[ \begin{array}{c} p_l^T \end{array} \right]_{l \in \mathcal{L}} \left[ \begin{array}{c} v_l^T \end{array} \right]_{l \in \mathcal{L}} \left[ \begin{array}{c} b_{x}^T \end{array} \right]_{l \in \mathcal{L}} b_{x},
$$

(2)

where $q^T$ is a unit quaternion denoting the rotation from the global frame $[G]$ to the IMU frame $[I]$, $p_l$ and $v_l$ are the IMU position and velocity, respectively. $b_x, b_a$ are the random walk biases for gyroscope and accelerometer, respectively.

With the state definition (1), we minimize the sum of prior estimate and the Mahalanobis norm of measurement residuals from IMU, heterogeneous feature, and structure priors to obtain a maximum posteriori estimation, as Equation (3).
where \( r^2 = r^\Sigma^{-1} r \) is defined as the squared Mahalanobis distance with covariance matrix \( \Sigma \). The first term of (3) is the cost on prior estimation residuals. The second term is the cost of IMU-based residual, and \( r_m \) defines the measurement residual between active frames \( m \) and \( m + 1 \). The IMU measurement between time step \( m \) and \( m + 1 \) is obtained by integrating high-frequency raw IMU measurements continuously with the technique called IMU preintegration (Forster et al., 2017), and \( \Sigma_m \) is the corresponding measurement covariance. The two terms are formulated the same way as in Qin et al. (2018). The second line of (3) represents the cost function of measurement residuals of point, line, and plane features and are weighted by their corresponding covariances. The third line is the cost of structure priors measurement residuals. A Huber loss \( \rho(\cdot) \) (Huber, 1992) is applied on each squared term to reduce potential mismatches between states and measurements.

The measurement residual is defined as the difference between the predicted measurement based on estimated state \( \hat{X} \) and a real measurement \( z \), as

\[
R(z, \hat{X}) = h(\hat{X}) - z,
\]  

where \( h(\cdot) \) is the measurement function for the estimated state between any two variables in the factor graph. The formulation of measurement functions \( h(\cdot) \) with regard to the geometric features and the structure priors are provided in Sections 4 and 5, respectively. The optimization of (3) is usually solved with an iterative least-squares solver through a linear approximation as detailed in Appendix A.

3.2 | Structure prior information

The structure information is ubiquitous, ranging from the fine-grained blueprint to common knowledge, such as parallel lines or planes. The greatest challenge to integrate the prior information in the optimization is to correctly associate the structure information with the observed features. Benefiting from using point, line, and plane features simultaneously, we can parameterize the spatial relationship between different geometric features as pairwise relative distances and angles in Section 5. The advantages of using such parameterization are mainly three folds.

1. First, the prior knowledge can be integrated in a simple fashion. With the distance- and angle-based formulation, the rigorous association process between the prior knowledge and current observation, which are normally based on high-dimensional and computational-demanding feature-matching processes, can be simplified to a scalar-level association matching processes based on thresholds.

2. Second, more extensive prior knowledge can be utilized to aid the navigation. The distances and angles can not only be extracted from prior maps with specific format, but also be obtained from structural common senses, hand-measured quantities, and so on.

3. As the angles and distances are stored as scalars, the storage can be reduced dramatically comparing to storing a map.

The structure priors can be extracted offline based on the following three steps:

1. extract structural primitives from various formats, for example, high-fidelity maps, local measurements, semantic information, and so forth,

2. measure and calculate the relative quantities between every two primitives (as formulated in Section 5), and

3. store salient structure quantities (distances/angles) as structure priors in a database.

The association process between offline structure prior information and online local observations is illustrated in Figure 3. Initially, finely estimated features within current optimization window are selected based on the estimation confidences. Then the structural quantities (angles/distances) based on the estimates can be calculated. Finally, a structure prior is integrated into the optimization when it is close enough to a structural quantity from estimates, based on a scalar threshold. Given two features’ estimates, \( \hat{x}_i, \hat{x}_j \), a distance/angle prior information \( z_{ij} \) is associated to \( i, j \) when

\[
|l_i(\hat{x}_i, \hat{x}_j) - z_{ij}| \leq \tau_i,
\]

where \( l_i(\cdot) \) is the function to extract distance/angles (see Section 5), and \( \tau_i \) is a scalar threshold. To prevent possible false association of structure prior information, we set a very small threshold \( \tau_i \), and allow structure priors to be assigned to only finely estimated features.

The above structure priors integration process should be carried out based on the following prerequisites. First, the structure priors should be quantities that are identifiable among all relative distances and angles. Second, the structure prior quantity should be representative of the most salient structural patterns of an environment. In many robotic operation environment, such as indoor navigation,
4 | HETEROGENEOUS GEOMETRIC FEATURES

In this section, we discuss how to model the point, line, and plane features based on onboard perception and incorporate them into the graph optimization.

4.1 | Point feature

As one of the most frequently implemented features in perception tasks, a point feature can be uniquely parameterized by its 3D coordinate, as shown in Figure 4. A point feature \( i \in P \) can be extracted from different sensors by measuring it in the local frame,

\[
z_i = h_i(x_i, x_0) + v_i = \bar{G}^T p_i - \bar{G} p_i + v_i,
\]

where \( \bar{G} p_i \in \mathbb{R}^3 \) and \( \bar{G} p_i \in \mathbb{R}^3 \) are the 3D positions of the point feature and the local vehicle in the global frame. \( v_i \in \mathbb{R}^3 \) is the measurement noise. \( \bar{G} \in \text{SO}(3) \) represents the rotation from the local frame to the global frame. By defining the state of point feature as \( x_i = \bar{G} p_i \), we have the Jacobians calculated in Appendix B.1.

4.2 | Line feature

One most commonly implemented representation of a line feature in 3D space is its Plücker coordinates. For an infinite line \( j \in L \), its Plücker coordinates are defined as \( l_j = \left[n_j^T, v_j^T\right]^T, \) \( n_j \) and \( v_j \) are the normal vector and directional vector calculated from any two distinct points on the line, as shown in Figure 4.

A measurement of an infinite line can be modeled as its Plücker coordinates in the local frame as

\[
z_j = h_j(c_{x_j}, q_j) + v_j = \begin{bmatrix} \frac{\partial R}{\partial \bar{G}} & -\frac{\partial R}{\partial \bar{G}} p_j \end{bmatrix} c_{q_j} + v_j,
\]

where \( v_j \in \mathbb{R}^6 \) is the measurement noise.

Obviously, the expression \( l_j \) is not a minimum parameterization of the line state. To calculate the Jacobian, we implement the CP approach described in Yang and Huang (2019b), which is formulated as \( p_j = d_j \hat{q} \in \mathbb{R}^4 \), where the unit quaternion \( \hat{q} \) and the closest distance of the line to the origin can be calculated from the Plücker coordinates, respectively, as

\[
R_j(\hat{q}) = \begin{bmatrix} \frac{n_j}{\|n_j\|} & v_j & \frac{n_j}{\|n_j\|} \times v_j \end{bmatrix},
\]

\[
d_j = \frac{\|n_j\|}{\|v_j\|}.
\]

where \( R_j(\hat{q}) \) is the corresponding rotation matrix to \( \hat{q} \). By defining the line state as \( x_j = \bar{G} p_j \), we have a minimum parameterization of the line in Euclidean space. The corresponding measurement Jacobians are provided in Appendix B.2.

4.3 | Plane feature

An infinite plane \( k \in \Omega \) can be minimally parameterized by the CP \( p_k = d_k n_k \in \mathbb{R}^3 \), as shown in Figure 4, where \( n_k \) is the plane’s unit normal vector, and \( d_k \) is the distance from the origin to the plane. The plane measurement here is modeled as the CP in the local frame as

\[
z_k = p_k + v_k = n_k d_k + v_k.
\]

where \( v_k \in \mathbb{R}^3 \) represents the plane measurement noise. The translation of the unit normal vector and the distance of a plane from the global frame to the local frame is

\[
\begin{bmatrix} n_k \\ d_k \end{bmatrix} = \begin{bmatrix} \bar{G} & 0_{3	imes1} \\ -\bar{G} p_i \bar{G}^T & 1 \end{bmatrix} n_k.
\]
Incorporating (10) into (9), the plane measurement can be expressed with the normal vector $\hat{n}_k$ and distance $d_k$ in the global frame as

$$z_k = \mathbf{I} \hat{n}_k + v_k = (\mathbf{G}^T \hat{n}_k + \mathbf{G} d_k) \mathbf{R}^T \hat{n}_k + v_k = -\mathbf{G}^T \hat{n}_k \mathbf{R}^T \hat{n}_k + \mathbf{G} d_k \mathbf{R}^T \hat{n}_k + v_k.$$ \hfill (11)

Define the state of a plane $k$ as $x_k = \mathbf{G} d_k = \mathbf{G} d_k \hat{n}_k$, which is the CP of the plane to the origin. We can calculate the measurement Jacobian in Appendix B.3.

**Remark 1.** In our paper, the raw sensor measurement models of different geometric features are not explicitly provided since they are highly dependent on the sensing mechanism of different sensors. Our purpose here is to provide a common framework implementing the heterogeneous features rather than considering a specific sensor.

### 5 | STRUCTURE PRIORS FORMULATION

In this part, the structure priors are formulated as the relative relationships between features. Specifically, the angles and distances are defined between different geometric primitives, including point, line, and plane.

#### 5.1 | Feature-to-feature prior modeling

The structure prior factors are plotted as blue edges in the factor graph, as shown in Figure 2b. Denote the topology set containing all the pairwise structure priors as $S$, then an edge $(a, b) \in S$ indicates that some quantitative measurements between two features $a, b \in \{P, L, \Pi\}$, denoted as $z_{ab}$, are known a priori. Let $h_{ab}$ denote the measurement function between $a$ and $b$, the structure prior residual can be obtained as

$$f_{ab} = h_{ab}(x_a, x_b) - z_{ab}. \hfill (12)$$

The residual cost is

$$\sum_{(a, b) \in S} \left\| f_{ab} \right\|^2_{\Sigma_{ab}}, \hfill (13)$$

where $\Sigma_{ab}$ is the covariance of the measurement noise which represents the fidelity of implementing specific structure prior constraints. The covariance is assumed to follow Gaussian distribution and is calculated statistically during the structure priors extraction process. The following are to model the pairwise measurements between point, line, and plane features described above.

#### 5.2 | Feature-to-feature factors

##### 5.2.1 | Point-to-point factor

When two salient points are detected, the possible structure prior information that characterizes the spatial relationship can be modeled as a 1D–3D measurement. Denote two point features $i, i' \in P$, and their relative translation $x_{i'} = x_i - x_i$, the point-to-point structure measurement

$$z_{ii'} = h_{ii'}(x_i), \hfill (14)$$

is to project the 3D displacement between the two points onto a specific 1D–3D metric in the global frame. Specifically, the distance measurement can be modeled as $z_{ii'} = ||x_i||$.

The measurement residual Jacobians with respect to the points state are provided in Appendix C.1.

**Remark 2.** In the point-to-point structure, the points should be salient in both texture and structure senses. In practice, most points are distributed according to the texture, and it may not be easy to endow structure information. Some examples of structurally salient points are intersection points, endpoints, and corner points. The integration of point-to-point structure prior information depends on the extraction and recognition of structural points, which may be challenging in practice.

##### 5.2.2 | Point-to-line factor

The spatial relationship between a point and an infinite line can be described with a 2D vector. With a point $i \in P$ and a line $j \in L$, we define a 2D displacement between them as

$$x_i = \begin{bmatrix} \hat{n}_j^T \hat{n}_j x_i + d_j \\ \hat{n}_j^T \hat{n}_j \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^3, \hfill (15)$$

where $\hat{n}_j = \frac{x_j}{||x_j||}$, $\hat{\mathbf{v}}_j = \frac{x_j}{||x_j||}$, and $d_j$ are the line $j$’s unit normal vector, unit directional vector, and the distance to the origin point, respectively. Denote the point-to-line measurement of $x_i$ as

$$z_{ij} = h_{ij}(x_i), \hfill (16)$$

The point-to-line distance can be calculated by letting $h_{ij}(\cdot)$ be a norm operator,

$$z_{ij}^2 = d_i = ||x_i||. \hfill (17)$$

Hereafter, the point-on-line constraint can be enforced as $d_i = 0$. The measurement residual Jacobian is calculated in Appendix C.2.

##### 5.2.3 | Point-to-plane factor

The relationship between a point and an infinite plane can be described with one scalar, that is, the distance from the point to the plane. With a point feature $i \in P$, and an infinite plane feature $k \in \Pi$, the distance between a point and a plane is defined as
Define a measurement function as $x_k^j = d_k$, then the point-on-plane constraint can be enforced by letting $d_k = 0$. The measurement residual Jacobian is provided in Appendix C.3.

### 5.2.4 Line-to-line factor

The relationship of two lines can be uniquely parameterized with a 3D translation vector and a rotation angle. Given two lines, denoted, respectively, as $j, j' \in L$, the rotation $\alpha_{ij}$ and translation $d_{ij} \in \mathbb{R}^3$ can be calculated as follows:

$$\alpha_{ij} = \vec{v}_j \hat{v}_j, \quad (19)$$

and

$$d_{ij} = \begin{cases} 0, & j \text{ and } j' \text{ intersection,} \\ \hat{d}_{ij}, & j \text{ and } j' \text{ are parallel,} \\ \left(\vec{v}_j \times \hat{v}_j\right) \hat{d}_{ij} \left(\vec{v}_j \times \hat{v}_j\right), & \text{otherwise}, \end{cases} \quad (20)$$

where $\hat{d}_{ij} = d_j \hat{n}_i \times \vec{v}_j - d_i \hat{n}_i \times \vec{v}_j$, $d_j \hat{n}_i \times \vec{v}_i$ and $d_i \hat{n}_i \times \vec{v}_j$ are the CP of line $j$ and $j'$ to the origin. It is straightforward to prove that as the relative angle $\alpha_{ij} \to \pm1$, the distances $d_{ij} = ||d_{ij}||$ denotes the relative distance between two parallel lines.

We first consider the rotation $\alpha_{ij}$ as a measurement between two lines. Further, when two lines are parallel, namely, $\alpha_{ij} = \pm1$, the distance $d_{ij} = ||d_{ij}||$ is considered as another measurement, namely,

$$z_{ij} = \begin{cases} [\alpha_{ij}, d_{ij}]^T, & \text{in parallel,} \\ \alpha_{ij}, & \text{otherwise}. \end{cases} \quad (21)$$

The Jacobian of line-line measurement residual is provided in C.4.

### 5.2.5 Line-to-plane factor

The spatial relationship between a line and a plane can be characterized by the dot product of the directional vector of a line $j \in L$ and the normal vector of a plane $k \in \Pi$, denoted as $\alpha_{jk}$:

$$\alpha_{jk} = \vec{v}_j \hat{n}_k. \quad (22)$$

Especially, when $\alpha_{jk} = 0$, namely, a line is parallel to a plane, a distance can be further calculated as

$$d_{jk} = n_k^T (\hat{n}_j \times \vec{v}_j) d_j - d_k. \quad (23)$$

The measurement therefore is

$$\begin{aligned}
z_k^j &= \begin{cases} [\alpha_{jk}, d_{jk}]^T, & \text{in parallel,} \\ \alpha_{jk}, & \text{otherwise.} \end{cases} \quad (24)
\end{aligned}$$

Specifically, the line-on-plane constraint is enforced as $\alpha_{jk} = 0$, and $d_k = 0$. The associated Jacobian is provided in Appendix C.5.

### 5.2.6 Plane-to-plane factor

Similar to the above formulations, a similar dot product between the unit normal vectors of two planes, can be calculated as

$$\alpha_{jk'} = \vec{n}_k \hat{n}_{k'}. \quad (25)$$

When two planes are parallel, the displacement can be calculated based on the CPs of two planes as

$$d_{jk'} = x_{k'} - x_k = n_k (\hat{n}_k - \hat{n}_{k'}) = n_k (\hat{n}_k - \hat{n}_{k'}). \quad (26)$$

Define the measurement of the relationship as

$$z_{jk} = \begin{cases} [\alpha_{jk}, d_{jk}]^T, & k \text{ parallel to } k', \\ \alpha_{jk}, & \text{otherwise.} \end{cases} \quad (27)$$

The measurement Jacobians are provided in Appendix C.6. With the above formulation of spatial relationships between features, the structure priors can be encoded into low-dimensional angles and/or distances, as summarized in Table 1. The low-dimensional encoding makes their associations with the structure priors database easy. Both the heterogeneous

| $\alpha_{ij}$ | $\alpha_{ij}$ | $\alpha_{ij}$ |
|----------------|----------------|----------------|
| $\alpha_{ij}$ | $\alpha_{ij}$ | $\alpha_{ij}$ |

### TABLE 1 Structure priors formulated as distances/angles.

| $i$ | $j$ | $k$ |
|-----|-----|-----|
| $d_{ij}$ | $d_{ij}$ | $d_{ik}$ |
| For example | For example | For example |
| Point-on-line: $d_j = 0$ | Point-on-plane: $d_k = 0$ | $d_{ik}$ if parallel $\alpha_{ik} = 0$ |
| $d_{ij}$ if parallel $\alpha_{ij} = 0$ | $d_{ik}$ if parallel $\alpha_{ik} = 0$ | \text{Orthogonality: } $\alpha_{ij} \pm 0$
| $\alpha_{ij}$ | $\alpha_{ik}$ | $\alpha_{ik}$ |
| $d_{ik}$ if parallel $\alpha_{ik} = \pm 1$ | $d_{ik}$ if parallel $\alpha_{ik} = \pm 1$ | \text{Orthogonality: } $\alpha_{ik} = \pm 1$
| For example | For example | For example |
geometric feature factors and the structure prior factors are further integrated with the graph optimization toolbox Georgia Tech Smoothing and Mapping (GTSAM; Dellaert, 2012).

5.3 Structure prior selection

Given the formulation of pairwise structure prior factors above, there are at most $n(n - 1)/2$ possible feature prior factors in a sliding window with $n$ finely estimated geometric features. In a graph optimization process, incorporating too many feature-to-feature structures prior information will severely damage the sparsity of the graph structure, and as a result will increase the burden on optimization computation.

Example 1. A synthetic environment with 100 points, 40 lines, and 40 planes is shown in Figure 5 with one platform trajectory. We can define as many as 16,110 structure prior factors. The optimization is carried out by model the features, trajectory, and different number of structure priors into one-factor graph. The position estimation error and calculation time based on different numbers of structure priors are plotted in Figure 5b.

As shown in Figure 5b, the localization error decreases as more structure priors are incorporated into the optimization, which indicates the effectiveness of reducing localization errors by using structure priors. Nevertheless, the computation overheads of the optimization grow dramatically when more than 500 structure priors are used. Therefore we should have a strategy to only integrate a limited number of structure prior factors in the factor graph, which is formulated as Problem 1 as follows.

Problem 1. Given a set of structure priors $S$, and a real value function representing the estimation accuracy improvement from structure priors $g : 2^S \rightarrow \mathbb{R}$. Then the selection strategy is to obtain a most helpful subset $S' \subseteq S$ that maximizes $g$:

$$
\max_{S' \subseteq S} g(X) \text{ s.t. } |S'| \leq N. \tag{28}
$$

Usually Problem 1 is NP-hard and cannot be solved easily for a large set $S$. However, by defining a proper function $g$, a suboptimal solution can be found more efficiently.

To improve the localization accuracy, we consider selecting a subset of the structure priors that minimize the estimation covariance. As different types of structure priors may we consider implementing the Fisher Information Matrix (FIM) to measure the contribution of a structure before the localization performance. Denote the belief of the state within the sliding window of time $t$ as $X_t \sim N(\tilde{X}_t, P_t)$, and one structure prior $s \in S$ of the current local map as a measurement, $h_s(X_t) \sim N(z_s, \Sigma_s)$, we have the following belief update equation according to Bayes’ rule

$$
P^{2}_{t} = P^{1}_{t} + \sum_{s \in S} l_s, \tag{29}
$$

where $l_s = J_s^{T}\Sigma_s^{-1}J_s$ is the FIM of a specific structure prior $s$. $J_s$ is the Jacobian of the structure prior $s$ with regard to the local pose. $P_t$ and $P_{t}^{1}$ denote the covariance before and after integrating the structure priors, respectively.

Further, we can define the function $g$ as a D-optimal criteria similar to Carlone and Karaman (2019),

$$
g(X) = \log \det \left( A P^{-1} A^{T} \right) = \log \det \left( A P^{-1} + \sum_{s \in S} l_s \right). \tag{30}
$$

By maximizing the $g$ function, the minimization of interest state, which is selected by the matrix $A$, can be achieved in the D-optimal fashion. As indicated in Shamaiah et al. (2010), the logdet() metric of the covariance is submodular w.r.t. the information gain of structure priors. This property indicates that a suboptimal solution with guaranteed optimization bound can be found based on more efficient greedy search strategy.

In this paper, we implement a stochastic greedy algorithm with the lazy evaluation as Algorithm 1. In the outer stochastic greedy loop, during each iteration, we select one most informative feature from a subset of candidate features with size $N$. The inner lazy evaluation is to speed up the searching with an upperBound() function, which is to calculate the upper bound of the logdet() function of a positive definite matrix using Hadamard’s inequality (Horn & Johnson, 2012).
Finally, we can use the marginalization technique (Carlone & Karaman, 2019) to obtain the localization uncertainty $P_t$ by marginalizing out other states.

### EXPERIMENT EVALUATION

In this part, the proposed SPINS is tested based on synthetic data, the public data sets, and, most importantly, on real flight data sets that are collected with a UAV during indoor and outdoor inspection tasks.

#### 6.1 Synthetic data

To evaluate the localization performance of the proposed framework, we create a customized 2.5D indoor simulation scenario with point, line, and plane features as presented in Figure 6. A 3D robot trajectory is generated within the simulation space based on spline functions as a red curve. An IMU is simulated according to the ADIS16448 IMU sensor specifications listed in Geneva et al. (2018). We assume that the 3D geometry information of the features is obtained according to the measurement function described in Section 4 with extra field-of-view limitations.

#### TABLE 2 Structure priors between features.

| Line     | Plane                  |
|----------|------------------------|
| Point    | Points on the line     |
| Line     | -Parallelism with distance |
| Line     | -Orthogonality         |
| Plane    | -Parallelism with distance |
| Plane    | -Orthogonality         |

#### FIGURE 6 The 2.5D simulation scenario with point, line, and plane features.

#### FIGURE 7 Translational and rotational RMSE under different feature and prior information configurations. P-INS, point feature; PL-INS, point and line features; PLP-INS, point, line, and plane features; RMSE, root-mean-square error; SP, structure prior; SPINS, Structure Priors aided Inertial Navigation System.
The optimization-based localization is solved with the iSAM2 solver (Kaess et al., 2012) from the GTSAM package, which we additionally integrate line factors, plane factors, and structure prior factors. Comparisons based on different features and structure prior factors are carried out. Specifically, we consider (1) point feature (P-INS), (2) point and line features (PL-INS), (3) point, line, and plane features (PLP-INS)-based methods, and (4) our structure prior aided method (SPINS). We select (1) 20 structure priors in each frame randomly (SPINS-Rand. 20 SP), (2) 20 structure priors according to Section 5.3 (SPINS-App. Opt. 20 SP), (3) the most informative 20 structure priors (SPINS-Info. Opt. 20 SP), and (4) all structure priors (SPINS-All SP). The structures information listed in Table 2 is adopted.

In addition, a prior-map-based localization method is used as a localization benchmark where perfect feature matching between local observations and map is assumed.

The root-mean-square errors (RMSEs) are plotted in Figure 7. The quantitative comparison between different strategies is also provided in Table 2. The localization results of using heterogeneous features are plotted as solid lines in different colors. It is apparent that, as more types of features are used, more accurate estimation can be achieved. More importantly, the integration of structure prior information can further improve the localization performance, as plotted in dashed lines. Specifically, integrating all structure prior information unsurprisingly achieves the closest localization performance to the prior map-based method. Nevertheless, the computation time for solving each round of the local optimization also significantly increases due to more structure factors that are integrated, as provided in Table 2. Among the 20 structure priors-based method, the optimal selection achieves the best performance, at the expense of greedy search computation overhead. Our proposed method achieved comparable result but with much less computation burden. Random selection-based localization result is the least accurate (Table 3).

### 6.2 Euroc data set

In this part, the proposed SPINS framework is tested on the public Euroc MAV Data set (Burri et al., 2016). The front-end detection and track point, line, and planes based on stereovision measurements. Specifically, the point features are detected and tracked with the Kanade-Lucas-Tomasi-based optical flow method similar to Qin et al. (2018). The line features are detected and tracked based on a modified line segment detector (LSD) as described in Fu et al. (2020). Moreover, the planes are extracted and tracked based on triangulated point features based on the method described in Nardi et al. (2019).

On the prior information part, we use the point cloud from the Vicon room to obtain the plane-related structure priors. The plane extraction based on the point cloud is shown in Figure 8. The corresponding distributions of angles between planes and the distances of parallel planes are plotted in Figure 9, which show the

### Table 3 Average RMSE over 100 Mento-Carlo simulations with different strategies.

| Strategies          | Translation errors (m) | Rotation errors (°) | Time per iteration (s) |
|---------------------|------------------------|---------------------|------------------------|
| P-INS               | 0.8186                 | 6.9399              | 0.0180                 |
| PL-INS              | 0.6103                 | 5.0499              | 0.0193                 |
| PLP-INS             | 0.4863                 | 3.9595              | 0.0274                 |
| SPINS-Rand. 20 SPs  | 0.3242                 | 3.1684              | 0.0259                 |
| SPINS-App. Opt. 20 SPs | 0.2082              | 3.1652              | 0.0301                 |
| SPINS-Opt. 20 SPs   | 0.1757                 | 2.0399              | 0.3175                 |
| SPINS-All SPs       | 0.1277                 | 1.6984              | 0.2463                 |
| Prior map           | 0.0547                 | 0.9414              | —                      |

Abbreviations: SP, structure prior; RMSE, root-mean-square error; P-INS, point feature; PL-INS, point and line features; PLP-INS, point, line, and plane features; SPINS, Structure Priors aided Inertial Navigation System.
repetition and sparsity angle and distance pattern in a man-made environment. On the basis of the prior information above, we extract the distance priors, such as point-on-plane, line-on-plane, and plane-to-plane distances, and angles priors, such as plane parallel, plane orthogonality, as structure priors to aid the localization.

The localization results by implementing the VINS FUSION (https://github.com/HKUST-Aerial-Robotics/VINS-Fusion.git), ORB-SLAM3 (https://github.com/UZ-SLAMLab/ORB_SLAM3.git), the heterogeneous features-based method (Yang & Huang, 2019b), and the proposed framework are listed in Table 4 based on the evaluation method described in Zhang and Scaramuzza (2018). We use the same experiment setup for all tests according to the Euroc data set parameters. As indicated, our proposed method can achieve the best performance in both translational and rotational RMSE in V1_01, V2_01, V2_02, and V2_03. Specifically, our proposed SPINS outperforms the PLP-based method (Yang & Huang, 2019b) in all data sets, which shows the effectiveness of incorporating structure prior information.

As one example, the results of estimated trajectories of V2_02 are plotted in Figure 10.

6.3 Field collected data

In this part, the proposed SPINS is further tested on large-scale inspection environment. A DJI M600 pro hexacopter carrying various sensors is utilized to detect features of the environment, as illustrated in Figure 11a. We considered two scenarios where geometric feature and structure patterns are rich, as shown in Figure 11b indoor navigation and Figure 11c building facade inspection, which is available as part of the VIRAL data set (Nguyen, Yuan, et al., 2021) (Figure 12).

We collect sensing data from visual cameras, LiDARs, and IMU sensors to properly detect the geometric features based on similar front-end processing to Section 6.2. The groundtruth (GT) is provided

| Data | Translational RMSE (m)/rotational RMSE (°) |
|------|------------------------------------------|
| V1_01 | 0.129/1.748 | 0.085/1.484 | 0.098/1.674 | 0.079/1.131 |
| V1_02 | 0.145/1.504 | 0.089/1.336 | 0.321/1.455 | 0.094/0.905 |
| V1_03 | 0.144/1.967 | 0.093/1.952 | 0.193/2.389 | 0.095/1.762 |
| V2_01 | 0.150/3.121 | 0.085/1.852 | 0.116/2.352 | 0.077/1.731 |
| V2_02 | 0.197/3.413 | 0.167/3.141 | 0.188/4.581 | 0.160/3.030 |
| V2_03 | 0.219/2.924 | 0.160/3.007 | 0.213/3.458 | 0.151/2.988 |

Note: Bold values indicate the best localization performance in the RMSE sense.
FIGURE 12  Point feature estimation before/after integrating the point-on-line constraint. When the constraint is imposed, the points on the wall are more accurately stick to the wall, lead to better point feature estimation.

| VINS-Mono | VINS-Stereo | A-LOAM | LVI-SAM | DLL | SPINS | ICP | GT |
|-----------|-------------|--------|---------|-----|-------|-----|----|

FIGURE 13  Estimated trajectories based on different methods. DLL, LiDAR localization; GT, groundtruth; LiDAR, light detection and ranging; LOAM, LiDAR Odometry and Mapping; LVI, lidar-visual-inertial; SAM, smoothing and mapping; SPINS, Structure Priors aided Inertial Navigation System; VINS, visual-inertial system.

FIGURE 14  Estimation errors of different methods. DLL, LiDAR localization; LiDAR, light detection and ranging; LOAM, LiDAR Odometry and Mapping; LVI, lidar-visual-inertial; SAM, smoothing and mapping; SPINS, Structure Priors aided Inertial Navigation System.
by a Leica Geosystem that measures the optical prism onboard. All tests are carried out based on the same parameter setup provided in https://ntu-aris.github.io/ntu_viral_dataset/.

### 6.3.1 Indoor navigation

In this part, the proposed structure prior information is further tested on an indoor auditorium. Similarly, we measure some potentially repetitive and salient features as the structure prior information. Moreover, we use the Leica system to obtain a point cloud map to extract plane-based structure priors similar to Section 6.2. Three different trajectories are generated to test the proposed methods.

We compare our results to methods based on monocular camera (VINS-Mono; Qin et al., 2018), stereo-camera (VINS-Stereo; Qin et al., 2019), LiDAR (LOAM; Zhang & Singh, 2014), LiDAR and Camera (LVI-SAM; Shan et al., 2021), a map-based localization method (DLL; Caballero & Merino, 2021), and our method (SPINS). The iterative closest point (ICP) method by registering LiDAR scans to point cloud map is also provided as a localization benchmark. More specifically, we use the LOAM odometry for ICP and DLL initialization. The parameters are set as 50 iterations and 0.05 m max correspondence distance in ICP. The estimated trajectories of three trials based on the above methods are plotted in Figure 13. The estimation errors of NYA03 are given in Figure 14. The position RMSE based on the methods is provided in Table 5. It is apparent, our proposed method, aided by the angle/distance priors (as plotted in Figure 15), can achieve the best performance among methods mentioned above. Moreover, the distances structure priors are more favored in most cases comparing to angles priors. One example of the effectiveness by imposing point-on-plane constraint, or zero point-to-plane distance constraints, is shown in Figure 12.

Additionally, we compare our method to the prior-map-based method DLL. Although the DLL method can achieve very close performance to the SPINS in both accuracy and time efficiency, it requires an odometer to provide the initialization for registration between map and local scan, which is an extra computation burden. In addition, the DLL method requires an initial position of the robot in the map, which may not be available in practical scenarios. The ICP-based method is also provided to indicate the best localization performance that map-based method can achieve. However, the ICP method uses more than 2 s for each registration process, and is hard to be implemented in real-time applications.

### 6.3.2 Building inspection

#### Scenario 1

In the larger-scale building inspection task, the UAV is driven to follow a trajectory covering the façade of a building. The geometric

---

**TABLE 5** RMSE of the estimation results based on different methods.

| Methods | Translational RMSE (m) | Time per iteration (s) |
|---------|------------------------|------------------------|
|         | NYA01                  | NYA02                  | NYA03                |                        |
| VINS-Mono |                        | 0.2576                  | 0.6118               | –                      |
| VINS-Stereo | 0.2427                  | 0.2424                  | 0.3808               | –                      |
| A-LOAM | 0.0768                  | 0.0902                  | 0.0797               | –                      |
| LVI-SAM | 0.0761                  | 0.0885                  | 0.0827               | –                      |
| DLL | 0.0734                  | 0.0663                  | 0.0607               | 0.0911                |
| SPINS | 0.0551                  | 0.0672                  | 0.0592               | 0.0798                |
| ICP | 0.0262                  | 0.0253                  | 0.0214               | 2.003                 |

Abbreviations: DLL, LiDAR localization; ICP, iterative closest point; LiDAR, light detection and ranging; LOAM, LiDAR Odometry and Mapping; LVI, lidar-visual-inertial; RMSE, root-mean-square error; SAM, smoothing and mapping; SPINS, Structure Priors aided Inertial Navigation System; VINS, visual-inertial system.

**FIGURE 15** Number of angle and distance structure priors. SP, structure prior.
feature extraction is shown in Figure 16. To utilize our proposed SPINS method, we manually measure some distances and angles, which we treated as the main patterns of the building, are summarized in Table 6.

The localization results of the inspection using different methods are presented in Figures 17, 18, and 19. A demonstrative video is provided at https://youtu.be/p-wca_WekvQ. The trajectories of different localization methods along with the GT are plotted in Figure 17, although all trajectories are initialized at the same \([0,0,0]\)^T, the trajectories from the VINS drifts as time goes on. Specifically, from Figure 18, it is clear that the trajectories from VINS drifts in the \(x\) and \(y\) directions due to the low feature density and variation during the horizontal movement, and on the other hand, the trajectory from A-LOAM drifts mainly in the \(z\) direction due to low depth variation during vertical movement in a 2.5D building. As indicated in Figure 19, the results obtained by using point, line, and plane features apparently have better performance in all three directions. Moreover, the integration of structure priors SPINS outperformed the PLP method with the provide structure prior information. The position estimation RMSE of the PLP and SPINS are 1.018 and 0.7452 m, respectively, with a significant position accuracy improvement by incorporating the structure priors.

**TABLE 6** Hand-measured distances for the building.

| Distance type            | Typical value (m) |
|--------------------------|-------------------|
| Parallel lines           | [0.3, 0.5, 1.2, 1.5, 2.5, 3.3, 4] |
| Parallel line to plane   | [0.5, 1.2, 2.7, 3, 4]  |
| Parallel planes          | [1.5, 3, 4.5, 6]    |

**FIGURE 16** Extraction of point, line, and plane features from both vision and LiDAR. LiDAR, light detection and ranging.

**FIGURE 17** Estimated trajectories from different methods. GT, groundtruth; LiDAR, light detection and ranging; LOAM, LiDAR Odometry and Mapping; PLP, point, line, and plane; VINS, visual-inertial system.

**FIGURE 18** Estimation on the \(X\), \(Y\), and \(Z\)-directions.
Scenario 2
To further illustrate the effectiveness of the proposed method in structure environment, we carry out experiments by driving the UAV to fly between buildings (as Figure 20a). A high accurate map is built by fusing LiDAR point clouds based on RTK GPS measurements (as shown in Figure 20b). Then we extract structure priors from the point clouds. We record two different trials. The UAV follows a relatively steady and smooth inspection pattern in the first trial, and it follows a more complicated and aggressive pattern in the second trial.

We test the localization performance based on the LiDAR odometry (A-LOAM), map-based localization, and our proposed method (SPINS) method. The localization results based on different methods are presented in Figures 21, 22, and 23. The trajectories estimated from different methods are plotted in Figure 21. Significant drifts of the LiDAR-based odometry can be observed from the trajectories in both trials. Further, from Figure 22 to Figure 23 we can see that the map-based method (DLL), as well as the heterogeneous feature fusion-based method (PLP), although can achieve better estimates in both trials, still have obvious localization errors, especially in the Z-direction in Trial 2. The proposed SPINS method can achieve stable trajectories tracking in both trials with the smallest estimation errors. The localization accuracy is compared in Table 7, which also indicates that our proposed method can achieve better performance in accuracy than other methods.

**FIGURE 19** Localization errors on the X-, Y-, and Z-directions.

**FIGURE 20** Building inspection scenario 2. (a) Buildings scenario 2 and (b) high accurate point clouds.
FIGURE 21  Estimated trajectories from different methods and the groundtruth (GT). (a) Trial 1 and (b) Trial 2. DLL, LiDAR localization; LiDAR, light detection and ranging; LOAM, LiDAR Odometry and Mapping; PLP, point, line, and plane; SPINS, Structure Priors aided Inertial Navigation System.

FIGURE 22  Estimated trajectories versus time. (a) Trial 1 and (b) Trial 2. DLL, LiDAR localization; GT, groundtruth; LiDAR, light detection and ranging; LOAM, LiDAR Odometry and Mapping; PLP, point, line, and plane; SPINS, Structure Priors aided Inertial Navigation System.

FIGURE 23  Estimated errors versus time. (a) Trial 1 and (b) Trial 2. DLL, LiDAR localization; LiDAR, light detection and ranging; LOAM, LiDAR Odometry and Mapping; PLP, point, line, and plane; SPINS, Structure Priors aided Inertial Navigation System.
TABLE 7 Translational RMSE of the estimation results based on different methods.

| Methods    | Translational RMSE (m) | Translational RMSE (m) |
|------------|------------------------|------------------------|
|            | Trial 1                | Trial 2                |
| A-LOAM     | 26.5461                | 7.9023                 |
| DLL        | 0.8629                 | 1.0883                 |
| PLP        | 0.8050                 | 1.0032                 |
| SPINS      | 0.4223                 | 0.5317                 |

Abbreviations: DLL, LiDAR localization; LiDAR, light detection and ranging; LOAM, LiDAR Odometry and Mapping; PLP, point, line, and plane; RMSE, root-mean-square error; SPINS, Structure Priors aided Inertial Navigation System.

7 | CONCLUSIONS

We propose a sliding-window optimization-based localization framework utilizing point, line, and plane features. Considering heterogeneous features, we further develop a structure priors integration method that can further improve localization robustness and accuracy. To alleviate the computation burden brought by extra structure factor edges in the factor graph, we adopt a screening mechanism to select the most informative structure priors. Although the proposed SPINS shows advantages in civilian environments in our experiment, its effectiveness in more generic environments is questionable. In the future, we will further investigate on using general semantic prior information to aid the localization to achieve more wide range of environment adaptation.

ACKNOWLEDGMENTS

The work is supported by National Research Foundation (NRF) Singapore, ST Engineering-NTU Corporate Lab under its NRF Corporate Lab@ University Scheme.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in NTU VIRAL: A Visual-Inertial-Ranging-LiDAR Data set for Auton at https://ntu-aris.github.io/ntu_viral_dataset/.

ORCID

Yang Lyu http://orcid.org/0000-0002-0702-908X

REFERENCES

Aloise, I., Corte, B. D., Nardi, F. & Grisetti, G. (2019) Systematic handling of heterogeneous geometric primitives in graph-SLAM optimization. *IEEE Robotics and Automation Letters*, 4(3), 2738–2745.

Burri, M., Nikolic, J., Gohl, P., Schneider, T., Rehder, J., Omari, S., Achtelik, M. W. & Siegwart, R. (2016) The EuRoC micro aerial vehicle datasets. *The International Journal of Robotics Research*, 35(10), 1157–1163.

Caballero, F. & Merino, L. (2021) DLL: direct LiDAR localization. A map-based localization approach for aerial robots. In: *2021 IEEE/RSJ international conference on intelligent robots and systems* (IROS), Prague, Czech Republic. Manhattan, Park Avenue: IEEE, pp. 5491–5498.

Cai, G., Dias, J. & Seneviratne, L. (2014) A survey of small-scale unmanned aerial vehicles: recent advances and future development trends. *Unmanned Systems*, 2(02), 175–199.

Carlone, L. & Karaman, S. (2019) Attention and anticipation in fast visual-inertial navigation. *IEEE Transactions on Robotics*, 35(1), 1–20.

Chavez, A., L’Heureux, D., Prabhakar, N., Clark, M., Law, W.-L. & Prazenica, R. J. (2017) Homography-based state estimation for autonomous UAV landing. In: *AIAA information systems-AIAA infotech aerospace*, Texas, USA. Reston, VA: American Institute of Aeronautics and Astronautics.

Dellaert, F. (2012) Factor graphs and GTSAM: a hands-on introduction. Georgia Institute of Technology. Technical Report.

Durrant-Whyte, H. & Bailey, T. (2006) Simultaneous localization and mapping: part I. *IEEE Robotics & Automation Magazine*, 13(2), 99–110.

Forster, C., Carlone, L., Dellaert, F. & Scaramuzza, D. (2017) On-manifold preintegration for real-time visual-inertial odometry. *IEEE Transactions on Robotics*, 33(1), 1–21.

Fu, Q., Wang, J., Yu, H., Ali, I., Guo, F. & Zhang, H. (2020) PL-VINS: real-time monocular visual-inertial SLAM with point and line. arXiv preprint, arXiv.

Geneva, P., Eckenhoff, K., Yang, Y. & Huang, G. (2018) LIPS: LiDAR-inertial 3D plane SLAM. In: *2018 IEEE/RSJ international conference on intelligent robots and systems* (IROS). IEEE, pp. 123–130.

Gomez-Ojeda, R., Moreno, F.-A., Zuniga-Noël, D., Scaramuzza, D. & Gonzalez-Jimenez, J. (2019) PL-SLAM: a stereo SLAM system through the combination of points and line segments. *IEEE Transactions on Robotics*, 35(3), 734–746.

Hasan, A., Qadir, A., Nordeng, I. & Neubert, J. (2017) Construction inspection through spatial database. In: *Proceedings of the international symposium on automation and robotics in construction* (ISARC), Taipei, Taiwan. Bratislava, Slovakia: IAARC, pp. 1–8.

He, Y., Zhao, J., Guo, Y., He, W. & Yuan, K. (2018) PL-VIO: tightly-coupled monocular visual-inertial odometry using point and line features. *Sensors*, 18(4), 1159.

Hodge, W. V. D., Hodge, W. & Pedoe, D. (1994) *Methods of algebraic geometry: Volume 1* (Vol. 2). Cambridge University Press.

Horner, R. A. & Johnson, C. R. (2012) *Matrix analysis*. Cambridge University Press.

Huber, P. J. (1992) Robust estimation of a location parameter. In: *Breakthroughs in statistics*. Berlin, Heidelberg: Springer, pp. 492–518.

Jovančević, I., Viana, I., Ortué, J.-J., Sentenac, T. & Larnier, S. (2016) Matching CAD model and image features for robot navigation and inspection of an aircraft. In: *5th international conference on pattern recognition applications and methods*, Rome, Italy, Setubal: SCITEPRESS—Science and Technology Publications, Lda, pp. 359–366.

Kaess, M., Johannsson, H., Robert, R., Ila, V., Leonard, J. J. & Dellaert, F. (2012) ISAM2: incremental smoothing and mapping using the Bayes tree. *The International Journal of Robotics Research*, 31(2), 216–235.

Kottas, D. G. & Roumeliotis, S. I. (2013) Efficient and consistent vision-aided inertial navigation using line observations. In: *2013 IEEE international conference on robotics and automation* (ICRA), Karlsruhe, Germany. Manhattan, Park Avenue: IEEE, pp. 1540–1547.

Lemaire, T., Berger, C., Jung, I.-K. & Lacroix, S. (2007) Vision-based SLAM: stereo and monocular approaches. *International Journal of Computer Vision*, 74(3), 343–364.

Li, X., He, Y., Lin, J. & Liu, X. (2020) Leveraging planar regularities for point line visual-inertial odometry. In: *2020 IEEE/RSJ international conference on intelligent robots and systems* (IROS), Las Vegas, NV, USA. Manhattan, Park Avenue: IEEE, pp. 5120–5127.

Lu, Y. & Song, D. (2015) Robust RGB-D odometry using point and line features. In: *The IEEE international conference on computer vision* (ICCV), Santiago, Chile. Manhattan, Park Avenue: IEEE, pp. 3934–3942.
Shan, T., Englot, B., Ratti, C. & Rus, D. (2021) LVI-SAM: D-coupled lidar-visual-inertial odometry via smoothing and mapping. In: 2021 IEEE international conference on robotics and automation (ICRA). Manhattan, Park Avenue: IEEE, pp. 5692–5698.

Strasdat, H., Montiel, J. & Davison, A. J. (2010) Scale drift-aware large scale monocular SLAM. Robotics: Science and Systems VI, 2(3), 7.

Sturm, J., Engelhard, N., Endres, F., Burgard, W. & Cremers, D. (2012) A benchmark for the evaluation of RGB-D SLAM systems. In: IEEE/RSJ international conference on intelligent robots and systems (IROS), Vilamoura-Algarve, Portugal, Manhattan, Park Avenue: IEEE, pp. 573–580.

Weiss, S., Scaramuzza, D. & Siegwart, R. (2011) Monocular-SLAM-based navigation for autonomous micro helicopters in GPS-denied environments. Journal of Field Robotics, 28(6), 854–874.

Yang, Y., Geneva, P., Eckenhoff, K. & Huang, G. (2019) Visual-inertial odometry with point and line features. In: 2019 IEEE/RSJ international conference on intelligent robots and systems (IROS), Macau, China. Manhattan, Park Avenue: IEEE, pp. 2447–2454.

Yang, Y., Geneva, P., Zuo, X., Eckenhoff, K., Liu, Y. & Huang, G. (2019a) Tight-coupled aided inertial navigation with point and plane features. In: 2019 International conference on robotics and automation (ICRA), Montreal, QC, Canada. Manhattan, Park Avenue: IEEE, pp. 6094–6100.

Yang, Y. & Huang, G. (2019b) Aided inertial navigation: unified feature representations and observability analysis. In: 2019 International Conference on Robotics and Automation (ICRA). IEEE, pp. 3528–3534.

Yang, Y. & Huang, G. (2019b) Observability analysis of aided INS with heterogeneous features of points, lines, and planes. IEEE Transactions on Robotics, 35(6), 1399–1418.

Yuan, S., Wang, H. & Xie, L. (2021) Survey on localization systems and algorithms for unmanned systems. Unmanned Systems, 9(2), 129–163.

Zhang, J. & Singh, S. (2014) LOAM: lidar odometry and mapping in real-time. In: Robotics: science and systems, Vol. 2.

Zhang, J. & Singh, S. (2015) Visual-inertial combined odometry system for aerial vehicles. Journal of Field Robotics, 32(8), 1043–1055.

Zhang, J., Zeng, G. & Zha, H. (2019) Structure-aware SLAM with planes and lines in man-made environment. Pattern Recognition Letters, 127, 181–190.

Zhang, Z. & Scaramuzza, D. (2018) A tutorial on quantitative trajectory evaluation for visual-inertial odometry. In: IEEE/RSJ international conference on intelligent robots and systems (IROS), Madrid, Spain. Manhattan, Park Avenue: IEEE, pp. 7244–7251.

Zheng, F., Tsai, G., Zhang, Z., Liu, S., Chi, C.-C. & Hu, H. (2018) Trifo-VIO: robust and efficient stereo visual inertial odometry using points and lines. In: 2018 IEEE/RSJ international conference on intelligent robots and systems (IROS). Madrid, Spain. Manhattan, Park Avenue: IEEE, pp. 3686–3693.

Zuo, X., Ye, W., Yang, Y., Zheng, R., Vidal-Callega, T., Huang, G. & Liu, Y. (2020) Multimodal localization: stereo over LiDAR map. Journal of Field Robotics, 37(6), 1003–1026.

How to cite this article: Lyu, Y., Nguyen, T.-M., Liu, L., Cao, M., Yuan, S., Nguyen, T.H. & Xie, L. (2023) SPINS: a structure priors aided inertial navigation system. Journal of Field Robotics, 40, 879–900. https://doi.org/10.1002/rob.22161.
APPENDIX A: LEAST-SQUARE SOLVER
In the non-Euclidean space $\mathcal{X}$, the approximation is achieved by expanding the residual around the origin of a chart computed at current estimation $\hat{\mathcal{X}}$ as

$$r(\mathcal{X} \oplus \Delta \mathcal{X}) = r(\hat{\mathcal{X}}) + \frac{\partial r(\mathcal{X} \oplus \Delta \mathcal{X})}{\partial \Delta \mathcal{X}}|_{\Delta \mathcal{X} = 0} \Delta \mathcal{X},$$

$$= r(\hat{\mathcal{X}}) + J \Delta \mathcal{X}.$$  

(A1)

The operator $\oplus$ applies a perturbation $\Delta \mathcal{X}$ to the manifold space $\mathcal{X}$. Specifically, for the Euclidean states, the $\oplus$ operator degenerates to the vector addition operator. $J$ is the sparse Jacobian matrix with only none zero block on the related two states. Incorporating the approximation (A1) into (3), we have the following form:

$$\Delta \mathcal{X}^T H \Delta \mathcal{X} + 2b^T \Delta \mathcal{X} + \text{const}(\hat{\mathcal{X}}),$$

(A2)

where

$$H = J^T J + \sum_{m \in \Omega} J^T_m J_m^T$$

and

$$b = r(\hat{\mathcal{X}}) + \sum_{m \in \Omega} r^T_m \hat{\mathcal{X}}_m + \sum_{i \in \mathcal{L}} \hat{\mathcal{X}}_i J^T_i \hat{\mathcal{X}}_i + \sum_{i \in \mathcal{L}_s} \hat{\mathcal{X}}_i J^T_i \hat{\mathcal{X}}_i$$

and const($\hat{\mathcal{X}}$) is a constant term depending only on $\hat{\mathcal{X}}$.

As a result, (A2) can be minimized by solving

$$(H + \lambda I) \Delta \mathcal{X} = -b,$$

(A3)

where $\lambda$ is a damping factor. The estimation is hereafter updated as

$$\hat{\mathcal{X}} = \hat{\mathcal{X}} \oplus \Delta \mathcal{X}.$$  

(A4)

The update iterates until convergence.

APPENDIX B: JACOBIANS OF HETEROGENEOUS FEATURE MEASUREMENTS

B.1 | Point measurement Jacobians
The Jacobians, based on the formulations in (A1), are related to the state of the point in Euclidean space $x_i = \hat{p}_i$ and the pose of the robot in $\text{SE}(3)$, and can be calculated as follows.

$$J^p_{\hat{p}_i} = \frac{\partial r(x_i + \Delta x_i)}{\partial \Delta x_i} |_{\Delta x_i = 0} = \frac{\partial r}{\partial \hat{R}},$$

(B1)

and

$$J^r_{\hat{R}} = \frac{\partial r(\hat{\mathcal{X}}) \oplus \Delta \mathcal{X}}{\partial \Delta \mathcal{X}} |_{\Delta \mathcal{X} = 0} = -\frac{\partial r}{\partial \hat{R}}.$$  

(B2)

where $\Delta x_i$ and $\Delta x_j$ are the perturbations on $x_i$ and $x_j$, respectively. Specifically, $[\delta \hat{R}^T, \delta \hat{R}]^T$ are the rotational and translational parts of the perturbation $\Delta x_i$. The measurement Jacobians over other parts of $x_i$ are 0.

B.2 | Line measurement Jacobians
The measurement residual Jacobian w.r.t. the local pose is

$$J^l_{\delta \mathcal{X}} = \frac{\partial r(l) \oplus \Delta l}{\partial \Delta l} |_{\Delta l = 0}$$

(B3)

$$= \left[ \begin{array}{c} \left[ R^G \ell \right]_x - \left[ R^G \hat{p}_j \right]_x \tilde{G}_x \left[ R^G \hat{e}_j \right]_x \\ -\left[ R^G \hat{e}_j \right]_x \end{array} \right].$$

The measurement residual Jacobians w.r.t. to line state can be calculated using the chain rule as

$$J^l_{\delta \mathcal{X}} = J^l_{\delta l} J^l_{\delta l},$$

(B4)

where $\Delta l = [\delta \hat{q}_j \delta d]^T$ is an intermediate state. Specifically, with perturbation $\delta l = [\delta \hat{b}_j \delta d]^T$ on manifold $l$, we have

$$J^l_{\delta l} = \frac{\partial r(l) \oplus \Delta l}{\partial \Delta l} |_{\Delta l = 0}$$

(B5)

$$= \left[ \begin{array}{c} \left[ R - R^G \hat{p}_j \right]_x \tilde{G}_x \left[ R^G \hat{e}_j \right]_x \\ 0_3 \\ -\left[ R^G \hat{e}_j \right]_x \end{array} \right]$$

and

$$J^l_{\delta \mathcal{X}} = \frac{\partial r(l) \oplus \Delta l}{\partial \Delta \mathcal{X}} |_{\Delta \mathcal{X} = 0}$$

where $\mathcal{X}$ is the vector part and scalar part of the quaternion $q_i$, respectively. Please refer to Yang and Huang (2019a) for a detailed derivation of (B6).

B.3 | Plane measurement Jacobians
By defining an intermediate state of the plane as $p_j = [\hat{b}_j \hat{d}_j]^T$, the measurement residual Jacobian w.r.t. the plane state $x_k = \hat{p}_k$ can be calculated using the chain rule as

$$J^p_{\hat{R}} = \frac{\partial r(p_j + \Delta p_j)}{\partial \Delta p_j} |_{\Delta p_j = 0}$$

and

$$J^p_{\hat{b}_j} = \frac{\partial r(p_j + \Delta p_j)}{\partial \Delta b_j} |_{\Delta b_j = 0}$$

$$J^p_{\hat{d}_j} = \frac{\partial r(p_j + \Delta p_j)}{\partial \Delta d_j} |_{\Delta d_j = 0}.$$
\[ f_k^i = f_{k_i}^i f_{k_i}^i, \]  

(B7) where

\[
\begin{align*}
J_{\delta\hat{p}_k}^i &= \frac{\partial (\bar{h}_k \mp \delta \hat{p}_k)}{\partial \delta \hat{p}_k} \bigg|_{\delta \hat{p}_k = 0} \\
&= [\partial^2 (G\alpha_k - \dot{G}\hat{n}_k \mp \delta \hat{p}_k)]_i \\
J_{\Delta \hat{x}_k}^i &= \frac{\partial \hat{x}_k (\bar{h}_k \mp \delta \hat{p}_k)}{\partial \Delta \hat{x}_k} \bigg|_{\Delta \hat{x}_k = 0} \\
&= \begin{bmatrix}
\bar{G}\dot{\hat{n}}_k \mp \delta \hat{p}_k \\
\hat{n}_k^T \\
\end{bmatrix},
\end{align*}
\]

(B8)

\[ \delta \hat{p}_k \text{ and } \Delta \hat{x}_k \text{ are perturbations on } \hat{p}_k \text{ and } \hat{x}_k, \text{ respectively.} \]

By injecting a small perturbation of local pose into (11), we have the measurement Jacobian w.r.t. the pose as follows:

\[
\begin{align*}
J_{\delta \hat{p}_i}^i &= \left. \frac{\partial (\bar{h}_i \mp \delta \hat{p}_i)}{\partial \delta \hat{p}_i} \right|_{\delta \hat{p}_i = 0} \\
&= \left. \left( G\alpha_i - \dot{G}\hat{n}_i \mp \delta \hat{p}_i \right) \right|_{\delta \hat{p}_i = 0} \\
J_{\Delta \hat{x}_i}^i &= \left. \frac{\partial \hat{x}_i (\bar{h}_i \mp \delta \hat{p}_i)}{\partial \Delta \hat{x}_i} \right|_{\Delta \hat{x}_i = 0} \\
&= \left. \left[ \hat{n}_i \mp \delta \hat{p}_i \right] \hat{n}_i^T \right|_{\Delta \hat{x}_i = 0},
\end{align*}
\]

(B9)

\[
\begin{align*}
J_{\delta \hat{p}_i}^i &= \frac{\partial h_i (\bar{h}_i \mp \delta \hat{p}_i)}{\partial \delta \hat{p}_i} \bigg|_{\delta \hat{p}_i = 0} \\
&= \left. \left( G\alpha_i - \dot{G}\hat{n}_i \mp \delta \hat{p}_i \right) \right|_{\delta \hat{p}_i = 0} \\
J_{\Delta \hat{x}_i}^i &= \frac{\partial \hat{x}_i (\bar{h}_i \mp \delta \hat{p}_i)}{\partial \Delta \hat{x}_i} \bigg|_{\Delta \hat{x}_i = 0} \\
&= \left. \left[ \hat{n}_i \mp \delta \hat{p}_i \right] \hat{n}_i^T \right|_{\Delta \hat{x}_i = 0}.
\end{align*}
\]

(B10)

APPENDIX C: JACOBIANS OF THE STRUCTURAL PRIORS MEASUREMENTS

C.1 | Point-point measurement Jacobians

For a pair of points \((i, i') \in S\), the Jacobians of the measurement residual, defined as \( r = (z - \bar{h}_{i'}(\hat{k}, \hat{x})) \), are derived as

\[
J_{k_j}^{i'} = J_{k_j}^{i'} J_{k_j}^{i'}, \quad \text{(C1)}
\]

where

\[
\begin{align*}
J_{k_j}^{i'} &= \left. \frac{\partial h_i (\bar{h}_i \mp \delta \hat{p}_i)}{\partial \delta \hat{p}_i} \right|_{\delta \hat{p}_i = 0} \\
&= \left. \left( G\alpha_i - \dot{G}\hat{n}_i \mp \delta \hat{p}_i \right) \right|_{\delta \hat{p}_i = 0} \\
J_{k_j}^{i'} &= \left. \frac{\partial \hat{x}_i (\bar{h}_i \mp \delta \hat{p}_i)}{\partial \Delta \hat{x}_i} \right|_{\Delta \hat{x}_i = 0} \\
&= \left. \left[ \hat{n}_i \mp \delta \hat{p}_i \right] \hat{n}_i^T \right|_{\Delta \hat{x}_i = 0}.
\end{align*}
\]

(C2)

Specifically, if we consider \( h_i(\cdot) \) as a distance measurement function, the Jacobians can be calculated as

\[
J_{k_j}^{i'} = \frac{\hat{x}_i}{||\hat{x}_i||},
\]

(C3)

C.2 | Point-line measurement Jacobians

The measurement residual Jacobian with respect to the point \( i \) is

\[
J_{k_j}^{i} = J_{k_j}^{i} J_{k_j}^{i},
\]

(C4)

where

\[
J_{k_j}^{i} = \begin{bmatrix}
\hat{n}_i^T \\
\hat{n}_i \times \hat{v}_j
\end{bmatrix}.
\]

The measurement residual Jacobian with respect to the line estimation error is

\[
J_{\delta}^{i} = J_{\delta}^{i} J_{\delta}^{i} J_{\delta}^{i},
\]

(C5)

\[ J_{\delta}^{i} = J_{\delta}^{i} J_{\delta}^{i} J_{\delta}^{i}, \]

(C6)

and \( J_{\delta}^{i} \) is calculated in (B6). When considering the measurement as point-line distance, \( J_{\delta}^{i} = \hat{r}_i \parallel \hat{n}_i \parallel \).

C.3 | Point-plane measurement Jacobians

The measurement residual Jacobian with respect to the point estimate error is

\[
J_{\delta}^{i} = \hat{G}\hat{n}_i^T.
\]

(C6)

The measurement residual Jacobian with respect to the plane estimate error is

\[
J_{\delta}^{i} = J_{\delta}^{i} J_{\delta}^{i},
\]

(C7)

where

\[
J_{\delta}^{i} = [\hat{G} \hat{n}_i^T]_i^T,
\]

and \( J_{\delta}^{i} \) is provided in (B8).

C.4 | Line-line measurement Jacobians

The Jacobians of the measurement error \( a_{ij} \) and \( d_{ij} \) w.r.t. the estimation error are

\[
J_{a_{ij}}^{ij} = J_{a_{ij}}^{ij} J_{a_{ij}}^{ij},
\]

(C8)

where

\[
J_{a_{ij}}^{ij} = \hat{v}_j [R_{e_2}]_i, \quad O_{2 \times 1}.
\]

The measurement residual Jacobian w.r.t. \( j' \) can be calculated similarly. Specifically, when two lines are parallel, namely, \( \hat{v}_j = \pm \hat{v}_j' \), the distance measurement Jacobian can be calculated as

\[
J_{d_{ij}}^{ij} = J_{d_{ij}}^{ij} J_{d_{ij}}^{ij},
\]

(C9)

where

\[
J_{d_{ij}}^{ij} = \hat{v}_j \times \left[ q [R_{e_2}]_i, \hat{R}_{e_2} \right].
\]

(C10)

C.5 | Line-plane measurement Jacobians

The directional correlation measurement Jacobian is as follows:

\[
J_{d_{ij}}^{ij} = J_{d_{ij}}^{ij} J_{d_{ij}}^{ij},
\]

(C11)
where
\[ J^b_j = [R_e e_3]_x \ 0_{3 \times 1} \]  \hspace{1cm} (C12)

and
\[ J^c_j = \hat{\psi}_j (I_3 - G_{\hat{\mathbf{n}_k}} G_{\hat{\mathbf{n}_k}}^T) \frac{\partial}{\partial \hat{\mathbf{n}_k}}. \]  \hspace{1cm} (C13)

When the line \( j \) is parallel to the plane, namely, \( \hat{\psi}_j \hat{n}_k = 0 \), the distance residual Jacobians
\[ J^d_{j_k} = J^b_j J^c_j, \]  \hspace{1cm} (C14)

where
\[ J^b_j = \hat{n}_k [R_e e_3]_x \ R_e e_2 \]  \hspace{1cm} (C15)

and
\[ J^c_j = \frac{(\hat{\mathbf{n}}_j \times \hat{\psi}_j^T) (I_3 - G_{\hat{\mathbf{n}_k}} G_{\hat{\mathbf{n}_k}}^T) G_j^T}{\partial \hat{\mathbf{n}_k}} + G_{\hat{\mathbf{n}_k}}^T. \]  \hspace{1cm} (C16)

C.6 | Plane-plane measurement Jacobians

The Jacobian of the angle measurement residual is
\[ J^a_{j_k} = (\hat{\mathbf{n}}_j \times \hat{\psi}_j) (I_3 - G_{\hat{\mathbf{n}_k}} G_{\hat{\mathbf{n}_k}}^T) G_j^T + G_{\hat{\mathbf{n}_k}}^T. \]  \hspace{1cm} (C17)

The distance measurement residual is calculated as
\[ J^d_{j_k} = \frac{G_{\hat{\mathbf{p}}_k} - G_{\hat{\mathbf{p}}_{k'}}}{\| \hat{\mathbf{p}}_k - \hat{\mathbf{p}}_{k'} \|.} \]  \hspace{1cm} (C18)

The Jacobian w.r.t. the plane \( k' \) can be calculated similarly.