Geometry-aware Subsampling and pole-enhanced Map Constraints for urban Localization of LiDAR-based Systems

Marvin Scherff, Dominik Ernst, Hamza Alkhatib

Geodetic Institute, Leibniz University Hanover, Germany - (scherff,ernst,alkhatib)@gih.uni-hannover.de

Keywords: LiDAR-based Localization, Error State Kalman Filter, Farthest Point Sampling, Pole-like Structures, Monte-Carlo Simulation

Abstract

Urban localization for autonomous driving requires accurate 6-DoF vehicle pose despite GNSS multipath, occlusions, and rapidly changing visibility. We fuse LiDAR, IMU, and GNSS in an error-state Kalman filter against a high-resolution (HR) map, aiming (i) to reduce LiDAR load without degrading accuracy and (ii) to improve robustness in building-sparse areas such as open junctions. The reference trajectory and HR map stem from a dedicated urban measurement campaign; Monte-Carlo simulations use ray-cast LiDAR, synthesized IMU, and GNSS tied to this trajectory so that only sensor noise is varied. A geometry-aware farthest-point sampling scheme prioritizes points informative for building/ground planes and pole-like structures, while an extended functional model introduces poles as additional vertical constraints. A retained-point rate of $10\ \%$ preserves trajectory-wide millimeter-level and sub-milliradian accuracy, meeting in theory automotive requirements. Filter runtime is reduced by about $82\ \%$ relative to the full LiDAR data. Compared with plane-only variants, the planes+poles configuration yields statistically significant but globally modest improvements in longitudinal, lateral, and yaw accuracy. More importantly, a sliding-window analysis reveals that it markedly stabilizes pose in plane-sparse junctions. Overall, the results suggest that task-aware subsampling preserves trajectory-wide performance while pole constraints add local robustness in challenging urban scenes; validation with real sensor logs remains necessary to confirm these accuracy margins, but the proposed filtering scheme shows promising potential for practical deployment.

1. Introduction

Reliable six-degree-of-freedom (6-DoF) vehicle pose in dense urban areas hinges on fusing complementary sensors with strong map priors to withstand Global Navigation Satellite Systems (GNSS) multipath and occlusions. For autonomous operation, accuracy targets are stringent: on local streets, vehicles should maintain $10\ cm$ lateral and longitudinal error and $3\ mrad$ heading error, each at the $95\ \%$ confidence level (Reid et al., 2019). Against this backdrop, reliable localization is foundational for public safety, efficient traffic management, and adherence to traffic regulations. It motivates (real-time) systems that estimate full 3D position and orientation online in a superordinate frame by combining onboard sensing with prior geographic knowledge.

Traditional navigation modalities alone are not sufficient in cities. GNSS can provide accurate positioning in open environments, yet performance degrades in ”urban canyons” due to signal blockage and multipath. Inertial sensing helps bridge outages but accumulates drift without absolute updates. These limitations have motivated map-based localization approaches that align onboard observations—most notably Light Detection and Ranging (LiDAR) and, in some cases, camera imagery—to a prior representation of the environment, while inertial cues support continuity and short-term dynamics.

High-definition/high-resolution (HD/HR) maps are particularly valuable in this context. They comprehensively encode stable structural references—such as building facades, road and curb geometry, lane markings, and traffic infrastructure—enabling centimeter-level localization when fused with onboard sensors (Ballardini et al., 2021, Zhang et al., 2022, Ghallabi et al., 2019). By providing semantically and geometrically consistent representations of the urban scene, HR maps help anchor the vehicle’s pose across changing viewpoints and partial occlusions and form the basis for the robust, map-centric localization studied in this work.

In this research, we refine an Error-State Kalman Filter (ESKF) for LiDAR–Inertial Measurement Unit (IMU)–GNSS fusion with the goal of meeting the above accuracy requirements (Ernst et al., 2023). The ESKF performs high-rate inertial propagation between updates and fuses asynchronous GNSS and LiDAR observations for correction, while carrying IMU bias states. Two contributions are central: Firstly, a geometry-aware farthest-point sampling (FPS) subsampling method is introduced, which prioritizes points that are most informative for the employed map constraints during the filter update. Secondly, the existing LiDAR-to-plane matching of buildings and ground is extended by adding pole-like landmarks as additional structural constraints. Together, these choices strengthen the geometric constraints exploited by the estimator and improve computational efficiency by reducing the number of non-informative LiDAR observations that must be compared against the utilized HR map. In our experiments, geometry-aware FPS is benchmarked against a simple voxel-grid prefiltering followed by random sampling to isolate the effect within the ESKF/map stack. A systematic comparison with other published subsampling strategies (including heuristic and learning-based methods) and with alternative localization pipelines is beyond the scope of this study.

The experimental evaluation is simulation-based but grounded in real data: a post-processed reference trajectory and an HR map from a dedicated urban measurement campaign, with

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer INST 187/742-1 FUGG and as part of the Research Training Group i.c.sens [RTG 2159].

a simulated sensor suite (GNSS, IMU, and a 64-line HESAI LiDAR) matching the deployed hardware. LiDAR returns are ray-cast into the map and consistency-filtered against nearby campaign points, yielding a stream that retains 78.6% of rays aligned with real measurements; Monte-Carlo (MC) analyses then vary sensor noise to probe estimator behavior. This design isolates estimation effects while preserving realistic geometry, yet it does not fully capture real-world complexities such as dynamic objects and their occlusions, adverse weather impacts on returns, or sensor-specific artifacts beyond simple Gaussian models. Consequently, the reported results should be read as an upper bound under controlled conditions, motivating subsequent validation with real sensor logs and richer stochastic models.

The remainder of the paper is organized as follows. Section 2 reviews related work; Section 3 details the refined ESKF together with the pole-based functional model and the advanced FPS strategy; Section 4 presents results and discussion based on MC simulations; and Section 5 concludes with an outlook.

2. Related Works

This section surveys the fundamentals and recent advances that inform the present study. The review commences with an examination of error-state Kalman filtering and implicit measurements. It then transitions to LiDAR–map localization with structural cues, contrasting lightweight maps with key–observation selection. The review subsequently summarizes point-cloud subsampling strategies and concludes with pole-like landmarks as complementary constraints.

2.1 Error-state filtering and implicit measurements

Error-state Kalman filtering has become a standard framework for inertial navigation and multi-sensor fusion, as it separates the nominal state from small perturbations, maintains a compact error representation, and handles quaternion-based attitude states consistently (Solà, 2017). In the context of vehicle localization, ESKF formulations have been employed, for instance, to integrate GNSS and IMU data in challenging environments. This approach utilizes a smoothed ESKF to generate robust pose estimates, particularly in the domain of intelligent transportation applications (Yin et al., 2023). ESKF variants also underpin tightly coupled LiDAR–IMU odometry and SLAM systems, where high-rate inertial propagation is combined with scan-matching residuals in an iterated error-state framework to achieve accurate motion estimation over long trajectories (You et al., 2023). In this general context, implicit functional measurement equations facilitate the direct incorporation of geometric constraints, such as distance-to-surface relations, into the update process. This approach is particularly advantageous for LiDAR- and vision-based systems that rely on point-to-map constraints. An implicit-measurement ESKF that integrates IMU propagation with LiDAR-to-plane constraints for position tracking has been demonstrated and evaluated in simulation, underscoring the potential for integrating dense geometric information asynchronously into the filter (Ernst et al., 2023). The methodology outlined in this paper is predicated on that formulation, with extensions to the functional model and observation handling for HR-map-based urban localization.

2.2 LiDAR–map localization with structural cues

The localization of LiDAR maps has been the subject of extensive research as a method for obtaining robust global poses in urban environments using prior 3D maps (Roriz et al., 2024). Beyond generic point-cloud matching, several approaches exploit explicit structural cues in the map. Building facades and ground surfaces are frequently represented as planes or mesh-like elements and utilized as stable anchors for scan-to-map registration. Road and curb boundaries have been extracted from mobile mapping data and used as constraints for accurate lane-level localization in complex urban scenes, demonstrating that detailed roadway geometry significantly supports vehicle pose estimation (Süleymanoğlu et al., 2024). Other methods encode structural information via voxelized deep feature maps or semantic layers, where LiDAR scans are matched against feature grids derived from HD maps (Endo and Kamijo, 2023). These approaches underscore the importance of geometry-rich, structurally organized map content for ensuring robust localization in the presence of occlusions and GNSS degradation.

2.3 Lightweight maps versus selecting key observations

Given the substantial size of dense LiDAR maps, matching efficiency has been pursued through two primary avenues: map compression and reduced observation utilization during runtime. As demonstrated in the research by Wiesmann et al., point-cloud compression can significantly reduce the memory requirements of dense outdoor topographies while maintaining registration fidelity for localization tasks (Wiesmann et al., 2021). A more extensive survey of automotive LiDAR compression techniques also emphasizes neural and classical schemes that balance storage, transmission bandwidth, and localization accuracy (Roriz et al., 2024). In addition to these map-centric strategies, several works have been developed to reduce the online load by streaming only submaps in vehicle-to-infrastructure settings (Kim et al., 2023). An alternative line of work involves maintaining a high-resolution map and instead selecting informative points from the live LiDAR scan for matching against this fixed representation. In this observation-centric perspective, the offline HR map is stored once in an efficient spatial index, and real-time workload is controlled by carefully choosing which LiDAR points to use in the filter update, rather than by downsampling or maintaining multiple map resolutions.

2.4 Point-cloud subsampling

The selection of observations in LiDAR-based localization frequently relies on point-cloud subsampling. Random sampling and voxel-grid downsampling are widely employed due to their simplicity and predictable density. However, these methods are agnostic to task-specific relevance, potentially resulting in the discarding of geometrically informative points (Rusu and Cousins, 2011). Farthest-point sampling (FPS) is frequently adopted as a coverage-oriented baseline, especially in point-cloud learning pipelines such as PointNet++ and related architectures, where FPS is used to build multi-scale sets of representative points (Qi et al., 2017). Recent studies have compared and benchmarked a range of downsampling techniques—including voxel grids, uniform random sampling, and FPS variants—with respect to geometric fidelity and downstream performance, underscoring the trade-off between computational cost and preservation of structural detail (Tzermia et al., 2025). Concurrently, the deep-learning literature has witnessed the proposal of weighted or loss-aware FPS schemes, wherein sampling probabilities are skewed towards points deemed more informative for the task at hand, signifying a prevailing tendency towards task-adaptive subsampling (Deng et al., 2023). Collectively,

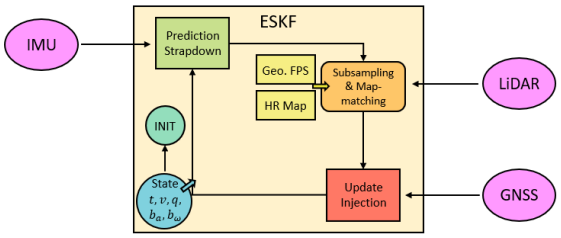

Figure 1. Adapted ESKF with geometry-aware LiDAR subsampling and HR map-matching, including building/ground planes and pole-like structures. IMU data is employed to predict the current epoch’s state, while LiDAR and GNSS misclosures perform the update.

these works demonstrate that subsampling is most effective when it accounts for both spatial coverage and the geometric or semantic constraints that will be enforced downstream.

2.5 Pole-like landmarks as complementary constraints

In addition to plane-based maps of facades and ground, vertical landmarks have been identified as valuable complementary cues for urban localization. Preliminary studies have shown that local pole patterns, derived from LiDAR scans, can facilitate global vehicle localization over urban networks by matching pole constellations against a landmark map (Brenner, 2009). Subsequent studies employed pole-like objects as sparse but stable landmarks, extracted from LiDAR or vision, and integrated them with odometry, GNSS, or lane-marking information in filtering or particle-filter frameworks to achieve robust localization in GNSS-challenged areas (Spangenberg et al., 2016, Chen et al., 2021). Long-term experiments with LiDAR-based pole mapping have further shown that pole landmarks can remain reliable over extended periods and under seasonal changes, especially at intersections and along corridors with limited facade visibility (Schaefer et al., 2021). Furthermore, matching detected poles and trunks to surveyed landmark inventories has been shown to provide absolute 2D constraints that suppress trajectory discontinuities and reduce cross-track errors in GNSS-degraded corridors (Meyer et al., 2023). When considered as a whole, these results suggest that poles and analogous vertical structures offer substantial horizontal and heading cues that effectively supplement building and ground planes in the context of urban map-based localization.

3. Methodology

This section briefly summarizes the ESKF developed by (Ernst et al., 2023) with its prediction and measurement update step. Then, the extended functional model of pole-like structures is demonstrated. Consequently, the sensor setup of the underlying measurement campaign for our studies is outlined, resulting in the employed HR map and LiDAR observations. Furthermore, the subsampling scheme for selecting key points in the rotating LiDAR is presented to reduce the computational effort in map-matching. Finally, the experimental setup conducted by MC simulations, incorporating the stochastic model, is described. An overview of the adapted ESKF is presented in Figure 1 to illustrate the order of individual steps.

3.1 ESKF state, error state, and prediction

The nominal filter state represented in the bottom left corner of the filtering scheme (c.f. Figure 1) collects the body pose in the superordinate frame $\{l\}$ and the IMU biases:

$$\mathbf{x} = \left[\, \mathbf{t}_b^{l\top},\ \mathbf{v}_b^{l\top},\ \mathbf{q}_b^{l\top},\ \mathbf{b}_a^{\top},\ \mathbf{b}_\omega^{\top} \,\right]^{\top}. \tag{1}$$

Here, $\mathbf{t}_b^l, \mathbf{v}_b^l \in \mathbb{R}^3$ are the body position and velocity expressed in $\{l\}$, $\mathbf{q}_b^l \in \mathbb{H}$ is the unit quaternion rotating vectors from $\{b\}$ to $\{l\}$, and $\mathbf{b}_a, \mathbf{b}_\omega \in \mathbb{R}^3$ are the accelerometer and gyroscope biases. For the entire paper can be assumed that the initial state (INIT) is well known as the estimation or derivation is not subject of our study.
The minimal error state is

$$\delta\mathbf{x} = \left[\, \delta\mathbf{t}^{\top},\ \delta\mathbf{v}^{\top},\ \delta\boldsymbol{\theta}^{\top},\ \delta\mathbf{b}_a^{\top},\ \delta\mathbf{b}_\omega^{\top} \,\right]^{\top}, \tag{2}$$

with $\delta\boldsymbol{\theta} \in \mathbb{R}^3$ the small-angle attitude perturbation used to update the quaternion reliably.

State update by error (injection). After computing $\delta\mathbf{x}$, the predicted state is corrected component-wise as

$$\begin{aligned}
\mathbf{t}_{b+}^l &= \mathbf{t}_{b-}^l + \delta\mathbf{t}, \\
\mathbf{v}_{b+}^l &= \mathbf{v}_{b-}^l + \delta\mathbf{v}, \\
\mathbf{q}_{b+}^l &= \delta\mathbf{q} \otimes \mathbf{q}_{b-}^l, \quad \delta\mathbf{q} \approx \begin{bmatrix} 1 \\ \frac{1}{2}\delta\boldsymbol{\theta} \end{bmatrix}, \\
\mathbf{b}_{a+} &= \mathbf{b}_{a-} + \delta\mathbf{a}_b, \\
\mathbf{b}_{\omega+} &= \mathbf{b}_{\omega-} + \delta\mathbf{b}_\omega,
\end{aligned} \tag{3}$$

followed by re-normalization of $\mathbf{q}_{b+}^l$ to unit norm. $\otimes$ denotes the (Hamilton) quaternion product. The state update or error injection step presents the final operation in each epoch (bottom right corner in ESKF block within Figure 1) before continuing with the next time stamp.

Filter epochs are defined by the IMU observations. Within each epoch k, all LiDAR and GNSS measurements with $t_{k-1} < t \leq t_k$ are incorporated in the state update. Because the platform moves during an epoch, it is crucial to interpolate the state (position and orientation) to each measurement time to obtain accurate misclosures. We therefore first detail the time parametrization and interpolation used per observation, and then describe the LiDAR–to–map residuals.

3.2 Time interpolation for points between IMU epochs

LiDAR points are acquired asynchronously within (t_{k-1}, t_k). Their misclosure evaluation relative to the HR map uses the pose interpolated between the two epoch boundaries via the relative timestamp

$$\Delta t = t_k - t_{k-1}, \qquad \tau_i = \frac{t_i - t_{k-1}}{\Delta t} \in [0, 1]. \tag{4}$$

Translation is interpolated linearly and attitude by SLERP (Shoemake, 1985):

$$\mathbf{t}_b^l(t_i) = (1 - \tau_i)\, \hat{\mathbf{t}}_b^l(t_{k-1}) + \tau_i\, \hat{\mathbf{t}}_b^l(t_k), \tag{5}$$

$$\begin{aligned}
\Delta\theta &= \arccos\!\left(\hat{\mathbf{q}}_b^l(t_{k-1})^{\top}\, \hat{\mathbf{q}}_b^l(t_k)\right) \\
\mathbf{q}_b^l(t_i) &= \frac{\sin\!\left((1-\tau_i)\Delta\theta\right)}{\sin(\Delta\theta)}\, \hat{\mathbf{q}}_b^l(t_{k-1}) + \frac{\sin\!\left(\tau_i\Delta\theta\right)}{\sin(\Delta\theta)}\, \hat{\mathbf{q}}_b^l(t_k).
\end{aligned} \tag{6}$$

The jacobians in the implicit state update utilizes the same information. The following subsections present the two map con-

straints and subsampling scheme to integrate the LiDAR sensor into the pose estimation (c.f. center of Figure 1).

3.3 Implicit LiDAR-to-plane update

Building and ground map elements are modeled as planar patches and integrated into the ESKF using the Hessian normal form

$$\mathbf{n}^\top \mathbf{x} + d = 0, \qquad \|\mathbf{n}\| = 1. \tag{7}$$

with $\mathbf{n}$ the unit plane normal and d the signed offset (both in $\{l\}$). To accelerate assignment, each plane is tessellated into triangles (Delaunay triangulation of the polygonal footprint), enabling fast point-to-triangle inclusion tests. Let a LiDAR return $\mathbf{p}_i^s$ be time-stamped at $t_i \in (t_{k-1}, t_k]$. With fixed sensor-to-body extrinsic calibration parameters $(\mathbf{R}_s^b, \mathbf{t}_s^b)$, the point transformed to the superordinate frame $\{l\}$ is

$$\mathbf{p}^l(t_i) = \mathbf{R}\big(\mathbf{q}_b^l(t_i)\big)\left(\mathbf{R}_s^b \mathbf{p}_i^s + \mathbf{t}_s^b\right) + \mathbf{t}_b^l(t_i). \tag{8}$$

If the orthogonal projection of $\mathbf{p}^l(t_i)$ lies inside a planar triangle (barycentric test), the implicit signed-distance residual $r_{\mathrm{pl},i}$ to $(\mathbf{n}, d$ is computed. A LiDAR point is considered a valid assignment to any planar element if

$$\left| r_{\mathrm{pl},i} \right| \le \tau_{\mathrm{pl}}, \tag{9}$$

where τ_{pl} represents the planar gating threshold. If multiple candidates satisfy (9), the element with the smallest absolute deviation $|r_{\mathrm{pl},i}|$ is selected (Ernst et al., 2023).

3.4 Pole extension of the functional model

As a main contribution, the functional model is extended to include pole-like structures, which provide stable, vertical landmarks and are frequently present at junctions and in open areas where building facades may be absent. Equally to the planar elements in $\{l\}$, poles are modeled in the superordinate frame by a center $\mathbf{c} \in \mathbb{R}^3$ and a unit principal axis $\mathbf{a} \in \mathbb{R}^3$ with $\|\mathbf{a}\| = 1$. For the transformed LiDAR point $\mathbf{p}^l(t_i)$ in (8), cylindrical coordinates aligned with $\mathbf{a}$ are derived.

$$\mathbf{v} = \mathbf{p}^l - \mathbf{c}, \qquad h = \mathbf{a}^\top \mathbf{v},$$
$$\mathbf{d} = \mathbf{v} - h\mathbf{a}^\top, \qquad r = \|\mathbf{d}\|. \tag{10}$$

Assignments are restricted to a height interval defined by the map-generating LiDAR points.

$$h_{\min} \le h \le h_{\max}, \tag{11}$$

In addition, only the front half of the cylindrical surface that is visible from the observing LiDAR scanner is used. Let $\mathbf{c} + h\,\mathbf{a}$ denote the closest point on the pole axis at height h, and $\mathbf{s}(t_i)$ the scanner origin at time t_i. The visibility condition is enforced by

$$\mathbf{d}^\top \big(\mathbf{s}(t_i) - (\mathbf{c} + h\,\mathbf{a})\big) \ge 0, \tag{12}$$

i.e., the radial vector $\mathbf{d}$ from the axis $\mathbf{a}$ to the cylinder surface or hull point and the viewing vector from the same axis to the scanner must form an acute angle. If the dot product is negative, the point lies on the backside of the pole with respect to the sensor and is discarded, which prevents assignments to pole surface regions that cannot be observed in the current scan.

Pole-like structures such as lamps and traffic lights cannot be represented accurately by a regular cylinder, as their masts typically show a mild, nearly linear taper over height as allowed by national standards. Hypothesis tests with higher-order polynomial radius models did also not yield a statistically significant improvement for the majority of objects. They rather tend to overfit local irregularities. Therefore we adopt the linear model

$$r_{\exp}(h) = r_0 + \kappa\, h, \tag{13}$$

where r_0 is the radius bias at $h = 0$ and κ the height-dependent (linear) scale. A sample pole primitive represented in l is generated for visualization purpose by tesselation in Figure 2.

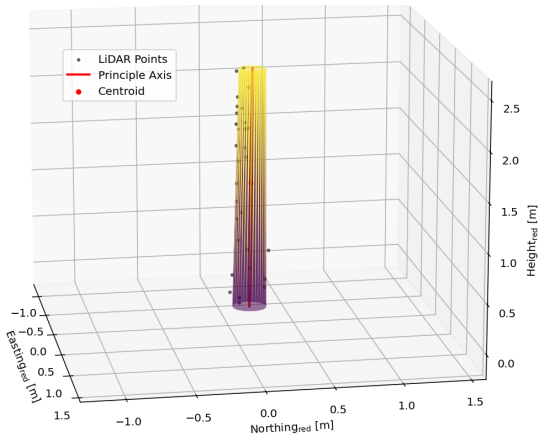

Figure 2. Estimated pole model with centroid $\mathbf{c}$, principal axis $\mathbf{a}$, and linearly tapering radius profile. The coordinate frame is shifted by the LiDAR point’s center for improved visual clarity.

Relative to the planar residuals, the comparison of measured and expected radius yields here the implicit pole misclosure

$$r_{\mathrm{po}} = r - r_{\exp}(h). \tag{14}$$

In analogy to the barycentric test, a LiDAR point has to fulfill the height constraint before the pole-specific thresholding $|r_{\mathrm{po}}| \le \tau_{\mathrm{po}}$ is performed.

3.5 Map generation

A dedicated measurement campaign was conducted in an urban environment. Two platforms were attached to a van’s roof: one consisted of a 64-line HESAI LiDAR, GNSS antenna, and low-/high-cost IMUs for the observation stream and evaluation process, and the other was a RIEGL Mobile Mapping System (MMS) used to build the HR map. All devices were hardware-synchronized by a GPS-disciplined PPS; GPS time served as the common time base. Sensor frames were calibrated in a laboratory using an installed plane configuration and a micrometer-level accurate laser tracker (Ernst et al., 2022); the resulting LiDAR extrinsics $\mathbf{R}_s^b$ and $\mathbf{t}_s^b$ are used in (8).

Several rounds of the trajectory employed in this work were driven to create a dense point cloud. Individual sections were aligned and georeferenced using extracted features and positioning data according to (Brenner, 2016). The fused cloud, represented in UTM with ellipsoidal heights, was manually annotated into *rest*, *building*, *street*, and *pole*, which are illustrated in Figure 3. In this work, *pole* denotes man-made, predominantly vertical and approximately cylindrical objects that can be assumed static in the map, such as masts of lamps, street signs, and traffic lights. As demonstrated in the figure attached objects to the mast are explicitly assigned to the *rest* class, enabling accurate pole parameter estimation (13).

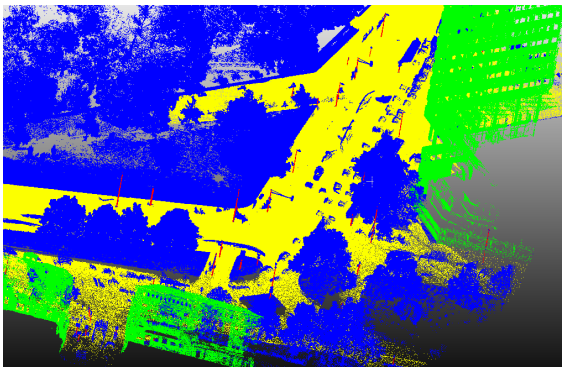

Figure 3. Oblique view on a junction area. The annotated point-cloud follow this scheme: blue = *rest*, green = *building*, yellow = *street*, red = *pole*.

The majority part of the HR map consists of planes (ground and facades), which were extracted via strict region growing: 5 cm distance threshold, limited expansion radius per iteration, and a tight normal-angle criterion; each plane was polygonalized and tessellated into triangles. Approximately 500 poles along the driven trajectory were estimated in two stages: first, the center $\mathbf{c}$ and principal axis $\mathbf{a}$ were identified using the cylinder-fitting method of (Eberly, 2003); subsequently, the radius profile $r_{\exp}(h) = r_0 + \kappa h$ and the bounds $h_{\min}, h_{\max}$ were modeled by robust fitting with a Huber M-estimator. For efficient map access, a bounding volume hierarchy (BVH) indexed plane triangles and pole primitives; queries return a compact candidate set for the assignment and residual tests in (9) and (14).

3.6 Geometry-aware subsampling

Only a small subset of LiDAR returns is actually needed for a strong update: points that support the map constraints we employ (building/ground planes and pole masts). We therefore select key points online instead of forwarding dense scans to the filter. A short moving window over two consecutive LiDAR slices (aligned with IMU epochs $k-1$ and k), covering approximately 20 % of the 10 Hz full-rotation scan, provides local spatial context during selection.

(1) Windowing and voxel prefilter. Let $\mathcal{P}_k = \{\mathbf{p}_i^s, t_i \mid t_{k-1} < t_i \leq t_k\}$ be the points in the current slice. A coarse voxel grid with cell size

$$\mathbf{v}_{\text{vox}} = (0.15\ \text{m},\ 0.15\ \text{m},\ 0.10\ \text{m})$$

keeps a single representative per occupied voxel, yielding $\mathcal{V}_k \subset \mathcal{P}_k$. The horizontal resolution of 15×15 cm suppresses redundant ground points close to the scanner while still resolving curb and facade structure at the scale relevant for localization (smallest planar map surfaces can still be matched with several points), and the smaller 10 cm vertical spacing maintains most of the nearby observed pole structures. For temporal continuity, candidates from the previous slice are temporarily pooled: $\mathcal{C}_k = \mathcal{V}_k \cup \mathcal{V}_{k-1}$.

(2) Local features for plane/pole cues. For each $\mathbf{x}_i \in \mathcal{C}_k$, a radius-aware neighbor set $\mathcal{N}_i$ guarantees at least K neighbors for feature computation. From the covariance of $\mathcal{N}_i$ with eigenpairs $\lambda_1 \geq \lambda_2 \geq \lambda_3$ and $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$, we compute

$$L_i = \tfrac{\lambda_1 - \lambda_2}{\lambda_1} \quad \text{(linearity)}, \qquad P_i = \tfrac{\lambda_2 - \lambda_3}{\lambda_1} \quad \text{(planarity)}$$

$$\mathbf{n}_i = \mathbf{v}_3 \quad \text{(normal)}, \qquad \mathbf{a}_i = \mathbf{v}_1 \quad \text{(principle axis)}, \tag{15}$$

and a pole verticality score $V_{a,i}$ relative to the scanner's $\mathbf{z}$ axis $|\mathbf{a}_i \cdot \mathbf{z}| \in [0, 1]$.

(3) Geometry-aware FPS. Classical FPS maximizes spatial coverage by repeatedly picking the point farthest from the current sample set in Euclidean distance. We retain the same greedy structure but bias distances toward samples that strengthen our map constraints. For a selected set S and candidate i, we define the per-pair weight

$$w_{i|j} \;=\; 1 \;+\alpha\, w_{N_{i|j}} \;+\beta\, w_{P_{i|j}}, \tag{16}$$

with

$$w_{N_{i|j}} = (1 - |\mathbf{n}_i \cdot \mathbf{n}_j|)\, P_i, \qquad w_{P_{i|j}} = L_i\, V_{a,i}.$$

The geometry-aware distance to S is

$$D(i \mid S) = \min_{j \in S} w_{i|j}\, \|\mathbf{x}_i - \mathbf{x}_j\|, \tag{17}$$

and the next sample is

$$i^\star = \arg\max_{i \in \mathcal{C}_k \setminus S} D(i \mid S). \tag{18}$$

To obtain reproducible results we start the subsampling algorithm by selecting the point farthest from the centroid of $\mathcal{C}_k$. After each pick, we update the running distances via element-wise minima, exactly as in classical FPS. Because planes vastly outnumber poles, we then tuned the pair weights in (16) via a noise-seeded grid search, sweeping $\alpha \in 0.5, 1.0$ and $\beta \in 0.5, 1.0, 2.0$. For comparability we fixed the retained-point rate at 10 % and—solely for this sensitivity study—disabled the per-object outlier gate in (20) (s. MC protocol and experiments) to amplify any differences attributable to α and β. Using trajectory-averaged Root Mean Square Error (RMSE) per 6-DoF element, we observed only minor variations between each other (typically $\leq$ 5 %). According to the overall results, we adopt the balanced setting $\alpha = \beta = 1$ for all subsequent experiments. This choice preserves FPS coverage while consistently prioritizing points that are most informative for the plane and pole residuals used in the filter update. Throughout the study, the subsampling rate roughly denotes the fraction of selected key points after the two-stage pipeline (voxel pre-filtering + geometry-aware FPS) relative to the maximum possible LiDAR observations available in the time window.

(4) Bookkeeping and carry-over. From the indices returned by (18), only those originating in the *current* slice $(t_{k-1}, t_k]$ are passed to the filter at epoch k. For the next epoch, we carry over $\mathcal{V}_k$ (the current voxel-reduced set) as context and discard all earlier information. This keeps selection lightweight while consistently supplying plane- and pole-informative key points to the ESKF.

3.7 Sensor measurements and stochastic model for MC simulations

This subsection specifies the data streams, timing, and noise models used in the MC evaluation that underpins our ablation studies on subsampling and pole integration.

Data sources. In the conducted measurement campaign the observing platform's key components were a HESAI Pandar

Figure 4. Reference trajectory projected onto a satellite base map. The route forms a rectangle with mostly two-lane roadways; green and red markers denote start and end points, respectively.

64 LiDAR, Multi-GNSS antenna plus receiver and Microstrain IMU. The post-processed reference trajectory was computed from a commercial solution using an additional high-grade iPRENA IMU and the multi-GNSS data with nearby reference station; given the two-lane urban setting and equipment, we expect only single-digit centimeter deviations. The 2.3 km rectangular course was driven in approximately 567 s, including an initial 0.5 s stationary phase and several stops at red lights (see Fig. 4).

Realistic LiDAR observations are then simulated by ray-casting from the sensor origin along this trajectory into the HR map. To account for map incompleteness (in particular missing or simplified objects), only rays whose intersections lie sufficiently close to actual campaign measurements are retained. This consistency check results in a simulated LiDAR stream containing 78.6 % of the original points; the remaining fraction is attributed mainly to dynamic or unmapped structures. The geometry-aware subsampling scheme is applied to this noise-free point cloud so that the MC variation is confined to sensor noise realizations rather than to changes in the set of map-consistent points.

The tightly-coupled reference trajectory is available at 10 Hz. The data is de-spiked by Hampel filter and smoothed before cubic splines $\mathcal{S}^3(\mathcal{T}) \in C^2$ upsample each pose component to 100 Hz, representing the evaluation trajectory. Subsequently, the signal is well-behaved to derive high-frequency IMU observations by differentiation, matching the effective observation frequency of the Microstrain 3DM-GX3-15 in the real-world experiment. For simulated GNSS positions we also followed the selected 10 Hz rate.

Measurement noise models. LiDAR measurements are modeled in polar form (range, elevation, azimuth). Stochastic properties are adopted from variance component estimation on a Velodyne VLP-16 as a conservative proxy for the HESAI Pandar 64. Independent, zero-mean Gaussian noise is added with standard deviations

$$\sigma_d = 8.5\ \text{mm}, \quad \sigma_E = 0.0296^\circ, \quad \sigma_{Az} = 0.0485^\circ,$$

before conversion to Cartesian coordinates.

IMU accelerations and turn rates are perturbed by zero-mean Gaussian noise and biases (initial bias and instability). The parameters are

$$\sigma_{a_n} = 4.91{\cdot}10^{-4}{\cdot}3\ \text{m/s}^2\sqrt{\text{s}}, \qquad \sigma_{\omega_n} = 3.49{\cdot}10^{-4}{\cdot}3\ \text{rad/s}\sqrt{\text{s}},$$

$$\sigma_{a_0} = 2 \cdot 10^{-3} \cdot g \cdot 3, \qquad \sigma_{\omega_0} = \deg(0.25) \cdot 3\ \text{rad/s},$$

$$\sigma_{a_w} = 1 \cdot 10^{-3} \cdot g \cdot 3, \qquad \sigma_{\omega_w} = \deg(18)/3600 \cdot 1\ \text{rad/s},$$

for accelerometer/gyroscope white-noise densities ($\sigma_{a_n}, \sigma_{\omega_n}$), initial biases ($\sigma_{a_0}, \sigma_{\omega_0}$) and bias instabilities ($\sigma_{a_w}, \sigma_{\omega_w}$). The factor 3 (except σ_{ω_w}) was chosen according to a prior calibration report of the sensor employed in the measurement campaign. The baseline values reflect the manufacturer’s datasheet.

The low-frequency GNSS positions are modeled with isotropic Gaussian noise of $\sigma_{\text{GNSS}} = 1$ m in the superordinate frame. This uncertainty signifies the relatively minor importance of GNSS in the fusion scheme. It represents a pragmatic compromise between open junctions and narrow, building-lined street canyons in urban environments. To better reflect real GNSS behavior, one could analyze the campaign data with a ray-tracing environment to label epochs by propagation condition—line-of-sight (LOS), shadowed/attenuated LOS, specular multipath, and non-LOS/diffuse—and then fit condition-specific stochastic models; since LiDAR–map matching carries most of the update, this refinement is secondary but would tighten GNSS weighting in mixed scenes. The calibrated lever arm between the antenna and the IMU is still accounted for in the measurement model.

MC protocol and experiments. The MC evaluation is formulated as a comparative simulation study to quantify the incremental pose-estimation benefit of adding pole constraints rather than to claim absolute real-world localization performance. To isolate this effect, only the sensor noise is varied while the map, nominal trajectory, and nominal sensor observations are kept fixed: for each run, the same set of ray-cast and subsampled LiDAR points is reused, but new polar LiDAR noise is drawn for these points, and different IMU and GNSS noise realizations (white noise, bias noise, and bias instability) are sampled along the reference trajectory. The simulated results should therefore be interpreted as best-case performance under this presented setup, shifting the focus to relative performance differences.

LiDAR–to–map assignments are controlled by pessimistic thresholds $\tau_{\text{pl}} = \tau_{\text{po}} = 0.15$ m and an adaptive, per-object outlier gate (cf. (Moftizadeh, 2024)), ensuring high inlier recall under pose uncertainty, and sensor noise while still suppressing wrong correspondences. Let $r_{\text{pl},i}$ and $r_{\text{po},i}$ denote the plane and pole misclosures, and $\mathcal{A}_e^{(k)}$ the set of points assigned in epoch k to triangle or pole e. The mean absolute misclosure for object e in epoch k is

$$\mu_e^{(k)} = \frac{1}{|\mathcal{A}_e^{(k)}|} \sum_{i \in \mathcal{A}_e^{(k)}} \left| r_{e,i}^{(k)} \right|. \tag{19}$$

Robust inliers are defined by

$$\mathcal{I}_e^{(k)} = \left\{ i \in \mathcal{A}_e^{(k)} \;\middle|\; \left| r_{e,i}^{(k)} \right| \le 1.5\, \mu_e^{(k)} \right\}, \tag{20}$$

and only points in $\mathcal{I}_e^{(k)}$ contribute to the implicit update. At each epoch, LiDAR observations are reduced in two stages,

starting with voxel-grid prefiltering. The MC study is structured in two experiments.

In Experiment 1, we sweep the subsampling rate over several retained fractions and use the geometry-aware FPS with both plane and pole terms active. The functional model includes ground, facade, and pole constraints. A configuration with 50 % retained points is used as reference, since it reproduces the trajectory obtained from the complete LiDAR point cloud to negligible sub-millimeter and sub-milliradian differences, while reducing the end-to-end runtime of a single MC simulation by 40 %, saving approximately 8 hours. Runtime was measured on an AMD EPYC 9554P CPU (64C/128T, 3.10 GHz) server. The extended ESKF is developed in Matlab with the focus on ease of use; however, the proposed subsampling and implicit-update scheme is in principle compatible with real-time operation given an optimized implementation. For each rate, a set of MC runs is evaluated per 6-DoF element, using the RMSE metric, to identify the smallest rate that maintains comparable pose accuracy.

In Experiment 2, the subsampling rate is fixed to the operating point identified in Experiment 1, and the effect of pole-like structures is examined. Two variants are compared: (i) the planes-only filter driven by the same geometry-aware FPS used in **Experiment 1**, where the subsampler still allocates points to poles, and (ii) the same filter combined with a plane-focused subsampler obtained by disabling the pole-likeness term (setting $\beta = 0$ in (16)). For each configuration, a significant amount of MC runs are performed and the pose accuracy is summarized over the complete course and selected segments to evaluate the localization performance in depth.

4. Results

In the first experiment we establish a point size-accuracy relation by sweeping the subsampling rate under voxel-grid prefiltering and geometry-aware FPS, using a 50 % retained-point configuration as baseline. Based on 10 simulations we identify the tipping point, followed by quantifying the effect of added pole-like structures relative to the plane-based functional model. The second experiment statistics aggregate 40 MC runs to representatively determine the effect for the full trajectory and we analyze statistically significant segments.

4.1 Experiment 1: Subsampling–accuracy trade-off

We identify the smallest subsampling rate that maintains the full–point cloud accuracy by sweeping over 50, 25, 20, 15, 10 and 5 % to identify the efficiency sweet spot for our simulation environment. A control run suggested that the pose difference to the full LiDAR point cloud is within sub-millimeters and sub-milliradian, therefore, we treat 50 % as a faithful proxy. For each retained fraction we compute, per pose element, the epoch-wise RMSE and then take the across–trajectory average. The tipping point is heuristically defined as the lowest rate whose mean RMSE for all 6-DoFs remains within an approximate 5 % window from the 50 % reference. We visualize the achieved accuracy by table 1, showing the ΔRMSE scores in percent.

The baseline (50 %) achieves mean pose RMSEs of 3.35, 3.14, 3.57 mm (F–L–U) and 0.28, 0.20, 0.18 mrad (roll–pitch–yaw), resembling the full–cloud performance. As the subsampling rate decreases from 25 % down to 10 %, the relative deviations in Table 1 remain within a few percent for all pose components, whereas 5 % shows a clear degradation, particularly in the lateral and yaw channels. We therefore select 10 % as the operating point for the subsequent analysis, and repeat this configuration until a total of 40 MC runs are available for the in-depth comparison carried out in Experiment 2. At equal implementation, this 10 % setup reduces the ESKF runtime by 70 % relative to the 50 % baseline (or 82 % relative to the complete LiDAR data), suggesting near-linear cost in the number of correspondences despite the computationally expensive point-to-map association and filter update step. The constant improvement of the up and roll component can be assigned to fewer associations of points that were generated from nearby car tires and falsely assigned to the adjacent ground planes - the cause of this effect is visible to a small degree in Fig. 3; elongated vehicle shapes in the map result in artifacts in the simulated LiDAR generation as they are close to real measurements. To eliminate this behavior, a class-agnostic filtering scheme has to be applied in the LiDAR generation. For context, a simple voxel-grid prefiltering followed by random sampling to a 10 % retained fraction yields trajectory-average performance that falls between the 10 and 5 % results in Table 1, underscoring the benefit of the proposed geometry-aware FPS in the examined scenario. The selected subsampling rate according to the accuracy condition may also only be valid for the employed HESAI 64-line LiDAR within the MC simulation and differ, probably towards a higher fraction, for the real experiment.

Table 1. Relative RMSE deviation (ΔRMSE [%]) to the 50 % subsampling baseline for each pose element in the F–L–U body frame. Positive scores denote higher mean deviation and vice-versa.

Pose element	**25%**	**20%**	**15%**	**10%**	**5%**
Longitudinal (x_F)	-0.4	-1.3	-1.9	-1.3	+3.8
Lateral (y_L)	-0.4	+1.9	+3.4	+5.1	+13.6
Up (z_U)	-2.0	-2.7	-4.0	-5.2	-6.0
Roll (ϕ_F)	-0.1	-0.6	-1.8	-3.5	-4.4
Pitch (θ_L)	+0.5	+0.8	+1.4	+1.8	+2.3
Yaw (ψ_U)	+0.7	+2.2	+3.6	+4.8	+7.6

4.2 Experiment 2: Effect of pole constraints

In Experiment 2, the subsampling rate is fixed to the operating point of 10 % identified in Experiment 1, and the impact of pole-like structures on localization performance is quantified. Three variants are compared: (i) a planes + poles estimator configuration using the complete geometry-aware subsampling, (ii) a planes-only filter run on the same subsampled point cloud, and (iii) a planes-only filter that uses a plane-focused subsampling ($\beta = 0$ in (16)), so that no samples are explicitly devoted to pole-like structures. For each variant we present firstly the trajectory-averaged deviation in terms of median and empirical 95 % confidence interval for the 6-DoFs in Table 2.

On this level, all three variants exhibit small biases and narrow confidence intervals, and we can confirm that all configurations satisfy the AD requirements of 10 cm longitudinal/lateral and 3 $mrad$ heading over the entire trajectory and across all MC runs. The most visible trajectory-wide difference occurs in longitudinal translation (c.f. Table 2), where variant (i) achieves slightly tighter 95 % tails than the plane-only baselines, whereas in lateral direction variants (ii)/(iii) show a marginally smaller lower tail. For z_U, roll, and pitch the overall spread and median errors are very similar across variants;

Table 2. Experiment 2: Trajectory-averaged deviation of variant (i) - (iii) described by median and 95 % confidence interval for the 6-DoFs in the F–L–U body frame over 40 MC runs at 10 % subsampling rate. Translational errors are in [mm], rotational errors in [mrad].

Var.	Stat	**Translation** [mm]			**Rotation** [mrad]		
		x_F	y_L	z_U	ϕ_F	θ_L	ψ_U
i	Med.	-0.43	-0.19	-3.07	-0.08	0.06	0.01
	$CI_{2.5}$	-6.12	-5.26	-5.13	-0.48	-0.29	-0.36
	$CI_{97.5}$	5.13	4.82	-1.00	0.31	0.40	0.37
ii	Med.	-0.41	0.05	-3.07	-0.08	0.06	0.02
	$CI_{2.5}$	-6.32	-4.78	-5.11	-0.48	-0.28	-0.34
	$CI_{97.5}$	5.38	4.78	-1.02	0.31	0.39	0.37
iii	Med.	-0.43	0.10	-3.07	-0.09	0.06	0.02
	$CI_{2.5}$	-6.40	-4.81	-5.14	-0.48	-0.28	-0.34
	$CI_{97.5}$	5.39	4.84	-1.02	0.31	0.39	0.38

Table 3. Paired t-tests on trajectory-averaged RMSE: mean difference $\overline{\Delta}_{\mathrm{Var},i} = \mathrm{RMSE}_{\mathrm{Var}} - \mathrm{RMSE}_{(i)}$ for variants (ii) and (iii) relative to (i). Translation in [mm], rotation in [μrad]. An asterisk marks $p_{\mathrm{Holm}} < 0.05$.

	Translation [mm]			**Rotation** [μrad]		
	x_F	y_L	z_U	ϕ_F	θ_L	ψ_U
$\overline{\Delta}_{(\mathrm{ii},i)}$	0.32^*	0.99^*	-0.01^*	1.52^*	-0.56	7.81^*
$\overline{\Delta}_{(\mathrm{iii},i)}$	0.28^*	0.98^*	0.00	1.99^*	-0.14	7.78^*

the small irregularities in the up component can be plausibly attributed to non-uniform ground coverage and occasional mis-associations of LiDAR points generated on elongated vehicle shapes within the HR map to nearby ground planes, rather than to systematic differences between the functional models. In yaw, poles should in principle provide heading information by constraining the projected radial vector around the mast, but this leverage is diluted in practice by the dominance of plane updates and by imperfect assignment of LiDAR points to the correct azimuth on the cylindrical surface.

Because the discussed metrics alone may hide consistent but small global trends, we further applied paired t-tests on trajectory-averaged RMSEs per MC run. For each DoF, we formed per-run mean RMSE over time and tested the null hypothesis of zero mean difference between variant (ii) and (i), and between (iii) and (i), using within-run pairing and Holm correction over the 6-DoFs. Table 3 allows to see the overall difference in subsampled and employed map geometries more clearly.

This analysis reveals that both plane-only variants (ii) and (iii) have significantly higher trajectory-averaged RMSE than (i) in the most critical channels for urban driving, namely longitudinal x_F, lateral y_L, and yaw ψ_U, with adjusted p-values well below 0.05 ($\leq 10^{-15}$). Differences in roll are statistically significant but numerically negligible, pitch shows no significant difference, and the up component appears slightly favorable for (ii) and indistinguishable for (iii), which is consistent with the above interpretation as a ground-coverage artifact. Overall, the paired tests support a systematic global advantage of the planes+poles configuration (i), even though the raw medians and CIs in Table 2 suggest only subtle differences.

To refine this trajectory-aggregated view, we perform a local

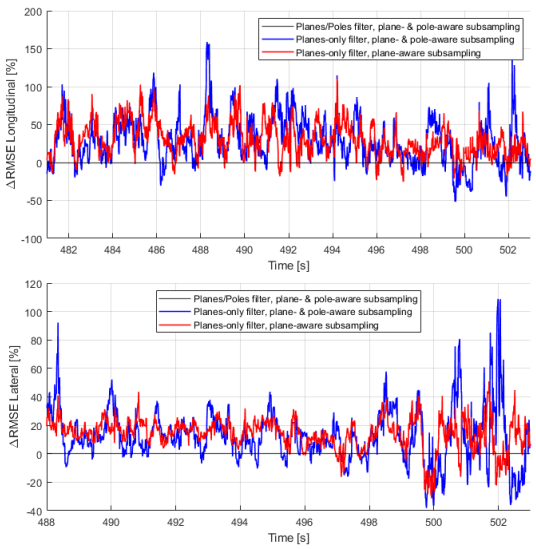

Figure 5. Relative RMSE deviation (ΔRMSE [%]) for the largest stressed segment to the planes + poles estimator with geometry-aware subsampling. The blue and red graphs represent planes-only filter model with and without pole-specific subsampling.

segment analysis using 1 s sliding windows to highlight where the filter configurations differ most. A segment is classified as significantly different if one variant shows a smaller median error in at least 80 % of the underlying epochs and, over more than 60 % of the window, both bounds of its empirical 95 % confidence interval are closer to zero. Under this criterion, detections appear only in the translational components. Variant (i) dominates in longitudinal and lateral directions and wins more often in height; most intervals lie just above the decision thresholds, but one merged segment against (ii) of about 7 s and one against (iii) of about 5.5 s stand out. A representative stressed segment is located immediately before the right turn through an open junction shown in Figure 5, where facade planes are sparse while the number of visible poles per LiDAR rotation is above trajectory-average. This 23 s segment is shown for the longitudinal and lateral component in $\Delta RMSE$ in percentage to highlight the differences.

In this scene, the planes+poles estimator (i) avoids the large deviations in longitudinal and lateral RMSE that occur with the plane-only variants, which either ignore poles in the functional model or in the subsampling. Over the stressed window, variant (iii) shows a sequence-wide mean (ΔRMSE) of about (+27 %) relative to (i), while variant (ii) exhibits sporadic outliers exceeding (+100 %), illustrating how entirely discarding pole information can critically under-constrain the lateral pose. Short intervals where (ii) or (iii) temporarily outperform (i) do occur, but they are associated with very small absolute errors on the order of a few millimeters and are therefore less critical from an AD viewpoint than the rare but substantially larger ($\approx$ 2 cm) deviations observed when poles are not fully exploited. Towards the end of the shown segment, where the number of matched poles per LiDAR rotation grows to roughly 50, very dense pole observations can also amplify non-optimal point-to-cylinder azimuth assignments and local map imperfections: under pose error, a pole return that is physically located on one side of a pole may be associated with a different side of the mapped cylinder when the assignment is based only on minimum radial distance, thereby inducing an incorrect correc-

tion direction. This suggests that future work should model pole observations explicitly as cylindrical objects and form pole misclosures with respect to the expected azimuth in the local pole frame (given the predicted pose) rather than relying solely on the nearest-distance criterion; in this regime, the purely plane-based update of variant (iii) appears relatively more resilient, but still remains systematically less accurate on average than the pole-augmented estimator (i).

Finally, we compare the computational cost at the selected $10\ \%$ subsampling rate. Relative to the robust plane-only configuration (iii), integrating pole assignments and their residuals in variant (i) increases the total processing time for matching and state updates by approximately $17.4\ \%$ over the entire trajectory. Given that this overhead is moderate compared with the roughly $70\ \%$ time savings already achieved by subsampling (Experiment 1), and in light of the demonstrated global and local robustness gains, the planes+poles estimator with geometry-aware subsampling emerges as the most convincing configuration for subsequent real-world validation and deployment.

5. Conclusion and Outlook

This work addressed urban 6-DoF vehicle localization by fusing LiDAR, IMU, and GNSS in an ESKF against an HR map, achieving the $10\ cm$ lateral/longitudinal and $3\ mrad$ heading requirements for autonomous driving (Reid et al., 2019). We pursued two objectives: first, to demonstrate that task-aware, geometry-guided subsampling can substantially reduce LiDAR load without degrading filter accuracy, and second, to quantify the incremental benefit of pole-like landmarks as complementary constraints to building/ground planes. MC analyses with simulated sensor streams tied to a post-processed reference trajectory from a dedicated measurement campaign indicate that retaining only $10\ \%$ of the LiDAR points preserves trajectory-wide RMSE at single-digit-millimeter and sub-milliradian level while reducing end-to-end runtime by about $82\ \%$ relative to the full point cloud. The planes+poles configuration achieves statistically significant but globally modest improvements over plane-only variants, on the order of $\approx 1\ \mathrm{mm}$ in position and $\approx 8\ \mu\mathrm{rad}$ in orientation when aggregated over the full trajectory. At the same time, the local segment analysis shows consistent short-term benefits in plane-sparse junction scenes. Over a representative stressed segment, discarding pole information increases the mean $\Delta\mathrm{RMSE}$ by about $+27\ \%$. It can also trigger sporadic outliers that exceed $+100\ \%$. By contrast, the pole-augmented estimator suppresses these rare but substantially larger deviations more reliably. In the most critical cases, the resulting errors can otherwise reach up to approximately $2\ \mathrm{cm}$ when the lateral pose becomes under-constrained.

In this simulation setting, variability stems purely from sensor-noise realizations, while map geometry and nominal observation geometry remain fixed. Consequently, the results support predominantly relative comparisons rather than absolute conclusions about real-world accuracy. The central next step is therefore validation on real multi-sensor logs from the same measurement campaign, where additional challenges — including map imperfections, dynamic objects, platform miscalibrations, and environment-dependent sensor effects — must be explicitly addressed. In particular, real deployments will likely require adaptive association thresholds and gating that account for increased uncertainty and systematic offsets (e.g., covariance-aware thresholds rather than fixed τ values), to avoid premature correspondence rejection or incorrect assignments under larger pose errors. Moreover, the dense-pole failure mode observed in stressed segments highlights that nearest-distance point-to-cylinder assignment can become azimuth-ambiguous under pose error, assigning a return to the wrong side of the mapped cylinder and thereby inducing an erroneous correction. This motivates extending the pole update towards an explicit cylindrical observation model with azimuth-consistent misclosures in the local pole frame (conditioned on the predicted pose), instead of relying solely on minimum point-to-cylinder distance. To further economize computation, we plan to semantically inform subsampling by blending geometric scores with class probabilities (building/street/pole). We also plan to adapt the plane/pole weights (α, β) online to the scene context, drawing on task-adaptive or loss-aware FPS ideas where sampling is skewed toward points most informative for the downstream objective (Deng et al., 2023). Finally, real-data validation will enable a critical re-assessment of whether the presented algorithmic pipeline and sensor configuration satisfy the autonomous-driving thresholds under operational conditions, or whether additional sensing, map quality, or model refinements are required to meet these requirements with sufficient margin.

References

Ballardini, A. L., Fontana, S., Cattaneo, D., Matteucci, M., Sorrenti, D. G., 2021. Vehicle Localization Using 3D Building Models and Point Cloud Matching. *Sensors*, 21(16). https://www.mdpi.com/1424-8220/21/16/5356.

Brenner, C., 2009. Global localization of vehicles using local pole patterns. *Pattern Recognition, DAGM 2009*, Lecture Notes in Computer Science, 5748, Springer, 61–70.

Brenner, C., 2016. Scalable estimation of precision maps in a mapreduce framework. M. Ali, S. Newsam, M. Renz, G. Trajcevski, S. Ravada (eds), *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ACM, New York, NY, USA, 1–10.

Chen, G., Lu, F., Li, Z., Liu, Y., Dong, J., Zhao, J., Yu, J., Knoll, A., 2021. Pole-Curb Fusion Based Robust and Efficient Autonomous Vehicle Localization System With Branch-and-Bound Global Optimization and Local Grid Map Method. *IEEE Transactions on Vehicular Technology*, 70(11), 11283-11294.

Deng, C., Peng, C., Liu, J., Wang, G., 2023. Point Cloud Deep Learning Network Based on Balanced Sampling and Self-Convolution. *Sensors*, 23(2), 981.

Eberly, D., 2003. Fitting 3d data with a cylinder. Technical report, Geometric Tools, LLC, Redmond, WA. Created Feb 25, 2003; Last Modified Feb 12, 2008. Formerly available as CylinderFitting.pdf on geometrictools.com.

Endo, Y., Kamijo, S., 2023. Deep Voxelized Feature Maps for Self-Localization in Autonomous Driving. *Sensors*, 23(12).

Ernst, D., Vogel, S., Neumann, I., Alkhatib, H., 2022. Analyse unterschiedlicher Positionskombinationen zur intrinsischen und extrinsischen Kalibrierung eines Velodyne VLP-16. *AVN Allgemeine Vermessungs-Nachrichten*, 129(6), 244–252.

Ernst, D., Vogel, S., Neumann, I., Alkhatib, H., 2023. Error state kalman filter with implicit measurement equations for position tracking of a multi-sensor system with imu and lidar. *2023*

13th International Conference on Indoor Positioning and Indoor Navigation (IPIN), 1–6.

Ghallabi, F., MITTET, M.-A., EL-HAJ-SHHADE, G., Nashashibi, F., 2019. Lidar-based high reflective landmarks (hrl)s for vehicle localization in an hd map. *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 4412–4418.

Kim, M.-j., Kwon, O., Kim, J., 2023. Vehicle to Infrastructure-Based LiDAR Localization Method for Autonomous Vehicles. *Electronics*, 12(12). https://www.mdpi.com/2079-9292/12/12/2684.

Meyer, J., Nebiker, S., Schürmann, S., Ferrari, E., Ammann, M., 2023. EXPLOITING POLE-LIKE OBJECTS FROM CADASTRES FOR SUB-METRE ACCURATE INTEGRATED GEOREFERENCING OF LOW-COST MOBILE MAPPING SYSTEMS. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-1/W1-2023, 303–310. https://isprs-archives.copernicus.org/articles/XLVIII-1-W1-2023/303/2023/.

Moftizadeh, R., 2024. Advanced Particle Filtering for Vehicle Navigation based on Collaborative Information. Dissertation, Bayerische Akademie der Wissenschaften, München. Identical with: Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik der Leibniz Universität Hannover, ISSN 0174-1454, Nr. 397, Hannover 2024.

Qi, C. R., Yi, L., Su, H., Guibas, L. J., 2017. Pointnet++: deep hierarchical feature learning on point sets in a metric space. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS’17, Curran Associates Inc., Red Hook, NY, USA, 5105–5114.

Reid, T. G., Houts, S. E., Cammarata, R., Mills, G., Agarwal, S., Vora, A., Pandey, G., 2019. Localization Requirements for Autonomous Vehicles. *SAE International Journal of Connected and Automated Vehicles*, 2(3). http://dx.doi.org/10.4271/12-02-03-0012.

Roriz, R., Vieira, D. P., Gomes, J. B., Silva, E. S., 2024. A Survey on Data Compression Techniques for Automotive LiDAR Point Clouds. *Sensors*, 24(10), 3185.

Rusu, R. B., Cousins, S., 2011. 3d is here: Point cloud library (pcl). *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, 1–4.

Schaefer, A., Büscher, D., Vertens, J., Luft, L., Burgard, W., 2021. Long-term vehicle localization in urban environments based on pole landmarks extracted from 3-D lidar scans. *Robotics and Autonomous Systems*, 136, 103709. https://www.sciencedirect.com/science/article/pii/S0921889020305492.

Shoemake, K., 1985. Animating rotation with quaternion curves. *Proceedings of the 12th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH ’85, Association for Computing Machinery, New York, NY, USA, 245–254.

Solà, J., 2017. Quaternion Kinematics for the Error-State Kalman Filter. *arXiv preprint arXiv:1711.02508*. https://arxiv.org/abs/1711.02508.

Spangenberg, R., Golle, T., Zuther, S., Rojas, R., 2016. Pole-based localization for autonomous vehicles in urban scenarios using local grid maps. *2016 IEEE Intelligent Vehicles Symposium (IV)*, 1060–1065.

Süleymanoğlu, B., Soycan, M., Toth, C., 2024. 3D Road Boundary Extraction Based on Machine Learning Strategy Using LiDAR and Image-Derived MMS Point Clouds. *Sensors*, 24(2), 503.

Tzermia, C., Psomopoulos, F., Chatzistamatis, S., Daras, P., 2025. Comparative analysis of downsampling techniques for machine learning on cultural heritage objects. *Proceedings of the 20th International Conference on Computer Graphics Theory and Applications (GRAPP)*. to appear; preprint available via ACM Digital Library.

Wiesmann, L., Milioto, A., Chen, X., Stachniss, C., Behley, J., 2021. Deep Compression for Dense Point Cloud Maps. *IEEE Robotics and Automation Letters*, 6(2), 2060–2067. https://www.ipb.uni-bonn.de/wp-content/papercite-data/pdf/wiesmann2021ral.pdf.

Yin, Y., Xu, X., tao Huo, L., Li, Y., Zhang, L., 2023. Sensor Fusion of GNSS and IMU Data for Robust Localization via Smoothed Error State Kalman Filter. *Sensors*, 23(7), 3676.

You, B., Li, S., Huang, Y., Fan, R., Liu, M., 2023. A Simultaneous Localization and Mapping System Using the Iterative Error State Kalman Filter Judgment Algorithm for Global Navigation Satellite System. *Sensors*, 23(13), 6000.

Zhang, Y., Wang, L., Jiang, X., Zeng, Y., Dai, Y., 2022. An efficient LiDAR-based localization method for self-driving cars in dynamic environments. *Robotica*, 40(1), 38–55.