All-Weather sub-50-cm Radar-Inertial Positioning

Lakshay Narula, *Student Member, IEEE,* Peter A. Iannucci, *Member, IEEE,*
and Todd E. Humphreys, *Member, IEEE*

***Abstract*—Deployment of automated ground vehicles beyond the confines of sunny and dry climes will require sub-lane-level positioning techniques based on radio waves rather than near-visible-light radiation. Like human sight, lidar and cameras perform poorly in low-visibility conditions. This paper develops and demonstrates a novel technique for robust sub-50-cm-accurate urban ground vehicle positioning based on all-weather sensors. The technique incorporates a computationally-efficient globally-optimal radar scan batch registration algorithm into a larger estimation pipeline that fuses data from commercially-available low-cost automotive radars, low-cost inertial sensors, vehicle motion constraints, and, when available, precise GNSS measurements. Performance is evaluated on an extensive and realistic urban data set. Comparison against ground truth shows that during $60\,\mathrm{min}$ of GNSS-denied driving in the urban center of Austin, TX, the technique maintains 95^{th}-percentile errors below $50\,\mathrm{cm}$ in horizontal position and 0.5° in heading.**

***Index Terms*—Radar, IMU, localization, all-weather, positioning, automated vehicles, inertial sensing, GNSS, automotive**

I. Introduction

DEVELOPMENT of automated ground vehicles (AGVs) has spurred research in lane-keeping assist systems, automated intersection management [1], tight-formation platooning, and cooperative sensing [2], [3], all of which demand accurate (e.g., 50-cm at 95%) ground vehicle positioning in an urban environment. But the majority of positioning techniques developed thus far depend on lidar or cameras, which perform poorly in low-visibility conditions such as snowy whiteout, dense fog, or heavy rain. Adoption of AGVs in many parts of the world will require all-weather localization techniques.

Radio-wave-based sensing techniques such as radar and GNSS (global navigation satellite system) remain operable even in extreme weather conditions [4] because their longer-wavelength electromagnetic radiation penetrates snow, fog, and rain. Carrier-phase-differential GNSS (CDGNSS) has been successfully applied for the past two decades as an all-weather decimeter-accurate localization technique in open-sky conditions. Proprioceptive sensors such as inertial measurement units (IMUs) also continue to operate regardless of external conditions. Coupling a CDGNSS receiver with a tactical-grade inertial sensor, as in [5]–[8] delivers robust high-accuracy positioning even during the extended signal outages common in the urban environment, but such systems are far too expensive for widespread deployment on AGVs. Recent work has shown that 20-cm-accurate (95%) CDGNSS positioning is possible at low cost even in dense urban areas, but solution availability remains below 90%, with occasional long gaps between high-accuracy solutions [9]. Moreover, the global trend of increasing radio interference in the GNSS bands, whether accidental or deliberate [10], underscores the need for GNSS-independent localization: GNSS jamming cannot be allowed to paralyze an area's automated vehicle networks.

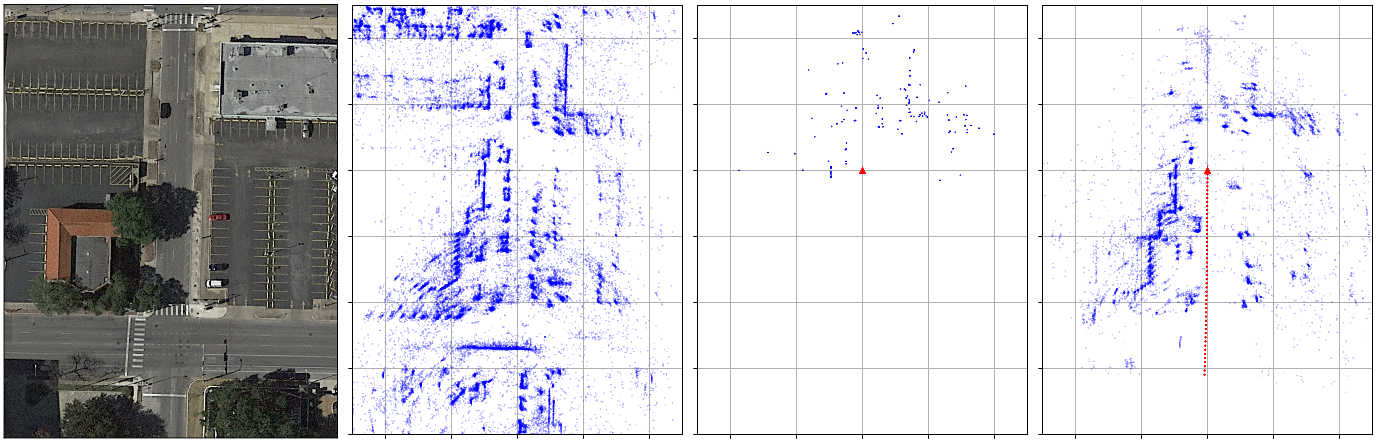
Clearly, there is a need for AGV localization that is low cost, accurate at the sub-50-cm level, robust to low-visibility conditions, and continuously available. This paper is the first to establish that low-cost inertial- and automotive-radar-based localization can meet these criteria.

Mass-market commercialization has brought the cost of automotive radar down enough that virtually all current production vehicles are equipped with at least one radar unit, which serves as the primary sensor for adaptive cruise control and automatic emergency braking. But use of automotive radar for localization faces the significant challenges of data sparsity and noise: an automotive radar scan has vastly lower resolution than a camera image or a dense lidar scan, and is subject to high rates of false detection (clutter) and missed detection. As such, it is nearly impossible to deduce semantic information or to extract distinctive environmental features from an individual radar scan. This is clear from Fig. 1c, which shows a sparse smattering of reflections from a single composite scan using three radar units. The key to localization is to aggregate sequential scans into a batch, as in Fig. 1d, where environmental structure is clearly evident. Even still, the data remain so sparse that localization based on traditional machine vision feature extraction and matching is not promising. Additionally, stable short-term odometry is a pre-requisite for aggregating radar scans, which in itself is a challenge when dealing with low-cost inertial sensing.

This paper proposes a two-step process for radar-based localization. The first is the mapping step: creation of a geo-referenced two-dimensional aggregated map of all radar targets across an area of interest. Fig. 1b shows such a map, hereafter referred to as a radar map. The full radar map used throughout this paper, of which Fig. 1b is a part, was constructed with the benefit of a highly stable inertial platform so that a trustworthy ground truth map would be available against which maps generated by other techniques could be compared. But an expensive inertial system or dedicated mobile mapping vehicle is not required to create a radar map. Instead, it can be crowd-sourced from the very user vehicles that will ultimately exploit the map for localization. During periods of favorable lighting conditions and good visibility, user vehicles can exploit a combination of low-cost CDGNSS, as in [9], and GNSS-aided visual simultaneous localization and mapping, as in [11], to achieve the continuous decimeter-and-sub-degree-accurate

L. Narula is with the Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX, 78712 USA e-mail: lakshay.narula@utexas.edu.

P.A. Iannucci and T.E. Humphreys are with the Department of Aerospace Engineering & Engineering Mechanics, The University of Texas at Austin, Austin, TX, 78712 USA.



(a) (b) (c) (d)

Fig. 1. Panel (a) shows a satellite view of the environment being mapped with automotive radar. Panel (b) shows the generated radar map point cloud with vehicle pose obtained from a reference localization system. Note the repeating structure along the road side due to parked vehicles. An individual radar scan obtained during localization is shown in panel (c), along with the red triangle denoting vehicle location and heading. The scan is sparse and contains significant clutter, making it challenging to register to the prior map. Panel (d) shows a batch of radar scans during localization, with the red dots denoting the vehicle trajectory over the past five seconds. The batch captures the underlying structure which can be registered to the prior map.

geo-referenced position and orientation (pose) required to lay down an accurate radar map. In other words, the radar map can be *created* when visibility is good and then *exploited* at any later time, such as during times of poor visibility.

Despite aggregation over multiple vehicle passes, the sparse and cluttered nature of automotive radar data is evident from the radar map shown in Fig. 1b: the generated point cloud is much less dense and has a substantially higher fraction of spurious returns than a typical lidar-derived point cloud, making automotive-radar-based localization a significantly more challenging problem.

The second step of this paper's technique is the localization step. Using a combination of all-weather odometric techniques such as inertial sensing, radar odometry, and ground vehicle dynamics constraints, a sensor fusion filter continually tracks the changes in vehicle pose over time. Over the latest short interval (e.g., $5\,\mathrm{s}$), pose estimates from the filter are used to spatially organize the multiple radar scans taken over the interval and generate what is hereafter referred to as a batch of scans, or batch for short. Fig. 1d shows a five-second batch terminating at the same location as the individual scan in Fig. 1c. In contrast to the individual scan, some environmental structure emerges in the batch of scans, making robust registration to the map feasible. Even so, the localization problem remains challenging due to the dynamic radar environment: note the absence of parked cars on the left side of the street during localization. The batch of scans is matched against the prior map of the surroundings to estimate the pose offset of the batch from the truth. This pose offset is then applied as a measurement to the sensor fusion filter to correct odometric drift.

Contributions. This paper's overall contribution is a robust pipeline for all-weather sub-50-cm urban ground vehicle positioning. This pipeline incorporates a computationally-efficient correlation-maximization-based globally-optimal radar scan registration algorithm that estimates a two-dimensional translational and a one-dimensional rotational offset between a prior radar map and a batch of current scans. Significantly, the registration algorithm can be applied to the highly sparse and cluttered data produced by commercially-available low-cost automotive radars. Maximization of correlation is shown to be equivalent to minimization of the L^2 distance between the prior map and the batch probability hypothesis densities. The pipeline supports the construction of the radar registration estimate and optimally fuses it with inertial measurements, radar range rate measurements, ground vehicle dynamics constraints, and cm-accurate GNSS measurements, when available. A novel technique for online estimation of the vehicle center of rotation is introduced, and calibration of various other extrinsic parameters necessary for optimal sensor fusion is described.

This paper also presents a thorough evaluation of the positioning pipeline on the large-scale dataset described in [12]. Data from automotive sensors are collected over two $1.5\,\mathrm{h}$ driving sessions through the urban center of Austin, TX on two separate days specifically chosen to provide variety in traffic and parking patterns. Comparison with a post-processed ground truth trajectory shows that proposed pipeline maintains 95^{th}-percentile errors below $35\,\mathrm{cm}$ in horizontal position and $0.5°$ in heading during $60\,\mathrm{min}$ of GNSS-denied driving.

A preliminary version of this paper describing the radar scan registration algorithm was published in [13]. The current version develops and tests a full sensor fusion pipeline that includes the radar batch estimation as a sub-component.

Organization of the rest of this paper. Sec. II establishes the significance of this contribution in view of the prior work in related fields. The radar batch-based pose estimation technique for the low-cost automotive radar sensor model is developed in Sec. III. Sec. IV describes the overall sensor fusion architecture involving inertial sensing, GNSS, motion constraints, and radar measurements. Implementation details and experimental results from field evaluation are presented in Sec. V, and Sec. VI provides concluding remarks.

II. Related Work

This section reviews a wide variety of literature on mapping and localization with radar and radar-inertial sensing. This includes prior work on point cloud alignment and image registration techniques, occupancy grid-based mapping and localization, random-finite-set-based mapping and localization, and inertial-aided mapping and localization.

Related work in point cloud alignment. A radar-based map can have many different representations. One obvious representation is to store all the radar measurements as a point cloud. Fig. 1b is an example of this representation. Localization within this map can be performed with point cloud registration techniques like the iterative closest point (ICP) algorithm. ICP is known to converge to local minima which may occur due to outlying points that do not have correspondences in the two point clouds being aligned. A number of variations and generalizations of ICP robust to such outliers have been proposed in the literature [14]–[20]. A few of these have been applied specifically to automotive radar data [15], [16]. But the technique in [15] is only evaluated on a $5\,\mathrm{min}$ dataset, while [16] performs poorly on datasets larger than $1\,\mathrm{km}$.

This paper steers away from ICP and its gradient-based variants because automotive radar data in urban areas exhibit another source of incorrect-but-plausible registration solutions which are not addressed in the above literature—repetitive structure, e.g., due to a series of parked cars, may result in multiple locally-optimal solutions within 2–$3\,\mathrm{m}$ of the globally-optimal solution. Gradient-based techniques which iteratively estimate correspondences based on the distance between pairs of points are susceptible to converge to such locally-optimal solutions. Accordingly, the batch-based pose estimator proposed in this paper is designed to approximate the globally-optimal solution.

In contrast to ICP and its variants, globally-optimal point cloud registration can be achieved by performing global point correspondence based on distinctive feature descriptors [21]–[23]. All of these works use a sophisticated mechanically-rotating radar unit that is not expected to be available on an AGV. Feature description and matching on the low-cost automotive radars used in this paper is likely to be fragile. Even when using the mechanically-rotating radar, [24] shows that a correlation-based approach, such as the one developed in this paper, outperforms other feature-descriptor-based point cloud methods.

Related work in image registration and occupancy grid techniques. Occupancy grid mapping and localization techniques have been traditionally applied for lidar-based systems, and recent work in [25], [26] has explored similar techniques with automotive radar data. In contrast to batch-based pose estimation described in this paper, both [25] and [26] perform particle-filter based localization with individual scans, as is typical for lidar-based systems. These methods were only evaluated on small-scale datasets collected in a parking lot, and even so, the reported meter-level localization accuracy is not sufficient for lane-level positioning.

Occupancy grid maps are similar to camera-based top-down images, and thus may be aligned with image registration techniques, that may be visual-descriptor-based [27], [28] or correlation-based [29]. Reliable extraction and matching of visual features, e.g., SIFT or SURF, is significantly more challenging with automotive radar data. Correlation-based registration is more robust [24], [29], and forms the basis of one of the components in this paper. In contrast to prior work [24], [29], this paper provides a probabilistic interpretation for the correlation operation. The mechanically-rotating radar of [24] allows correlation-based pose estimation based on a single scan of radar data. But for the low-cost automotive radars used in this paper, it becomes necessary to accumulate radar scans over time, which requires integration with other odometric sensors. This paper develops and demonstrates a complete sensor fusion pipeline around radar-based pose estimation and evaluates its performance on a large urban dataset.

Related work in random finite set mapping and localization. The occupancy grid representation commonly used in robotics is an approximation to the probability hypothesis density (PHD) function [30], [31]: a concept first introduced in the random finite set (RFS) based target tracking literature. Unsurprisingly, PHD- and RFS-based mapping and localization have been previously studied in [32]–[34]. In contrast to occupancy grid-based methods, techniques in [32]–[34] make the point target assumption where no target may generate more than one measurement in a single scan, and no target may occlude another target. However, in reality, planar and extended targets such as walls and building fronts are commonplace in the urban AGV environment. Mapping of ellipsoidal extended targets has recently been proposed in [35], but occlusions are not modeled and only simulation results are presented.

Related work in inertial-aided mapping and localization. This paper couples radar batch-based pose estimation with other all-weather automotive sensors such as IMU and GNSS. Inertial aiding has been widely applied in visual- and lidar-based mapping and localization [36]–[42]. This paper extends inertial-aiding to sensors that can operate under harsh weather conditions. Recently, radar measurements have been applied to constrain IMU position drift in [43]. Radar-inertial odometry for indoor robots has been proposed in [44], [45]. This paper is the first to integrate low-cost automotive radars with inertial sensing, GNSS, and ground vehicle dynamics for lane-level accurate positioning in challenging urban environments.

III. Radar-Batch-Based Pose Estimation

This section describes the formulation of the radar-batch-based pose estimation method introduced in this paper. It first details the statistical motivation behind the method, and then develops an efficient approximation to the globally-optimal estimator. The output of this estimator acts as one of the measurements provided to the overall localization system presented later in Sec. IV.

A. Pose Estimation using Probability Hypothesis Density

For the purpose of radar-based pose estimation, an AGVs environment can be described as a collection of arbitrarily shaped radar reflectors in a specific spatial arrangement. Assuming sufficient temporal permanence of this environment, radar-equipped AGVs make sample measurements of the underlying structure over time.

1) The Probability Hypothesis Density Function: A probabilistic description of the radar environment is required to set up radar-based pose estimation as an optimization problem. This paper chooses the PHD function [30] representation of the radar environment. The PHD at a given location gives the density of the expected number of radar reflectors at that location. For a static radar environment, the PHD $D(\boldsymbol{x})$ at a location $\boldsymbol{x} \in \mathcal{X}$ can be written as

$$D(\boldsymbol{x}) = I \cdot p(\boldsymbol{x})$$

where $\mathcal{X}$ is the set of all locations in the environment, $p(\boldsymbol{x})$ is a probability density function such that $\int p(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = 1$, and I, the intensity, is the total number of radar reflectors in the environment. For a time-varying radar environment, both I and $p(\boldsymbol{x})$ are functions of time. For $\mathcal{A} \subset \mathcal{X}$, the expected number of radar reflectors in $\mathcal{A}$ is given as

$$I_{\mathcal{A}} = \int_{\mathcal{A}} D(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$$

2) Estimating Vehicle State from PHDs: Let $D_{\mathrm{m}}(\boldsymbol{x})$ denote the "map" PHD function representing the distribution of radar reflectors in an environment, estimated as a result of mapping with known vehicle poses. During localization, the vehicle makes a radar scan, or a series of consecutive radar scans. A natural solution to the pose estimation problem may be stated as the vehicle pose which maximizes the likelihood of the observed batch of scans, given that the scan was drawn from $D_{\mathrm{m}}(\boldsymbol{x})$ [19]. This maximum likelihood estimate (MLE) has many desirable properties such as asymptotic efficiency. However, the MLE solution is known to be sensitive to outliers that may occur if the batch of scans was sampled from a slightly different PHD, e.g., due to variations in the radar environment between mapping and localization [18].

A more robust solution to the PHD-based pose estimation problem may be stated as follows. Let $\boldsymbol{\Theta}$ denote the vector of parameters of the rigid or non-rigid transformation $\mathcal{T}$ between the vehicle's prior belief of its pose, and its true pose. For example, in case of a two-dimensional rigid transformation, $\boldsymbol{\Theta} = [\Delta x, \Delta y, \Delta \phi]^{\top}$, where Δx and Δy denote a two-dimensional position and $\Delta \phi$ denotes heading. Also, let $D_{\mathrm{b}}(\boldsymbol{x}')$ denote a local "batch" PHD function estimated from a batch of scans during localization, defined over $\boldsymbol{x}' \in \mathcal{A} \subset \mathcal{X}$. This PHD is represented in the coordinate system consistent with vehicle's prior belief, such that $\boldsymbol{x}' = \mathcal{T}_{\boldsymbol{\Theta}}(\boldsymbol{x})$. Estimating the vehicle pose during localization is defined as estimating $\boldsymbol{\Theta}$ such that some distance metric between the PHDs $D_{\mathrm{m}}(\boldsymbol{x})$ and $D_{\mathrm{b}}(\boldsymbol{x}')$ is minimized.

This paper chooses the L^2 distance between $D_{\mathrm{m}}(\boldsymbol{x})$ and $D_{\mathrm{f}}(\boldsymbol{x}')$ as the distance metric to be minimized. As compared to the MLE which minimizes Kullback-Leibler divergence, L^2 minimization trades off asymptotic efficiency for robustness to measurement model inaccuracy [18]. The L^2 distance $d_{L^2}(\boldsymbol{\Theta})$ to be minimized is given as

$$d_{L^2}(\boldsymbol{\Theta}) = \int_{\mathcal{A}} \left(D_{\mathrm{m}}(\boldsymbol{x}) - D_{\mathrm{b}}(\mathcal{T}_{\boldsymbol{\Theta}}(\boldsymbol{x}))\right)^2 \mathrm{d}\boldsymbol{x}$$

For rigid two-dimensional transformations, it can be shown as follows that minimizing the L^2 distance between the PHDs is equivalent to maximization of the cross-correlation between the PHDs.

$$\begin{aligned}
\widehat{\boldsymbol{\Theta}} &= \underset{\boldsymbol{\Theta}'}{\operatorname{argmin}} \int_{\mathcal{A}} \left(D_{\mathrm{m}}(\boldsymbol{x}) - D_{\mathrm{b}}(\mathcal{T}_{\boldsymbol{\Theta}'}(\boldsymbol{x}))\right)^2 \mathrm{d}\boldsymbol{x} \\
&= \underset{\boldsymbol{\Theta}'}{\operatorname{argmin}} \Bigg[\int_{\mathcal{A}} D_{\mathrm{m}}^2(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \int_{\mathcal{A}} D_{\mathrm{b}}^2(\mathcal{T}_{\boldsymbol{\Theta}'}(\boldsymbol{x}))\mathrm{d}\boldsymbol{x} \\
&\qquad\qquad -2\int_{\mathcal{A}} D_{\mathrm{m}}(\boldsymbol{x}) D_{\mathrm{b}}(\mathcal{T}_{\boldsymbol{\Theta}'}(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}\Bigg]
\end{aligned}$$

Note that the first term above is fixed during optimization, while the second term is invariant under rigid transformation. As a result, the above optimization is equivalent to maximizing the cross-correlation:

$$\widehat{\boldsymbol{\Theta}} = \underset{\boldsymbol{\Theta}'}{\operatorname{argmax}} \int_{\mathcal{A}} D_{\mathrm{m}}(\boldsymbol{x}) D_{\mathrm{b}}(\mathcal{T}_{\boldsymbol{\Theta}'}(\boldsymbol{x}))\mathrm{d}\boldsymbol{x} \tag{1}$$

For differentiable D_{m} and D_{b}, the above optimization can be solved with gradient-based methods. However, the cross-correlation maximization problem in the urban AGV environment may have locally optimal solutions in the vicinity of the global minimum due to repetitive structure of radar reflectors. In applications with high integrity requirements, a search for the globally optimal solution is necessary. This paper notes that if the PHDs in (1) were to be discretized in $\boldsymbol{x}$, then the cross-correlation values can be evaluated exhaustively with computationally efficient techniques. Let $\boldsymbol{x}_{pq}$ denote the location at the (p, q) translational offset in discretized $\mathcal{A}$. Then

$$\widehat{\boldsymbol{\Theta}} = \underset{\boldsymbol{\Theta}'}{\operatorname{argmax}} \sum_{p=0}^{P-1} \sum_{q=0}^{Q-1} D_{\mathrm{m}}(\boldsymbol{x}_{pq}) D_{\mathrm{b}}(\lfloor \mathcal{T}_{\boldsymbol{\Theta}'}(\boldsymbol{x}_{pq}) \rceil) \tag{2}$$

where $\lfloor . \rceil$ denotes the nearest grid point in the discretized space.

The technique developed above relies on the PHDs D_{m} and D_{b}. The next subsections detail the recipe for estimating these PHDs from the radar observations.

B. Estimating the map PHD from measurements

This section addresses the procedure to estimate the map PHD $D_{\mathrm{m}}(\boldsymbol{x})$ from radar measurements. This paper works with an occupancy grid map (OGM) approximation to the continuous PHD function. In [31], it has been shown that the PHD representation is a limiting case of the OGM as the grid cell size becomes vanishingly small. Intuitively, let c_{pq} denote the grid cell region with center $\boldsymbol{x}_{pq}$, and let δc_{pq} denote the area of this grid cell, which is small enough such that no more than one reflector may be found in any cell. Let $p_{pq}(O)$ denote the occupancy probability of c_{pq}, and let $\mathcal{A}$ be defined as the region formed by the union of all c_{pq} whose centers $\boldsymbol{x}_{pq}$ fall within $\mathcal{A}$. Then, the expected number of radar reflectors $\mathbb{E}[|\mathcal{A}|]$ in $\mathcal{A}$ is given by

$$\begin{aligned}\mathbb{E}[|\mathcal{A}|] = \sum_{c_{pq}\in\mathcal{A}} p_{pq}(O) &= \sum_{c_{pq}\in\mathcal{A}} \frac{p_{pq}(O)}{\delta c_{pq}}\delta c_{pq} \\ &\triangleq \sum_{c_{pq}\in\mathcal{A}} \bar{D}(\boldsymbol{x}_{pq})\delta c_{pq} \\ &= \int_{\mathcal{A}} \bar{D}(\boldsymbol{x}_{pq})\mathrm{d}\boldsymbol{x}, \quad \text{as} \lim_{\delta c_{pq}\to 0}\end{aligned}$$

where $\bar{D}(\boldsymbol{x}_{pq}) \equiv \frac{p_{pq}(O)}{\delta c_{pq}}$ can be considered to be an approximation of the PHD $\bar{D}(\boldsymbol{x})$ for $\boldsymbol{x} \in c_{pq}$ since its integral over $\mathcal{A}$ is equal to the expected number of reflectors in $\mathcal{A}$.

The advantage of working with an OGM approximation of the PHD is two-fold: first, since the OGM does not attempt to model individual objects, it is straightforward to represent arbitrarily-shaped objects, and second, in contrast to the "point target" measurement model assumption in standard PHD filtering, the OGM can straightforwardly model occlusions due to extended objects.

At this point, the task of estimating $D_{\mathrm{m}}(\boldsymbol{x})$ has been reduced to estimating the occupancy probability of each grid cell in discretized $\mathcal{A}$. Each grid cell c_{pq} takes up one of two states: occupied (O) or free (F). Based on the radar measurement $\boldsymbol{z}_k$ at each time k, the Bernoulli probability distribution of such binary state cells may be recursively updated with the binary Bayes filter. In particular, let $\boldsymbol{z}_{1:k}$ denote all radar measurements made up to time k, and let

$$l_{pq}^{k}(O) \equiv \log \frac{p_{pq}(O \mid \boldsymbol{z}_{1:k})}{1 - p_{pq}(O \mid \boldsymbol{z}_{1:k})} \tag{3}$$

denote the *log odds ratio* of c_{pq} being in state O. Also define $l_{pq}^{0}(O)$ as

$$l_{pq}^{0}(O) \equiv \log \frac{p_{pq}(O)}{1 - p_{pq}(O)}$$

with $p_{pq}(O)$ being the prior belief on the occupancy state of c_{pq} before any measurements are made. With these definitions, the binary Bayes filter update is given by [46]

$$l_{pq}^{k}(O) = \log \frac{p_{pq}(O \mid \boldsymbol{z}_{k})}{1 - p_{pq}(O \mid \boldsymbol{z}_{k})} - l_{pq}^{0}(O) + l_{pq}^{k-1}(O) \tag{4}$$

where $p_{pq}(O \mid \boldsymbol{z}_k)$ is known as the *inverse* sensor model: it describes the probability of c_{pq} being in state O, given only the latest radar scan $\boldsymbol{z}_k$.

The required occupancy probability $p_{pq}(O \mid \boldsymbol{z}_{1:k})$ is easy to compute from the log odds ratio in (3). Observe that the inverse sensor model $p_{pq}(O \mid \boldsymbol{z}_k)$, in addition to the prior occupancy belief $p_{pq}(O)$, completely describes the procedure for estimating the OGM from radar measurements, and hence approximating the PHD. Adapting $p_{pq}(O \mid \boldsymbol{z}_k)$ to the characteristics of the automotive radar sensors, however, is not straightforward, and is discussed next.

C. Automotive Radar Inverse Sensor Model

This section addresses the challenge of adapting the inverse sensor model $p_{pq}(O \mid \boldsymbol{z}_k)$ to the measurement characteristics of automotive radar sensors. Fig. 2 shows a simplified radar scan $\boldsymbol{z}_k$ of an underlying occupancy grid. For clarity of exposition, four distinct categories of grid cells in Fig. 2 are defined below:

- *Type A*: Grid cells in the vicinity of a radar range-azimuth return.
- *Type B*: Grid cells along the path between the radar sensor and *Type A* grid cells.
- *Type C*: Grid cells in the "viewshed" of the radar sensor, i.e., within the radar field-of-view and not shadowed by another object, but not of *Type A* or *Type B*.
- *Type D*: Grid cells outside the field-of-view of the radar (*Type D1*) or shadowed by other objects closer to the radar (*Type D2*).

The inverse sensor model must choose a $p_{pq}(O \mid \boldsymbol{z}_k)$ value for each of these types of grid cells. In the following, the subscript pq is dropped for cleaner notation.

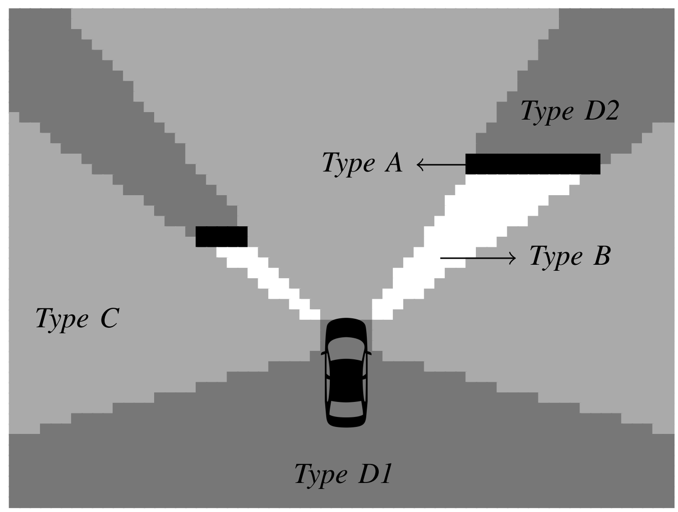


Fig. 2. Schematic diagram showing four types of grid cells.

1) Conventional Choices for the Inverse Sensor Model: Since $\boldsymbol{z}_k$ provides no additional information on *Type D* grid cells, the occupancy in these cells is conditionally independent of $\boldsymbol{z}_k$, that is

$$p^{D}(O \mid \boldsymbol{z}_k) = p(O)$$

where $p(O)$ is the prior probability of occupancy defined earlier in Sec. III-A.

Grid cells of *Type B* and *Type C* may be hypothesized to have low occupancy probability, since these grid cells were scanned by the sensor but no return was obtained. As a result, conventionally

$$p^{B}(O \mid \boldsymbol{z}_k) \leq p(O)$$

and

$$p^{C}(O \mid \boldsymbol{z}_k) \leq p(O)$$

Finally, grid cells of *Type A* may be hypothesized to have higher occupancy probability, since a return has been observed in the vicinity of these cells. Conventionally,

$$p^{A}(O \mid \boldsymbol{z}_k) \geq p(O)$$

In the limit, if the OGM grid cell size is comparable to the sensor range and angle uncertainty, or if the number of scans is large enough such that the uncertainty is captured empirically, only the grid cells that contain the sensor measurement may be considered to be of *Type A*.

2) Automotive Radar Sensor Characteristics: Intense clutter properties and sparsity of the automotive radar data complicate the choice of the inverse sensor model.

Sparsity. First, sparsity of the radar scan implies that many occupied *Type A* grid cells in the radar environment might be incorrectly categorized as free *Type C* cells. This can be observed in Fig. 1. As evidenced by the batch of scans in Fig. 1d, the radar environment is "dense" in that many grid cells contain radar reflectors. However, any individual radar scan, such as the one shown in Fig. 1c, suggests a much more sparse radar environment. As a result, a grid cell which is occupied in truth will be incorrectly categorized as *Type C* in many scans, and correctly as *Type A* in a few scans.

The sparsity of radar returns also makes it challenging to distinguish *Type C* cells from cells of *Type D2*. Since many occluding obstacles are not detected in each scan, the occluded cells of *Type D2* are conflated with free cells of *Type C*.

In context of the inverse sensor model, as the radar scan becomes more sparse

$$p^C(O \mid \boldsymbol{z}_k) \rightarrow p^D(O \mid \boldsymbol{z}_k)^-$$

where the superscript $-$ denotes a limit approaching from below. Intuitively, approaching $p^D(O \mid \boldsymbol{z}_k)$ implies that the measurement $\boldsymbol{z}_k$ is very sparse in comparison to the true occupancy, and thus does not provide much information on lack of occupancy.

Clutter. Second, there is the matter of clutter. The grid cells in the vicinity of a clutter measurement may be incorrectly categorized as *Type A*, and the grid cells along the path between the radar and clutter measurement may be incorrectly categorized as *Type B*.

In context of the inverse sensor model, as the radar scan becomes more cluttered

$$p^B(O \mid \boldsymbol{z}_k) \rightarrow p^D(O \mid \boldsymbol{z}_k)^-$$
$$p^A(O \mid \boldsymbol{z}_k) \rightarrow p^D(O \mid \boldsymbol{z}_k)^+$$

where the superscript $+$ denotes a limit approaching from above.

3) A Pessimistic Inverse Sensor Model: The results presented in Sec. V are based on a pessimistic sensor model, such that $p^B(O \mid \boldsymbol{z}_k) = p^C(O \mid \boldsymbol{z}_k) = p^D(O \mid \boldsymbol{z}_k)$. This model assumes that the radar measurements provide no information about free space in the radar environment.

In particular, the inverse sensor model assumes

$$p^B(O \mid \boldsymbol{z}_k) = p^C(O \mid \boldsymbol{z}_k) = p^D(O \mid \boldsymbol{z}_k) = p(O) = 0.1$$

and

$$p^A(O \mid \boldsymbol{z}_k) = 0.2$$

D. Estimating the batch PHD from measurements

The procedure for generating an approximation to $D_{\mathrm{b}}(\boldsymbol{x}')$ from a batch of radar measurements is identical to the procedure for generating $D_{\mathrm{m}}(\boldsymbol{x})$ from mapping vehicle data, except that precise, absolute location and orientation data is not available during localization. Instead, pose estimates from the sensor fusion filter described in Sec. IV are used to estimate the relative locations and orientations of each radar scan in the batch, and the scans are transformed into a common coordinate frame before updating the occupancy state of grid cells.

Once the map and batch PHDs have been approximated from radar measurements, the correlation-maximization technique developed in Sec. III-A can be applied to obtain the estimate $\widehat{\boldsymbol{\Theta}}$. This estimate is handed back to the sensor fusion filter as a pose offset measurement to constrain the odometric drift during absence of other sources of absolute localization, e.g., GNSS.

IV. State Estimation with Sensor Fusion

Thus far, Sec. III has developed the theory and implementation of the radar batch correlation measurement, which provides an estimate $\widehat{\boldsymbol{\Theta}}$ of the 3-DoF (degrees-of-freedom) pose offset relative to the prior map. This section details a localization pipeline that incorporates the batch measurement update along with an array of other automotive all-weather sensing modalities to track the full 6-DoF vehicle pose trajectory. The high-rate pose estimates from this pipeline are also used to spatially organize individual scans to form the batch of radar scans used in the batch correlation update.

The choice of sensors available for all-weather localization is limited to radio-frequency sensors such as GNSS and automotive radars, and to proprioceptive sensors such as IMUs and wheel encoders. Any additional domain knowledge, such as properties of ground vehicle dynamics, may also be combined with these sensor measurements.

The localization pipeline in this paper is developed around a low-cost MEMS IMU. Fig. 3 shows a block diagram of the overall pipeline. The error-state multiplicative extended Kalman filter (EKF) makes use of cm-accurate CDGNSS position measurements whenever such measurements are available, e.g., in clear-sky GNSS environments. Radial velocity and bearing measurements from low-cost automotive radars are combined with nearly-zero sideslip and vertical speed constraints of a ground vehicle to continually track and limit the errors in inertial navigation. Smoothed batches of radar scans are correlated with a prior map to limit odometric position drift during CDGNSS outages. The following subsections outline the formulation of the estimator, the nonlinear state dynamics, the various measurement models, and the necessary calibration procedures.

A. Sensor Platform & Coordinate Frames

To facilitate the discussion on measurement models and calibration, the sensor-instrumented vehicle and a few related coordinate frames are introduced here. An integrated perception platform called the *Sensorium*, shown schematically in Fig. 4, brings together the various low-cost automotive sensors considered in this paper. Many of these sensors provide measurements in their respective local frames, leading to a number of different coordinate frames that must be considered.

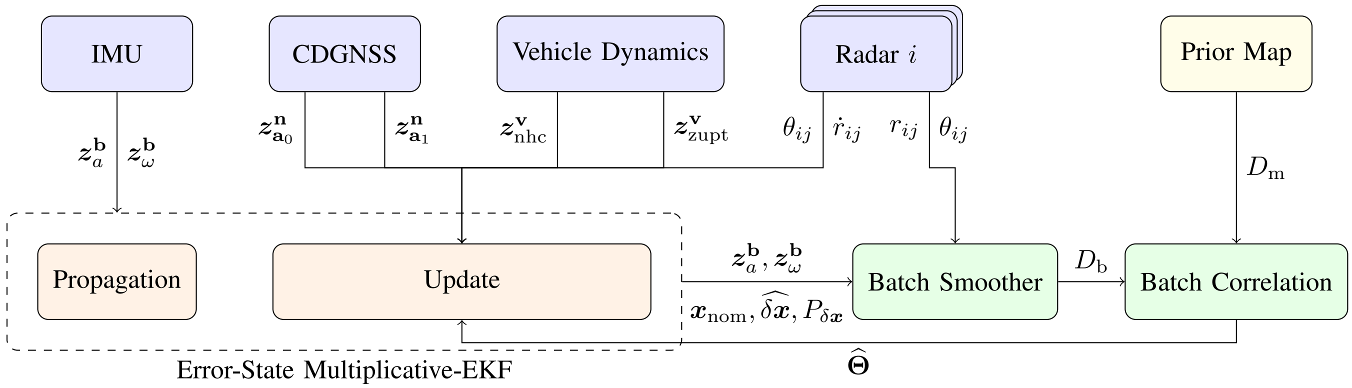


Fig. 3. Block diagram of the localization pipeline. A low-cost MEMS IMU provides high-rate specific force and angular rate measurements. The error-state multiplicative extended Kalman filter (EKF) makes use of cm-accurate CDGNSS position measurements whenever such measurements are available, e.g., in clear-sky GNSS environments. Radial velocity and bearing measurements from low-cost automotive radars are combined with nearly-zero sideslip and vertical speed constraints of a ground vehicle to continually track and limit the errors in inertial navigation. Smoothed batches of radar scans are correlated with a prior map to limit odometric position drift during CDGNSS outages.

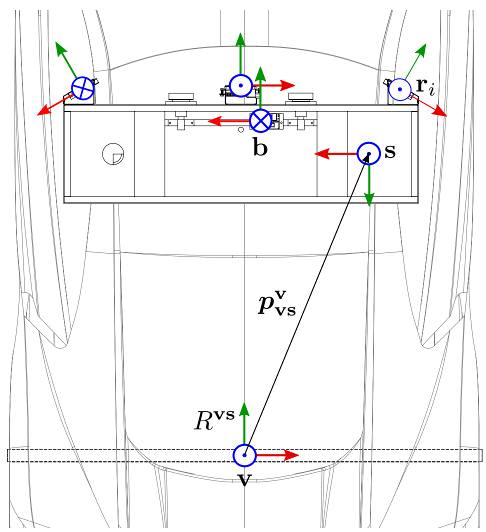


Fig. 4. The University of Texas Sensorium is an integrated platform for automated and connected vehicle perception research. It includes three automotive radar units, one electronically-scanning radar (ESR) and two short-range radars (SRR2s); stereo visible light cameras; automotive- and industrial-grade inertial measurement units (IMUs); a dual-antenna, multi-frequency software-defined GNSS receiver; and an internal computer. An iXblue ATLANS-C CDGNSS-disciplined inertial navigation system (INS) (not shown) is mounted at the rear of the platform to provide the ground truth trajectory. The vehicle frame $\mathbf{v}$ is located approximately at the center of the line connecting the rear axles.

The IMU *body frame*, denoted $\mathbf{b}$, is the frame defined by the IMU's accelerometer triad.

The *navigation frame*, denoted $\mathbf{n}$, is a local geographical reference frame, e.g., an ENU frame. The estimator wishes to track the pose trajectory of $\mathbf{b}$ with respect to $\mathbf{n}$.

The *radar frames*, denoted $\mathbf{r}_i$ for the ith radar, are local frames in which the radar sensors report range, range rate, and bearing to a number of targets.

The *vehicle frame*, denoted $\mathbf{v}$, is characterized by the direction in which the vehicle travels when the commanded steering angle is zero. This direction defines the y-axis of $\mathbf{v}$, as shown in Fig. 4. The origin of $\mathbf{v}$ is located at the center of rotation of the vehicle.

The *Sensorium frame*, denoted $\mathbf{s}$, is defined by the physical structure of the Sensorium. It is essentially a convenience reference frame in which the nominal lever arm and orientation between different sensors are available per the mechanical specifications of the Sensorium. The origin of $\mathbf{s}$ is arbitrarily chosen to be co-located with one of the GNSS antennas.

B. Error-State Filtering

The localization system of Fig. 3 estimates the following 16-element state vector:

$$\boldsymbol{x}_k = \left[\boldsymbol{p}_k^{\mathbf{n}}, \boldsymbol{v}_k^{\mathbf{n}}, \boldsymbol{q}_k^{\mathbf{nb}}, \boldsymbol{b}_{a,k}^{\mathbf{b}}, \boldsymbol{b}_{\omega,k}^{\mathbf{b}}\right]$$

where $\boldsymbol{p}_k^{\mathbf{n}}$ is the vector from $\mathbf{n}$ to $\mathbf{b}$ at time k expressed in $\mathbf{n}$, $\boldsymbol{v}_k^{\mathbf{n}}$ is the velocity of $\mathbf{b}$ relative to $\mathbf{n}$ at time k expressed in the $\mathbf{n}$ frame, $\boldsymbol{q}_k^{\mathbf{nb}}$ is the quaternion that rotates a vector from $\mathbf{b}$ to $\mathbf{n}$ at time k, and $\boldsymbol{b}_{a,k}^{\mathbf{b}}$ and $\boldsymbol{b}_{\omega,k}^{\mathbf{b}}$ are the accelerometer and gyroscope biases of the IMU at time k, expressed in $\mathbf{b}$.

Note that the vehicle orientation only has three effective degrees-of-freedom since $\boldsymbol{q}_k^{\mathbf{nb}}$ is constrained to be a unit quaternion. Enforcing such a constraint may result in a singular covariance matrix. This issue is typically dealt with an error-state filter [47] where the true state is split into a nominal-state vector

$$\boldsymbol{x}_{\text{nom},k} = \left[\tilde{\boldsymbol{p}}_k^{\mathbf{n}}, \tilde{\boldsymbol{v}}_k^{\mathbf{n}}, \tilde{\boldsymbol{q}}_k^{\mathbf{nb}}, \tilde{\boldsymbol{b}}_{a,k}^{\mathbf{b}}, \tilde{\boldsymbol{b}}_{\omega,k}^{\mathbf{b}}\right]$$

and an error-state vector $\delta\boldsymbol{x}_k$, related by the generalized addition operator $\oplus$ as follows:

$$\boldsymbol{x}_k = \boldsymbol{x}_{\text{nom},k} \oplus \delta\boldsymbol{x}_k$$

where the error-state vector $\delta\boldsymbol{x}_k$ is the minimal 15-element state representation denoted component-wise as follows:

$$\delta\boldsymbol{x}_k = \left[\delta\boldsymbol{p}_k^{\mathbf{n}}, \delta\boldsymbol{v}_k^{\mathbf{n}}, \boldsymbol{\eta}_k^{\mathbf{n}}, \delta\boldsymbol{b}_{a,k}^{\mathbf{b}}, \delta\boldsymbol{b}_{\omega,k}^{\mathbf{b}}\right]$$

The $\oplus$ operator corresponds to usual vector addition for the position, velocity, and bias states. For the orientation state, $\oplus$ is defined as

$$\begin{aligned} \boldsymbol{q}_k^{\mathbf{nb}} &= \tilde{\boldsymbol{q}}_k^{\mathbf{nb}} \oplus \boldsymbol{\eta}_k^{\mathbf{n}} \\ &= \exp_q\left(\frac{\boldsymbol{\eta}_k^{\mathbf{n}}}{2}\right) \odot \tilde{\boldsymbol{q}}_k^{\mathbf{nb}} \end{aligned}$$

where $\exp_q$ denotes the exponential map from $\mathfrak{so}(3)$ to $SO(3)$ [48], represented as a quaternion, and $\odot$ denotes quaternion multiplication. Note that $\boldsymbol{\eta}_k^{\mathbf{n}}$ is parametrized as an orientation deviation in $\mathbf{n}$. A similar formulation may be derived with the orientation deviation expressed in $\mathbf{b}$ [47].

The nonlinear error-state is tracked with an error-state EKF. Owing to the multiplicative orientation dynamics and update, this filter is sometimes referred to as the multiplicative-EKF [49].

C. State Dynamics

Inertial measurements, collectively denoted $\boldsymbol{u}_k$, are interpreted as control inputs during the state propagation step. The true-state dynamics function $f_k(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k)$ is modeled as

$$\begin{aligned} \boldsymbol{p}_{k+1}^{\mathbf{n}} &= \boldsymbol{p}_k^{\mathbf{n}} + T\boldsymbol{v}_k^{\mathbf{n}} + \frac{T^2}{2}\big(R_k^{\mathbf{nb}}\big(\boldsymbol{z}_{a,k}^{\mathbf{b}} - \boldsymbol{b}_{a,k}^{\mathbf{b}} - \boldsymbol{w}_{a,k}^{\mathbf{b}}\big) + \boldsymbol{g}^{\mathbf{n}}\big) \\ \boldsymbol{v}_{k+1}^{\mathbf{n}} &= \boldsymbol{v}_k^{\mathbf{n}} + T\big(R_k^{\mathbf{nb}}\big(\boldsymbol{z}_{a,k}^{\mathbf{b}} - \boldsymbol{b}_{a,k}^{\mathbf{b}} - \boldsymbol{w}_{a,k}^{\mathbf{b}}\big) + \boldsymbol{g}^{\mathbf{n}}\big) \\ \boldsymbol{q}_{k+1}^{\mathbf{nb}} &= \boldsymbol{q}_k^{\mathbf{nb}} \odot \exp_q\left(\frac{T}{2}\big(\boldsymbol{z}_{\omega,k}^{\mathbf{b}} - \boldsymbol{b}_{\omega,k}^{\mathbf{b}} - R_k^{\mathbf{bn}}\boldsymbol{\omega}_{\mathbf{e}}^{\mathbf{n}} - \boldsymbol{w}_{\omega,k}^{\mathbf{b}}\big)\right) \\ \boldsymbol{b}_{a,k+1}^{\mathbf{b}} &= \boldsymbol{b}_{a,k}^{\mathbf{b}} + \boldsymbol{w}_{\boldsymbol{b}_a,k}^{\mathbf{b}} \\ \boldsymbol{b}_{\omega,k+1}^{\mathbf{b}} &= \boldsymbol{b}_{\omega,k}^{\mathbf{b}} + \boldsymbol{w}_{\boldsymbol{b}_\omega,k}^{\mathbf{b}} \end{aligned}$$

where T is the propagation duration, $R_k^{\mathbf{nb}}$ is the rotation matrix representation of $\boldsymbol{q}_k^{\mathbf{nb}}$, $\boldsymbol{z}_{a,k}^{\mathbf{b}}$ and $\boldsymbol{z}_{\omega,k}^{\mathbf{b}}$ are the IMU specific force and angular rate measurements, respectively, $\boldsymbol{w}_{a,k}$ and $\boldsymbol{w}_{\omega,k}$ are the IMU specific force and angular rate white noise, respectively, $\boldsymbol{g}^{\mathbf{n}} \approx \left[0, 0, -9.8\,\mathrm{m\,s^{-2}}\right]$ is the acceleration due to gravity after compensation for the centripetal force due to earth's rotation, and $\boldsymbol{\omega}_{\mathbf{e}}^{\mathbf{n}}$ is the angular rate of the earth with respect to an inertial frame. The accelerometer and gyroscope biases are modeled as random walk processes driven by white noise $\boldsymbol{w}_{\boldsymbol{b}_a,k}^{\mathbf{b}}$ and $\boldsymbol{w}_{\boldsymbol{b}_\omega,k}^{\mathbf{b}}$, respectively, whose variances are derived from the IMU bias instability parameters [50].

The nominal-state dynamics function $f_{\mathrm{nom},k}(\boldsymbol{x}_{\mathrm{nom},k}, \boldsymbol{u}_k, \boldsymbol{w}_k)$ is similar to $f_k(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k)$:

$$\begin{aligned} \tilde{\boldsymbol{p}}_{k+1}^{\mathbf{n}} &= \tilde{\boldsymbol{p}}_k^{\mathbf{n}} + T\tilde{\boldsymbol{v}}_k^{\mathbf{n}} + \frac{T^2}{2}\left(\tilde{R}_k^{\mathbf{nb}}\left(\boldsymbol{z}_{a,k}^{\mathbf{b}} - \tilde{\boldsymbol{b}}_{a,k}^{\mathbf{b}}\right) + \boldsymbol{g}^{\mathbf{n}}\right) \\ \tilde{\boldsymbol{v}}_{k+1}^{\mathbf{n}} &= \tilde{\boldsymbol{v}}_k^{\mathbf{n}} + T\left(\tilde{R}_k^{\mathbf{nb}}\left(\boldsymbol{z}_{a,k}^{\mathbf{b}} - \tilde{\boldsymbol{b}}_{a,k}^{\mathbf{b}}\right) + \boldsymbol{g}^{\mathbf{n}}\right) \\ \tilde{\boldsymbol{q}}_{k+1}^{\mathbf{nb}} &= \tilde{\boldsymbol{q}}_k^{\mathbf{nb}} \odot \exp_q\left(\frac{T}{2}\left(\boldsymbol{z}_{\omega,k}^{\mathbf{b}} - \tilde{\boldsymbol{b}}_{\omega,k}^{\mathbf{b}} - \tilde{R}_k^{\mathbf{bn}}\boldsymbol{\omega}_{\mathbf{e}}^{\mathbf{n}}\right)\right) \\ \tilde{\boldsymbol{b}}_{a,k+1}^{\mathbf{b}} &= \tilde{\boldsymbol{b}}_{a,k}^{\mathbf{b}} \\ \tilde{\boldsymbol{b}}_{\omega,k+1}^{\mathbf{b}} &= \tilde{\boldsymbol{b}}_{\omega,k}^{\mathbf{b}} \end{aligned}$$

The error-state dynamics function $f_{\mathrm{err},k}(\delta\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k)$, is straightforwardly defined as

$$f_{\mathrm{err},k} \triangleq f_k \ominus f_{\mathrm{nom},k}$$

where $\ominus$ denotes a generalized subtraction operator similar to $\oplus$ defined earlier.

The linearized covariance propagation step of the EKF requires computation of the following Jacobians.

$$F_k = \frac{\partial \boldsymbol{f}_{\mathrm{err},k}(\delta\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k)}{\partial \delta\boldsymbol{x}_k}\bigg|_{\substack{\delta\boldsymbol{x}_k=0 \\ \boldsymbol{w}_k=0}} \tag{5}$$

$$G_k = \frac{\partial \boldsymbol{f}_{\mathrm{err},k}(\delta\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k)}{\partial \boldsymbol{w}_k}\bigg|_{\substack{\delta\boldsymbol{x}_k=0 \\ \boldsymbol{w}_k=0}} \tag{6}$$

This involves calculus of rotations. The interested reader is referred to [47], [48] for further details. The nontrivial sub-blocks of F_k and G_k are documented in Appendix A.

D. Measurement Models & Calibration

This section details the measurement models for the various measurements applied to the error-state EKF, along with the calibration procedures necessary for the application of these measurements.

1) Inertial Measurements: IMUs measure the specific force and angular rate experienced by $\mathbf{b}$ relative to an inertial frame. If the centripetal force due to earth's rotation is absorbed in $\boldsymbol{g}^{\mathbf{n}}$, then the accelerometer and gyroscope measurements $\boldsymbol{z}_{a,k}^{\mathbf{b}}$ and $\boldsymbol{z}_{\omega,k}^{\mathbf{b}}$, respectively, are modeled as

$$\begin{aligned} \boldsymbol{z}_{a,k}^{\mathbf{b}} &= R_k^{\mathbf{bn}}(\boldsymbol{a}_k^{\mathbf{n}} - \boldsymbol{g}^{\mathbf{n}}) + \boldsymbol{b}_{a,k}^{\mathbf{b}} + \boldsymbol{w}_{a,k}^{\mathbf{b}} \\ \boldsymbol{z}_{\omega,k}^{\mathbf{b}} &= \boldsymbol{\omega}_k^{\mathbf{b}} + R_k^{\mathbf{bn}}\boldsymbol{\omega}_{\mathbf{e}}^{\mathbf{n}} + \boldsymbol{b}_{\omega,k}^{\mathbf{b}} + \boldsymbol{w}_{\omega,k}^{\mathbf{b}} \end{aligned}$$

where $\boldsymbol{a}_k^{\mathbf{n}}$ is the true acceleration of the IMU in the $\mathbf{n}$ frame, which double-integrates to position deviation, and $\boldsymbol{\omega}_k^{\mathbf{b}}$ is the true angular rate of the IMU in the $\mathbf{n}$ frame, which integrates to orientation deviation. For low-quality IMUs, accelerometer and gyroscope scale factors may also need to be modeled. For the MEMS IMU used in this work, it was observed that modeling the scale factors did not yield any performance benefit.

The stochastic models for IMU white noise and random walk process are derived from the IMU specifications. In addition to such intrinsic calibration, extrinsic calibration of the IMU with respect to $\mathbf{s}$ is necessary for the application of other measurements expressed in $\mathbf{s}$. The vector $\boldsymbol{p}_{\mathbf{sb}}^{\mathbf{s}}$ from $\mathbf{s}$ to $\mathbf{b}$ is taken to be known from the mechanical specification since this is not strongly observable from the available measurements. It is, however, important to estimate any deviations from the mechanically specified orientation $\bar{\boldsymbol{q}}^{\mathbf{sb}}$ between $\mathbf{b}$ and $\mathbf{s}$, since even sub-degree errors in the IMU orientation relative to $\mathbf{s}$ may lead to substantial errors when multiplied with the lever arm to another sensor.

The orientation deviation of $\bar{\boldsymbol{q}}^{\mathbf{sb}}$ from truth, denoted $\boldsymbol{\eta}_{\mathbf{sb}}^{\mathbf{s}}$, can be effectively estimated when CDGNSS measurements from multiple antennas are available to the EKF, as will be discussed in Sec. IV-D2. Accordingly, the state vector $\delta\boldsymbol{x}_k$ is augmented with $\boldsymbol{\eta}_{\mathbf{sb}}^{\mathbf{s}}$ during clear-sky periods. It must be noted, however, that since the IMU is mounted near the line connecting the Sensorium's two GNSS antennas, only two of the three elements in $\boldsymbol{\eta}_{\mathbf{sb}}^{\mathbf{s}}$ are strongly observable. Any orientation deviation about the vector joining the two antennas is poorly unobservable, and must be constrained by construction. Also note that estimation of $\boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{sb}}$ only need be performed once as long as all sensors are rigidly mounted, and may not even be necessary if the mechanical tolerances are acceptably small.

2) CDGNSS Measurements: CDGNSS offers cm-accurate position measurements under all weather conditions, but typically offers reduced solution availability in deep urban environments. This paper takes the approach of incorporating CDGNSS measurements in the localization engine whenever they are available, while being capable of maintaining the required lane-level accuracy over long CDGNSS outages in deep urban canyons. In essence, the approach developed in this paper leverages CDGNSS for periodic or one-time intrinsic and extrinsic calibration of other on-board sensors, and relies on these sensors for accurate localization when CDGNSS is unavailable.

Signals captured from the two GNSS antennas on the Sensorium are processed together with those from a nearby reference station to provide nearly-independent three-dimensional position measurements of the antennas in the $\mathbf{n}$ frame. The position measurement for antenna $\mathbf{a}_i,\ i \in \{0, 1\}$ is modeled as

$$\boldsymbol{z}^{\mathbf{n}}_{\mathbf{a}_i,k} = \boldsymbol{p}^{\mathbf{n}}_k + R^{\mathbf{nb}}_k R^{\mathbf{bs}} \boldsymbol{p}^{\mathbf{s}}_{\mathbf{ba}_i} + \boldsymbol{e}_{\mathbf{a}_i,k} \tag{7}$$

where $\boldsymbol{e}_{\mathbf{a}_i,k}$ is the CDGNSS measurement noise. The vector $\boldsymbol{p}^{\mathbf{s}}_{\mathbf{ba}_i}$ from $\mathbf{b}$ to the antenna $\mathbf{a}_i$, expressed in $\mathbf{s}$, is available from the mechanical specification. As discussed above, $R^{\mathbf{bs}}$ may be taken to be the same as $\bar{R}^{\mathbf{bs}}$ from the mechanical specification, or may be further calibrated by augmenting the state with $\boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{sb}}$.

Additionally, the error-state EKF requires the Jacobian of the measurement model with respect to the error state:

$$H_{\mathbf{a}_i,k} \triangleq \frac{\partial \boldsymbol{z}^{\mathbf{n}}_{\mathbf{a}_i,k}}{\partial \delta \boldsymbol{x}_k}\bigg|_{\substack{\delta \boldsymbol{x}_k = 0 \\ \boldsymbol{e}_{\mathbf{a}_i,k} = 0}} = \frac{\partial \boldsymbol{z}^{\mathbf{n}}_{\mathbf{a}_i,k}}{\partial \boldsymbol{x}_k}\bigg|_{\substack{\boldsymbol{x}_k = \boldsymbol{x}_{\mathrm{nom},k} \\ \boldsymbol{e}_{\mathbf{a}_i,k} = 0}} \cdot \frac{\partial \boldsymbol{x}_k}{\partial \delta \boldsymbol{x}_k}\bigg|_{\substack{\delta \boldsymbol{x}_k = 0 \\ \boldsymbol{e}_{\mathbf{a}_i,k} = 0}}$$

The nontrivial sub-blocks of $H_{\mathbf{a}_i,k}$ are documented in Appendix A.

3) Radar Range Rate & Bearing Measurements: The range rate and bearing measurements from automotive radars provide a valuable velocity constraint for inertial navigation. Importantly, the frequency modulated continuous wave (FMCW) signal used in automotive radars provides instantaneous range rate measurements to the detected targets, i.e., target tracking and/or matching across cluttered radar scans is not necessary to obtain and apply this measurement.

The relative velocity of a stationary target with respect to $\mathbf{r}_i$ is given by the negative of the velocity with respect to $\mathbf{n}$ of the ith radar, expressed in $\mathbf{r}_i$, written $-\boldsymbol{v}^{\mathbf{r}_i}_{\mathbf{r}_i,k}$, as shown in Fig. 5. Assuming that the radar only detects targets in the two-dimensional plane of the linear phased array, the range rate measurement is modeled as

$$\dot{r}_{ij,k} = \begin{bmatrix} \sin\theta_{ij,k} \\ -\cos\theta_{ij,k} \\ 0 \end{bmatrix}^{\top} R^{\mathbf{r}_i\mathbf{s}} R^{\mathbf{sb}} \big(R^{\mathbf{bn}}_k \boldsymbol{v}^{\mathbf{n}}_k + \big(\boldsymbol{\omega}^{\mathbf{b}}_k \times R^{\mathbf{bs}} \boldsymbol{p}^{\mathbf{s}}_{\mathbf{br}_i} \big) \big) \tag{8}$$

where the vector $\boldsymbol{p}^{\mathbf{s}}_{\mathbf{br}_i}$ and the radar orientation $R^{\mathbf{r}_i\mathbf{s}}$ may be taken from the mechanical specifications. Note that unlike typical measurement models where the right-hand side is

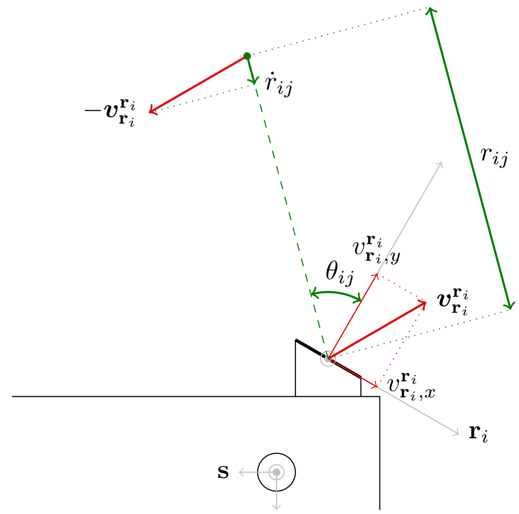


Fig. 5. A visual description of the radar range rate measurement model. Quantities labeled in green are measured by the radar. The relative velocity of a stationary target with respect to $\mathbf{r}_i$ is the negative of the velocity with respect to $\mathbf{n}$ of the ith radar, expressed in $\mathbf{r}_i$, written $-\boldsymbol{v}^{\mathbf{r}_i}_{\mathbf{r}_i,k}$. The measured radial velocity $\dot{r}_{ij}$ of the jth stationary target is the projection of $-\boldsymbol{v}^{\mathbf{r}_i}_{\mathbf{r}_i,k}$ onto the line-of-sight direction between the ith radar and the jth target.

composed of quantities that are either known or are being estimated, 8 has *measured* quantities $\theta_{ij,k}$ on the right-hand side of the equation. This implies that any errors in the bearing measurements will not be accounted for if the range rate measurements are modeled in the EKF as shown.

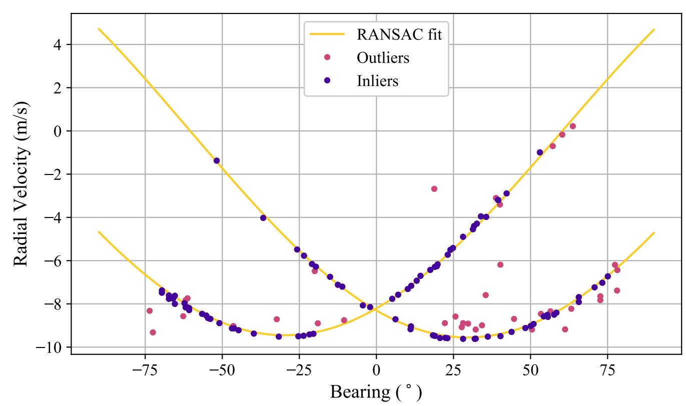


Fig. 6. Example results of the RANSAC operation on radar range rate and bearing measurements. The two yellow sinusoidal curves represent the RANSAC-predicted radial velocities for the port and starboard radars from Fig. 4 as a function of the bearing. With a threshold of $0.2\,\mathrm{m\,s^{-1}}$, RANSAC considers violet dots as inliers and magenta dots as outliers. Note that the radial velocity magnitude is maximized at $-30°$ and $30°$ for the port and starboard radars, respectively, in agreement with the mounting angles of these radars on the vehicle.

The application of range rate constraints comes with two major challenges. First, individual radar scans contain a number of spurious targets as discussed in Sec. I. Second, automotive phased-array radars exhibit poor bearing resolution and accuracy, and this is further exacerbated by the unusual range rate measurement model described above. Both of these challenges are addressed by pre-processing the range rate and bearing measurements with a RANSAC routine that estimates a best-fit two-dimensional radar velocity model to the radar measurements. In particular, with N detected targets, the RANSAC operation finds a robust solution to the following system of equations:

$$\begin{bmatrix} \dot{r}_{i0} \\ \vdots \\ \dot{r}_{iN} \end{bmatrix} = \begin{bmatrix} \sin\theta_{i0} & -\cos\theta_{i0} \\ \vdots & \vdots \\ \sin\theta_{iN} & -\cos\theta_{iN} \end{bmatrix} \begin{bmatrix} v^{\mathbf{r}_i}_{\mathbf{r}_i,x} \\ v^{\mathbf{r}_i}_{\mathbf{r}_i,y} \end{bmatrix} \tag{9}$$

while eliminating the $(\dot{r}_{ij}, \theta_{ij})$ pairs that may be outliers. Example results from the RANSAC procedure are shown in Fig. 6. Ultimately, the solution to 9 is applied as a measurement to the EKF with the following measurement model:

$$\begin{aligned} \boldsymbol{z}^{\mathbf{r}_i}_{\mathbf{r}_i,k} &\triangleq \begin{bmatrix} v^{\mathbf{r}_i}_{\mathbf{r}_i,x} \\ v^{\mathbf{r}_i}_{\mathbf{r}_i,y} \end{bmatrix}_k \\ &= \left[R^{\mathbf{r}_i\mathbf{s}} R^{\mathbf{sb}} \left(R^{\mathbf{bn}}_k \boldsymbol{v}^{\mathbf{n}}_k + \left(\boldsymbol{\omega}^{\mathbf{b}}_k \times R^{\mathbf{bs}} \boldsymbol{p}^{\mathbf{s}}_{\mathbf{br}_i} \right) \right) \right]_{[0,1]} + \boldsymbol{e}_{\mathbf{r}_i,k} \end{aligned}$$

where the subscript $[0,1]$ denotes the first two elements of the three-element vector. Parts of the Jacobian of this measurement model with respect to the EKF error-state are documented in Appendix A.

4) Ground Vehicle Dynamics Constraints: Under nominal driving conditions, a ground vehicle respects dynamical constraints which can be leveraged as measurements to the EKF. This paper incorporates near-zero sideslip and vertical velocity constraints, commonly referred to as nonholonomic constraints (NHC), as well as zero-speed updates (ZUPT). The measurement models for these constraints are described below.

a) Nonholonomic Constraints (NHC): The application of NHC is based on the following assumptions:

1) There exists a fixed center of rotation, taken to be the origin of $\mathbf{v}$, about which the vehicle rotates when a steering control input is applied.
2) When a zero steering input is applied, the vehicle only moves in the $\mathbf{v}_y$ direction. This holds by definition of $\mathbf{v}$.
3) The vehicle does not slip sideways or leave the surface of the road.

When the above assumptions hold, it follows that the velocity of the vehicle, when expressed in $\mathbf{v}$, is zero in the $\mathbf{v}_x$ and $\mathbf{v}_z$ directions at all times. In practice, however, these assumptions only hold approximately. Accordingly, the zero sideslip and vertical velocity constraints are applied as *soft* constraints in the form of measurements with an associated measurement error covariance. The NHC is modeled as

$$\mathbf{0}_{2\times1} \triangleq \boldsymbol{z}^{\mathbf{v}}_{\mathrm{nhc},k} \tag{10}$$

$$\begin{aligned} &= [\boldsymbol{v}^{\mathbf{v}}_k]_{[0,2]} + \boldsymbol{e}_{\mathrm{nhc},k} \\ &= \left[R^{\mathbf{vs}} R^{\mathbf{sb}} \left(R^{\mathbf{bn}}_k \boldsymbol{v}^{\mathbf{n}}_k + \left(\boldsymbol{\omega}^{\mathbf{b}}_k \times R^{\mathbf{bs}} \boldsymbol{p}^{\mathbf{s}}_{\mathbf{bv}} \right) \right) \right]_{[0,2]} + \boldsymbol{e}_{\mathrm{nhc},k} \end{aligned} \tag{11}$$

where $\boldsymbol{p}^{\mathbf{s}}_{\mathbf{bv}} = \boldsymbol{p}^{\mathbf{s}}_{\mathbf{bs}} + \boldsymbol{p}^{\mathbf{s}}_{\mathbf{sv}}$ and $R^{\mathbf{vs}}$ are parts of the extrinsic calibration between $\mathbf{v}$ and $\mathbf{s}$. Precise manual measurement of $\boldsymbol{p}^{\mathbf{s}}_{\mathbf{sv}}$ and $R^{\mathbf{vs}}$ is challenging. First, it is not obvious where the origin of $\mathbf{v}$ lies, though the center of line connecting the two rear axles might be a reasonable guess. Second, it would be challenging to measure, for example, the pitch of the Sensorium relative to the plane of the vehicle chassis. Accordingly, this paper takes a data-driven approach to extrinsic calibration of $\mathbf{v}$.

Once again, the extrinsic calibration technique relies on clear-sky periods with good CDGNSS availability, such that the nominal state estimates of $\boldsymbol{v}^{\mathbf{n}}_k$, $\boldsymbol{q}^{\mathbf{nb}}_k$, and $\boldsymbol{b}^{\mathbf{b}}_{\omega,k}$ are close to their true values. Furthermore, calibration begins with coarse initial guesses of $R^{\mathbf{vs}}$ and $\boldsymbol{p}^{\mathbf{s}}_{\mathbf{sv}}$, denoted $\bar{R}^{\mathbf{vs}}$ and $\bar{\boldsymbol{p}}^{\mathbf{s}}_{\mathbf{sv}}$, respectively, and attempts to estimate the orientation deviation $\boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{vs}}$ and lever arm deviation $\delta\boldsymbol{p}^{\mathbf{s}}_{\mathbf{sv}}$ with respect to these. With other quantities assumed known, 11 may be rewritten as

$$\begin{aligned} \boldsymbol{e}_{\mathrm{nhc},k} &= \left[\left(\bar{R}^{\mathbf{vs}} \oplus \boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{vs}} \right) \left(\boldsymbol{v}^{\mathbf{s}}_k + \left(\boldsymbol{\omega}^{\mathbf{s}}_k \times \left(\bar{\boldsymbol{p}}^{\mathbf{s}}_{\mathbf{bv}} + \delta\boldsymbol{p}^{\mathbf{s}}_{\mathbf{bv}} \right) \right) \right) \right]_{[0,2]} \\ &\triangleq \boldsymbol{h}_{\mathrm{nhc},k}(\boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{vs}}, \delta\boldsymbol{p}^{\mathbf{s}}_{\mathbf{bv}}) \end{aligned}$$

This model is nonlinear in $\boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{vs}}$, and may be solved as a nonlinear least squares problem, e.g., with the Gauss-Newton method. The Jacobian of $\boldsymbol{h}_{\mathrm{nhc},k}$ evaluated at $\boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{vs}} = \mathbf{0}$ and $\delta\boldsymbol{p}^{\mathbf{s}}_{\mathbf{bv}} = \mathbf{0}$ is composed of

$$\begin{aligned} \frac{\partial \boldsymbol{h}_{\mathrm{nhc},k}}{\partial \boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{vs}}} &= \left[\left(\boldsymbol{v}^{\mathbf{s}}_k + \boldsymbol{\omega}^{\mathbf{s}}_k \times \bar{\boldsymbol{p}}^{\mathbf{s}}_{\mathbf{bv}} \right)^\top \otimes \left[\bar{R}^{\mathbf{vs}} \right]_{[(0,2),(:)]} \right] \begin{bmatrix} [-\hat{\mathbf{i}}]_\times \\ [-\hat{\mathbf{j}}]_\times \\ [-\hat{\mathbf{k}}]_\times \end{bmatrix} \\ \frac{\partial \boldsymbol{h}_{\mathrm{nhc},k}}{\partial \delta\boldsymbol{p}^{\mathbf{s}}_{\mathbf{bv}}} &= \left[\bar{R}^{\mathbf{vs}} \right]_{[(0,2),(:)]} [\boldsymbol{\omega}^{\mathbf{s}}_k]_\times \end{aligned}$$

where $\otimes$ denotes the Kronecker product, subscript $[(0,2),(:)]$ denotes selection of the first and third rows of a matrix, $[\cdot]_\times$ denotes the skew-symmetric cross-product matrix corresponding to the 3-element argument, and $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ denote the cardinal unit vectors. To make the system observable, measurements from multiple epochs must be stacked and solved as a batch. Additionally, the nonlinear problem must be iteratively linearized and solved until convergence.

b) Zero-Speed Update (ZUPT): The ZUPT constraint is another valuable measurement that limits odometric drift, especially in situations where the platform makes frequent stops. The measurement model for ZUPT is trivially written as

$$\begin{aligned} \mathbf{0}_{3\times1} &\triangleq \boldsymbol{z}^{\mathbf{v}}_{\mathrm{zupt},k} \\ &= R^{\mathbf{vs}} R^{\mathbf{sb}} R^{\mathbf{bn}}_k \boldsymbol{v}^{\mathbf{n}}_k + \boldsymbol{e}_{\mathrm{zupt},k} \end{aligned} \tag{12}$$

The primary challenge of applying ZUPT is detection of epochs where this constraint is valid. Importantly, this condition must be detected independently from the EKF state estimate, e.g., by inspection of the raw IMU measurements. In theory, it is not possible to make any claims about zero speed based on acceleration and/or angular rate data, since IMU measurements of a vehicle moving with a constant velocity and orientation must be indistinguishable from those of a stationary vehicle. In practice, however, the IMU measurements exhibit a distinct behavior when the vehicle is in motion, e.g., due to road roughness and vehicle vibrations. Prior work has made use of these *artifacts* to detect stationary periods. This paper follows the angular rate energy method from [51] for ZUPT detection. In practice, if wheel odometry data are available from the vehicle CAN bus, as is common in most modern vehicles, then ZUPT detection can be performed trivially and with high reliability.

An observant reader might wonder why ZUPT is not applied directly to $\boldsymbol{v}_k^{\mathbf{n}}$ in 12. The advantage of applying ZUPT in $\mathbf{v}$ is that a tighter zero-speed constraint can be reliably applied in the lateral and vertical directions.

E. Batch Smoothing & Update

Real-time estimates of the vehicle pose trajectory obtained from the EKF may be used to string together individual scans and perform a radar batch measurement update. However, since these data are processed batches, it is desirable to perform backward smoothing over the short duration of the batch. Backward smoothing enforces the dynamics function backwards in time, ironing out any large jumps that may have occurred in the EKF forward pass.

Accordingly, the batch smoother component in Fig. 3 stacks all inertial measurements and snapshots of the estimator state over the duration of the batch. When the batch is ready to be processed for correlation, backward smoothing is enforced with the inertial measurements as control inputs. The smoothing formulation in this case is somewhat more complicated than usual [52] due to nonlinear backward dynamics and the error-state formulation. Details on nonlinear error-state Rauch-Tung-Striebel smoothing are provided in Appendix B.

The correlation peak search region is taken to be $\pm 5\,\mathrm{m}$ and $\pm 3°$. The 3-DoF pose offset $\widehat{\boldsymbol{\Theta}}$ from radar batch correlation is applied as horizontal position and heading measurements to the EKF. Outliers from batch correlation are excluded in the EKF based on a χ^2-test on the normalized innovation squared (NIS) [53].

V. Experimental Results

The radar-inertial positioning system of Fig. 3 was evaluated experimentally using the dataset described in [12], collected during approximately $1.5\,\mathrm{h}$ of driving on two separate days in and around the urban center of Austin, TX. This section presents the evaluation results.

A. Dataset

Fig. 7 shows the route followed by the sensor-instrumented vehicle on Thursday, May 9, 2019 (in blue) and Sunday, May 12, 2019 (in red). The test route combs through every street in the Austin, TX downtown area, since such environments are the most challenging for CDGNSS-based positioning [9] and would benefit the most from multi-sensor all-weather positioning. The route was driven once on a weekday and again on the weekend to evaluate robustness of the proposed map-based approach to changes in the traffic and parking patterns. Note that the final part of the route (the north-east segment) was different on the two days, preventing the use of a map-based positioning approach. This section of the test route has been omitted from the evaluation results.

1) Sensors: The Sensorium, shown in Fig. 4, features two types of automotive radars: one Delphi electronically-scanning radar (ESR) in the middle and two Delphi short-range radars (SRR2s) on the two sides. Both the ESR and the SRR2 are commercially available; similar radars are available on economy-class consumer vehicles. The ESR provides simultaneous sensing in a narrow ($\pm 10°$) long-range ($175\,\mathrm{m}$) coverage area and a wider ($\pm 45°$) medium-range ($60\,\mathrm{m}$) area. The SRR2 units each have a coverage area of $\pm 75°$ and $80\,\mathrm{m}$ (see [13, Fig. 6]). Each SRR2 is installed facing outward from the center-line at an angle of $30°$. The Sensorium's onboard computer timestamps and logs the radar returns from the three radar units.

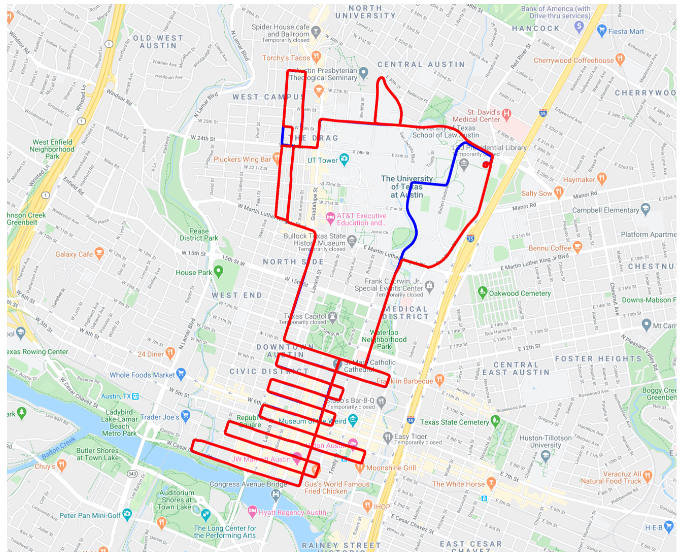


Fig. 7. Test route through The University of Texas west campus and Austin downtown. These areas are the most challenging for precise GNSS-based positioning and thus would benefit the most from radar-based positioning. The route was driven once on a weekday and again on the weekend to evaluate robustness of the radar map to changes in traffic and parking patterns. Red is the mapping run (May 12), blue is the localization run (May 9). A prior map is not available in the visible blue areas.

The LORD MicroStrain 3DM-GX5-25 MEMS IMU is an industrial-grade inertial sensor that acts as the core sensor of the localization pipeline. The IMU provides temperature-compensated accelerometer and gyroscope readings at $100\,\mathrm{Hz}$. Two Antcom G8Ant-3A4TNB1 high performance GNSS patch antennas pull in signals from all three GNSS frequency bands and include a $40\,\mathrm{dB}$ active low-noise amplifier.

2) Ground-Truth Trajectory: The ground-truth position and orientation trajectory for the data are generated with the iXblue ATLANS-C, a high-performance CDGNSS coupled fiber-optic gyroscope INS. The post-processed position solution obtained from the ATLANS-C is decimeter-accurate throughout the dataset.

3) Dataset Splits: With a limited amount of field data available for development and evaluation, it is critical to ensure that the proposed positioning technique does not overfit this particular dataset. Accordingly, the data used in the development of the algorithms were restricted to a fixed $30\,\mathrm{min}$ segment, where the prior radar map was constructed with radar measurements from May 9 and localization was performed with radar, inertial, and CDGNSS measurements from May 12. In contrast, during evaluation the full $62\,\mathrm{min}$ of data were used, and the mapping and localization datasets were inverted, i.e., the prior map was constructed with radar measurements from May 12, and localization was performed with all sensor data from May 9. The algorithms have not been modified to maximize the performance over the evaluation dataset.

B. Prior Radar Mapping

The first step to radar-map-based localization is the generation of a radar map point cloud. Radar scans collected from the May 12, 2019 drive were aggregated to create a map with the benefit of the ATLANS-C ground-truth trajectory. In a practical system, the radar map may be generated during favorable conditions for optical sensors such as cameras and lidar, such that the mapping vehicle can accurately track its pose. Additionally, the mapping process may be crowed-sourced from consumer vehicles [11], [54]. The map point cloud is stored in a k-d tree for efficient querying during localization.

Two implementation notes are in order here. First, automotive radar clutter is especially intense when the vehicle is stationary. Accordingly, radar range measurements obtained when the vehicle was moving slower than $1\,\mathrm{m\,s^{-1}}$ were discarded for both mapping and localization. This implies that radar correlation measurements were only available during periods when the vehicle was moving faster than $1\,\mathrm{m\,s^{-1}}$. Second, it was observed that radar returns far from the vehicle are mostly clutter and have negligible resemblance to the surrounding structure. Radar returns with range larger than $50\,\mathrm{m}$ were discarded for both the map and batch PHDs. It is noted that these two parameters have not been optimized to produce the smallest estimation errors; instead they have been fixed based on visual inspection.

C. Offline Calibration

Extrinsic calibration among the IMU frame $\mathbf{b}$, the Sensorium frame $\mathbf{s}$, and the vehicle frame $\mathbf{v}$ was performed offline with $125\,\mathrm{s}$ of sensor data with CDGNSS availability. While it is possible to estimate the calibration parameters online, it may not be desirable to do so if these parameters are not expected to change over time.

The orientation deviation $\boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{sb}}$ between the IMU body frame and the Sensorium frame was calibrated for the localization dataset, as described in Sec. IV-D1. With two GNSS antennas, only two out of the three DoFs in $\boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{sb}}$ are observable. Accordingly, the orientation deviation around $\mathbf{b}_x$, which is mostly unobservable, was tightly constrained to the initial guess of zero. The deviations around $\mathbf{b}_y$ and $\mathbf{b}_z$ rapidly converged to sub-degree offsets from the mechanical specification.

Extrinsic calibration between $\mathbf{v}$ and $\mathbf{s}$ was similarly estimated over the $125\,\mathrm{s}$ period as detailed in Sec. IV-D4.

The commercial automotive radars on the Sensorium do not offer any mechanism to synchronize their scans with an external reference clock. Analysis of the radar range rate residuals in the EKF showed clear evidence of latency between the data logging timestamp and the true scan times. Accordingly, radar latency calibration was performed offline with a best fit approach.

D. Implementation Notes

A few implementation- and dataset-specific notes relating to the localization pipeline are documented below.

a) CDGNSS Measurements & Outages: The CDGNSS position measurements used in this evaluation are in fact the output of the post-processed ground-truth system, i.e., these measurements have not been obtained from an unaided CDGNSS receiver. While this is not ideal for realistic evaluation, the evaluation results presented herein do not mislead because, first, CDGNSS measurements are only applied for a $125\,\mathrm{s}$ period for initial calibration, and second, any commercial CDGNSS receiver would be able to generate similar cm-accurate position solutions in the clear-sky region where the CDGNSS measurements were applied.

b) Measurement Noise Correlation: Observations from the field data revealed that the measurement noise in the radar range rate measurements is not independent between consecutive radar scans. This is problematic since the EKF applied assumes each measurement to have errors that are uncorrelated in time. Accordingly, the radar range rate measurements were decimated to $1\,\mathrm{Hz}$ such that the measurements were spaced out by roughly the decorrelation time of the measurement noise. A more principled approach to this problem is to augment the state vector with states to pre-whiten the measurements. But this approach was empirically observed to not outperform the straightforward measurement decimation, while introducing additional complexity and tuning parameters.

Similarly, the NHC and ZUPT measurements can in theory be applied at every applicable IMU epoch. But to prevent correlated errors in these constraints (e.g., due to sideslip experienced while cornering) from making the EKF inconsistent, they are only applied at $1\,\mathrm{Hz}$.

c) Filter Tuning Parameters: The process noise covariance used in the EKF is derived from the IMU datasheet parameters [50], [55]. The measurement noise covariance associated with CDGNSS measurements is available directly from the ATLANS-C receiver. A few other measurement noise standard deviations and tuning parameters are documented in Table I.

TABLE I
A LIST OF PARAMETERS INVOLVED IN THE LOCALIZATION PIPELINE

Minimum speed for valid radar range	$1\,\mathrm{m\,s^{-1}}$
Maximum valid radar range	$50\,\mathrm{m}$
Minimum RANSAC inliers	10
Minimum fraction of RANSAC inliers	0.65
$v^{\mathbf{r}_i}_{\mathbf{r}_i,x}$ (broadside) standard deviation	$0.2\,\mathrm{m\,s^{-1}}$
$v^{\mathbf{r}_i}_{\mathbf{r}_i,y}$ (boresight) standard deviation	$0.1\,\mathrm{m\,s^{-1}}$
$v^{\mathbf{v}}_{\mathrm{nhc},x}$ (lateral) standard deviation	$0.1\,\mathrm{m\,s^{-1}}$
$v^{\mathbf{v}}_{\mathrm{nhc},z}$ (vertical) standard deviation	$0.2\,\mathrm{m\,s^{-1}}$

E. Localization Results

This section presents empirical error statistics obtained from field evaluation of the proposed method. The test scenario evaluated in this section is an extreme one: the vehicle starts off in a clear-sky environment with $125\,\mathrm{s}$ of CDGNSS availability, and subsequently all CDGNSS measurements are cut off for the next $3600\,\mathrm{s}$ of driving, during which the system must rely on radar and inertial sensing along with vehicle dynamical constraints to maintain an accurate estimate of its pose.

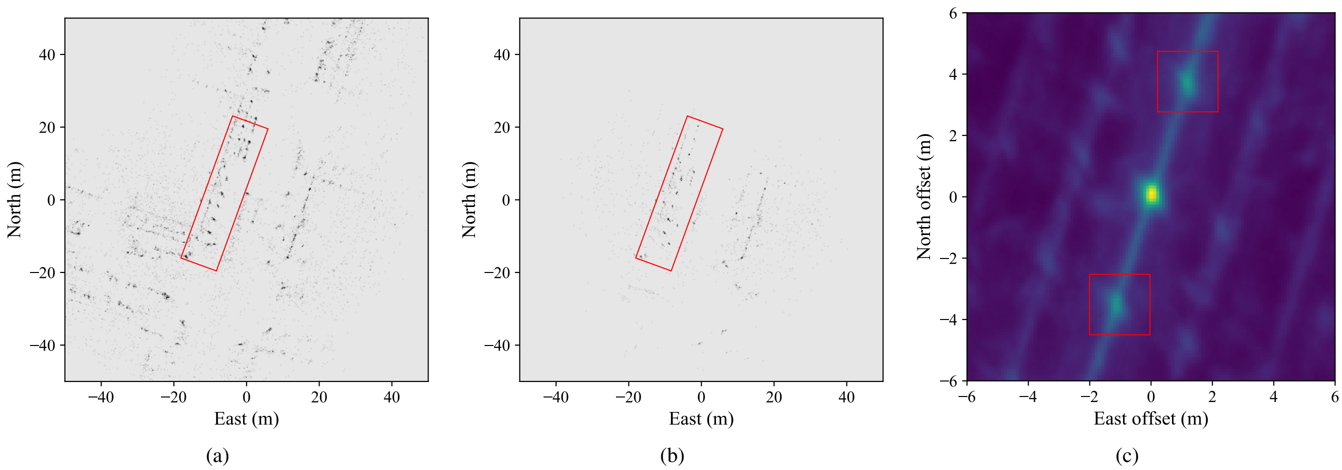


Fig. 8. This figure shows an interesting example of radar-based urban positioning with the proposed method. Panel (a) shows the occupancy grid estimated from the prior map point cloud. Panel (b) shows the same for a $5\,\mathrm{s}$ batch of scans collected in the same region. For ease of visualization, the batch occupancy grid has already been aligned with the map occupancy grid. Panel (c) shows the cross-correlation between the batch and map occupancy grids at $\Delta\phi = 0°$. Given that no rotational or translational offset error has been applied to the batch, the correlation peak should appear at $(0, 0)$. The offset of the peak in panel (c) from $(0, 0)$ is the translational estimate error of the proposed method. Also note the increased positioning uncertainty in the along-track direction, and the two local correlation peaks (marked with red squares in panel (c)) due to the repeating periodic pattern of radar reflectors in the map and the batch (marked with red rectangles in panels (a) and (b)).

Before diving into the quantitative analysis, it is interesting to inspect the example of a radar batch update shown in Fig. 8. For ease of visualization, the batch point cloud to be localized has already been adjusted for any translational or rotational offset from the ground truth. The occupancy grid estimated from the $5\,\mathrm{s}$ batch of scans is shown in Fig. 8b. Similarly, Fig. 8a shows the occupancy grid estimated from the map point cloud retrieved from the map database. Fig. 8c shows the cross-correlation between the batch and map occupancy grids. Given that the batch is already aligned with ground truth, one should expect the correlation peak to appear at $(0, 0)$ in Fig. 8c. The offset of the peak from $(0, 0)$ in this case would be the translational estimate error.

Two interesting features of the cross-correlation in Fig. 8c are worth noting. First, the correlation peak decays slower in the along-track direction—in this case approximately aligned with the south-southwest direction. This is a general feature observed throughout the dataset, since most of the radar reflectors are aligned along the sides of the streets. Second, there emerge two local correlation peaks offset by $\approx 4\,\mathrm{m}$ along the direction of travel. These local peaks are due to the repeating periodic structure of radar reflectors in both the map and the batch occupancy grids. In other words, shifting the batch occupancy grid forward or backward along the vehicle trajectory by $\approx 4\,\mathrm{m}$ aligns the periodically-repeating reflectors in an off-by-one manner, leading to another plausible solution. Importantly, the uncertainty envelope of the initial position estimate can span several meters, encompassing both the global optimum and one or more local optima. This explains why gradient-based methods, which seek the nearest optimum, are poorly suited for use in the urban automotive radar environment.

1) Performance with $4\,\mathrm{s}$ Radar Batches: Fig. 9 shows the east and north position error time histories from the test scenario described above. For the results presented in Fig. 9 and 10, a $4\,\mathrm{s}$ radar batch duration is chosen. In the first $125\,\mathrm{s}$ of clear-sky conditions with CDGNSS availability, the east and north position errors with respect to the ground truth are sub-decimeter, as expected. Over the subsequent $60\,\mathrm{min}$ of driving in and around the urban center of the city, the proposed method maintains sub-35-cm horizontal position errors (95%). The horizontal position estimation errors are consistent with the predicted standard deviation from the EKF. This is a remarkable result which shows that, given a prior radar map, lane-level-accurate horizontal positioning is achievable under zero-visibility GNSS-denied conditions with the types of sensors that are already available on commercial vehicles. Vertical position errors are not shown in Fig. 9 since these are not constrained by the two-dimensional radar batch correlation update. For ground vehicle applications, a digital elevation map can effectively constrain errors in altitude, if necessary.

Vehicle orientation estimation errors for the same scenario are shown in Fig. 10. Heading estimation error, shown in the bottom panel, is most important for ground vehicle applications. The proposed technique maintains vehicle heading estimates to within $0.5°$ of the ground truth throughout most of the dataset, and the errors are consistent with the predicted uncertainty. Roll and pitch estimation errors are smaller and stay within $0.2°$ of the ground truth. Better estimation of roll and pitch is expected since these are directly observable with the accelerometer measurements. The same phenomenon explains the substantially shorter decorrelation times for roll and pitch errors as compared to the heading error. Finally, it is noted that the EKF is mildly inconsistent in regards to roll and pitch estimation errors. This suggests that the accelerometer white noise and bias stability characteristics claimed in the IMU datasheet [55] may be optimistic in field application.

2) Choosing a Radar Batch Length: The problem of choosing the duration of a radar batch during localization presents an

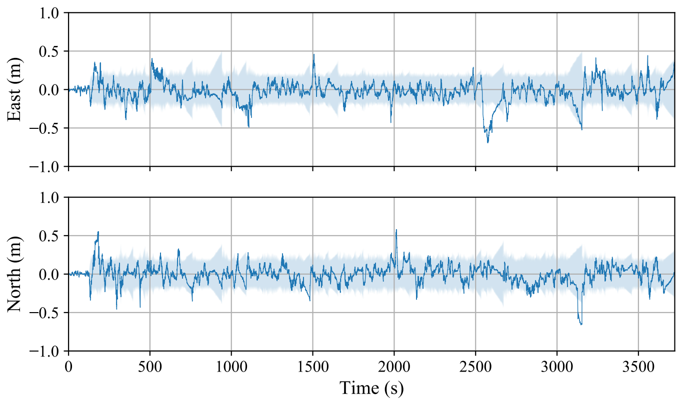


Fig. 9. East and north position error time histories from field evaluation. In the first 125 s of clear-sky conditions with CDGNSS availability, the east and north position errors with respect to the ground truth are sub-decimeter, as expected. Over the subsequent 60 min of driving in and around the urban center of the city, the proposed method maintains sub-35-cm (95%) horizontal position errors. The horizontal position estimation errors are consistent with the predicted standard deviation from the EKF.

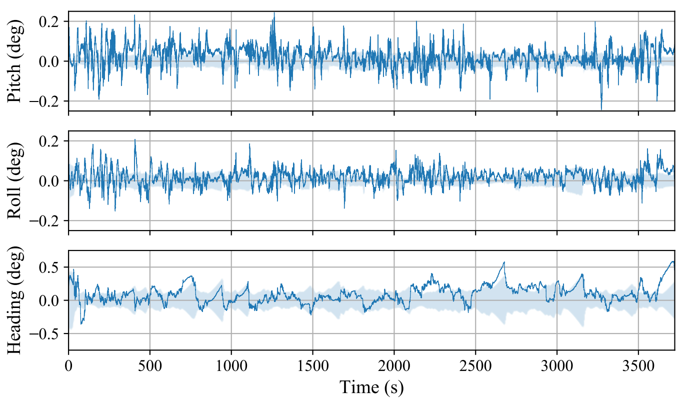


Fig. 10. Vehicle orientation estimation errors from field evaluation. The proposed technique maintains vehicle heading estimates to within $0.5°$ of the ground truth throughout most of the dataset, and the errors are consistent with the predicted uncertainty. Roll and pitch estimation errors are smaller and stay within $0.2°$ of the ground truth.

interesting trade-off. On the one hand, longer batch durations are preferable because, intuitively, cross-correlation using a larger *patch* of the radar environment is more likely to produce a strong and unambiguous correlation peak. Fig. 11 shows results from an empirical test of this intuition. In this test, radar batches of different durations between 1 s and 8 s were generated with ground-truth odometry and correlated against a prior map to obtain the estimated offset from the ground-truth pose. The complementary cumulative distribution function (CCDF) of the horizontal position estimation errors is shown in 11. It is interesting to note that up to the 70^{th} percentile, errors are similar for different batch lengths. The difference between the CCDFs becomes more pronounced at higher percentiles, implying that errors for shorter batch lengths have heavy tails. Recall that in the overall localization pipeline of Fig. 3, these errors will act as measurement errors in $\widehat{\boldsymbol{\Theta}}$. An EKF models measurement errors to be Gaussian, which is not a good model for heavy-tailed distributions. Accordingly, longer batch durations would appear preferable.

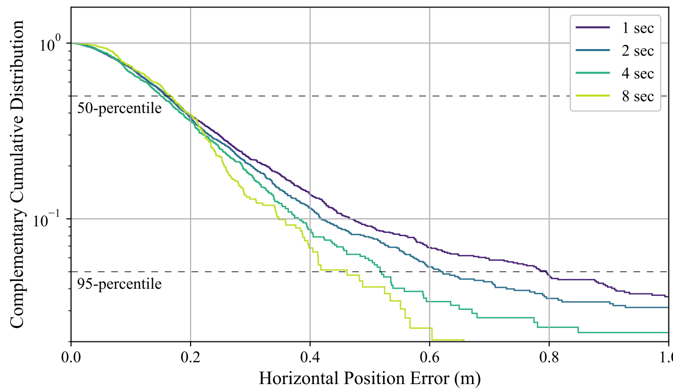


Fig. 11. CCDFs for different batch lengths between 1 s and 8 s. The 50-percentile errors are similar for shorter and longer batch lengths, but the difference becomes more noticeable at higher percentiles.

On the other hand, longer batches have several disadvantages. First, longer durations between batch measurement updates leads to larger odometric drift *between* updates, as well as poorer reconstruction of the radar batch itself. Second, some of the worst outliers due to shorter batch lengths may be rejected in the EKF based on the χ^2 NIS test, thus blunting the relative advantage of longer batches. Shorter batch lengths allow for a larger number of measurement updates to be performed per unit time, even if a few of those measurements may have to be rejected as outliers.

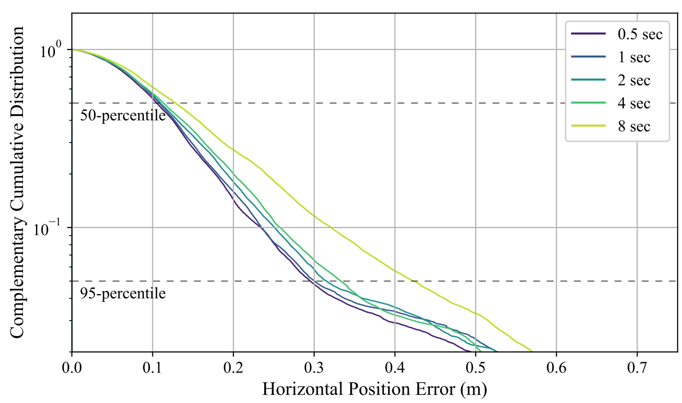


Fig. 12. End-to-end effect of different batch lengths on horizontal positioning performance. Other than the longest batch length of 8 s, most batch lengths appear to perform similarly well, with 95^{th}-percentile horizontal position errors near 30 cm.

Fig. 12 reveals the end-to-end effect of different batch lengths. For a given batch length, its measurement error standard deviation was obtained from the corresponding CCDF in Fig. 11, i.e., the $\widehat{\boldsymbol{\Theta}}$ measurement standard deviation is smaller for longer batches. Interestingly, other than the longest batch length of 8 s, most batch lengths appear to perform similarly well, with 95-percentile horizontal position errors near 30 cm. Given the heavy-tailed nature of measurement noise distributions when working with very short batches (from Fig. 11), batch lengths from 2 to 4 s may be taken to be a good compromise.

VI. Conclusion

A robust pipeline for all-weather sub-50-cm urban ground vehicle positioning has been proposed and evaluated. The positioning engine is based on commercially-available low-cost automotive radars, MEMS IMU, ground vehicle dynamics constraints, and, when available, precise GNSS measurements. Remarkably, it has been shown that given a prior radar map, lane-level-accurate horizontal positioning is achievable under zero-visibility GNSS-denied conditions with the types of sensors that are already available on commercial vehicles. In comparison with a post-processed ground truth trajectory, it was shown that during 60 min of GNSS-denied driving in the urban center of Austin, TX, the proposed pipeline has 95^{th}-percentile errors of 35 cm in horizontal position and $0.5°$ in heading. This is a significant development in the field of AGV localization, which has traditionally been based on sensors such as lidar and cameras that perform poorly in bad weather conditions.

Acknowledgments

This work has been supported by Honda R&D Americas through The University of Texas Situation-Aware Vehicular Engineering Systems (SAVES) Center (http://utsaves.org/), an initiative of the UT Wireless Networking and Communications Group; by the National Science Foundation under Grant No. 1454474 (CAREER); by the Data-supported Transportation Operations and Planning Center (DSTOP), a Tier 1 USDOT University Transportation Center; and by Sandia National Laboratories.

Appendix A
Partial Derivatives

A. Linearized Forward Dynamics

A few block components of F_k and G_k from (5) and (6) are listed below.

$$\frac{\partial \delta \boldsymbol{p}^{\mathbf{n}}_{k+1}}{\partial \boldsymbol{\eta}^{\mathbf{n}}_k}\bigg|_{\substack{\delta\boldsymbol{x}_k=0\\ \boldsymbol{w}_k=0}} = -\frac{T^2}{2}\left(\tilde{R}^{\mathbf{nb}}_k\left[\boldsymbol{z}^{\mathbf{b}}_{a,k}-\tilde{\boldsymbol{b}}^{\mathbf{b}}_{a,k}\right]_\times\right)$$
$$\frac{\partial \delta \boldsymbol{p}^{\mathbf{n}}_{k+1}}{\partial \delta\boldsymbol{b}^{\mathbf{b}}_{a,k}}\bigg|_{\substack{\delta\boldsymbol{x}_k=0\\ \boldsymbol{w}_k=0}} = -\frac{T^2}{2}\tilde{R}^{\mathbf{nb}}_k$$

The partial derivates of $\delta\boldsymbol{v}^{\mathbf{n}}_{k+1}$ with respect to $\delta\boldsymbol{x}_k$ follow similarly.

$$\frac{\partial \boldsymbol{\eta}^{\mathbf{n}}_{k+1}}{\partial \boldsymbol{\eta}^{\mathbf{n}}_k}\bigg|_{\substack{\delta\boldsymbol{x}_k=0\\ \boldsymbol{w}_k=0}} \approx I_{3\times3}$$
$$\begin{aligned}\frac{\partial \boldsymbol{\eta}^{\mathbf{n}}_{k+1}}{\partial \delta\boldsymbol{b}^{\mathbf{b}}_{\omega,k}}\bigg|_{\substack{\delta\boldsymbol{x}_k=0\\ \boldsymbol{w}_k=0}} &\approx -T\tilde{R}^{\mathbf{nb}}_{k+1}\mathrm{J}_r\left(\frac{T}{2}\left(\boldsymbol{z}^{\mathbf{b}}_{\omega,k}-\tilde{\boldsymbol{b}}^{\mathbf{b}}_{\omega,k}-\tilde{R}^{\mathbf{bn}}_k\boldsymbol{\omega}^{\mathbf{n}}_{\mathbf{e}}\right)\right)\\ &\approx -T\tilde{R}^{\mathbf{nb}}_{k+1}\end{aligned}$$

where

$$\mathrm{J}_r(\boldsymbol{\theta}) = I_{3\times3} - \frac{1-\cos\|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^2}[\boldsymbol{\theta}]_\times + \frac{\|\boldsymbol{\theta}\|-\sin\|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^3}[\boldsymbol{\theta}]^2_\times$$

is the right Jacobian of $SO(3)$ [47].

$$\begin{aligned}\frac{\partial \boldsymbol{\eta}^{\mathbf{n}}_{k+1}}{\partial \boldsymbol{w}_{\omega,k}}\bigg|_{\substack{\delta\boldsymbol{x}_k=0\\ \boldsymbol{w}_k=0}} &\approx T\tilde{R}^{\mathbf{nb}}_{k+1}\mathrm{J}_r\left(\frac{T}{2}\left(\boldsymbol{z}^{\mathbf{b}}_{\omega,k}-\tilde{\boldsymbol{b}}^{\mathbf{b}}_{\omega,k}-\tilde{R}^{\mathbf{bn}}_k\boldsymbol{\omega}^{\mathbf{n}}_{\mathbf{e}}\right)\right)\\ &\approx T\tilde{R}^{\mathbf{nb}}_{k+1}\end{aligned}$$

B. Linearized Measurement Models

The partial derivative of the measurement $\boldsymbol{z}^{\mathbf{n}}_{\mathbf{a}_i,k}$ from (7) can be expressed as

$$\frac{\partial \boldsymbol{z}^{\mathbf{n}}_{\mathbf{a}_i,k}}{\partial \delta\boldsymbol{x}_k}\bigg|_{\substack{\delta\boldsymbol{x}_k=0\\ \boldsymbol{e}_{\mathbf{a}_i,k}=0}} = \frac{\partial \boldsymbol{z}^{\mathbf{n}}_{\mathbf{a}_i,k}}{\partial \boldsymbol{x}_k}\bigg|_{\substack{\delta\boldsymbol{x}_k=0\\ \boldsymbol{e}_{\mathbf{a}_i,k}=0}} \cdot \frac{\partial \boldsymbol{x}_k}{\partial \delta\boldsymbol{x}_k}\bigg|_{\substack{\delta\boldsymbol{x}_k=0\\ \boldsymbol{e}_{\mathbf{a}_i,k}=0}}$$

where the non-trivial block matrices are as follows:

$$\frac{\partial \boldsymbol{z}^{\mathbf{n}}_{\mathbf{a}_i,k}}{\partial \boldsymbol{q}^{\mathbf{nb}}_k}\bigg|_{\substack{\delta\boldsymbol{x}_k=0\\ \boldsymbol{e}_{\mathbf{a}_i,k}=0}} = \frac{\partial\left(\boldsymbol{q}^{\mathbf{nb}}_k \odot R^{\mathbf{bs}}\boldsymbol{p}^{\mathbf{s}}_{\mathbf{ba}_i}\odot \boldsymbol{q}^{\mathbf{bn}}_k\right)}{\partial \boldsymbol{q}^{\mathbf{nb}}_k}$$
$$\frac{\partial \boldsymbol{q}^{\mathbf{nb}}_k}{\partial \boldsymbol{\eta}^{\mathbf{n}}_k}\bigg|_{\substack{\delta\boldsymbol{x}_k=0\\ \boldsymbol{e}_{\mathbf{a}_i,k}=0}} = \frac{1}{2}\begin{bmatrix}-q_x & -q_y & -q_z\\ q_w & q_z & -q_y\\ -q_z & q_w & q_x\\ q_y & -q_x & q_w\end{bmatrix}$$

with $\tilde{\boldsymbol{q}}^{\mathbf{nb}}_k = [q_w, q_x, q_y, q_z]$. The expression for derivative of the rotation with respect to the quaternion can be found in [47, Sec. 4.3.2].

For the radar range rate measurement $\boldsymbol{z}^{\mathbf{r}_i}_{\mathbf{r}_i,k}$

$$\frac{\partial \boldsymbol{z}^{\mathbf{r}_i}_{\mathbf{r}_i,k}}{\partial \boldsymbol{v}^{\mathbf{n}}_k} = R^{\mathbf{r}_i\mathbf{s}}R^{\mathbf{sb}}\tilde{R}^{\mathbf{bn}}_k$$
$$\frac{\partial \boldsymbol{z}^{\mathbf{r}_i}_{\mathbf{r}_i,k}}{\partial \boldsymbol{q}^{\mathbf{nb}}_k} = R^{\mathbf{r}_i\mathbf{s}}R^{\mathbf{sb}}\frac{\partial\left(\boldsymbol{q}^{\mathbf{bn}}_k\odot\tilde{\boldsymbol{v}}^{\mathbf{n}}_k\odot\boldsymbol{q}^{\mathbf{nb}}_k\right)}{\partial \boldsymbol{q}^{\mathbf{nb}}_k}$$
$$\frac{\partial \boldsymbol{z}^{\mathbf{r}_i}_{\mathbf{r}_i,k}}{\partial \boldsymbol{b}^{\mathbf{b}}_{\omega,k}} = -R^{\mathbf{r}_i\mathbf{s}}R^{\mathbf{sb}}\left[R^{\mathbf{bs}}\boldsymbol{p}^{\mathbf{s}}_{\mathbf{br}_i}\right]_\times$$

The partial derivatives of $\boldsymbol{z}^{\mathbf{v}}_{\text{nhc},k}$ and $\boldsymbol{z}^{\mathbf{v}}_{\text{zupt},k}$ follow similarly.

Appendix B
Nonlinear Error-State Rauch-Tung-Striebel Smoother

The conventional expression for the extended Rauch-Tung-Striebel (RTS) smoother is given as [52, Chap. 9]

$$\boldsymbol{x}^\star_k = \hat{\boldsymbol{x}}_k + C_k\left(\boldsymbol{x}^\star_{k+1} - f_k(\hat{\boldsymbol{x}}_k)\right)$$
$$P^\star_k = P_k + C_k\left(P^\star_{k+1} - F_kP_kF_k^\top - G_kQ_kG_k^\top\right)C_k^\top$$

with

$$C_k = P_kF_k^\top\left(F_kP_kF_k^\top + G_kQ_kG_k^\top\right)^{-1}$$

where $\star$ indicates the smoothed estimate and $\hat{}$ indicates the filtered estimate. This expression is derived by linearizing the dynamics at the *filtered* state estimate during the *backward smoothing* pass.

In contrast, this paper prefers to linearize the dynamics at the predicted smoothed estimate $\bar{\boldsymbol{x}}^\star_k$ instead

$$\bar{\boldsymbol{x}}^\star_k \triangleq f_k^{-1}\left(\boldsymbol{x}^\star_{k+1}, \boldsymbol{u}_k, \mathbf{0}\right)$$

This formulation results in a similar but slightly modified expression for the extended RTS smoother

$$\boldsymbol{x}^\star_k = \hat{\boldsymbol{x}}_k + C^\star_kF^\star_k(\bar{\boldsymbol{x}}^\star_k - \hat{\boldsymbol{x}}_k)$$
$$P^\star_k = P_k + C^\star_k\left(P^\star_{k+1} - F^\star_kP_kF^{\star\top}_k - G^\star_kQ_kG^{\star\top}_k\right)C^{\star\top}_k$$

with

$$C^\star_k = P_kF^{\star\top}_k\left(F^\star_kP_kF^{\star\top}_k + G^\star_kQ_kG^{\star\top}_k\right)^{-1}$$

where $F^\star_k$ and $G^\star_k$ denote linearized forward dynamics around $\bar{\boldsymbol{x}}^\star_k$.

References

[1] D. Fajardo, T.-C. Au, S. Waller, P. Stone, and D. Yang, "Automated intersection control: Performance of future innovation versus current traffic signal control," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2259, pp. 223–232, 2011.

[2] J. Choi, V. Va, N. Gonzalez-Prelcic, R. Daniels, C. R. Bhat, and R. W. Heath, "Millimeter-wave vehicular communication to support massive automotive sensing," *IEEE Communications Magazine*, vol. 54, no. 12, pp. 160–167, December 2016.

[3] D. LaChapelle, T. E. Humphreys, L. Narula, P. A. Iannucci, and E. Moradi-Pari, "Automotive collision risk estimation under cooperative sensing," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Barcelona, Spain, 2020.

[4] K. S. Yen, C. Shankwitz, B. Newstrom, T. A. Lasky, and B. Ravani, "Evaluation of the University of Minnesota GPS Snowplow Driver Assistance Program," California Department of Transportation, Tech. Rep., 2015.

[5] M. Petovello, M. Cannon, and G. Lachapelle, "Benefits of using a tactical-grade IMU for high-accuracy positioning," *Navigation, Journal of the Institute of Navigation*, vol. 51, no. 1, pp. 1–12, 2004.

[6] B. M. Scherzinger, "Precise robust positioning with inertially aided RTK," *Navigation*, vol. 53, no. 2, pp. 73–83, 2006.

[7] H. T. Zhang, "Performance comparison on kinematic GPS integrated with different tactical-grade IMUs," Master's thesis, The University of Calgary, Jan. 2006.

[8] S. Kennedy, J. Hamilton, and H. Martell, "Architecture and system performance of SPAN—NovAtel's GPS/INS solution," in *Position, Location, And Navigation Symposium, 2006 IEEE/ION*. IEEE, 2006, p. 266.

[9] T. E. Humphreys, M. J. Murrian, and L. Narula, "Deep urban unaided precise Global Navigation Satellite System vehicle positioning," *IEEE Intelligent Transportation Systems Magazine*, 2020.

[10] T. E. Humphreys, *Springer Handbook of Global Navigation Satellite Systems*. Springer, 2017, ch. Interference, pp. 469–504.

[11] L. Narula, J. M. Wooten, M. J. Murrian, D. M. LaChapelle, and T. E. Humphreys, "Accurate collaborative globally-referenced digital mapping with standard GNSS," *Sensors*, vol. 18, no. 8, 2018. [Online]. Available: http://www.mdpi.com/1424-8220/18/8/2452

[12] L. Narula, D. M. LaChapelle, M. J. Murrian, J. M. Wooten, T. E. Humphreys, J.-B. Lacambre, E. de Toldi, and G. Morvant, "TEX-CUP: The University of Texas Challenge for Urban Positioning," in *Proceedings of the IEEE/ION PLANSx Meeting*, 2020.

[13] L. Narula, P. A. Iannucci, and T. E. Humphreys, "Automotive-radar-based 50-cm urban positioning," in *Proceedings of the IEEE/ION PLANSx Meeting*, 2020.

[14] D. Chetverikov, D. Svirko, D. Stepanov, and P. Krsek, "The trimmed iterative closest point algorithm," in *Object recognition supported by user interaction for service robots*, vol. 3. IEEE, 2002, pp. 545–548.

[15] E. Ward and J. Folkesson, "Vehicle localization with low cost radar sensors," in *2016 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2016, pp. 864–870.

[16] M. Holder, S. Hellwig, and H. Winner, "Real-time pose graph SLAM based on radar," in *2019 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2019, pp. 1145–1151.

[17] Y. Tsin and T. Kanade, "A correlation-based approach to robust point set registration," in *European conference on computer vision*. Springer, 2004, pp. 558–569.

[18] B. Jian and B. C. Vemuri, "Robust point set registration using Gaussian mixture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1633–1645, 2010.

[19] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 12, pp. 2262–2275, 2010.

[20] W. Gao and R. Tedrake, "FilterReg: Robust and efficient probabilistic point-set registration using gaussian filter and twist parameterization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 11 095–11 104.

[21] S. H. Cen and P. Newman, "Precise ego-motion estimation with millimeter-wave radar under diverse and challenging conditions," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 1–8.

[22] ——, "Radar-only ego-motion estimation in difficult settings via graph matching," *arXiv preprint arXiv:1904.11476*, 2019.

[23] D. Barnes and I. Posner, "Under the radar: Learning to predict robust keypoints for odometry estimation and metric localisation in radar," *arXiv preprint arXiv:2001.10789*, 2020.

[24] D. Barnes, R. Weston, and I. Posner, "Masking by moving: Learning distraction-free radar odometry from pose information," *arXiv preprint arXiv:1909.03752*, 2019.

[25] F. Schuster, C. G. Keller, M. Rapp, M. Haueis, and C. Curio, "Landmark based radar SLAM using graph optimization," in *Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on*. IEEE, 2016, pp. 2559–2564.

[26] M. Schoen, M. Horn, M. Hahn, and J. Dickmann, "Real-time radar SLAM."

[27] J. Callmer, D. Törnqvist, F. Gustafsson, H. Svensson, and P. Carlbom, "Radar SLAM using visual features," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, no. 1, p. 71, 2011.

[28] Z. Hong, Y. Petillot, and S. Wang, "RadarSLAM: Radar based large-scale SLAM in all weathers," *arXiv preprint arXiv:2005.02198*, 2020.

[29] K. Yoneda, N. Hashimoto, R. Yanase, M. Aldibaja, and N. Suganuma, "Vehicle localization using 76GHz omnidirectional millimeter-wave radar for winter automated driving," in *2018 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2018, pp. 971–977.

[30] R. P. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Transactions on Aerospace and Electronic systems*, vol. 39, no. 4, pp. 1152–1178, 2003.

[31] O. Erdinc, P. Willett, and Y. Bar-Shalom, "The bin-occupancy filter and its connection to the PHD filters," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4232–4246, 2009.

[32] J. Mullane, B.-N. Vo, M. D. Adams, and B.-T. Vo, "A random-finite-set approach to Bayesian SLAM," *IEEE Transactions on Robotics*, vol. 27, no. 2, pp. 268–282, 2011.

[33] H. Deusch, S. Reuter, and K. Dietmayer, "The labeled multi-Bernoulli SLAM filter," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1561–1565, 2015.

[34] M. Stübler, S. Reuter, and K. Dietmayer, "A continuously learning feature-based map using a Bernoulli filtering approach," in *2017 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*. IEEE, 2017, pp. 1–6.

[35] M. Fatemi, K. Granström, L. Svensson, F. J. Ruiz, and L. Hammarstrand, "Poisson multi-bernoulli mapping using Gibbs sampling," *IEEE Transactions on Signal Processing*, vol. 65, no. 11, pp. 2814–2827, 2017.

[36] T. Qin, P. Li, and S. Shen, "VINS-mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, 2018.

[37] R. Mur-Artal and J. D. Tardós, "Visual-inertial monocular SLAM with map reuse," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 796–803, 2017.

[38] K.-W. Chiang, G.-J. Tsai, Y.-H. Li, Y. Li, and N. El-Sheimy, "Navigation engine design for automated driving using INS/GNSS/3D LiDAR-SLAM and integrity assessment," *Remote Sensing*, vol. 12, no. 10, p. 1564, 2020.

[39] C. Forster, S. Lynen, L. Kneip, and D. Scaramuzza, "Collaborative monocular slam with multiple micro aerial vehicles," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Nov 2013, pp. 3962–3970.

[40] B. Steder, G. Grisetti, C. Stachniss, and W. Burgard, "Visual SLAM for flying vehicles," *IEEE Transactions on Robotics*, vol. 24, no. 5, pp. 1088–1093, 2008.

[41] H. Ye, Y. Chen, and M. Liu, "Tightly coupled 3d lidar inertial odometry and mapping," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 3144–3150.

[42] R. Li, J. Liu, L. Zhang, and Y. Hang, "Lidar/mems imu integrated navigation (slam) method for a small uav in indoor environments," in *2014 DGON Inertial Sensors and Systems (ISS)*. IEEE, 2014, pp. 1–15.

[43] J. Barra, S. Lesecq, M. Zarudniev, O. Debicki, N. Mareau, and L. Ouvry, "Localization system in GPS-denied environments using radar and IMU measurements: Application to a smart white cane," in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 1201–1206.

[44] Y. Almalioglu, M. Turan, C. X. Lu, N. Trigoni, and A. Markham, "Milli-RIO: Ego-motion estimation with low-cost millimetre-wave radar," *arXiv preprint arXiv:1909.05774*, 2019.

[45] A. Kramer, C. Stahoviak, A. Santamaria-Navarro, A.-A. Agha-Mohammadi, and C. Heckman, "Radar-inertial ego-velocity estimation for visually degraded environments," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020.

[46] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. MIT press, 2005.

[47] J. Sola, "Quaternion kinematics for the error-state Kalman filter," *arXiv preprint arXiv:1711.02508*, 2017.

[48] M. Kok, J. D. Hol, and T. B. Schön, "Using inertial sensors for position and orientation estimation," *arXiv preprint arXiv:1704.06053*, 2017.

[49] J. L. Crassidis, F. L. Markley, and Y. Cheng, "Survey of nonlinear attitude estimation methods," *Journal of guidance control and dynamics*, vol. 30, no. 1, p. 12, 2007.
[50] O. Woodman, "An introduction to inertial navigation," *University of Cambridge, Computer Laboratory, Tech. Rep. UCAMCL-TR-696*, 2007.
[51] I. Skog, P. Handel, J.-O. Nilsson, and J. Rantakokko, "Zero-velocity detection–An algorithm evaluation," *IEEE transactions on biomedical engineering*, vol. 57, no. 11, pp. 2657–2666, 2010.
[52] S. Särkkä, *Bayesian filtering and smoothing*. Cambridge University Press, 2013, vol. 3.
[53] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. New York: John Wiley and Sons, 2001.
[54] L. Narula, M. J. Murrian, and T. E. Humphreys, "Accuracy limits for globally-referenced digital mapping using standard GNSS," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2018, pp. 3075–3082.
[55] LORD Sensing MicroStrain, "3DM-GX5-25 Attitude and Heading Reference System," https://bit.ly/32CKIaO, accessed 2020-08-31.