ION

Robust GPS-Vision Localization via Integrity-Driven Landmark Attention

Sriramya Bhamidipati | **Grace Gao**

Department of Aeronautics and Astronautics
Stanford University
Stanford, CA, USA

Correspondence
Grace Gao, Department of Aeronautics and Astronautics, Stanford University, Stanford, CA, USA, 94305
Email: gracegao@stanford.edu

Abstract

For robust GPS-vision navigation in urban areas, we propose an integrity-driven landmark attention (ILA) technique via stochastic reachability. Inspired by cognitive attention in humans, we perform convex optimization to select a subset of landmarks from GPS and vision measurements that maximizes integrity-driven performance. Given known measurement error bounds in non-faulty conditions, our ILA technique follows a unified approach to address both GPS and vision faults and is compatible with any off-the-shelf estimator. We analyze measurement deviation to estimate the stochastic reachable set of positions associated with each landmark, which is parameterized via probabilistic zonotope (p-zonotope). We apply set union to formulate a p-zonotopic cost that represents the size of position bounds based on landmark inclusion/exclusion. We jointly minimize the p-zonotopic cost and maximize the number of landmarks via convex relaxation. For an urban data set, we demonstrate improved localization accuracy and robust predicted availability for a pre-defined risk and alert limit.

1 | INTRODUCTION

Sensor fusion, a widely-implemented framework for urban navigation, combines measurements from complementary sensors, such as GPS and vision, to address the limitations of individual contributors. These limitations induce time-varying bias in GPS and vision data that are termed *measurement faults*. For instance, GPS signals may be blocked or reflected by surrounding infrastructure, such as tall buildings and thick foliage, which significantly degrades the satellite visibility and induces a multipath effect in the received measurements (Zhu et al., 2018). In contrast, while visual odometry performs well in urban areas due to feature-rich surroundings, it suffers from data association errors due to similarities in building infrastructure, dynamic occlusions, and illumination variations (Gil et al., 2006). Furthermore, urban areas are susceptible to multiple measurement faults.

Recently, *integrity* is emerging as an important safety metric to assess urban navigation performance. Integrity is defined as the measure of trust that can be placed in the correctness of the computed position estimate (Tossaint et al., 2007). In particular, the measures of integrity that are of interest to us include *position error bounds* and *system availability*. The position error bound guarantees that the probability of absolute position error exceeding a certain value is smaller than or

equal to the pre-defined risk (Ochieng et al., 2003). The term position error bound is analogous to *protection level*, which is commonly used in the GNSS community as an integrity metric. The *alert limit* (AL) is the maximum tolerable position error beyond which the system is to be declared unavailable.

Traditionally, integrity is developed for GPS receivers operating in the aviation sector (Tossaint et al., 2007; Ochieng et al., 2003). These techniques utilize the redundancy in number of GPS signals available under open-sky conditions to mitigate signal-in-space (SIS) anomalies in satellite broadcast navigation messages (Misra & Enge, 2006). However, given the presence of multiple measurement faults, which are time-varying, applying integrity to sensor-fusion-based urban navigation is not straightforward (Joerger & Spenko, 2017). This is because in urban areas, there exists major challenges to addressing multiple measurement faults in both GPS and vision, which are summarized as follows:

- Given the availability of a large number of measurements (from GPS and vision) and the potential for multiple faults, evaluating all multi-fault hypotheses to perform fault detection and exclusion (FDE) leads to a combinatorial explosion (Woeginger, 2003), and is therefore, impractical to solve. For instance, with 30 available measurements, the number of multi-fault hypotheses to be evaluated already equals $30! \approx 1e^{32}$, where the symbol ! denotes factorial. In reality, the number of measurements in GPS-vision localization is even higher (in the order of hundreds or thousands).

- The measurement fault distributions are unknown, unbounded, multi-modal, and time-varying. Therefore, approximating the measurement faults by known distributions (Larson et al., 2019; Raboui et al., 2011; Gaussian-Pareto, Rayleigh, mixture model, etc.) is not reliable enough for safety-critical applications. The need for tailored FDE approaches for each sensor, in this case GPS and vision, further limits their practical applicability.

- While FDE is important, exclusion solely based on the analysis in the measurement domain may lead to poor dilution of precision (DOP) and thereby, degrade accuracy in the position domain (Won et al., 2014).

1.1 │ Related Work

A few FDE approaches have been explored in the literature for addressing measurement faults in urban areas. We divide the relevant literature into four broad categories as follows: a) FDE approaches for GNSS-only setup; b) FDE approaches for GNSS faults via aid from external sensors; c) FDE approaches for multiple fault modalities in sensor fusion; and d) feature selection techniques.

1.1.1 │ *FDE Approaches for GNSS-only Setup*

Prior works (Bhamidipati & Gao, 2018; Zhu et al., 2018; Ziebold et al., 2017) combined the *global test*, which is based on the normalized sum of squared measurement residuals, and the *local test*, which leverages the normalized measurement residuals of each satellite independently, to perform fault detection and fault exclusion, respectively. The authors of the Bayesian receiver autonomous integrity monitoring (RAIM) approach (Zhang & Gui, 2015) represented each GNSS satellite by a classification variable that indicated its faulty/non-faulty status, and thereafter, evaluated the posterior probabilities of the classification variable using Gibbs'

sampler. In the particle RAIM approach (Gupta & Gao, 2019), the authors utilized a particle filter to formulate the multi-modal probability distribution over the position and updated the associated particle weights using a traditional residual-based RAIM (Angus, 2006; Joerger et al., 2014). However, these approaches approximated the GNSS faults with known distributions such as Gaussian and mixture model, thereby limiting the urban navigation performance.

Schroth et al. (2008) developed a range consensus approach that computed the position estimate from a subset of four satellites, and thereafter evaluated the pseudorange residuals of the GNSS satellites that did not contribute to this solution. The subset with the most inliers (i.e., the pseudorange residuals below a certain threshold) was consequently utilized for the detection of faulty GNSS measurements. Similarly, based on the fact that non-faulty measurements are consistent with each other and the faulty ones are inconsistent, Hsu et al. (2017) performed FDE by formulating this consistency check as a greedy optimization problem that minimized the weighted sum of squared pseudorange residuals. A major limitation of these works is that their FDE analysis was executed in the measurement domain, and therefore did not guarantee the localization performance, which depends on the position error bound associated with the measurements detected as non-faulty.

Another line of work (Zhang et al., 2017) addressed an interesting perspective that, given the restricted satellite geometry in urban areas, the influence of excluding a faulty satellite (which is detected via traditional residual-based RAIM [Angus, 2006; Joerger et al., 2014]), would be greater than the fault itself. The authors developed an exclusion accuracy filter that compared the position estimates before and after exclusion to analyze the trade-off between the effect of fault magnitude and the influence of exclusion on the resulting DOP. However, this work is only designed to analyze the inclusion/exclusion of single faults at a time, and is therefore not scalable with an increase in the number of available measurements (e.g., GPS-vision sensor fusion) and multiple measurement faults.

In Kaddour et al. (2015), the authors performed FDE by projecting the GNSS pseudoranges from the measurement space to the information space of a Kalman filter. This work performed fault exclusion by projecting the measurements into the position domain, and thereafter, clustering the ones most likely to be non-faulty. However, clustering point-valued estimates in the presence of multiple measurement faults is quite challenging, and this work did not provide performance guarantees of their FDE technique.

In addition, there exists a plethora of FDE research (Joerger & Spenko, 2017; Zhu et al., 2018) on line-of-sight (LOS)/non-line-of-sight (NLOS) classifications that include dual-polarization antennae, code discriminator design, carrier-to-noise density-based weighting models, and so on.

1.1.2 | FDE Approaches for GNSS Faults via Aiding from External Sensors

Some existing works rely on external information sources to perform FDE for GNSS measurements in urban areas. Prior work (Binjammez et al., 2013) developed three phases of integrity checks that included assessing the position quality via traditional RAIM (Angus, 2006; Joerger et al., 2014), speed integrity via GPS Doppler, and map-matching accuracy via fuzzy inference systems. Another work (Zhang et al., 2018) developed a multipath detection algorithm that utilized a reference 3D map to predict and exclude the multipath signals using ray-tracing, and

thereafter, performed sensitivity analysis to estimate the confidence level of the estimated multipath prediction. However, these approaches have practical limitations because the offline map database was not always available, and its accuracy could not be guaranteed due to the dynamic changes in the urban surroundings.

In Ghinamo et al. (2010), the authors performed FDE via additional information provided by the Public Land Mobile Network (PLMN). In particular, the pseudorange residuals were formulated using the coarse position estimate obtained by analyzing the received signal power from the known cell tower locations in the PLMN database. However, this framework required access to additional infrastructure. Furthermore, this framework required unobstructed LOS signals from a sufficient number of cell towers, which is challenging in an urban setting given the surrounding tall buildings.

Shytermeja et al. (2014) utilized a fish-eye camera for isolating GPS faults with an assumption that the faults in vision were negligible. However, this approach did not account for the faults in fish-eye cameras that occur due to urban artifacts. Other research (El-Mowafy & Kubo, 2017; Li et al., 2017) performed FDE by utilizing odometry data from other sensors, such as inertial measurement units and speedometers.

1.1.3 | FDE Approaches for Multiple Fault Modalities in Sensor Fusion

Prior works (Al Hage et al., 2018, 2020) utilized an information filter for GNSS/INS data fusion and thereby assessed the Kullback-Leibler (KL) divergence between the predicted and updated distributions for fault detection. Other works (Maaref & Kassas, 2019; Maaref et al., 2020) fused terrestrial signals of opportunity (SOPs) with GPS signals to improve redundancy, and thereby performed FDE across GPS and SOP measurements using the traditional residual-based test statistic that considered the distribution to be centered chi-square in the absence of faults and non-centered chi-square otherwise.

Tmazirte et al. (2012) developed a distributed information filter for a multi-sensor framework, wherein faults were detected by examining the consistency through a log-likelihood ratio of the measurement residuals using a mutual information concept. However, the main drawback of these FDE approaches is that they modeled the measurement faults as Gaussian distributions, and are thus, not suitable for GPS-vision sensor fusion.

While research on GPS-vision navigation (Heredia et al., 2009; Shepard & Humphreys, 2014) is gaining momentum in recent times, to the best of our knowledge, there is only limited literature on FDE that accounts for multiple sensor fault modalities. Mohanty et al. (2020) proposed a GPS-vision sensor fusion via particle filter that accounted for vision faults by deriving a distribution of positions from camera images using map matching, and thereafter, formulated a KL-divergence metric to assess the consistency of GPS and vision measurements. However, analyzing vision faults requires access to a reference map database, and therefore, the reliability of this approach depends on the quality and accuracy of the map considered. In our prior works (Bhamidipati & Gao, 2019, 2020a), we performed FDE using analysis of temporal correlation across GPS measurement residuals and spatial correlation across vision intensity residuals. While these methodologies tackled both GPS and vision faults, the FDE analysis was conducted in the measurement domain, and moreover, separate FDE approaches are still to be designed for GPS and vision.

1.1.4 | *Feature Selection Techniques*

The algorithm developed in Lerner et al. (2007) constructed a requirement matrix that pre-defined the relative importance of each state (based on the navigation task), and thereafter searched for the subset of the minimum required measurements that collectively minimized the size of state covariance matrix. A prior work (Mousavi & Motee, 2020) developed a randomized algorithm for visual feature selection over a fixed-length moving time horizon by associating a sampling probability to each candidate feature, randomly sampling a subset of features according to these probabilities for a number of independent experiments, and selecting the outcome with the best estimation quality.

Considering the availability of only limited computational resources, Carlone and Karaman (2017) retained the most relevant visual cues that maximized the performance of visual-inertial navigation using a greedy optimization objective. While these optimization algorithms demonstrate satisfactory performance in their convergence, the cost function is only based on the state covariance that does not account for multiple measurement faults (non-zero bias) in urban areas.

In Joerger et al. (2017), the authors developed a multiple-hypothesis innovation-based data association for lidar-based navigation to select landmark data and detect unwanted obstacles. This was achieved by deriving a lower-bound on the mean innovation vector, which is used in the integrity risk equation to quantify the probability of correct association. While this analysis is one of the few prior works that account for measurement faults and integrity measures in their feature selection, their work is explicitly designed for detecting the data association errors in light detection and ranging (LiDAR) point-clouds across consecutive times and is therefore, not generalizable to a sensor-fusion framework. Another limitation is that they perform multiple-hypotheses solution separation for detecting unwanted obstacles that leads to computational intractability when a large number of measurements are considered (such as GPS and vision).

To address the earlier-described challenges of GPS-vision localization in urban areas and the gaps in existing literature, we formulate the following three research objectives for this work: a) to select a desirable subset of measurements from the available GPS and vision data that minimizes the position error bound; b) to design a unified approach that accounts for multiple faults in GPS and vision while ensuring computational tractability; and c) to predict integrity-driven measures of expected navigation performance for the measurement subset.

1.2 | Overview and Key Features of the Proposed Work: Integrity-Driven Landmark Attention

Our current work is inspired by the cognitive processes in humans (Torralba et al., 2006). A human brain can seamlessly process large amounts of sensory cues, such as visual, auditory, olfactory, haptic, and environmental stimuli, to ensure robust navigation (Caduff & Timpf, 2008). For instance, to navigate from Point A to Point B, our brain focuses on distinctive features, such as street names, stop signs, lane markings, building facade(s), etc. Furthermore, our brain ignores other cues like clouds, dry leaves, malfunctioning road signs, and accessories of other people. Therefore, humans navigate by giving *attention* to certain aspects in the surroundings, while *filtering out* the irrelevant or faulty ones.
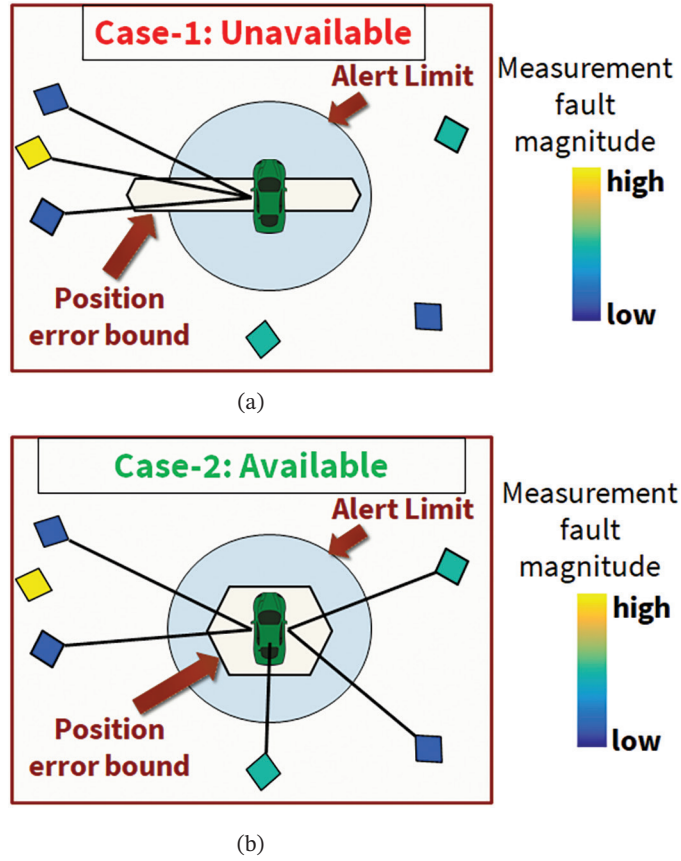
FIGURE 1 Maximizing integrity-driven navigation performance by varying the selected landmarks as follows: (a) unavailable, as AL is violated; and (b) unavailable, as AL is satisfied. The blue translucent circle denotes the AL, which is pre-defined for a particular navigation task. The diamonds denote landmarks, and their measurement quality is indicated by the Parula colormap (Nunez et al., 2018) that ranges from blue to yellow, with blue representing low fault magnitude and yellow representing high faults.

We propose an integrity-driven landmark attention (ILA) technique for GPS-vision navigation in urban areas. This work is based on our ION GNSS+ 2020 conference paper (Bhamidipati & Gao, 2020b). We define two terms: *landmark* and *attention*. Landmarks are the 3D features in urban surroundings that are used for navigation (Lerner et al., 2007). We consider two types of landmarks: GPS satellites, whose locations are known and calculated from broadcast ephemeris (Misra & Enge, 2006) and 3D visual features in urban infrastructure, whose locations are unknown and need to be triangulated via techniques such as simultaneous localization and mapping (SLAM; Cadena et al., 2016). Attention is the process of parsing a large number of measurements from GPS and vision to select landmarks that maximize the integrity-driven navigation performance in urban areas while ensuring computational tractability.

Figure 1 provides an intuitive understanding of maximizing integrity-driven navigation performance given the available landmarks. The subset of landmarks selected for localization in Figure 1(a) includes a high measurement fault magnitude and poor DOP. This causes the estimated position error bound based on a pre-defined risk to violate the pre-defined AL, and therefore, system is unavailable. In contrast, Figure 1(b) shows a desirable subset of landmarks that exclude the faults necessary to minimize the position error bound, thereby ensuring
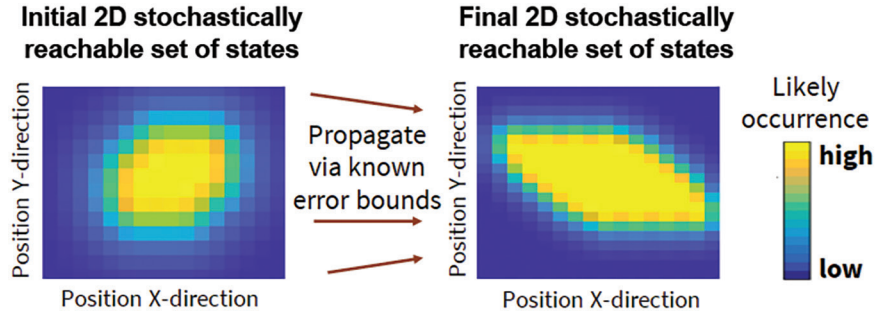
FIGURE 2 Illustration of SR that computes a final set of stochastic reachable set of 2D positions given an initial stochastic set and known error bounds, where blue represents the state with low likely occurrence and yellow indicates high likelihood.

compliance with the pre-defined AL. Our proposed ILA technique is referred to as *integrity-driven* because we compute the landmark subset that minimizes the position error bound for a pre-defined risk, wherein position error bound is analogous to an integrity measure called *protection level* that was described earlier.

Analyzing the position error bound accounts for multiple measurement faults in GPS and vision while ensuring statistical guarantees on the expected localization accuracy in urban areas. Note that the proposed ILA technique emphasizes the notion of integrity-driven FDE and therefore does not perform integrity monitoring in the traditional sense that was developed for aviation-grade GNSS receivers (Ochieng et al., 2003; Tossaint et al., 2007). Nevertheless, this proposed work draws multiple parallels with the various integrity measures commonly used in the GNSS community, such as protection levels, integrity risk, and system availability which were described earlier.

Compared to the existing integrity works (Joerger et al., 2014; Ochieng et al., 2003; Tossaint et al., 2007) that utilize probabilistic analysis, the formulation of position error bound in the proposed ILA technique is inspired by a set-valued approach known as *stochastic reachability* (SR; Althoff et al., 2009). SR, which is designed for stochastic processes, has been widely used in robotics for applications such as path planning and collision avoidance, but has not been applied to the field of fault resilience for sensor fusion.

As seen in Figure 2, SR computes a set of stochastic reachable states given an initial stochastic set of states, and the known stochastic sets of system disturbances and measurement errors. In this context, as seen in Figure 2, each state in the stochastic reachable set is associated with a stochastic measure that indicates the likelihood of its occurrence. In our proposed work, we represent the system states as a stochastic reachable set, and thereafter estimate its size to compute the position error bound for a pre-defined risk.

The three key features of the proposed ILA technique that are designed to meet the research objectives listed earlier at the end of Section 1.1 are a) adaptive measurement error bounds via probabilistic zonotopes (p-zonotopes); b) optimization framework; and c) SR-based cost function using p-zonotopes.

1.2.1 | *Adaptive Measurement Error Bounds via P-Zonotopes*

Compared to the prior work discussed in Section 1.1 that approximates the unknown and multi-modal measurement faults via known distributions, the proposed ILA technique only requires the known measurement error bounds of GPS

and vision landmarks in non-faulty conditions. Our SR-based ILA technique follows a unified set representation known as a *p-zonotope* to account for GPS and vision faults, and therefore does not require separate FDE approaches for each sensor. A p-zonotope is an enclosing probabilistic hull of all possible probability density functions (Althoff et al., 2009). A p-zonotope can enclose a variety of distributions such as mixture models, distributions with uncertain or time-dependent means, and non-Gaussian or unknown distributions. More details regarding the properties of p-zonotopes and their mathematical notations are given later in Figure 3 and Section 2.

Among the existing set representations (Althoff & Dolan, 2014; zonotopes, ellipsoids, polytopes, etc.) and well-studied probabilistic distributions (Larson et al., 2019; Raboui et al., 2011; Gaussian-Pareto, Gaussian mixture model, etc.), we choose one-dimensional (1D) p-zonotopes to independently represent the measurement error bounds of each landmark (among GPS and vision). This is because p-zonotopes can efficiently enclose the measurement errors of GPS and vision in non-faulty conditions irrespective of the nature of their distribution.

Utilizing these p-zonotopes of measurement errors in non-faulty conditions (i.e., known measurement error bounds), the proposed ILA independently analyzes the deviation of the most recent history of measurement residuals at each landmark to adaptively scale the error bounds at each time iteration. Therefore, the proposed ILA technique utilizes SR formulation to account for the measurement faults (biases in measurements caused by urban artifacts). We pre-compute the 1D p-zonotopes of measurement error bounds in non-faulty conditions during initialization, which is explained later in Section 4.2.

1.2.2 | *Optimization Framework*

To maximize the integrity-driven navigation performance of GPS-vision localization, we formulated an optimization framework that estimates a binary decision for each landmark indicating whether it is to be included or excluded during localization. Our optimization objective jointly maximizes the number of landmarks to be included while minimizing the SR-based position error bound.

Mathematically, we represent inclusion of a landmark by a binary value of 1 and exclusion by 0. Our ILA technique works with any off-the-shelf GPS-vision estimator, and is thus, modular and add-on in nature. Details regarding the off-the-shelf GPS-vision estimator used to validate our ILA technique are given later in Section 4.1.

1.2.3 | *SR-Based Cost Function using P-Zonotopes*

We designed an SR-based cost function for the optimization framework by parameterizing the stochastic reachable set of system states via p-zonotopes. As discussed in Section 1.1, the cost functions of existing feature selection techniques require computing the inverse of the state covariance matrix for each landmark. In contrast, our SR-based cost, termed *p-zonotopic cost*, leverages set properties (Althoff et al., 2009), such as Minkowski sums, unions, and convex hulls, to enable computationally tractable operations.

By using p-zonotopes to formulate our cost, we not only estimate a desirable subset of landmarks, but also a predicted measure of integrity termed as *system availability* for a pre-defined AL. We mention the term *predicted* because the proposed

ILA technique provides a conservative estimate of system availability expected when its estimated subset of landmarks are used for localization (valid for any off-the-shelf GPS-vision estimator). Note that the exact system availability depends on the specifics of the off-the-shelf estimator and, thus, estimating this metric is beyond the scope of the current work.

For a pre-defined risk, the proposed ILA technique provides a conservative (or predicted) estimate of the position error bound by evaluating our cost function at the estimated subset of landmarks. If the predicted position error bound is within the predefined AL, the system is predicted to be available when the off-the-shelf GPS-vision estimator utilizes the estimated landmark subset for localization.

1.3 | Key Contributions

The main contributions of this paper are listed as follows:

1. **Landmark Attention via Optimization:** We designed an optimization framework to select a desirable subset of landmarks from among available GPS and vision. We maximized the integrity-driven performance of GPS-vision localization in urban areas by formulating a total cost that jointly maximized the number of included (or selected) landmarks while minimizing the associated position error bound.

2. **P-zonotopic Cost via SR:** Parameterizing the known measurement error bounds of landmarks in non-faulty conditions using p-zonotopes (Althoff et al., 2009), we adaptively scaled these bounds based on a history of measurement residuals. By utilizing the set properties of SR, we computed the stochastic reachable set of expected states for each landmark via p-zonotopes (Althoff et al., 2009). We designed the p-zonotopic cost, which represents the position error bound, as a union function across landmarks, wherein each landmark was associated with a binary variable that indicated its inclusion or exclusion.

3. **Optimization via Convex Relaxation:** We formulated a convex optimization problem (Boyd & Vandenberghe, 2004) by relaxing the constraints on each landmark such that they layed in a continuous domain between [0, 1] instead of binary domain {0, 1}. We derived the sub-optimality gap that evaluated the convergence performance of convex relaxation. We also evaluated the p-zonotopic cost using the estimated desirable subset of landmarks to predict the system availability for a pre-defined risk and AL.

4. **Experimental Validation:** We have validated the performance of the proposed ILA technique on an urban sequence from a monocular visual odometry data set (Engel et al., 2016) that is combined with simulated GPS data. In the presence of multiple GPS and vision faults, the proposed ILA technique demonstrated an improved localization accuracy compared to the existing methods on landmark selection and FDE. For a quantitative analysis, we have also computed a robust measure of the predicted system availability across different pre-defined values of risk and AL.

The rest of the paper is organized as follows: Section 2 describes the preliminaries of SR and p-zonotopes; Section 3 describes the proposed ILA technique; Section 4 experimentally demonstrates the improved localization accuracy and robust measure of predicted system availability; and Section 5 concludes the paper.

2 | PRELIMINARIES OF SR AND P-ZONOTOPES

A *p-zonotope*, as mentioned earlier in Section 1.2, is an enclosing probabilistic hull of all possible probability density functions (Althoff et al., 2009) that can enclose a variety of distributions such as mixture models, distributions with uncertain or time-dependent mean, and non-Gaussian or unknown distributions. A p-zonotope $\mathcal{L}$, as seen in Equation (1) and Figure 3(a), is characterized by three parameters (Althoff et al., 2009), namely: (a) the mean of its center, denoted by $\boldsymbol{c} \in \mathbb{R}^n$; (b) the uncertainty in the center, represented by the generator matrix $G \in \mathbb{R}^{n \times e}$; and (c) the over-bounding covariance of the Gaussian tails, denoted by $\Sigma \in \mathbb{R}^{n \times n}$:

$$\mathcal{L} = (\boldsymbol{c}, G, \Sigma) \tag{1}$$

Intuitively, a p-zonotope with a certain center mean (i.e., zero center uncertainty $G = \boldsymbol{0}$), represents a Gaussian distribution that overbounds multiple probability density functions (PDFs) with the same center mean $\boldsymbol{c}$. This is represented by $\mathcal{Z}$ in Equation (2):

$$\mathcal{Z} = \left\{ \boldsymbol{x} \in \mathbb{R}^n \,\middle|\, \boldsymbol{x} = \boldsymbol{c} + \sum_{i=1}^{l} n_i \, \underline{g}^i, \; n_i = \mathcal{N}(0,1) \right\} \tag{2}$$

where $\{\cdot\}$ denotes the set representation and $n_i \, \forall i \in \{1, \cdots, l\}$ denotes pairwise independent Gaussian distributed random variables. $\underline{g}^i \in \mathbb{R}^n$ denotes the associated generator vectors of $\mathcal{Z}$, such that $\underline{G} = [\underline{g}^1, \cdots, \underline{g}^l]$ with $[\cdot, \cdot]$ as the horizontal concatenation operator. Note that, the covariance Σ of a p-zonotope defined earlier in Equation (1) is equivalent to $\Sigma = \underline{G}\,\underline{G}^\top$.

Similarly, a p-zonotope with zero covariance $\Sigma = \boldsymbol{0}$ simplifies to a zonotope (Althoff & Dolan, 2014) that encompasses all possible values of the center with the generator matrix of Z as $G = [g^1, \cdots, g^e]$, where $g^i \in \mathbb{R}^n \, \forall i \in \{1, \cdots, e\}$ denotes the associated generator vectors. This case is mathematically represented by Z in Equation (3):

$$Z = \langle \boldsymbol{c}, G \rangle = \left\{ \boldsymbol{x} \in \mathbb{R}^n \,\middle|\, \boldsymbol{x} = \boldsymbol{c} + \sum_{i=1}^{e} \beta_i \, g^i, \; -1 \le \beta_i \le 1 \right\} \tag{3}$$

where $\beta_i \, \forall i \in \{1, \cdots, e\}$ denotes independent random numbers. Note that the parameter G defined in Equations (1) and (3) is not the same as $\underline{G}$ defined in Equation (2).

The p-zonotope is a combination of two sets $\mathcal{Z}_0 = (0, \underline{G})$ and $Z = \langle \boldsymbol{c}, G \rangle$, such that $\mathcal{L} = \mathcal{Z}_0 \boxplus Z$, where $\boxplus$ is a set-addition operator. In other words, $\mathcal{L} = \sup\left\{ f_{\mathcal{Z}_0} \,\middle|\, \mathbb{E}[\mathcal{Z}_0] \in Z \right\}$, where $f_{\mathcal{Z}_0}$ is the probability density function (PDF) of the distributions represented by $\mathcal{Z}_0$ and $\mathbb{E}[\cdot]$ denotes the expectation operator. From Equations (1)–(3), note that $\mathcal{L}$ depends on both G and $\underline{G}$. Additionally, note that unlike a PDF, the area enclosed by a p-zonotope does not equal one. The illustrations of 1D and 2D p-zonotopes are shown in Figures 3(a) and 3(b), respectively. More details regarding the characteristics of zonotopes and p-zonotopes can be found in the prior literature (Althoff et al., 2009).

To perform set operations on stochastic reachable sets using p-zonotopes, we require six important properties (Althoff et al., 2009; Le Guernic & Girard, 2010):

- **Minkowski sum of addition:** The addition of two p-zonotopes $\mathcal{L}_1 = (\boldsymbol{c}_1, G_1, \Sigma_1)$ and $\mathcal{L}_2 = (\boldsymbol{c}_2, G_2, \Sigma_2)$ is given by:
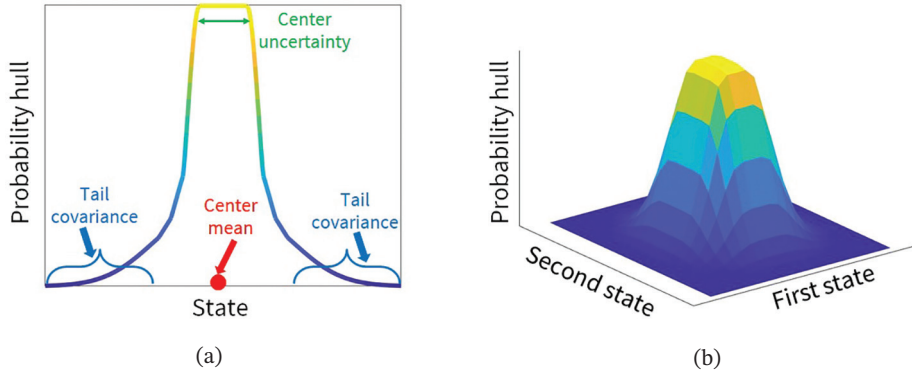
FIGURE 3 Illustration of p-zonotopes: (a) shows an example of a 1D p-zonotope whose probability is indicated by the Parula colormap (Nunez et al., 2018) while (b) shows an example of a 2D p-zonotope and is indicated by the Parula colormap (Nunez et al., 2018).

$$\mathcal{L}_1 \oplus \mathcal{L}_2 = \left\{ \boldsymbol{x}_1 + \boldsymbol{x}_2 \,\middle|\, \boldsymbol{x}_1 \in \mathcal{L}_1, \boldsymbol{x}_2 \in \mathcal{L}_2 \right\} = \left(\boldsymbol{c}_1 + \boldsymbol{c}_2, \left[G_1, G_2 \right], \Sigma_1 + \Sigma_2 \right) \tag{4a}$$

where $\oplus$ is known as Minkowski sum operator.

- **Linear map:** Given a matrix $A \in \mathbb{R}^{n \times n}$ and a p-zonotope $\mathcal{L} = (\boldsymbol{c}, G, \Sigma)$,

$$A\mathcal{L} = \left\{ Ax \,\middle|\, \boldsymbol{x} \in \mathcal{L} \right\} = (A\boldsymbol{c}, AG, A\Sigma A^{\top}) \tag{4b}$$

- **Translation:** Given a vector $\mu \in \mathbb{R}^n$ and a p-zonotope $\mathcal{L} = (\boldsymbol{c}, G, \Sigma)$,

$$\mu + \mathcal{L} = (\mu + \boldsymbol{c}, G, \Sigma) \tag{4c}$$

- **Union:** $\quad \mathcal{L}_1 \cup \mathcal{L}_2 = \left\{ \kappa \boldsymbol{x}_1 + (1 - \kappa) \boldsymbol{x}_2 \,\middle|\, \boldsymbol{x}_1 \in \mathcal{L}_1, \boldsymbol{x}_2 \in \mathcal{L}_2 \text{ where } 0 \le \kappa \le 1 \right\} \tag{4d}$

- **Intersection:** $\quad \mathcal{L}_1 \cap \mathcal{L}_2 = \left\{ \boldsymbol{x} \,\middle|\, \boldsymbol{x} \in \mathcal{L}_1, \boldsymbol{x} \in \mathcal{L}_2 \right\} \tag{4e}$

- **γ-confidence:** For $\mathcal{L} = \mathcal{Z}_0 \boxplus Z$ with $\mathcal{Z}_0 = \left(0, \underline{G} \right)$ and $Z = \left\langle \boldsymbol{c}, G \right\rangle$, transforming a p-zonotope $\mathcal{L}$ to a confidence zonotope, which is denoted by $Z_{\mathcal{L}}$, is given by:

$$Z_{\mathcal{L}} = Z \boxplus \left\langle 0, \gamma \underline{G} \right\rangle \tag{4f}$$

where γ denotes the pre-defined parameter that thresholds the tail of a p-zonotope. This property is used in the proposed ILA technique to estimate the size of the p-zonotope for a given γ. Note that $P\left[-\gamma \le \mathcal{N}(0,1) \le \gamma \right] = \mathrm{erf}\left(\dfrac{\gamma}{\sqrt{2}} \right)$ is well known, where erf denotes the error function (Oldham et al., 2009). Given l probabilistic generators in $\underline{G}$, the probability of lying outside the γ-confidence set is $p_R = 1 - \mathrm{erf}\left(\dfrac{\gamma}{\sqrt{2}} \right)^{2l}$. We consider this value p_R to be analogous to the term *risk* that is commonly used in the GNSS community as an integrity metric (Misra & Enge, 2006). As an example, for a pre-defined parameter $\gamma = 3.5$ and number of generators $l = 2$, the value of risk $p_R = 1.9e^{-3}$.

3 | PROPOSED ILA TECHNIQUE

The proposed ILA technique shown in Figure 4 consists of two modules: cost formulation via SR and optimization via convex relaxation. We performed an offline empirical analysis to compute the measurement error bounds of GPS and vision under non-faulty conditions, which is explained later in Section 4.2. At each time iteration, the following steps were executed.

1. We considered measurements from GPS, vision, and a known motion model. We used the motion model input to facilitate global convergence of our ILA and to perform measurement fault analysis.
2. For each landmark, we independently analyzed the deviation of received measurements against their known error bounds to estimate the p-zonotope of expected states. Thereafter, we applied the set union property of p-zonotopes to compute an over-approximated p-zonotope of expected states that was a function of landmarks to be selected.
3. We jointly optimized the p-zonotopic cost, which represents the size of the over-approximated p-zonotope for a pre-defined risk, and the number of landmarks to be selected via convex relaxation, and then estimated a desirable subset of landmarks. Thereafter, given a pre-defined AL, we computed a predicted measure of system availability. Details regarding the need for over-approximation in set unions and the size computation of a p-zonotope are explained later in Section 3.2.
4. The estimated desirable subset of landmarks using the proposed ILA technique are later given to any off-the-shelf GPS-vision estimator to perform localization. More details regarding the off-the-shelf estimator are given in Section 4.1.

Denoting the number of GPS landmarks by N and the number of vision landmarks by L, we defined an attention set Q that consisted of binary variables to be estimated via convex optimization, such that $Q = \left\{ q_{\text{gps}}^1, \cdots, q_{\text{gps}}^N, q_{\text{vis}}^1, \cdots, q_{\text{vis}}^L \right\}$, where $q_{\text{gps}}^i \in \{0,1\} \ \forall i \in \{1, \cdots, N\}$ and $q_{\text{vis}}^j \in \{0,1\} \ \forall j \in \{1, \cdots, L\}$.
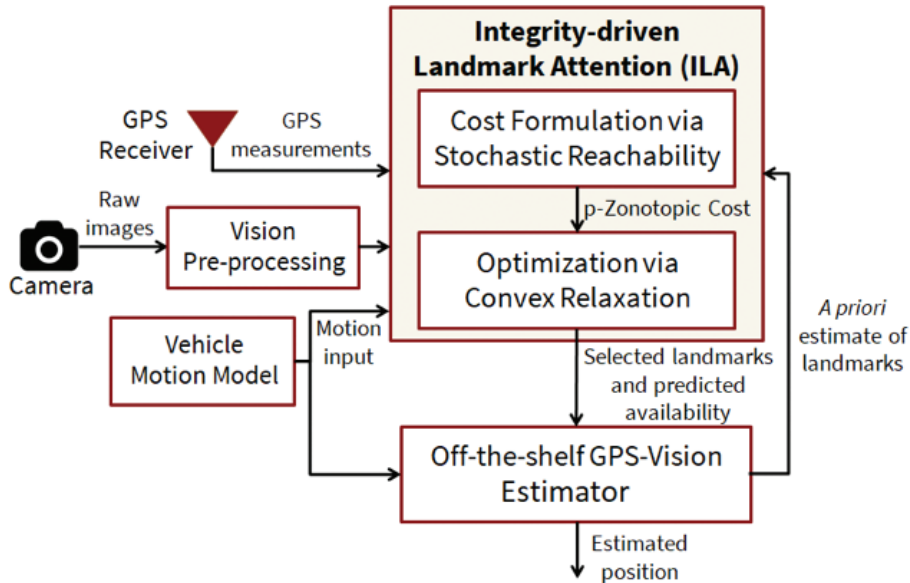

FIGURE 4 Architecture of the proposed ILA technique comprised of cost formulation via SR (Althoff et al., 2009) and optimization via convex relaxation (Boyd et al., 2004).

As mentioned earlier, a binary value of one indicates that the landmark is included/selected and zero indicates it is excluded. Let's consider the optimal solution of the attention set for this minimization problem as $Q^\dagger$. We formulated the total convex cost, as seen in Equation (5), by jointly minimizing the p-zonotopic cost and maximizing the number of landmarks to be selected. Furthermore, we defined inequality constraints to ensure that a sufficient number of landmarks were available for localization by the off-the-shelf GPS-vision estimator.

$$\underset{q_{\text{gps}}^1,\cdots,q_{\text{gps}}^N,q_{\text{vis}}^1,\cdots,q_{\text{vis}}^L}{\text{minimize}} \left(\frac{1}{\sum_{i=1}^N q_{\text{gps}}^i + \sum_{j=1}^L q_{\text{vis}}^j} \right) f\left(q_{\text{gps}}^1,\cdots,q_{\text{gps}}^N,q_{\text{vis}}^1,\cdots,q_{\text{vis}}^L\right) \quad (5)$$

$$\text{subject to} \sum_{i=1}^N q_{\text{gps}}^i > N_{\min}$$

$$\sum_{j=1}^L q_{\text{vis}}^j > L_{\min}$$

$$q_{\text{gps}}^i \in \{0,1\} \forall i \in \{1,\cdots,N\}$$

$$q_{\text{vis}}^j \in \{0,1\} \forall j \in \{1,\cdots,L\}$$

where $f(Q) \in \mathbb{R}$ denotes the scalar p-zonotopic cost as a function of the attention set Q and is constructed to be convex later in Section 3.2. This p-zonotopic cost $f(Q)$ depends on the p-zonotope of expected states $\mathcal{L}(Q)$ that represents the set union of landmarks to be included (or selected); $N_{\min}$ and $L_{\min}$ are the pre-defined constants that represent the minimum number of GPS and vision landmarks to be included and are set during initialization, which is explained later in Section 4.1.

A lower p-zonotopic cost $f(Q)$ indicates better navigation performance as the position error bound is lower. Also, the larger the number of included/selected landmarks, the higher the confidence in the associated position error bound. In the proposed ILA technique, we computed the attention set Q^*, which indicated the landmarks to be included, and the predicted system availability for a pre-defined risk and AL, which is associated with the p-zonotope $\mathcal{L}(Q^*)$.

3.1 | Input Measurements

We considered the measurements from the GPS, vision, and motion model as inputs to the proposed ILA technique. As explained earlier, the positions of GPS satellite landmarks are known and calculated from the ephemeris (Misra & Enge, 2006), while the positions of visual landmarks (i.e., 3D urban features) are unknown and localized in the off-the-shelf GPS-vision estimator.

We denote the state vector at the k-th time iteration by x_k, such that $x_k = [\mathbf{x}, \psi, c\delta t]_k$, where $\mathbf{x}$ and ψ denote the 3D position and 3D orientation of the system with respect to a global reference frame, respectively, and $c\delta t$ denotes the clock bias of the GPS receiver. The global reference frame is set during initialization of the off-the-shelf GPS-vision estimator, which is explained later in Section 4.1. Note that the number of GPS and vision landmarks (i.e., N and L, respectively) are not required to be constant across time; we simply drop the subscript k for notational simplicity.

In the GPS module, given N visible satellite landmarks, the pseudorange measurement received from the i-th satellite at the k-th time iteration is denoted by $z_{\text{gps},k}^i$, such that:

$$z_{\mathrm{gps},k}^i = h^i\left(\boldsymbol{x}_k\right) + f_{\mathrm{gps},k}^i + \eta_{\mathrm{gps},k}^i \qquad (6)$$

where h^i represents the GPS measurement model (Misra & Enge, 2006) associated with the i-th satellite, with $h^i(\boldsymbol{x}_k) = \left\|\mathbf{y}^i - \mathbf{x}\right\| + \left(c\delta t - c\delta t^i\right)$; $\mathbf{y}^i$ and $c\delta t^i$ denote the 3D position in the global reference frame and clock corrections in the i-th satellite, respectively; $f_{\mathrm{gps},k}^i$ and $\eta_{\mathrm{gps},k}^i$ represent the measurement fault (due to multipath) and stochastic noise associated with the i-th satellite, respectively.

In the vision module, we performed direct image alignment (Engel et al., 2018; Forster et al., 2014) to formulate the photometric difference between pixels across two image frames. As seen in Equation (7), we projected a 2D pixel coordinate (denoted by $\mathbf{u}$) from the keyframe to a current image frame at the k-th time iteration. In this context, the keyframe denotes the reference image obtained from the monocular camera relative to which the position and orientation of the current image is calculated. In a keyframe, the selected pixels with high intensity gradient are termed *key pixels*. To compute the semi-dense depth map of key pixels, we executed a short-temporal baseline matching of key pixels across images to replicate the notion of stereo-matching. Detailed explanations regarding the keyframe selection and estimation of semi-dense depth maps is given in prior literature (Engel et al., 2014, 2018; Forster et al., 2014).

$$\text{For } \mathbf{u} \in \Pi_{\mathrm{kf}},\ I_k(\mathbf{u}) = I_{\mathrm{kf}}\left(\pi\left(\omega\left(\boldsymbol{x}_k,\mathbf{u}\right)\right)\right) + f_I(\mathbf{u}) + \xi_I(\mathbf{u}) \qquad (7)$$

with

$$\omega(\boldsymbol{x}_k,\mathbf{u}) = \mathbf{R}\left(\boldsymbol{x}_k - \boldsymbol{x}_{\mathrm{kf}}\right)\pi^{-1}\left(\mathbf{u},d_{\mathrm{kf}}(\mathbf{u})\right) + \mathbf{t}\left(\boldsymbol{x}_k - \boldsymbol{x}_{\mathrm{kf}}\right)$$

$$\pi(\mathbf{p}) = \begin{bmatrix} f_x & 0 \\ 0 & f_y \end{bmatrix}\begin{bmatrix} \dfrac{p_x}{p_z} \\ \dfrac{p_y}{p_z} \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \end{bmatrix} \quad \text{from Forster et al. (2014)}$$

where the subscript kf refers to keyframe (Engel et al., 2014):
- $\Pi_{\mathrm{kf}} \subset \mathbb{R}^2$ and $\Pi \subset \mathbb{R}^2$ denote the domain of the keyframe and current image at the k-th time iteration, respectively. The image domain of the keyframe consists of only the key pixels;
- $I_{\mathrm{kf}}(\mathbf{u}): \Pi_{\mathrm{kf}} \to \mathbb{R}$ and $I_k(\mathbf{u}): \Pi \to \mathbb{R}$ denote the intensity of 2D pixel coordinates $\mathbf{u}$ in the keyframe and current image at the k-th time iteration, respectively;
- $\pi : \mathbb{R}^3 \to \Pi$ is the projection function that maps the 3D coordinates of camera frame, denoted by $\mathbf{p} = [p_x, p_y, p_z]^\top$, to 2D pixels. The origin of the 3D camera frame is at its optical center and its x- and y-axis are parallel to the u- and v-axis of the 2D image plane, respectively. From existing literature (Engel et al., 2014), we defined the projection function of the monocular camera via a pinhole model. We calibrated the intrinsic parameters, namely f_x, f_y, c_x, and c_y, and radial distortion parameters during initialization, which is explained later in Section 4.1;
- $\omega(\boldsymbol{x}_k, \mathbf{u})$ denotes the 3D warp function that unprojects the pixel coordinates $\mathbf{u}$ and transforms it by a relative change in the state vector. This relative change is the difference between the current state vector, given by $\boldsymbol{x}_k$ with respect to that of the keyframe, given by $\boldsymbol{x}_{\mathrm{kf}}$;

- $\mathbf{R}(\cdot) \in \mathbb{R}^{3\times3}$ and $\mathbf{t}(\cdot) \in \mathbb{R}^3$ denote the rotation matrix and translation vector, respectively, of frame transformation from the global reference frame to current 3D camera frame;
- $\pi^{-1}(\mathbf{u}, d_{kf}(\mathbf{u})): \Pi_{kf} \times \mathbb{R}^+ \rightarrow \mathbb{R}^3$ denotes the unprojection function (Engel et al., 2014) that maps the 2D pixel $\mathbf{u}$ to 3D camera coordinates via an inverse-depth, which is denoted by $d^{kf}(\mathbf{u})$; and
- $f_I(\mathbf{u})$ and $\xi_I(\mathbf{u})$ indicate the measurement fault due to data association of pixel $\mathbf{u}$ and measurement noise, respectively.

The number of visual landmarks is $L = |\Pi_{kf}|$, where $|\Pi_{kf}|$ is the cardinality of image domain Π_{kf}. From Equation (7), we formulate the non-linear vision measurement model as:

$$z_{vis,k}^j = b^j\left(\mathbf{x}_k, \mathbf{p}^j\right) + f_{vis,k}^j + \xi_{vis,k}^j \tag{8}$$

where $\mathbf{p}^j \ \forall j \in \{1, \cdots, L\}$ represents the 3D position of the j-th vision landmark in the global reference frame. The mapping from key pixels to the vision landmarks is bijective, i.e., $\mathbf{p}^j = \pi^{-1}\left(\mathbf{u}, d_{kf}(\mathbf{u})\right) \forall \mathbf{u} \in \Pi_{kf}$; b^j represents the vision measurement model, such that $b^j\left(\mathbf{x}_k, \mathbf{p}_k^j\right) = I_k\left(\pi\left(\omega\left(\mathbf{x}_k, \pi\left(\mathbf{p}^j\right)\right)\right)\right)$; $f_{vis,t}^j$ and $\xi_{vis,t}^j$ represent the measurement fault (due to data association errors) and intensity noise associated with the j-th vision landmark in keyframe and image frame at the k-th iteration, respectively.

In the motion model, we considered a linear state transition model to compute the state vector at the k-th time iteration, as seen in Equation (9). Note that the proposed ILA technique is generalizable to any motion model, both linear or non-linear. Given that the focus of the proposed ILA technique is on GPS and vision faults, we considered no measurement faults in the motion model. Refer to our prior work (Bhamidipati & Gao, 2018) for addressing faults in motion inputs.

$$z_{mm,k} = F\mathbf{x}_{k-1} = \mathbf{x}_k + \mathbf{v}_k \tag{9}$$

where $z_{mm,k}$ denotes the estimated state vector using a known motion model F; and $\mathbf{v}_k$ represents the noise vector; $\mathbf{x}_{k-1}$ denotes the state vector computed at the previous time iteration via the off-the-shelf estimator, whose details are given later in Section 3.3.

3.2 | Cost Formulation via SR

Utilizing the motion model, history of received measurements from GPS and vision, and their known error bounds in non-faulty conditions, we formulated the p-zonotopic cost $f(Q_k)$, defined earlier in Equation (5), as a function of the attention set $Q_k = \left\{q_{gps,k}^1, \cdots, q_{gps,k}^N, q_{vis,k}^1, \cdots, q_{vis,k}^L\right\}$. We denote the known error bounds as follows: GPS measurement noise $\eta_{gps,k}^i$ by a p-zonotope $\mathcal{L}_{\eta_{gps},k}^i$, vision measurement noise $\xi_{vis,k}^j$ by a p-zonotope $\mathcal{L}_{\xi_{vis},k}^j$, and noise associated with the motion model $\mathbf{v}_k$ by a p-zonotope $\mathcal{L}_{\mathbf{v}_{mm},k}$. The details regarding the estimation of measurement error bounds of GPS and vision in non-faulty conditions are given later in Section 4.2. We performed the following steps at the k-th iteration:

- **Step 1:** Using the properties of SR, we formulated the p-zonotope of expected states by considering non-faulty conditions for all landmarks (i.e., $f_{gps,k}^i = 0 \ \forall i, f_{vis,k}^j = 0 \ \forall j$). For this, as seen in Equations (10a)–(10c), we linearize the non-linear measurement models of GPS and vision that

are given in Equations (6) and (8). Although some variants of SR explicitly handle non-linear state-space equations (August et al., 2012), we exclude them from the scope of the current work.

$$\Delta z_{\text{gps},k}^i = H_k^i \Delta \boldsymbol{x}_k + \eta_{\text{gps},k}^i \ \forall i \in 1, \cdots, N \tag{10a}$$

$$\Delta z_{\text{vis},k}^j = B_{\boldsymbol{x},k}^j \Delta \boldsymbol{x}_k + B_{\mathbf{p},k}^j \mathbf{p}_{\text{apriori}}^j + \xi_{\text{vis},k}^j \ \forall j \in 1, \cdots, L \tag{10b}$$

$$\Delta z_{\text{mm},k} = \Delta \boldsymbol{x}_k + \nu_k \tag{10c}$$

where $\Delta \boldsymbol{x}_k = \boldsymbol{x}_k - \boldsymbol{x}_{\text{apriori}}$, such that $\boldsymbol{x}_{\text{apriori}}$ is the a priori guess of the state vector used for linearization; $\mathbf{p}_{\text{apriori}}^j \ \forall j \in \{1, \cdots, L\}$ is the a priori estimate of 3D visual landmarks, such that $\mathbf{p}_{\text{apriori}}^j = \bar{\mathbf{p}}_{k-1}^j$ (assuming the visual landmarks to be stationary), where $\bar{\mathbf{p}}_{k-1}^j$ is obtained from the off-the-shelf GPS-vision estimator; H_k^i denotes the Jacobian matrix of GPS measurement model, such that $H_k^i = \dfrac{dh^i(\boldsymbol{x}_k)}{d\boldsymbol{x}}\Big|_{\boldsymbol{x}_k = \boldsymbol{x}_{\text{apriori}}}; B_{\boldsymbol{x},k}^j$ and $B_{\mathbf{p},k}^j$ denote the Jacobian matrices for vision, such that $B_{\boldsymbol{x},k}^j = \dfrac{db_k^j(\boldsymbol{x},\mathbf{p}^j)}{d\boldsymbol{x}}\Big|_{\boldsymbol{x}=\boldsymbol{x}_{\text{apriori}}, \mathbf{p}^j = \mathbf{p}_{\text{apriori}}^j}$ and $B_{\mathbf{p},k}^j = \dfrac{db_k^j(\boldsymbol{x},\mathbf{p}^j)}{d\mathbf{p}^j}\Big|_{\boldsymbol{x}=\boldsymbol{x}_{\text{apriori}}, \mathbf{p}^j = \mathbf{p}_{\text{apriori}}^j}$, respectively.

We applied set properties in Equations (4a)–(4c) to compute the p-zonotope of expected state estimation errors that are associated with the motion model, GPS landmarks for $i \in \{1, \cdots, N\}$, and vision landmarks for $j \in \{1, \cdots, L\}$. These p-zonotopes of expected state estimation errors are defined as follows: $\mathcal{L}_{\Delta \boldsymbol{x}_{\text{gps}},k}^i$ for the i-th GPS landmark in Equation (11a), $\mathcal{L}_{\Delta \boldsymbol{x}_{\text{vis}},k}^j$ for the j-th 3D visual landmark in Equation (11b), and $\mathcal{L}_{\Delta \boldsymbol{x}_{\text{mm}},k}$ for the motion model in Equation (11c). Note that during the non-faulty case, the errors in received landmark measurements lie within the known error bounds.

$$\mathcal{L}_{\Delta \boldsymbol{x}_{\text{gps}},k}^i = \left(\left(H_k^i \right)^\top H_k^i \right)^{-1} \left(H_k^i \right)^\top \left(\Delta z_{\text{gps},k} + \mathcal{L}_{\eta_{\text{gps}},k}^i \right) \tag{11a}$$
$$\forall i \text{ using Equation (4c)}$$

$$\mathcal{L}_{\Delta \boldsymbol{x}_{\text{vis}},k}^j = \left(\left(B_{\boldsymbol{x},k}^j \right)^\top B_{\boldsymbol{x},k}^j \right)^{-1} \left(B_{\boldsymbol{x},k}^j \right)^\top \left(B_{\mathbf{p},k}^j \mathcal{L}_{\mathbf{p},k}^j \oplus \left(\Delta z_{\text{vis},k} + \mathcal{L}_{\xi_{\text{vis}},k}^j \right) \right) \tag{11b}$$
$$\forall j \text{ using Equations (4a)–(4c)}$$

$$\mathcal{L}_{\Delta \boldsymbol{x}_{\text{mm}},k} = \Delta \boldsymbol{z}_{\text{mm},k} + \mathcal{L}_{\nu_{\text{mm}},k} \ \text{ using Equation (4c)} \tag{11c}$$

where $\mathcal{L}_{\mathbf{p},k}^j$ is the p-zonotope of the j-th visual landmark that is formulated from its estimated position and uncertainty using the off-the-shelf GPS-vision estimator. Observing Equations (11a)–(11b), note that compared to GPS landmarks, vision landmark formulation has an additional p-zonotope term indicating its localization uncertainty.

- **Step 2:** We utilized the set union property of p-zonotopes, which was defined earlier in Equation (4d), to combine the p-zonotopes of expected states from the motion model, as well as the included GPS and vision landmarks. Therefore, the set union of p-zonotopes, as seen in Equation (12), is a linear function of the attention set Q_k. Note that the exact state estimation error in GPS-vision localization depends on the associated off-the-shelf estimator, which is described later in Section 4.1.
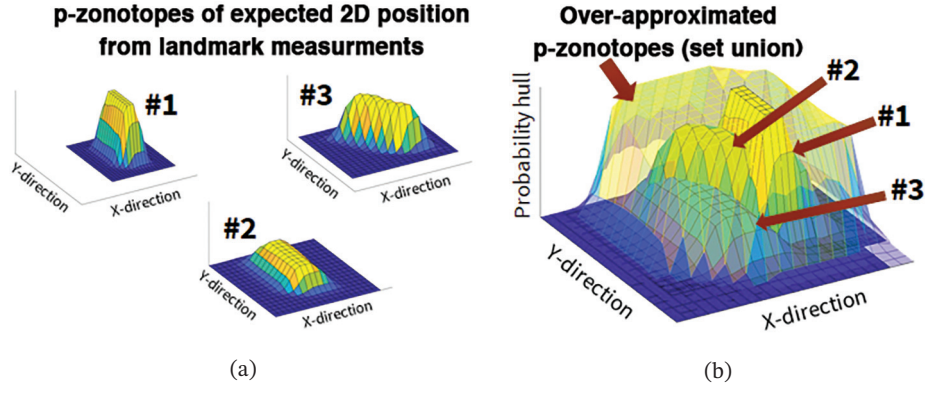
FIGURE 5 An example illustrating the set union operation and over-approximation; (a) shows p-zonotopes of expected 2D position from three landmarks, and (b) shows the over-approximated 2D p-zonotope that encloses the union of the p-zonotopes.

For given GPS and vision measurement models at any k-th time iteration, the set union of p-zonotopes computes an over-approximated (or predicted) position error bound that is valid for any off-the-shelf estimator. Given that set union is not closed under p-zonotopes, we over-approximate the union of p-zonotopes via a p-zonotope (Girard & Le Guernic, 2010). Based on Equation (4d), we denote the over-approximated p-zonotope in Equation (12) by $\mathcal{L}(Q)$.

For an intuitive understanding of set union, Figure 5(a) shows an example of p-zonotopes of expected 2D position from three landmarks #1 to #3, and Figure 5(b) demonstrates the set union of these three p-zonotopes that is over-approximated by another p-zonotope. During implementation, we utilized the *enclose function* of the MATLAB Continuous Reachability Analyzer (CORA) toolbox (Althoff, 2016) to perform set union operation on p-zonotopes. Note that Equation (12) considers the measurement errors to lie within the known error bounds, but, this assumption is invalid during the presence of faults.

$$\mathcal{L}(Q_k) = \mathcal{L}_{\Delta \boldsymbol{x}_{\mathrm{mm}},k} \bigcup_{i \in N} q_{\mathrm{gps}}^i \mathcal{L}_{\Delta \boldsymbol{x}_{\mathrm{gps}},k}^i \bigcup_{j \in L} q_{\mathrm{vis}}^j \mathcal{L}_{\Delta \boldsymbol{x}_{\mathrm{vis}},k}^j \tag{12}$$

- **Step 3:** We designed a unified approach to account for GPS and vision measurement faults by leveraging the stochastic nature of p-zonotopes. We utilized the received measurements and motion model to estimate the measurement innovation for GPS and vision landmarks as $\epsilon_{\mathrm{gps}}^i = \Delta \boldsymbol{z}_{\mathrm{gps}}^i - H^i \Delta \boldsymbol{x}_{\mathrm{mm}}$ and $\epsilon_{\mathrm{vis}}^j = \Delta \boldsymbol{z}_{\mathrm{vis}}^j - B_{\boldsymbol{x}}^j \Delta \boldsymbol{x}_{\mathrm{mm}} - B_{\mathbf{p}}^j \mathbf{p}_{\mathrm{apriori}}^j$, respectively, with $\Delta \boldsymbol{x}_{\mathrm{mm}} = \boldsymbol{z}_{\mathrm{mm}} - \boldsymbol{x}_{\mathrm{apriori}}$. Utilizing the p-zonotope of expected states from motion model $\mathcal{L}_{\Delta \boldsymbol{x}_{\mathrm{mm}}}$ defined earlier in Equation (11c), we applied the set properties of p-zonotopes in Equations (4a)–(4b) on the measurement innovation to compute the p-zonotope of expected measurement innovation for GPS and vision landmarks as follows:

$$\mathcal{L}_{\epsilon_{\mathrm{gps}}}^i = H^i \mathcal{L}_{\Delta \boldsymbol{x}_{\mathrm{mm}}} \oplus \mathcal{L}_{\eta_{\mathrm{gps}}}^i \tag{13}$$

$$\mathcal{L}_{\epsilon_{\mathrm{vis}}}^j = B_{\boldsymbol{x},k}^j \mathcal{L}_{\Delta \boldsymbol{x}_{\mathrm{mm}}} \oplus B_{\mathbf{p},k}^j \mathcal{L}_{\mathbf{p}_{\mathrm{apriori}}}^j \oplus \mathcal{L}_{\xi_{\mathrm{vis}}}^j \tag{14}$$

For each landmark, we independently analyzed the deviation of estimated measurement innovation from the p-zonotope of expected measurement innovation that is indicative of non-faulty measurement errors. The

p-zonotope of expected measurement innovation provides a probability value corresponding to each estimated measurement innovation and this probability is denoted by $\alpha_{\text{gps}}^i \ \forall i \in \{1, \cdots, N\}$ for GPS landmarks and $\alpha_{\text{vis}}^j \ \forall j \in \{1, \cdots, L\}$ for vision landmarks.

Based on this, for each available landmark, we independently computed the fault status of received measurements by normalizing the obtained probability value with the probability value corresponding to the center mean of the p-zonotope and subtracting it from one. Each measurement fault status $\tilde{\alpha}_{\text{gps}}^i \in \mathbb{R} \ \forall i$ and $\tilde{\alpha}_{\text{vis}}^j \in \mathbb{R} \ \forall j$ lies between [0, 1].

In a non-faulty condition, the estimated measurement innovation and its p-zonotope of expected measurement innovation are in close agreement, and hence a low fault status ≈ 0 is obtained. However, in the presence of measurement faults, the estimated measurement innovation does not comply with this expected p-zonotope, and therefore a high value of fault status ≈ 1 is observed.

We utilized the history of past K measurements to compute a joint fault status that represented the temporal confidence in the landmark. This also satisfies the need for temporal measurements of a visual landmark to triangulate its unknown position. At the k-th time iteration, for each GPS landmark $i \in \{1, \cdots, N\}$, we utilized $\boldsymbol{z}_{\text{gps},K}^i = \left\{ z_{\text{gps},k-K}^i, \cdots, z_{\text{gps},k}^i \right\} \in \mathbb{R}^K$ to compute the corresponding joint fault status $\alpha_{\text{gps},k}^i \in \mathbb{R}$, and for each visual landmark $j \in \{1, \cdots, N\}$, we utilize $\boldsymbol{z}_{\text{vis},K}^j = \left\{ z_{\text{vis},k-K}^j, \cdots, z_{\text{vis},k}^j \right\} \in \mathbb{R}^K$ to compute the corresponding joint fault status $\alpha_{\text{vis},k}^j \in \mathbb{R}$. We utilized the inverse of one minus estimated fault status for each landmark to adaptively scale the p-zonotope of measurement error bound in non-faulty conditions. Intuitively, to minimize the size of set union, we performed a trade-off between including the scaled measurement error bounds and analyzing the impact of its exclusion on the size of set union. Thereafter, we modified the union formulation of p-zonotopes in Equation (12) to adaptively account for measurement faults, as seen in Equation (15).

$$\mathcal{L}(Q_k) = \mathcal{L}_{\Delta \boldsymbol{x}_{\text{mm}},k} \bigcup_{i \in N} \left(q_{\text{gps}}^i \left(1 - \tilde{\alpha}_{\text{gps},k}^i \right)^{-1} \mathcal{L}_{\Delta \boldsymbol{x}_{\text{gps}},k}^i \right) \bigcup_{j \in L} \left(q_{\text{vis}}^j \left(1 - \tilde{\alpha}_{\text{vis},k}^j \right)^{-1} \mathcal{L}_{\Delta \boldsymbol{x}_{\text{vis}},k}^j \right) \quad (15)$$

- **Step 4:** We formulated the scalar p-zonotopic cost $f(Q_k)$ defined earlier in Equation (5) to denote the size of the over-approximated p-zonotope $\mathcal{L}(Q_k)$. Based on Section 2, for computing a finite size of the p-zonotope, we threshold the over-approximated p-zonotope of expected states $\mathcal{L}(Q_k)$ based on a γ-confidence, which was described earlier in Equation (4f). An example of this thresholding is seen in Figure 6.

  Note that the value of γ is chosen based on the pre-defined risk p_R for a navigation task. The relationship between γ and p_R is described earlier in Equation (4f). The cut-off p-zonotope for a pre-defined risk is represented by a confidence zonotope, denoted by $Z(Q_k)$. From Equation (3), we defined $Z(Q_k) = \left\langle c_Z(Q_k), G_Z(Q_k) \right\rangle$, where $c_Z(Q_k)$ and $G_Z(Q_k)$ denote the center and generator matrix of the confidence zonotope, respectively. From Althoff and Dolan (2014), the size of the confidence zonotope $Z(Q_k)$ is given by $G_Z^\top(Q_k) G_Z(Q_k)$.

  As seen in Equation (16), we defined the p-zonotopic cost as the weighted sum of eigenvalues of $G_Z^\top(Q_k) G_Z(Q_k)$ along each axis. We perform a weighted sum to account for the perturbation sensitivity of the state estimation error
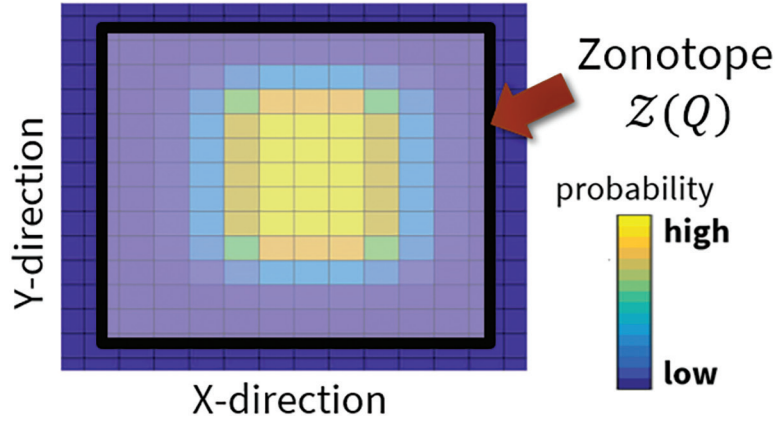
FIGURE 6 Illustration representing the top view of an over-approximated p-zonotope of expected 2D position, denoted by a Parula colormap (Nunez et al., 2018). For a given pre-defined value of risk, the over-approximated p-zonotope is transformed to a confidence zonotope $Z(Q)$, which is indicated by a translucent white box with a black border.

across each axis. The weights also account for the linearized approximation of measurement models considered for applying SR in Step 1.

$$f(Q_k) = \sum_{n=1}^{M} w_n \lambda_n \left(G_Z^\top(Q_k) G_Z(Q_k) \right) \text{ from Equation (4f)} \tag{16}$$

where λ_n denotes the eigenvalue of $G_Z^\top(Q_k) G_Z(Q_k)$ and w_n represents the associated weights, such that $\sum_{n=1}^{M} w_n = 1$ and M denotes the number of states in the vector $\boldsymbol{x}_k$.

3.3 | Optimization via Convex Relaxation

We utilized the p-zonotopic cost derived in Equation (16) to formulate the binary convex optimization problem represented earlier in Equation (5). However, given the binary constraints on the attention set Q_k, computing the optimal solution, which is defined to be $Q_k^\dagger$ earlier in Equation (5), is still a non-deterministic polynomial-time (NP) hard problem, and therefore, not computationally tractable. Therefore, we performed *convex relaxation* (Boyd et al., 2004; Carlone & Karaman, 2017) in which the binary constraints of attention set {0, 1} were replaced by convex constraints that lie in a continuous domain between [0, 1]. The modified optimization problem of the proposed ILA technique via convex relaxation is:

$$\underset{Q=\{q_{\text{gps}}^1, \cdots, q_{\text{gps}}^N, q_{\text{vis}}^1, \cdots, q_{\text{vis}}^L\}}{\text{minimize}} \left(\frac{1}{\sum_{i \in N} q_{\text{gps}}^i + \sum_{j \in L} q_{\text{vis}}^j} \right) f\left(q_{\text{gps}}^1, \cdots, q_{\text{gps}}^N, q_{\text{vis}}^1, \cdots, q_{\text{vis}}^L \right) \tag{17}$$

$$\text{subject to} \sum_{i=1}^{N} q_{\text{gps}}^i > N_{\min}$$

$$\sum_{j=1}^{L} q_{\text{vis}}^j > L_{\min}$$

$$q_{\text{gps}}^i \in [0,1] \, \forall i \in \{1, \cdots, N\}$$

$$q_{\text{vis}}^j \in [0,1] \, \forall j \in \{1, \cdots, L\}$$

We utilized the MATLAB CVX toolbox (Grant & Boyd, 2014) to perform the above convex optimization and estimate the relaxed attention set, which is denoted by Q_k^*. Given that the solution was not binary, we performed a simple rounding procedure, such that the landmarks above a pre-defined threshold β were included while the others were excluded. This pre-defined threshold β was set during initialization, which is explained later in Sections 3.4 and 4.1. This final rounded attention set is denoted by Q_k°. Note that other sophisticated methods to perform the rounding procedure are found in existing literature (Carlone & Karaman, 2017; Lerner et al., 2007).

From the rounded attention set Q°, the landmarks to be included were identified (i.e., elements with binary value of 1), and were later used to estimate the over-approximated p-zonotope comprised of the set union of included landmarks, which is denoted by $\mathcal{L}(Q_k^\circ)$.

To compute the predicted system availability, we performed the following set operations: a) using the same γ-confidence considered earlier in Step 4, we computed a rounded attention zonotope, denoted by $Z(Q_k^\circ)$, that was evaluated using the included landmarks identified from the rounded attention set Q_k°; b) we transformed the rounded attention zonotope $Z(Q_k^\circ)$ to a rounded attention polytope $\mathcal{P}(Q_k^\circ)$, which was an exact conversion (Althoff, 2016); and c) by representing the AL by a general vertex representation of a polytope $\mathcal{P}_{AL}$, we computed set intersection of the rounded attention polytope $\mathcal{P}(Q_k^\circ)$ and AL polytope $\mathcal{P}_{AL}$ using Equation (4e).

We performed set intersection in polytope space for computational efficiency and also because polytopes can represent any arbitrary or non-regular sets, such as circles, in an exact manner. If the set intersection is the same as the rounded attention polytope (i.e., $\mathcal{P}(Q_k^\circ) \cap \mathcal{P}_{AL} = \mathcal{P}(Q_k^\circ)$), then the position error bounds associated with the rounded attention polytope are within the AL, and therefore the predicted system availability is 1. If this condition fails, then the predicted system availability is 0. An intuitive understanding for this set intersection is provided earlier in Figure 1.

The output from the proposed ILA technique (i.e., rounded attention set) and predicted system availability are given to the off-the-shelf estimator for GPS-vision localization. More details regarding the off-the-shelf GPS-vision estimator chosen for validating the proposed ILA technique are described later in Section 4.1.

3.4 | Performance Guarantees via Sub-Optimality Analysis

We derived the performance guarantees of the rounded attention set Q_k°, which was obtained by performing a rounding procedure on the relaxed attention set Q_k^* computed via convex relaxation in Equation (17). For our original minimization problem defined earlier in Equation (5), we represented the optimal solution of attention set as $Q_k^\dagger$. Similarly, for our modified minimization problem that performs optimization via convex relaxation, we represented the estimate of relaxed attention set as Q_k^* and rounded attention set as Q_k°. We derived the performance guarantees of the rounded attention set Q_k° obtained by executing a rounding procedure on the relaxed attention set Q_k^* computed via convex relaxation.

$$g(Q^*) \leq g(Q^\dagger) \leq g(Q^\circ) \tag{18}$$

where $\quad g(Q) = g\left(q_{\text{gps}}^1, \cdots, q_{\text{gps}}^N, q_{\text{vis}}^1, \cdots, q_{\text{vis}}^L\right) = \left(\dfrac{1}{\sum_{i \in N} q_{\text{gps}}^i + \sum_{j \in L} q_{\text{vis}}^j}\right) f\left(q_{\text{gps}}^1, \cdots, q_{\text{gps}}^N,\right.$

$\left. q_{\text{vis}}^1, \cdots, q_{\text{vis}}^L\right)$ based on Equations (5) and (17). Note that $g(Q)$ is always strictly greater than zero. The first inequality was based on the theory of convex relaxation (Boyd et al., 2004; Carlone & Karaman, 2017) that states that the solution of our

modified minimization problem in Equation (17) is at most as large as the optimal solution (i.e., $Q^\dagger$) of the original problem in Equation (5). The second inequality followed from the optimality of $Q^\dagger$ based on Equation (5), according to which no subset of landmarks can have a cost less than $g(Q^\dagger)$.

We rearranged the chain of inequality defined in Equation (18) as Equation (19) to provide an a-posteriori performance guarantee on the quality of the solution estimated by rounding the solution from convex relaxation.

$$g(Q^\circ) - g(Q^\dagger) \le g(Q^\circ) - g(Q^*) \tag{19}$$

where $g(Q^\circ) - g(Q^\dagger)$ denotes the sub-optimality gap (Chlamtac & Tulsiani, 2012; Joshi & Boyd, 2009) of the rounded attention subset Q°. The sub-optimality gap is upper-bound by the difference $g(Q^\circ) - g(Q^*)$ that can be calculated in an a-posteriori sense.

This upper-bound serves as a criterion for choosing the pre-defined threshold β that is used to compute the rounded attention set Q_k° in our proposed ILA technique. By performing a heuristic analysis during initialization, which is explained in Section 4.1, we set a pre-defined value of β that ensures the upper-bound on the sub-optimality gap is small. Furthermore, this upper-bound could also be utilized for adaptively estimating the pre-defined threshold or for assessing other sophisticated rounding procedures.

4 | EXPERIMENTAL RESULTS

We validated the performance of the proposed ILA technique that accounts for multiple GPS and vision faults in a unified manner to estimate the desirable subset of landmarks. Furthermore, the proposed ILA technique also is proven to predict the expected navigation performance of GPS-vision localization (i.e., the predicted system availability for a pre-defined risk and AL).

4.1 | Implementation and Initialization Details

We considered an urban sequence from the monocular visual odometry data set (Engel et al., 2016) that came with a high-fidelity reference ground truth and camera calibration files. Figure 7 shows the trajectory visualization for an experimental duration of 60 s, wherein the surroundings include a narrow alleyway, open sky, and many tall buildings. As explained in Section 3.1, we pre-processed the image frames via direct image alignment to extract key pixels in each frame.

To adapt this urban data set to our input measurements, we simulated GPS data using the ground truth 3D position and online-available satellite ephemeris. In particular, we used a C++ language-based software-defined GPS simulator known as GPS-SIM-SDR (Bhamidipati et al., 2019; Ebinuma, 2020) to generate the raw GPS signals. Later, we post-processed the simulated GPS signals using a MATLAB-based software-defined radio known as SoftGNSS (Paul, 2020).

When navigating through a narrow alleyway (i.e., during $t = 9 - 24$ s), we induced simulated GPS faults (i.e., multipath and satellite blockage) based on the elevation and azimuth of the satellites with respect to the direction of travel. We added simulated multipath effects in the low-elevation GPS satellites (i.e., $20° - 45°$) and within the azimuth range of $45° - 135°$ and $225° - 315°$. Similarly, for the same range of azimuth, we blocked the GPS satellites with elevation angles $< 20°$. Based on existing
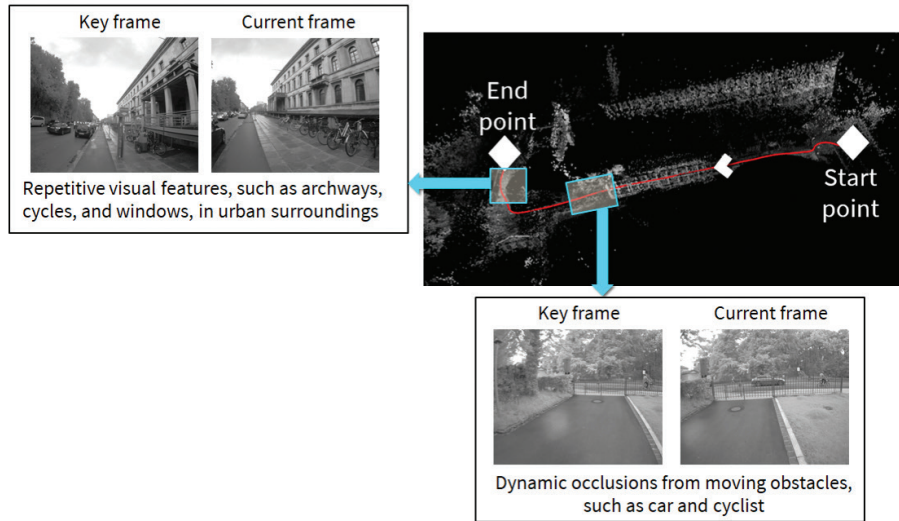
FIGURE 7 Trajectory visualization for an experiment duration of 60 s. This utilizes camera images of an urban sequence from the monocular visual odometry data set (Engel et al., 2016) and simulated GPS data. The surroundings included a narrow alleyway, open sky, and tall buildings, and therefore, the GPS and vision measurements were susceptible to faults at multiple time iterations. Two scenarios of vision faults are highlighted: one is due to dynamic occlusions from moving obstacles (i.e., with non-zero relative velocity such as a car or cyclist); the other is due to repetitive urban features (such as archways, cycles, and windows) in the surroundings.

literature (Phelts & Enge, 2000), we induced multipath errors that ranged between $20 - 65$ m.

We considered an off-the-shelf estimator from the existing literature (Bhamidipati & Gao, 2020a; Shepard & Humphreys, 2014) on GPS-vision localization in urban areas. The off-the-shelf GPS-vision estimator considers the same measurement models as that of the proposed ILA technique discussed earlier in Section 3.1. Utilizing the most recent history of input measurements from GPS and vision, we designed a cost function using the Huber M-estimator (Huber, 2004) that is comprised of a summation of the following residuals: GPS pseudoranges, pixel intensities, and system motion dynamics. Thereafter, we executed graph optimization using the Levenberg Marquardt method (Laurakis, 2005) to simultaneously localize both the system and key image pixels. From heuristic analysis, we chose the Huber constant for the M-estimator as 4.32. To maintain uniformity, both the off-the-shelf estimator and the proposed ILA technique shared the same initialization procedure.

As seen in Figure 8, the initialization procedure of the proposed ILA technique was performed in two stages: camera calibration and GPS-vision calibration. Referring to the existing works (Bhamidipati & Gao, 2020a; Engel et al., 2018), we followed a standard calibration procedure for the camera comprised of intrinsic calibration, removal of radial distortion, and photometric calibration. Using several images taken from different positions, we estimated the intrinsic parameters (i.e., f_x, f_y, c_x, and c_y) that mapped the 2D image frame to the 3D coordinates of the camera frame, and vice-versa, which are denoted by π and π^{-1} in Equation (7), respectively.

During the camera calibration stage of initialization, we also estimated the radial distortion coefficients associated with the field-of-view (FOV) distortion model (Li et al., 2015) that converts the 2D pixel coordinates of the image frame from a distorted space to an undistorted one. These undistorted 2D pixel coordinates are given to the vision measurement model described earlier in Equations (7) and (8). Furthermore, we performed a photometric calibration that
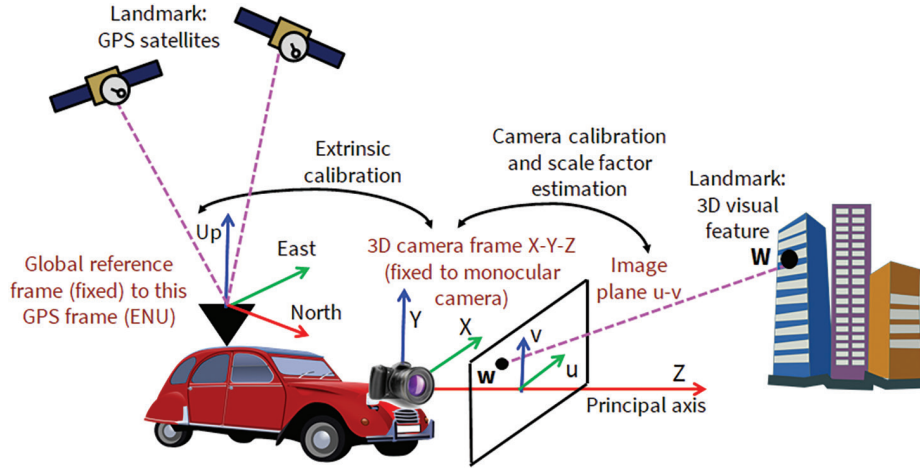
FIGURE 8 Initialization of the proposed ILA via camera calibration and GPS-vision calibration. For the camera, we performed intrinsic calibration, removal of radial distortion, and photometric calibration. The global reference frame was set to the ENU coordinate system with the center fixed to be the GPS receiver location at the start of initialization. In GPS-vision calibration, we performed extrinsic calibration and scale factor estimation.

mapped the real-world energy received by a pixel (irradiance) to the respective intensity value. This photometric camera calibration (Engel et al., 2018), which accounted for lens attenuation, gamma correction, and known exposure times, increased the accuracy and robustness of our vision measurement model that is based on pixel intensity analysis.

During the GPS-vision calibration step (the second stage of initialization), the system executes a random motion both in terms of rotation and translation for a couple of K_{calib} time iterations, say from $k = -K_{calib}$ to $k = -1$. In this step, we performed an extrinsic calibration and scale factor estimation. This step is conducted in fairly open-sky settings (i.e., no simulated multipath in GPS measurements) with the presence of feature-rich conditions for the camera, such as a fence, billboards, trees, etc. Referring to Bhamidipati and Gao (2020a) and Chen et al. (2018), we set the global reference frame to be the GPS frame at the start of the calibration step, which denoted the East-North-Up (ENU) coordinate system whose center is fixed to the receiver location.

While the GPS and camera frames were fixed to the GPS receiver and monocular camera, respectively, and therefore, were able to change based on the system dynamics, the relative translation and rotation between the GPS receiver and camera was still fixed and known. Using this information in extrinsic calibration, we estimated the frame transformation between the GPS frame and the camera frame.

To resolve the scale factor, we first independently estimated the system positions in the global reference frame of the executed random motion via two variants of the off-the-shelf GPS-vision estimator (Bhamidipati & Gao, 2020a; Engel et al., 2018): one using GPS-only measurements, and the other using vision-only measurements. For vision-only measurements, we estimated the system positions in the global reference frame by setting an initial scale factor of one and using the parameters computed from camera calibration, which is the first stage of initialization, and extrinsic calibration, which was explained earlier.

Referring to Chen et al. (2018), we updated the scale factor as the square root of the ratio between the weighted root-mean-square-error (RMSE) associated with the estimated system positions from GPS-only and vision-only. Note that the traditional monocular visual odometry techniques, which represent a relative

localization setup, can estimate system translation only up to a scale that is later scaled correctly via methods (Engel et al., 2014, 2018; Forster et al., 2014) such as place recognition and loop closure.

While scale factor estimation is a major challenge in traditional visual odometry that uses only monocular cameras, our sensor fusion framework that combines the relative measurement source (vision) with an absolute localization source (GPS) is more robust against ambiguities in the scale factor. After initialization, we periodically updated the vision measurement model of the proposed ILA technique with the scale factor estimated by the mapping module of the off-the-shelf GPS-vision estimator. Note that accounting for the errors in scale factor estimated via mapping modules is beyond the scope of this current work.

We performed integrity-driven convex optimization to compute a desirable subset of landmarks using the open-source MATLAB CVX toolbox (Grant & Boyd, 2014). We then performed the set operations of SR and transformed across various set representations (such as polytopes, zonotopes, p-zonotopes, etc.) using the open-source MATLAB CORA toolbox (Althoff, 2016). We considered a history of past measurements, such that $K = 8$, for the proposed ILA technique. If the predicted system availability at $(k-1)$-th time iteration is 1, then we update the motion model x_{k-1} seen in Equation (9) using the state vector estimated via the off-the-shelf GPS-vision estimator (Bhamidipati & Gao, 2020a), otherwise $x_{k-1} = z_{\mathrm{mm},k-1}$. We pre-defined the following parameters based on heuristic analysis as follows: $\beta = 0.75$, $N_{\min} = 5$, $L_{\min} = 600$, and $\gamma = 0.999$.

4.2 | Offline Empirical Analysis of Measurement Error Bounds in Non-Fault Conditions

As mentioned earlier, we pre-defined the measurement error bounds in non-faulty conditions using a naive offline empirical analysis of the openly available GPS and vision data sets. While we describe the analysis conducted for GPS and vision separately, notice that the underlying steps-to-execute remain the same irrespective of the sensor. This procedure is explained as follows:

1. **For GPS**, the measurement error bounds of simulated data in non-faulty conditions (i.e., no induced multipath effects) are inspired by the Monte Carlo runs performed using open-source data sets and their ground truths available from websites, such as Continuously Operating Reference Stations (CORS; Snay & Soler, 2008). In each Monte Carlo run, we estimated the measurement residuals by considering the GPS measurement model described in Equation (6) and the available ground truth 3D position. Figure 9 shows some example distributions of measurement errors in non-faulty conditions obtained from different Monte Carlo runs and are indicated in gray.

    While the distribution from these different Monte Carlo runs exhibit Gaussian characteristics, in addition to the changes in variance, we also observed non-zero values of the mean. The mean depends on the LOS vector between receiver and GPS satellite, and its non-zero values can be attributed to the accuracy limitations in modeling the receiver clock bias, ionospheric corrections, and tropospheric corrections. Based on this, we defined the simulated measurement errors observed during non-faulty conditions, which are denoted by η_{gps}^i in Equation (10a), to have time-varying Gaussian characteristics but with the following bounds: the mean and covariance lie between [−5 m, 5 m] and [0 m, 5 m], respectively. We efficiently enclosed
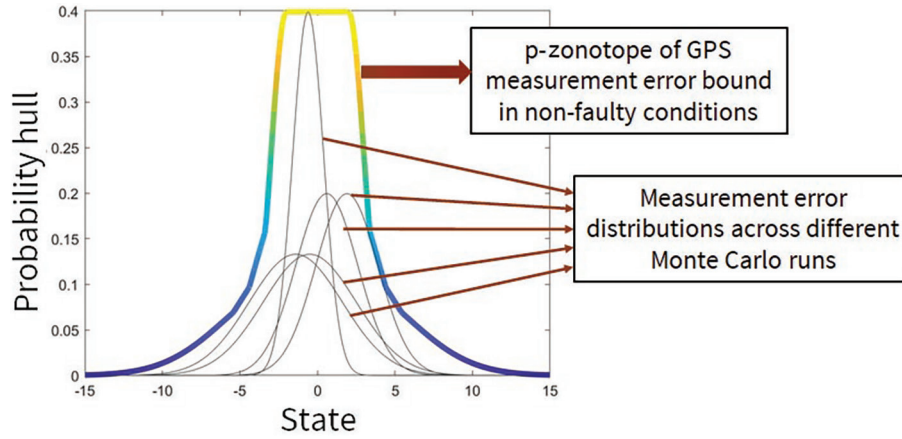
FIGURE 9 Illustration showing offline empirical analysis for estimating p-zonotopes that represents the GPS measurement error bound in non-faulty conditions and is indicated by the Parula colormap (Nunez et al., 2018). The p-zonotope encloses the time-varying Gaussian characteristics of all the measurement errors obtained during Monte Carlo runs, which are indicated in gray.

all the Gaussian distributions obtained from Monte Carlo runs using a p-zonotope, which is indicated by a Parula colormap (Nunez et al., 2018) in Figure 9. For any i-th satellite at k-th time iteration, this p-zonotope is represented by $\mathcal{L}^i_{\eta_{\mathrm{gps}},k}$. Note that we assign the same p-zonotope, which represents the GPS measurement error bound in non-faulty conditions, to all the satellite landmarks so that this bound is valid, irrespective of the LOS vector considered.

2. **For vision**, similar to GPS, we performed Monte Carlo runs that utilized selected sequences (i.e., with no urban artifacts that cause data-association errors) from the monocular visual odometry data set (Engel et al., 2016) along with the available ground truth regarding the 3D position and orientation of the system. In each Monte Carlo run, we first selected a keyframe, and thereafter, considered a set of relevant image frames that were in a temporal sequence with the keyframe and exhibit a minimum spatial overlap with the keyframe.

Considering the ground truth and vision measurement model described in Equations (7) and (8), we stored the intensity residuals of the overlap region between the keyframe and each image frame in the set. The intensity residuals associated with an example Monte Carlo run are plotted as a histogram shown in Figure 10(a).

In addition, Figure 10(b) illustrates example histograms associated with the measurement error analysis of different keyframes (i.e., during different Monte Carlo runs). Note that all the plotted histograms have the same limits on the x- and y-axis, which denote bins and frequency, respectively. We demonstrated that fitting the data with a Gaussian distribution, which is indicated in red, is not always suitable for modeling the vision measurement errors in non-faulty conditions.

However, irrespective of the underlying distribution, we observed that the measurement errors, which are denoted by ξ^j_{vis} in Equation (10b), are bounded in non-faulty conditions. Therefore, we utilized a p-zonotope, which is described earlier in Section 2 and Figure 3(a), to efficiently enclose all the histograms obtained from Monte Carlo runs.

For any j-th landmark at k-th time iteration, this p-zonotope is denoted by $\mathcal{L}^j_{\xi_{\text{vis}},k}$. This process of formulating the p-zonotope from histograms is similar to that executed for GPS measurement errors indicated earlier in Figure 9. Note that we assigned the same p-zonotope, which represents the vision measurement error bound in non-faulty conditions, to all the visual landmarks so that this bound was invariant to the surroundings.

These pre-defined measurement error bounds of GPS and vision in non-faulty conditions are later adaptively scaled in Step 3 of Section 3.2 to account for measurement faults. This work demonstrates a naive implementation of the offline empirical analysis, wherein we assigned the same measurement error bound in non-faulty conditions to all the landmarks associated with a particular sensor. However, note that the framework of proposed ILA technique is flexible to account for more sophisticated procedures that assign different p-zonotopes of measurement error bounds in non-faulty conditions to each landmark.

(a)

(b)
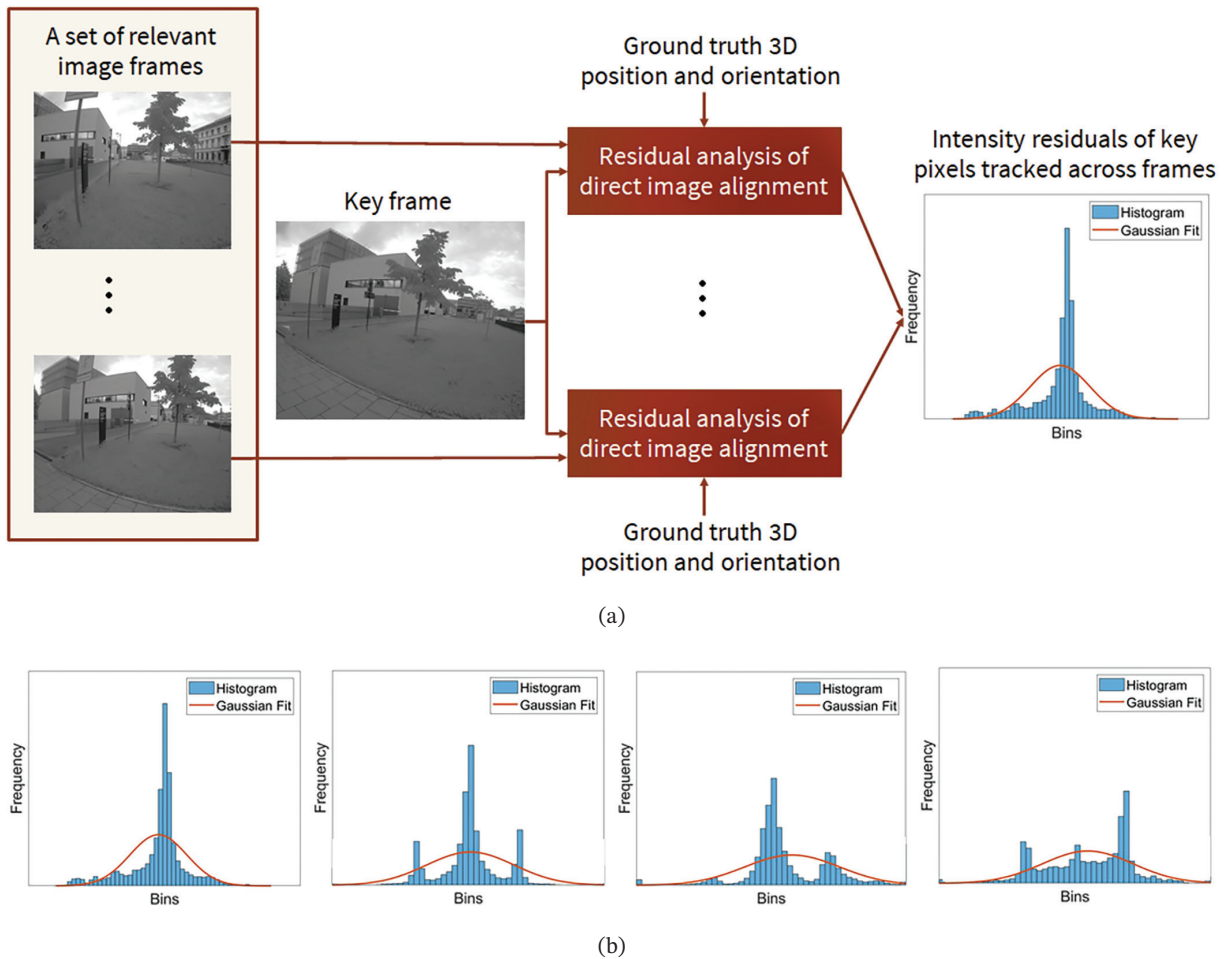
FIGURE 10 Illustration of offline empirical analysis for estimating the vision measurement error bounds in non-faulty conditions. (a) Procedure executed for vision during each Monte Carlo run where a keyframe and a set of relevant images were analyzed to compute the histogram of intensity residuals, which is indicated in blue. For reference, the histogram data is also fitted with a Gaussian distribution that is indicated in red; (b) Example histograms obtained from different Monte Carlo runs that depict the distribution of intensity residuals associated with each keyframe. All the plotted histograms have the same limits on the x- and y-axis, which denotes bins and frequency, respectively. P-zonotopes are apt for enclosing (i.e., upper-bound) these complex distributions of vision measurement errors that are bounded in nature but there the underlying distribution does not necessarily mimic a Gaussian distribution.

In addition to p-zonotopes being apt for bounding complex distributions in non-faulty conditions, these set representations can also efficiently enclose various sensor measurement error profiles, including the ones inducing constant offsets and the ones following Gaussian distributions with fixed mean and covariance. More details regarding these preliminaries were explained earlier in Equations (2) and (3) of Section 2.

4.3 | Results and Discussion

In Figure 11, we demonstrate an improved 2D position accuracy using the proposed ILA technique compared to other subsets of landmarks that included only GPS, all available GPS and vision, and random selection. In the random selection option, while fixing the number of selected landmarks to be the same as that of the proposed ILA technique, we ran 300 Monte Carlo runs at each time iteration, wherein for each run, we randomly selected from among all the available GPS and vision landmarks.

Thereafter, we retained the minimum localization error from among the 300 runs executed at each time iteration. Our reasoning for the random selection was to compare the performance of the proposed ILA technique against a computationally tractable alternative to a conventional FDE (Joerger et al., 2014) that would evaluate all multi-fault hypotheses in a brute-force manner. More details regarding the conventional FDE via multi-fault hypotheses evaluation and their computational intractability when applied to a large number of measurements was described earlier in Section 1. Note that in a heuristic sense, for a case where none of the GPS and vision measurements are faulty, the minimum error from random selection converges to a case of uniformly distributed landmarks.

As mentioned earlier, we processed each of these landmark subsets using an off-the-shelf GPS-vision estimator that formulated a Huber M-estimator-based cost function (Huber, 2004). In this context, the M-estimator (Shepard & Humphreys, 2014) represents a robust regression method that accounted for data containing outliers and extreme observations. Therefore, the off-the-shelf GPS-vision estimator
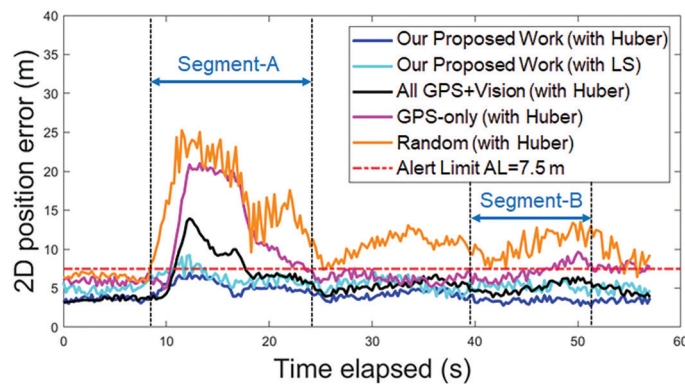
FIGURE 11 Maximum position error achieved using the off-the-shelf GPS-vision estimator (based on the Huber M-estimator) when the following landmark subsets are selected: the proposed ILA technique is denoted by blue; only GPS landmarks are indicated by magenta; all GPS and vision landmarks are represented by black; random selection is indicated in orange; and the proposed ILA technique with conventional LS-based off-the-shelf GPS-vision estimator is denoted by cyan. Our ILA (with both Huber and LS) shows a smaller localization error as compared to other methods.

TABLE 1

Quantitative comparison of maximum position error for different subsets of landmarks.

Landmarks Selected	Max Error
Proposed ILA (with Huber)	6.6 m
Proposed ILA (with LS)	9.3 m
Only GPS	20.2 m
All GPS and vision	13.9 m
Random (300 runs)	25.2 m

implemented another FDE approach (Bhamidipati & Gao, 2020a) by de-weighting the landmarks based on only the magnitude of their measurement residuals.

To compare the performance of the proposed ILA technique, we analyzed two variants of cost function used in the off-the-shelf GPS-vision estimator: one using the Huber M-estimator, where the selected landmarks are weighed based on their measurement residuals, and the other using a conventional least squares (LS), where all selected landmarks are given equal weight. These results are also included in Figure 11.

We identified two challenging segments in the experimental trajectory via post-processing that are labeled as *Segment A* and *Segment B* in Figure 11. Segment B corresponds to the region where vision landmarks satisfy the following heuristic condition (i.e., when compared against ground truth positions more than 50% of the available landmarks showcase large measurement residuals for at least 4 s time duration). Similarly, Segment A corresponds to the narrow alleyway region where both GPS and vision landmarks satisfy the above-stated condition.

As seen in Table 1, the proposed ILA technique demonstrates a maximum 2D position error of only 6.6 m and 9.3 m for the off-the-shelf GPS-vision estimator with the Huber and LS methods, respectively. Given that there are no standardized safety metrics as of September 2020, we set the pre-defined $AL = 7.5$ m based on the general street specifications.

We observed that while the proposed ILA technique (with the Huber estimator) lies below the AL for the entire experiment duration, the proposed ILA technique (with the LS estimator) exceeded AL at times. This behavior is explained more in-depth later in Figure 13. The proposed ILA technique (via both the Huber and LS-based off-the-shelf GPS-vision estimators) validated the improved GPS-vision localization in urban areas compared to using other landmark subsets (i.e., GPS-only, all GPS and vision, and random) with maximum position errors exceeding >13 m.

On an Intel i7 processor, the processing times for different comparison methods are as follows: (a) for the proposed ILA technique, optimization via convex relaxation, which was described earlier in Section 3.3, takes 2.2 s and off-the-shelf GPS-vision estimator with Huber takes 1.6 s; and (b) for reference, executing 300 Monte Carlo runs of random selection takes around 453.2 s. Through these statistics, we demonstrate the computational tractability of the proposed ILA technique compared to directly evaluating the multi-fault hypotheses to perform FDE.

At a time iteration $k = 54$ s, we compared the following statistics across the above-listed four techniques for landmark selection: position error and the number of selected landmarks. Unlike the other methods seen in Table 2, a low position error of 2.9 m (with 1.4 m along and 2.5 m perpendicular to the direction of motion) was achieved using our ILA technique (with the Huber estimator) that selected six

TABLE 2
Comparison of position error and number of landmarks at $k = 54$ s across different comparison methods with G : GPS landmarks, and V : vision landmarks.

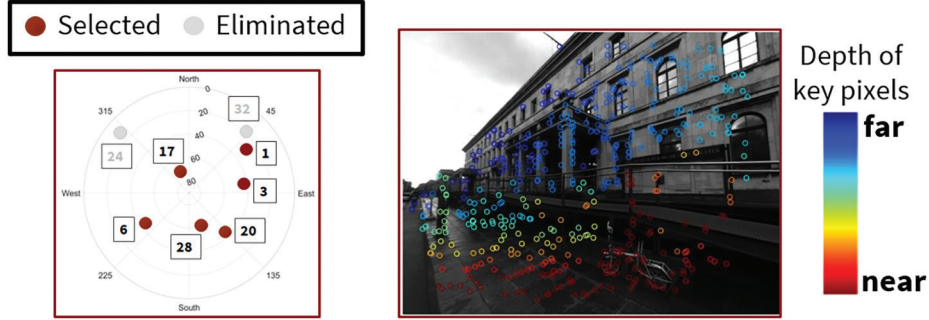| Methods	Position Error	Landmarks
Proposed ILA (with Huber)	2.9 m	6 (G), 956 (V)
Proposed ILA (with LS)	4.6 m	6 (G), 956 (V)
Only GPS	6.9 m	6 (G), 0 (V)
All GPS and vision	5.4 m	6 (G), 4722 (V)
Random (300 runs)	9.1 m	962 ($G + V$)

FIGURE 12 Selected GPS and vision landmarks via the proposed ILA technique at $k = 54$ s.

out of eight GPS landmarks and 956 out of 4,722 visual landmarks. Similarly, a position error of 4.6 m was achieved using the ILA technique (with the LS estimator).

An illustration of the selected landmarks in GPS and vision (rounded attention set) at time iteration $k = 54$ s is shown in Figure 12. Note that the proposed work does not emphasize direct detection and exclusion of all measurement faults, but instead analyzes a trade-off between the effect of fault magnitude and the impact of its exclusion on the position error bound.

In addition to estimating the rounded attention set that comprised the desired subset of landmarks, for a pre-defined risk $p_R = 1.9e^{-3}$ (which is equivalent to γ-confidence = 3.5 based on Equation [4f]), the proposed ILA technique also estimated the predicted 2D position error bound as 5.1 m with 2.7 m along and 4.3 m perpendicular to the direction of motion. Refer to Section 2 for more details regarding the property of γ-confidence and risk.

We observed two takeaways from this snapshot analysis at $k = 54$ s. First, the predicted position error bound (where position error bound is analogous to a measure of integrity called protection levels, i.e., 5.1 m), successfully bounded the estimated position error (i.e., 2.9 m with the Huber and 4.6 m with the LS estimator). Second, since the predicted position error bound is less than the pre-defined $AL = 7.5$, the system that leverages the proposed ILA technique for landmark selection is predicted to be available.

Extending our earlier analysis regarding the predicted position error bounds and predicted system availability at $k = 54$ s, we plotted the predicted system availability for the entire time duration of 60 s in Figure 13. The predicted availability of one indicates that the predicted position error bounds via the proposed ILA technique are less than an $AL = 7.5$ m, and zero indicates otherwise. More details regarding the predicted system availability and how it is computed during the optimization via convex relaxation was given earlier in Section 3.3.

Note that the 2D position error shown earlier in Figure 11 is calculated by comparing the position estimate from the off-the-shelf GPS-vision estimator with
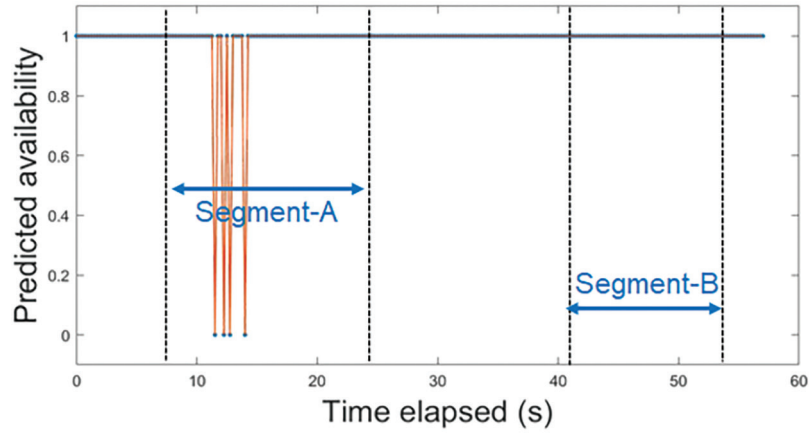
FIGURE 13 The predicted system availability using the proposed ILA technique demonstrates that the system is available for 98.2% of the entire time for a pre-defined risk $p_R = 1.9e^{-3}$ and $AL = 7.5$ m.

TABLE 3
Predicted system availability across the entire experiment duration of 60 s for different pre-defined values of p_R and AL.

p_R / AL	6 m	7.5 m	10 m
$1.9e^{-3}$	90.3%	98.2%	98.7%
$2.6e^{-4}$	88.5%	92.3%	93.5%

that of the ground truth position. In contrast, the predicted system availability is obtained as a by-product from the optimization by via convex relaxation step of the proposed ILA technique and does not depend on the off-the-shelf estimator utilized.

Given a pre-defined risk $p_R = 1.9e^{-3}$ (which is equivalent to γ-confidence $= 3.5$ based on Equation [4f]) and $AL = 7.5$ m, we demonstrated that the proposed ILA technique provides a robust measure of predicted availability with the system being available for 98.2% of the entire experiment. In a narrow alleyway, the proposed ILA technique successfully detected degraded localization accuracy (i.e., predicted availability = 0) due to large faults in both GPS and vision, and is therefore reflective of the measurement quality in urban surroundings. In an intuitive sense, the predicted system availability serves as a performance metric of the proposed ILA technique, wherein a value of one indicates that the landmarks selected provide a position error less than $AL = 7.5$ m, irrespective of the off-the-shelf GPS-vision estimator utilized.

However, when the predicted system availability is zero, it implies that the associated predicted position error bound estimated via the proposed ILA technique is greater than $AL = 7.5$ m, and therefore, by utilizing the associated landmarks, the position error estimated via the off-the-shelf GPS-vision estimator cannot be guaranteed to lie within an $AL = 7.5$ m. This was reflected earlier in Figure 11 by analyzing the position error of the proposed ILA technique with Huber compared to that with LS. At other time iterations, the proposed ILA technique performed a robust landmark selection (i.e., among GPS and vision) to ensure compliance with the pre-defined AL.

In Table 3, we showcase the predicted system availability for the entire time across different pre-defined values of risk and AL. Intuitively, we observed that as the

pre-defined AL increases for a fixed risk p_R, the predicted system availability also increases, which is because a larger predicted position error bound is acceptable for achieving the expected navigation performance. Similarly, as the pre-defined p_R increases (which increases the associated value of γ-confidence defined earlier in Equation [4f]) for a fixed AL, the system is available for a longer time duration as smaller estimates of the predicted position error bound is obtained.

5 | CONCLUSION

In summary, we developed an integrity-driven landmark attention (ILA) technique for GPS-vision navigation that was inspired by cognitive attention in humans. We performed a two-tiered approach: cost formulation via stochastic reachability (SR) and optimization via convex relaxation to select the subset of desired landmark measurements from the available, namely GPS satellites and 3D visual features.

Given the known measurement error bounds for each landmark in non-faulty conditions, we independently analyzed the received measurements to estimate the stochastic reachable set of expected positions. By parameterizing the stochastic set using probabilistic zonotope (p-zonotope), we applied the set union property to compute the position error bound for a pre-defined risk as a function of included landmarks. Our ILA technique works with any off-the-shelf GPS-vision estimator and follows a unified approach to account for multiple faults in GPS and vision measurements.

We validated the proposed ILA technique using an urban data set with real-world camera images and simulated GPS data. We demonstrated the improved localization accuracy of the proposed ILA technique with maximum position errors of only 6.6 m compared to other techniques for landmark selection that showed maximum position errors of > 13 m, thereby violating the pre-defined alert limit $AL = 7.5$ m. We also showcased that our ILA technique provides a robust measure of predicted availability with the system being available for 98.24% of the entire duration. We also provided a quantitative comparison of the predicted system availability across the entire experiment for different pre-defined values of risk and AL.

REFERENCES

Al Hage, J., Xu, P., & Bonnifait, P. (2018). Bounding localization errors with student distribution for road vehicles. *International Technical Symposium on Navigation and Timing*, Toulouse, France. https://www.doi.org/10.31701/itsnt2018.11

Al Hage, J., Xu, P., Bonnifait, P., & Ibanez-Guzman, J. (2020). Localization integrity for intelligent vehicles through fault detection and position error characterization. *IEEE Transactions on Intelligent Transportation Systems*, 1–13. https://www.doi.org/10.1109/tits.2020.3027433

Althoff, M. (2016). *CORA 2016 manual*. Technische Universität München.

Althoff, M., & Dolan, J. M. (2014). Online verification of automated road vehicles using reachability analysis. *IEEE Transactions on Robotics*, 30(4), 903–918. https://www.doi.org/10.1109/tro.2014.2312453

Althoff, M., Stursberg, O., & Buss, M. (2009). Safety assessment for stochastic linear systems using enclosing hulls of probability density functions. *2009 European Control Conference (ECC)*, Budapest, Hungary. https://www.doi.org/10.23919/ecc.2009.7074473

Angus, J. E. (2006). RAIM with multiple faults. *NAVIGATION,* 53(4), 249–257. https://www.doi.org/10.1002/j.2161-4296.2006.tb00387.x

August, E., Lu, J., & Koeppl, H. (2012). Trajectory enclosures for nonlinear systems with uncertain initial conditions and parameters. *2012 American Control Conference (ACC)*, Montreal, QC. https://www.doi.org/10.1109/acc.2012.6314741

Bhamidipati, S., & Gao, G. X. (2018). Multiple GPS fault detection and isolation using a Graph-SLAM framework. *Proc. of the 31st International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2018)*, Miami, FL. https://www.doi.org/10.33012/2018.16030

Bhamidipati, S. & Gao, G. X. (2019). Distributed cooperative SLAM-based integrity monitoring via a network of receivers. *Proc. of the 32nd International Technical Meeting of the Satellite*

Division of the Institute of Navigation (ION GNSS+ 2019), Miami, FL, 2023–2034. https://www.doi.org/10.33012/2019.16882

Bhamidipati, S., & Gao, G. X. (2020a). Integrity monitoring of Graph-SLAM using GPS and fish-eye camera. *NAVIGATION, 67*(3), 583–600. https://www.doi.org/10.1002/navi.381

Bhamidipati, S., & Gao, G. X. (2020b). Integrity-driven landmark attention for GPS-vision navigation via stochastic reachability. *Proc. of the 33rd International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2020)*, 2311–2326. https://www.doi.org/10.33012/2020.17546

Bhamidipati, S., Kim, K. J., Sun, H., & Orlik, P. V. (2019). GPS spoofing detection and mitigation in PMUs using distributed multiple directional antennas. *2019 IEEE International Conference on Communications (ICC)*, Shanghai, China. https://www.doi.org/10.1109/icc.2019.8761208

Binjammaz, T., Al-Bayatti, A., & Al-Hargan, A. (2013) GPS integrity monitoring for an intelligent transport system. *2013 10th Workshop on Positioning, Navigation and Communication (WPNC)*, Dresden, Germany. https://www.doi.org/10.1109/wpnc.2013.6533268

Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press. https://www.doi.org/10.1017/cbo9780511804441.005

Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., & Leonard, J. J. (2016). Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Transactions on Robotics*, 32(6), 1309–1332. https://www.doi.org/10.1109/tro.2016.2624754

Caduff, D., & Timpf, S. (2008). On the assessment of landmark salience for human navigation. *Cognitive Processing*, 9(4), 249–267. https://www.doi.org/10.1007/s10339-007-0199-2

Carlone, L., & Karaman, S. (2017). Attention and anticipation in fast visual-inertial navigation. *2017 IEEE International Conference on Robotics and Automation (ICRA),* Singapore. https://www.doi.org/10.1109/icra.2017.7989448

Chen, X., Hu, W., Zhang, L., Shi, Z., & Li, M. (2018). Integration of low-cost GNSS and monocular cameras for simultaneous localization and mapping. *Sensors*, 18(7). https://doi.org/10.3390/s18072193

Chlamtac, E., & Tulsiani, M. (2012) Convex relaxations and integrality gaps. In M. F. Anjos & J. B. Lasserre (Eds.), *Handbook on Semidefinite, Conic and Polynomial Optimization* (pp. 139–169). Springer. https://www.doi. org/10.1007/978-1-4614-0769-0_6

Ebinuma, T. (2020). *GPS-SDR-SIM*. Github. https://github.com/osqzss/gps-sdr-sim

El-Mowafy, A., & Kubo, N. (2017). Integrity monitoring of vehicle positioning in urban environment using RTK-GNSS, IMU and speedometer. *Measurement Science and Technology*, 28(5). https://www.doi.org/10.1088/1361-6501/aa5c66

Engel, J., Koltun, V., & Cremers, D. (2018). Direct sparse odometry. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(3), 611–625. https://www.doi.org/10.1109/tpami.2017.2658577

Engel, J., Schöps, T., & Cremers, D. (2014). LSD-SLAM: Large-scale direct monocular SLAM. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014* (pp. 834–849). Springer. https://www.doi.org/10.1007/978-3-319-10605-2_54

Engel, J., Usenko, V., & Cremers, D. (2016). A photometrically calibrated benchmark for monocular visual odometry.

Forster, C., Pizzoli, M., & Scaramuzza, D. (2014). SVO: Fast semi-direct monocular visual odometry. *2014 IEEE international Conference on Robotics and Automation (ICRA)*, Hong Kong, China. https://www.doi.org/10.1109/icra.2014.6906584

Ghinamo, G., Boiero, G., Lovisolo, P., Dalla Torre, A., & Detoma, E. (2010). Hybrid fault detection technique in assisted GNSS. *IEEE/ION Position, Location and Navigation Symposium,* Indian Wells, CA. https://www.doi.org/10.1109/plans.2010.5507262

Gil, A., Reinoso, O., Mozos, O. M., Stachniss, C., & Burgard, W. (2006). Improving data association in vision-based SLAM. *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Beijing, China. https://www.doi.org/10.1109/iros.2006.282483

Grant, M., & Boyd, S. (2014). *CVX: Matlab software for disciplined convex programming* (version 2.1). http://cvxr.com/cvx

Gupta, S. & Gao, G. X. (2019). Particle RAIM for Integrity Monitoring. *Proc. of the 32nd International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2019)*, Miami, FL, 811–826. https://www.doi.org/10.33012/2019.16939

Heredia, G., Caballero, F., Maza, I., Merino, L., Viguria, A., & Ollero, A. (2009). Multi-unmanned aerial vehicle (UAV) cooperative fault detection employing differential global positioning (DGPS), inertial and vision sensors. *Sensors*, 9(9), 7566–7579. https://www.doi.org/10.3390/s90907566

Hsu, L.-T., Tokura, H., Kubo, N., Gu, Y., & Kamijo, S. (2017). Multiple faulty GNSS measurement exclusion based on consistency check in urban canyons. *IEEE Sensors Journal*, 17(6), 1909–1917. https://www.doi.org/10.1109/jsen.2017.2654359

Huber, P. J. (2004). *Robust statistics*. John Wiley & Sons. https://www.doi.org/10.1002/0471725250

Joerger, M., Arana, G. D., Spenko, M., & Pervan, B. (2017). Landmark data selection and unmapped obstacle detection in lidar-based navigation. *Proc. of the 30th International Technical Meeting of*

the Satellite Division of the Institute of Navigation (ION GNSS+ 2017), Portland, OR, 1886–1903. https://www.doi.org/10.33012/2017.15406

Joerger, M., Chan, F.-C., & Pervan, B. (2014). Solution separation versus residual-based RAIM. *NAVIGATION*, 61(4), 273–291. https://www.doi.org/10.1002/navi.71

Joerger, M., & Spenko, M. (2017). Towards navigation safety for autonomous cars. *Inside GNSS*. https://insidegnss.com/towards-navigation-safety-for-autonomous-cars

Joshi, S., & Boyd, S. (2009). Sensor selection via convex optimization. *IEEE Transactions on Signal Processing*, 57(2), 451–462. https://www.doi.org/10.1109/tsp.2008.2007095

Kaddour, M., El Najjar, M. E., Naja, Z., Tmazirte, N. A., & Moubayed, N. (2015). Fault detection and exclusion for GNSS measurements using observations projection on information space. *2015 Fifth International Conference on Digital Information and Communication Technology and its Applications (DICTAP)*, Beirut, Lebanon. https://www.doi.org/10.1109/DICTAP.2015.7113199

Larson, J. D., Gebre-Egziabher, D., & Rife, J. H. (2019). Gaussian-Pareto overbounding of DGNSS pseudoranges from CORS. *NAVIGATION*, 66(1), 139–150. https://www.doi.org/10.1002/navi.276

Le Guernic, C., & Girard, A. (2010). Reachability analysis of linear systems using support functions. *Nonlinear Analysis: Hybrid Systems*, 4(2), 250–262. https://doi.org/10.1016/j.nahs.2009.03.002

Lerner, R., Rivlin, E., & Shimshoni, I. (2007). Landmark selection for task-oriented navigation. *IEEE Transactions on Robotics*, 23(3), 494–505. https://www.doi.org/10.1109/tro.2007.895070

Li, F., Bonnifait, P., Ibanez-Guzman, J., & Zinoune, C. (2017). Lane-level map-matching with integrity on high-definition maps. *2017 IEEE Intelligent Vehicles Symposium (IV)*, Los Angeles, CA. https://www.doi.org/10.1109/ivs.2017.7995872

Li, A., Wu, Y., Xia, X., Huang, Y., Feng, C., & Zheng, Z. (2015). Computational method for correcting complex optical distortion based on FOV division. *Applied Optics*, 54(9), 2441–2449. https://www.doi.org/10.1364/ao.54.002441

Lourakis, M. I. A. (2005). A brief description of the Levenberg-Marquardt algorithm implemented by levmar. *Foundation of Research and Technology*, 4(1), 1–6. https://www.researchgate.net/publication/239328019_A_Brief_Description_of_the_Levenberg-Marquardt_Algorithm_Implemened_by_levmar

Maaref, M., & Kassas, Z. M. (2019). UAV integrity monitoring measure improvement using terrestrial signals of opportunity. *Proc. of the 32nd International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2019)*, Miami, FL, 3045–3056. https://www.doi.org/10.33012/2019.17009

Maaref, M., Khalife, J., & Kassas, Z. M. (2020). Enhanced safety of autonomous driving by incorporating terrestrial signals of opportunity. *2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain. https://www.doi.org/10.1109/icassp40776.2020.9053017

Misra, P. & Enge, P. (2006) *Global positioning system: Signals, measurements, and performance* (2nd ed.). Ganga-Jamuna Press.

Mohanty, A., Gupta, S., & Gao, G. X. (2020). A particle filtering framework for integrity risk of GNSS-camera sensor fusion. *Proc. of the 33rd International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2020)*, 2440–2452. https://www.doi.org/10.33012/2020.17660

Mousavi, H. K. & Motee, N. (2020). Estimation with fast feature selection in robot visual navigation. *IEEE Robotics and Automation Letters*, 5(2), 3572–3579. https://www.doi.org/10.1109/lra.2020.2974654

Nuñez, J. R., Anderton, C. R., & Renslow, R. S. (2018). Optimizing colormaps with consideration for color vision deficiency to enable accurate interpretation of scientific data. *PloS ONE*, 13(7). https://www.doi.org/10.1371/journal.pone.0199239

Ochieng, W. Y., Sauer, K., Walsh, D., Brodin, G., Griffin, S., & Denney, M. (2003). GPS integrity and potential impact on aviation safety. *The Journal of Navigation*, 56(1), 51–65. https://www.doi.org/10.1017/s0373463302002096

Oldham, K. B., Myland, J. C., & Spanier, J. (2009). The error function erf (x) and its complement erfc (x). In *An Atlas of Functions* (pp. 405–415). Springer. https://doi.org/10.1007/978-0-387-48807-3_41

Paul, K. (2020). *Soft GNSS*. Github. https://github.com/kristianpaul/SoftGNSS

Phelts, R. E., & Enge, P. (2000). Multipath mitigation for narrowband receivers. *IEEE 2000 Position, Location and Navigation Symposium,* San Diego, CA. https://www.doi.org/10.1109/plans.2000.838280

Rabaoui, A., Viandier, N., Duflos, E., Marais, J., & Vanheeghe, P. (2011). Dirichlet process mixtures for density estimation in dynamic nonlinear modeling: Application to GPS positioning in urban canyons. *IEEE Transactions on Signal Processing*, 60(4), 1638–1655. https://www.doi.org/10.1109/tsp.2011.2180901

Schroth, G., Ene, A., Blanch, J., Walter, T., & Enge, P. (2008). Failure detection and exclusion via range consensus.

Shepard, D. P., & Humphreys, T. E. (2014). *2014 IEEE/ION Position, Location and Navigation Symposium (PLANS 2014)*, Monterey, CA. https://www.doi.org/10.1109/plans.2014.6851506

Shytermeja, E., Garcia-Pena, A., & Julien, O. (2014). Proposed architecture for integrity monitoring of a GNSS/MEMS system with a fisheye camera in urban environment. *International Conference on Localization and GNSS 2014 (ICL-GNSS 2014)*, Helsinki, Finland. https://www.doi.org/10.1109/icl-gnss.2014.6934179

Snay, R. A., & Soler, T. (2008). Continuously operating reference station (CORS): History, applications, and future enhancements. *Journal of Surveying Engineering*, 134(4), 95–104. https://www.doi.org/10.1061/9780784411643.ch01

Tmazirte, N. A., El Najjar, M. E., Smaili, C., & Pomorski, D. (2012). Multi-sensor data fusion based on information theory. Application to GNSS positioning and integrity monitoring. *2012 15th International Conference on Information Fusion*, Singapore.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, 113(4), 766–786. https://www.doi.org/10.1037/0033-295x.113.4.766

Tossaint, M., Samson, J., Toran, F., Ventura-Traveset, J., Hernandez-Pajares, M., Juan, J. M., Sanz, J., & Ramos-Bosch, P. (2007). The Stanford-ESA integrity diagram: A new tool for the user domain SBAS integrity assessment. *NAVIGATION,* 54(2), 153–162.

Woeginger, G. J. (2003). Exact algorithms for NP-hard problems: A survey. In M. Jünger, G. Reinelt, & G. Rinaldi (Eds.), *Combinatorial optimization—Eureka, you shrink!* (pp. 185–207). Springer.

Won, D. H., Lee, E., Heo, M., Lee, S.-W., Lee, J., Kim, J., Sung, S., & Lee, Y. J. (2014). Selective integration of GNSS, vision sensor, and INS using weighted DOP under GNSS-challenged environments. *IEEE Transactions on Instrumentation and Measurement*, 63(9), 2288–2298. https://www.doi.org/10.1109/tim.2014.2304365

Zhang, L., Li, J., Cui, T., and Liu, S. (2017). An adapted RAIM algorithm for urban canyon environment. *2017 Forum on Cooperative Positioning and Service (CPGPS),* Harbin, China. https://www.doi.org/10.1109/cpgps.2017.8075108

Zhang, S., Lo, S., Chen, Y.-H., Walter, T., & Enge, P. (2018). GNSS multipath detection in urban environment using 3D building model. *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, Monterey, CA. https://www.doi.org/10.1109/plans.2018.8373486

Zhang, Q., & Gui, Q. (2015). A new Bayesian RAIM for multiple faults detection and exclusion in GNSS. *The Journal of Navigation*, 68(3), 465–479. https://www.doi.org/10.1017/s0373463314000721

Zhu, N., Marais, J., Bétaille, D., & Berbineau, M. (2018). GNSS position integrity in urban environments: A review of literature. *IEEE Transactions on Intelligent Transportation Systems*, 19(9), 2762–2778. https://www.doi.org/10.1109/tits.2017.2766768

Ziebold, R., Lanca, L., & Romanovas, M. (2017). On fault detection and exclusion in snapshot and recursive positioning algorithms for maritime applications. *European Transport Research Review*, 9(1). https://www.doi.org/10.1007/s12544-016-0217-5