Landmark Data Selection and Unmapped Obstacle Detection in Lidar-Based Navigation

Mathieu Joerger, *The University of Arizona*
Guillermo Duenas Arana, Matthew Spenko, Boris Pervan, *Illinois Institute of Technology*

BIOGRAPHIES

Dr. Mathieu Joerger obtained a Master in Mechatronics from the National Institute of Applied Sciences in Strasbourg, France, in 2002, and a M.S. and a Ph.D. in Mechanical and Aerospace Engineering from the Illinois Institute of Technology (IIT), in 2002 and 2009 respectively. He is the 2009 recipient of the Institute of Navigation (ION) Bradford Parkinson award, and the 2014 recipient of the ION Early Achievement Award. He is currently a research assistant professor at IIT, working on multi-sensor integration, on sequential fault-detection for multi-constellation navigation systems, and on safety of sense and avoid for unmanned aircraft systems.

Guillermo Duenas Arana obtained his "Licenciatura" (M.S. degree) in Mechanical Engineering from the Escuela Técnica Superior de Ingeniería de Bilbao, Spain, and his M.Eng. in Mechanical and Aerospace Engineering from the Illinois Institute of Technology (IIT). He is currently a Ph.D. student in Mechanical and Aerospace Engineering at IIT working on methods to evaluate navigation safety for mobile robots.

Dr. Matthew Spenko is an associate professor in the Mechanical, Materials, and Aerospace Engineering Department at the Illinois Institute of Technology. He earned the B.S. degree in Mechanical Engineering from Northwestern University in 1999 and the M.S. and Ph.D. degrees in Mechanical Engineering from Massachusetts Institute of Technology in 2001 and 2005, respectively. He was an Intelligence Community Postdoctoral Scholar in the Center for Design Research at Stanford University from 2005 to 2007. He has been a faculty member at the Illinois Institute of Technology since 2007. His research is in robotics with specific attention to mobility in challenging environments.

Dr. Boris Pervan is a Professor of Mechanical and Aerospace Engineering at IIT, where he conducts research on advanced navigation systems. Prior to joining IIT, he was a spacecraft mission analyst at Hughes Aircraft Company and a postdoctoral research associate at Stanford University. Prof. Pervan received his B.S. from the University of Notre Dame, M.S. from the California Institute of Technology, and Ph.D. from Stanford University. He is an Associate Fellow of the AIAA, a Fellow of the ION, and Editor-in-Chief of the ION journal NAVIGATION. He was the recipient of the IIT Sigma Xi Excellence in University Research Award (2011, 2002), Ralph Barnett Mechanical and Aerospace Dept. Outstanding Teaching Award (2009, 2002), Mechanical and Aerospace Dept. Excellence in Research Award (2007), University Excellence in Teaching Award (2005), IEEE Aerospace and Electronic Systems Society M. Barry Carlton Award (1999), RTCA William E. Jackson Award (1996), Guggenheim Fellowship (Caltech 1987), and Albert J. Zahm Prize in Aeronautics (Notre Dame 1986).

ABSTRACT

This research establishes new methods to quantify lidar-based navigation safety in highly automated vehicle (HAV) applications. Lidar navigation requires feature extraction (FE) and data association (DA). In prior work, an FE and DA risk prediction process was developed assuming that the set of extracted features matched the set of mapped landmarks. This paper addresses these limiting assumptions by first providing the means to select a subset of feature measurements (to be used in the estimator) while accounting for all existing landmarks in the surroundings. This is achieved by employing a probabilistic lower-bound on the mean innovation vector's norm. This measure of landmark separation is used in an analytical integrity risk bound that accounts for all possible association hypotheses. Then, a solution separation algorithm is employed to detect unmapped obstacles and wrong extractions. The integrity risk bound is modified to incorporate the risk of not detecting an unwanted obstacle (UO) when one might be present. Covariance analysis, direct simulation, and preliminary testing show that selecting fewer extracted features can significantly reduce integrity risk, but can also decrease landmark redundancy, thereby reducing UO detection capability.

INTRODUCTION

This paper describes the design, analysis, and preliminary testing of a new method to maximize safety by selecting landmark data in lidar-based navigation systems. In addition, an integrity risk bound is derived, which accounts for failures to detect undesirable, unmapped and wrongly-extracted obstacles. This work is intended for highly automated vehicles (HAV) operating in changing environments where unknown, moving obstacles (cars, buses, and trucks) are not wanted as landmarks for localization, and may also occlude other useful, mapped landmarks.

This paper leverages prior analytical work carried out in civilian aviation navigation where safety is assessed in terms of integrity and continuity [1]. These top-level quantifiable performance metrics are sensor- and platform-independent. Integrity is a measure of trust in sensor information: integrity risk is the probability of undetected sensor errors causing unacceptably large positioning uncertainty [1]. Continuity is a measure of the navigation system's ability to operate without unscheduled interruption. Both loss of integrity and loss of continuity can place the HAV in hazardous situations [1].

Several methods have been established to predict integrity and continuity risks in GNSS-based aviation applications [2-4]. Unfortunately, the same methods do not directly apply to HAVs, because ground vehicles operate under sky-obstructed areas where GNSS signals can be altered or blocked by buildings and trees.

HAVs require sensors in addition to GNSS, including lidars, cameras, or radars. This paper focuses on lidars because of their prevalence in HAVs, of their market availability, and of our prior experience. A raw lidar scan is made of thousands of data points, each of which individually does not carry any useful information. Raw measurements must be pre-processed before they can be used for navigation. One widely-implemented approach is to achieve sensor localization by tracking recognizable, static features in the perceived environment [5-19].

Features can include, for example, lines or planes corresponding to building walls in two- or three-dimensional scans, respectively. Previous knowledge of feature parameters can be provided either from a landmark map, or from past-time estimation in Simultaneous Localization and Mapping (SLAM) [20-22]. The resulting information can then be iteratively processed using sequential estimators in SLAM (e.g., Extended Kalman filter or EKF), which is convenient in practical implementations. To estimate the HAV's pose starting from a raw lidar scan, two intermediary, pre-estimator procedures must be carried out: feature extraction (FE), and data association (DA).

First, FE aims at finding the few most consistently recognizable, viewpoint-invariant, and mutually distinguishable landmarks in the raw sensor data. Second, DA aims at assigning the extracted features to the corresponding feature parameters assumed in the estimation process, i.e., at finding the ordering of mapped landmarks that matches the ordering of extracted features over successive observations. Incorrect association is a well-known problem that can lead to large navigation errors [23], thereby representing a threat to navigation integrity. FE and DA can be challenging in the presence of sensor uncertainty. This is why many sophisticated algorithms have been devised [5-19]. But, how can we prove whether FE and DA are safe for life-critical HAV navigation applications?

This research question is mostly unexplored. Several publications on multi-target tracking describe relevant approaches to evaluate the probability of correct association in the presence of measurement uncertainty [23, 24]. However, these algorithms are not well suited for safety-critical HAV applications due to their lack of prediction capability, to approximations that do not necessarily upper-bound risks, and to high computational loads. Also, the risk of FE is not addressed. Overall, research on integrity and continuity of FE and DA is sparse.

This paper builds upon prior work in [25-27], where we developed an analytical integrity risk prediction method using a multiple-hypothesis innovation-based DA process. We established a compact expression for the integrity risk of lidar-based pose estimation over successive iterations. However, references [25-27] made simplifying assumptions that limit the applicability of these prior results. For example, we assumed that the set of landmarks in the a-priori map was exactly the same as the one being extracted. In practice, unexpected obstacles such as other vehicles passing by are likely to be extracted, and may even occlude other mapped landmarks. The safety risk of using such unknown obstacles as landmarks for navigation has yet to be quantified.

In response, in this paper, we derive new methods to select the set of extracted features that minimize safety risk, and to quantify the integrity risk caused by failures to detect unwanted obstacles (UO), while guaranteeing a predefined continuity risk requirement.

The next section of the paper provides an overview of the methods developed in [25-27], and of their limitations. We quantify risks involved with using a nearest-neighbor DA criterion [5], defined by the minimum normalized norm of the innovation vectors over all possible landmark permutations. In the third section, an integrity-risk-minimizing data-selection procedure is presented, which considers the following key tradeoff: on the one hand, increasing the number of extracted landmarks reduces the pose estimation error assuming correct association, but on the other hand, it also increases the risk of incorrect association. Thus, if the integrity risk using the full set of landmarks is unacceptably large, the selection process systematically screens all landmark subsets to find the one that meets the integrity risk requirement. The fourth section deals with the situation where a mapped landmark is not extracted, but another unknown obstacle is extracted instead (e.g., case of an obstacle masking a mapped landmark, or failure to extract the intended subset of selected landmark data). A solution separation algorithm is employed to detect UO. The integrity risk bound is modified to incorporate the risk of not detecting a UO when one might be present. Navigation integrity performance using these new methods is assessed in the fifth section of the paper. Covariance analysis, direct simulation, and preliminary testing show that selecting fewer extracted features significantly reduces integrity risk, but also decreases landmark redundancy, thereby reducing UO detection capability.

BACKGROUND: INTEGRITY RISK BOUND ACCOUNTING FOR INCORRECT ASSOCIATIONS

This section presents an overview of the integrity risk evaluation method described in [25, 27], which uses a multiple-hypothesis innovation-based DA process.

Integrity Risk Definition and Integrity Risk Bound

The integrity risk, or probability of hazardous misleading information (HMI), is defined as illustrated in Figure 1. In prior work [25, 27], we established an analytical bound on the integrity risk, which accounts for the risk of incorrect association. This bound is expressed as:

$$P(HMI_k) \le 1 - \left[1 - P(HMI_k \mid CA_K)\right] P(CA_k) \tag{1}$$

with

$$P(HMI_k \mid CA_K) = 2Q\left\{\ell / \sigma_k\right\} \tag{2}$$

$$P(CA_k) \ge \prod_{l=1}^{k}\left[1 - P\left(q_l^2 \ge \min_{i=1,\ldots n_A}\left\{ y_{i,l}^2 / 4\right\}\right)\right] \tag{3}$$

where

k : is an index identifying a time step

K : designates a range of indices: $K \equiv \{0,\ldots k\}$, from filter initiation to time k.

CA_K : is the correct association hypothesis for all landmarks at all times 0, ..., k

$Q\{\ \}$: is the tail probability function of the standard normal distribution

ℓ : is the specified alert limit that defines a hazardous situation [1, 4]

σ_k : is the standard deviation of the estimation error for the vehicle state (or linear combination of states) of interest

q_l^2 : is a chi-square distributed random variable with a number of degrees of freedom that is the sum of the number of measurements and of states at time step l

$y_{i,l}^2$: is the norm of the mean innovation vector weighted by the inverse of the innovation covariance matrix for candidate association i at time step l.

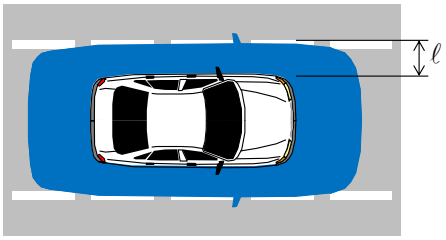

Figure 1. Defining Integrity Risk for Automotive Applications. The integrity risk is the probability of the car being outside the alert limit requirement box (blue shaded area) when it was estimated to be inside the box. If the focus is on the lateral deviation, then the alert limit is the distance ℓ between the edge of the car and the edge of the lane.

The integrity risk bound in equation (1) will be refined in the next sections of this paper when considering additional sources of uncertainty affecting $y_{i,l}^2$, and when accounting for failures to detect unwanted obstacles. The safety criterion is:

$$P(HMI_k) \le I_{REQ,k} \tag{4}$$

where

$I_{REQ,k}$: is a predefined integrity risk requirement, which is set by a certification authority (similar to requirements set for aviation applications in [1, 4])

$P(HMI_k)$ is unknown, but we can compare the upper bound on $P(HMI_k)$ given in equations (1) to (3) to $I_{REQ,k}$. All terms in (1) to (3) are known, except for $y_{i,l}^2$, which is lower-bounded in the next sections.

Multiple-Hypothesis Innovation-Based Data Association

Equation (1) is derived assuming the innovation-based DA process described in the following paragraphs. Let n_L be the total number of visible landmarks, and n_F the number of estimated feature parameters per landmark. Feature parameters can include landmark position, size, orientation, surface properties, etc. The total number of feature parameters within the visible landmark set is: $n \equiv n_L n_F$. We can stack the actual (true) values of the extracted feature parameters for all landmarks in an $n \times 1$ vector $\mathbf{z}_k$. Let $\hat{\mathbf{z}}_k$ be an estimate of $\mathbf{z}_k$. We assume that the cumulative distribution function of $\hat{\mathbf{z}}_k$ can be bounded by a Gaussian function with mean $\mathbf{z}_k$ and covariance matrix $\mathbf{V}_k$ [28, 29]. We use the notation: $\hat{\mathbf{z}}_k \sim N(\mathbf{z}_k\ ,\ \mathbf{V}_k)$.

Let m be the number of system state parameters to be estimated. The non-linear measurement equation can be written in terms of the $m \times 1$ state parameter vector $\mathbf{x}_k$ as:

$$\hat{\mathbf{z}}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k \tag{5}$$

where

$\mathbf{x}_k$ includes vehicle pose parameters, and may also include landmark feature parameters (for SLAM-type approaches)

$\mathbf{v}_k$ is the extracted measurement noise vector: $\mathbf{v}_k \sim N(\mathbf{0}_{n\times 1},\ \mathbf{V}_k)$ where $\mathbf{0}_{a\times b}$ is an $a \times b$ matrix of zeros.

The mean of $\hat{\mathbf{z}}_k$ is: $\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k)$. Equation (4) can be linearized about an estimate $\bar{\mathbf{x}}_k$ of $\mathbf{x}_k$:

$$\hat{\mathbf{z}}_k \approx \mathbf{h}_k(\bar{\mathbf{x}}_k) + \mathbf{H}_k(\mathbf{x}_k - \bar{\mathbf{x}}_k) + \mathbf{v}_k \quad \text{where} \quad \mathbf{H}_k \equiv \left.\frac{\partial \mathbf{h}_k(\mathbf{x}_k)}{\partial \mathbf{x}_k}\right|_{\bar{\mathbf{x}}_k} \tag{6}$$

The ordering of landmarks in $\hat{\mathbf{z}}_k$ is arbitrary and unknown. A nearest neighbor approach (described below) is used to determine the ordering measurement-to-state coefficients in $\mathbf{h}_k(\bar{\mathbf{x}}_k)$ and $\mathbf{H}_k$. Failing to find the landmark ordering that matches that of $\hat{\mathbf{z}}_k$ causes estimation errors called 'incorrect associations' (IA).

If n_L landmarks are extracted, there are $(n_L!)$ ways to arrange measurements in $\hat{\mathbf{z}}_k$, which we call $(n_L!)$ candidate associations. We first assume that the total number of mapped landmarks –or of previously observed landmarks when using SLAM– is also the number n_L of extracted landmarks. Let subscript i designates association hypotheses, for $i = 0,\ldots,n_A$, where $n_A = n_L!-1$. We define $i = 0$ the fault-free, correct association (CA) hypothesis, the other n_A hypotheses are IA. IA impacts the EKF estimation process through the innovation vector $\boldsymbol{\gamma}_{i,k}$. Vector $\boldsymbol{\gamma}_{i,k}$ is an effective indicator of CA because it is zero mean *only for the correct association*.

In all IA cases, the mean of $\boldsymbol{\gamma}_{i,k}$ is not zero, and is expressed in terms of $n \times n$ permutation matrices $\mathbf{A}_{i,k}$, for $i = 1,\ldots,n_A$, as:

$$\begin{aligned} \boldsymbol{\gamma}_{i,k} &= \hat{\mathbf{z}}_k - \mathbf{A}_{i,k}\mathbf{h}_k(\bar{\mathbf{x}}_k) \\ &= \mathbf{y}_{i,k} + \mathbf{v}_k - \mathbf{A}_{i,k}\mathbf{H}_k\bar{\boldsymbol{\varepsilon}}_k \end{aligned} \tag{7}$$

where $$\mathbf{y}_{i,k} \equiv \mathbf{h}_k(\mathbf{x}_k) - \mathbf{A}_{i,k}\mathbf{h}_k(\mathbf{x}_k) = (\mathbf{I}_n - \mathbf{A}_{i,k})\mathbf{z}_k \quad , \quad \mathbf{y}_{0,k} = \mathbf{0} \tag{8}$$

where $\bar{\boldsymbol{\varepsilon}}_k$ is the EKF state prediction error vector ($\bar{\boldsymbol{\varepsilon}}_k \equiv \bar{\mathbf{x}}_k - \mathbf{x}_k$), and where $\mathbf{I}_a$ is the $a \times a$ identity matrix.

Let $\bar{\mathbf{P}}_k$ be the EKF state prediction error covariance matrix. The nearest neighbor association criterion [5] is defined as:

$$\min_{i=0,\ldots,n_A} \boldsymbol{\gamma}_{i,k}^T \mathbf{Y}_{i,k}^{-1} \boldsymbol{\gamma}_{i,k} \tag{9}$$

where $$\mathbf{Y}_{i,k} = \mathbf{A}_{i,k}\mathbf{H}_k\bar{\mathbf{P}}_k\mathbf{H}_k^T\mathbf{A}_{i,k}^T + \mathbf{V}_k \quad \text{for} \quad i = 0,\ldots,\, n_A \tag{10}$$

The problem with the integrity risk bound in (1) to (3) is that the mean innovation's norm $y_{i,l}^2$ ($y_{i,l}^2 = \mathbf{y}_{i,l}^T\mathbf{Y}_{i,l}^{-1}\mathbf{y}_{i,l}$), is unknown. The only information available is the single sample of innovation norm $\boldsymbol{\gamma}_{i,k}^T\mathbf{Y}_{i,k}^{-1}\boldsymbol{\gamma}_{i,k}$, whose mean is unknown. To get an upper bound on $P(HMI_k)$, several lower-bounds on $y_{i,l}^2$ were attempted, which involved: (a) the use of a heuristically-set design parameter in [25], (b) the use of a threshold set to meet a continuity risk requirement in [26], (c) the use of a Bayesian approach in [30], which makes safety analysis and prediction more challenging. An new approach is employed in this work, which provides the means to select extracted features.

To lighten notations in the next sections, we drop the time subscript k.

FEATURE DATA SELECTION METHOD

This section aims at defining a separation metric that guarantees, with quantifiable integrity, that there is a minimum separation between landmark features [27]. Features that are separated are distinguishable, hence easier to associate. Within a set of features, this minimum separation is determined by considering all possible permutations. This multiple-hypothesis approach purposely mirrors that used in the previous section, and is adopted to exploit the relationship between mean separation vectors and mean innovation vectors. This approach is in part described in [27], and is modified here to enable landmark data selection.

Map-Based Landmark Separation Metric

The minimum feature separation is evaluated using mapped (or previously observed) features $\bar{\mathbf{z}}$. Vector $\bar{\mathbf{z}}$ is an $n \times 1$ vector obtained by stacking all previously-observed feature parameters: $\bar{\mathbf{z}} \sim N(\mathbf{z}\ ,\ \overline{\mathbf{V}})$.

We consider a comprehensive set of non-identity landmark permutation matrices $\mathbf{A}_l$, and we use matrices $\mathbf{B}_l \equiv \mathbf{I}_n - \mathbf{A}_l$, for $l = 1, \ldots, n_A$. The set of $\mathbf{A}_l$ matrices is the same as the set of $\mathbf{A}_i$ in (7), but we are excluding the identity permutation noted $\mathbf{A}_0$ for $l = 0$. For a permutation l, the landmark 'separation' vector is defined as:

$$\bar{\mathbf{d}}_l \equiv \mathbf{B}_l \bar{\mathbf{z}} \quad , \quad \bar{\mathbf{d}}_l \sim N(\mathbf{B}_l \mathbf{z}\ ,\ \mathbf{B}_l \overline{\mathbf{V}} \mathbf{B}_l^T) \tag{11}$$

$\mathbf{B}_l$ is rank deficient, with rank values ranging from m_F to $n - m_F$ depending on how many landmarks are involved in permutation l, for $l = 1, \ldots, n_A$. Let r_l be the rank of $\mathbf{B}_l$. An orthogonal decomposition of the symmetric positive semi definite matrix $\mathbf{D}_l$ is expressed as:

$$\mathbf{D}_l = \begin{bmatrix} \mathbf{U}_l & \mathbf{U}_{l,0} \end{bmatrix} \begin{bmatrix} \mathbf{S}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_l^T \\ \mathbf{U}_{l,0}^T \end{bmatrix} = \mathbf{U}_l \mathbf{S}_l \mathbf{U}_l^T$$

where $\mathbf{S}_l$ is a $r_l \times r_l$ diagonal matrix of non-zero singular values, and $\mathbf{U}_l$ is a $n \times r_l$ matrix such that $\mathbf{U}_l^T \mathbf{U}_l = \mathbf{I}_{r_l}$. The weighted norm squared $\bar{d}_l^2$ is a measure of separation for permutation l, and can be written as:

$$\bar{d}_l^2 \equiv \bar{\mathbf{d}}_l^T \mathbf{U}_l \mathbf{S}_l^{-1} \mathbf{U}_l^T \bar{\mathbf{d}}_l$$

$\bar{d}_l^2$ is non-centrally chi-square distributed with r_l degrees of freedom and with non-centrality parameter d_l^2 defined as: $\bar{d}_l^2 \equiv \mathbf{d}_l^T \mathbf{U}_l \mathbf{S}_l^{-1} \mathbf{U}_l^T \mathbf{d}_l$. We use the notation: $\bar{d}_l^2 \sim \chi^2(r_l\ ,\ d_l^2)$. The minimum feature separation metric is defined as:

$$\bar{d} \equiv \min_{l=1,\ldots n_A} (\bar{d}_l)\,. \tag{12}$$

Given a sample estimate $\bar{d}$ ($\bar{d}$ is considered prior knowledge), we want to find an integrity bound L on the true separation d. In mathematical terms, we want to find L to satisfy: $P(d < L) \le I_{ALLOC}$, where I_{ALLOC} is a chosen, predefined allocation such that $I_{ALLOC} \ll I_{REQ}$. Reference [27] shows that:

$$L = \bar{d} - \sqrt{\chi_{m_F}^{-2}(1 - I_{ALLOC})}\ , \tag{13}$$

where $\chi_{m_F}^{-2}(\ )$ is the inverse chi-square distribution with m_F degrees of freedom.

Probabilistic Bound on Mean Innovation Norm

Equations (11) to (12) can be used to address the fact that y_i^2 is unknown ($y_i^2 = \mathbf{y}_i^T \mathbf{Y}_i^{-1} \mathbf{y}_i$). y_i^2 can be rewritten as:

$$y_i^2 = \mathbf{y}_i^T \mathbf{U}_i \mathbf{S}_i^{-1/2} \mathbf{U}_i^T (\mathbf{U}_i \mathbf{S}_i^{1/2} \mathbf{U}_i^T \mathbf{Y}_i^{-1} \mathbf{U}_i \mathbf{S}_i^{1/2} \mathbf{U}_i^T) \mathbf{U}_i \mathbf{S}_i^{-1/2} \mathbf{U}_i^T \mathbf{y}_i$$

y_i^2 can then be lower-bounded using the following inequality:

$$y_i^2 \geq \mathbf{y}_i^T \mathbf{U}_i \mathbf{S}_i^{-1} \mathbf{U}_i^T \mathbf{y}_i \lambda^2 \quad \text{for } i = 0,\ldots,n_A, \quad \text{with} \quad \lambda^2 = \min_{i=0,\ldots,n_A} \lambda_{MIN,i}^2 \tag{14}$$

where $\lambda_{MIN,i}^2$ is the minimum eigenvalue of $(\mathbf{U}_i \mathbf{S}_i^{1/2} \mathbf{U}_i^T \mathbf{Y}_i^{-1} \mathbf{U}_i \mathbf{S}_i^{1/2} \mathbf{U}_i^T)$. Equations (7) and (11) show that the minimum mean innovation vector and the minimum separation vector are equal:

$$\mathbf{y}_i \equiv (\mathbf{I}_n - \mathbf{A}_i)\mathbf{z} \text{ for } i = 1,\ldots,n_A \quad \text{and} \quad \mathbf{d}_l \equiv (\mathbf{I}_n - \mathbf{A}_l)\mathbf{z} \text{ for } l = 1,\ldots,n_A$$

Thus, (13) establishes with probability larger or equal to $1 - I_{ALLOC}$ that the following bound holds:

$$d_i^2 = \mathbf{y}_i^T \mathbf{U}_i \mathbf{S}_i^{-1} \mathbf{U}_i^T \mathbf{y}_i \geq d^2 \geq L^2 \tag{15}$$

Substituting, (15) into (14), and the result into (3) provides a lower-bound on $P(CA_k)$. The first section of Appendix and [27] show that the bound that we need, and that we actually obtain, is on $P(CA_k \mid d \geq L)$. It is expressed in terms of known quantities as:

$$P(CA_k \mid d \geq L) \geq \prod_{l=1}^{k} \left[1 - P\left(q_l^2 \geq \frac{L_l^2 \lambda_{j,l}^2}{4} \right) \right] \tag{16}$$

Landmark Data Selection Procedure

Equation (1) captures a key tradeoff in landmark data selection: on the one hand, selecting fewer measurements increases the nominal estimation error and hence $P(HMI_k \mid CA_K)$, but on the other hand, it improves $P(CA_k)$ because measurements become more distinguishable.

The bound in (16) is relevant to landmark data selection because it accounts for the minimum separation L between *all* landmarks in the map, but at the same time, opens the possibility to only associate a subset of feature measurements. (Ignoring existing mapped landmarks would give optimistic risk estimates, which is unsafe [30].) To capture this idea, subscript j is included in $\lambda_{j,l}^2$ to indicate a candidate measurement subset. The integrity risk bound in (1) to (3) becomes:

$$P(HMI_{j,k}) \leq 1 - \left[1 - 2Q\left\{ \frac{\ell}{\sigma_{j,k}} \right\} \right] \prod_{l=1}^{k} \left[1 - P\left(q_l^2 \geq \frac{L_l^2 \lambda_{j,l}^2}{4} \right) \right] + I_{ALLOC} \tag{17}$$

where $\sigma_{j,k}$ is the standard deviation of the estimation error using the selected feature measurement subset j, and the allocation I_{ALLOC} is incorporated using the derivation in the first section of Appendix.

We can distinguish two approaches to bound $P(HMI_{j,k})$ after selecting a landmark subset.

- The first approach aims at fixing the number n_S of extracted landmarks, and at considering measurements for *any* subset of n_S landmarks. With this approach, equation (17) applies directly, but the number of association candidates n_A changes from $n_L!$ to $n_L!/(n_L - n_S)!$. (This number is derived from $C(n_L, n_S) n_S!$, where $C(n_L, n_S)$ is the number of combinations of n_S landmarks within a total set of n_L landmarks.)
- The second approach focuses on selecting *one specific* subset of landmark feature measurements. The method provides more control over measurements to be selected, but also introduces a risk of the extracted measurement subset not being the intended one. Wrong extraction can be addressed using the unwanted obstacle (UO) detection process described in the next section. (In addition to wrong extractions, UO detection mitigates risks caused by unmapped obstacles.)

Following the second approach, we select the optimal measurement subset as the one that minimizes integrity risk, or more precisely, that minimizes the maximum integrity risk over the mission duration. In this paper, we use a brute force technique to find the subset measurement j that satisfies:

$$\min_{j}\left(\max_{k}\left\{ P(HMI_{j,k})\right\} \right) \tag{18}$$

UNMAPPED OBSTACLE DETECTION USING SOLUTION SEPARATION

In the presence of a wrong extraction or unmapped obstacle, the innovation vector's norm in (9) is always non-zero. In this scenario, the correct association hypothesis must be redefined. We call the correct association (CA) the one that minimizes the mean innovation norm squared $y_i^2 \equiv \mathbf{y}_i^T \mathbf{Y}_i^{-1} \mathbf{y}_i$. All we have is a noisy sample measure of y_i^2, noted: $\gamma_i^2 \equiv \boldsymbol{\gamma}_i^T \mathbf{Y}_i^{-1} \boldsymbol{\gamma}_i$.

An Innovation-Based Approach Impractical

If an unwanted obstacle (UO) is present, γ_i is not zero mean even under CA, which provides an opportunity for fault detection. Thus, our initial idea for detecting unwanted obstacles (UO) was to set a threshold T_k on the minimum innovation norm ($\min_{i=0,\ldots n_A} \gamma_i^2$). If T_k is exceeded, we can either select a different subset of measurements, or interrupt the mission. If T_k is not exceeded, we must quantify the risk of having an undetected UO.

The challenge with using innovation-based detection is in bounding the mean estimation error, because the impact of undetected UO accumulates over time. For example, let us express the state estimation error $\hat{\varepsilon}_k$ at time k as:

$$\hat{\varepsilon}_k = \bar{\varepsilon}_k + \mathbf{k}_{i,k}^T \boldsymbol{\gamma}_{i,k}$$

where

$\bar{\varepsilon}_k$: is the state prediction error for the state of interest (or linear combination of states) at time k

$\mathbf{k}_{i,k}$: is the $n \times 1$ Kalman gain vector for the state of interest, for candidate association i

A detection test at time k guarantees that the weighted norm of $\boldsymbol{\gamma}_{i,k}$ does not exceed T_k. This ensures an upper bound on the mean of $|\hat{\varepsilon}_k|$, which through the EKF time update also ensures an upper bound on $|\bar{\varepsilon}_{k+1}|$. At time $k+1$, the detection test guarantees again that the norm of $\boldsymbol{\gamma}_{i,k+1}$ is bounded, and so is the mean of $|\bar{\varepsilon}_{k+1}|$, but these bounds add up considering the measurement update equation $\hat{\varepsilon}_{k+1} = \bar{\varepsilon}_{k+1} + \mathbf{k}_{i,k+1}^T \boldsymbol{\gamma}_{i,k+1}$. Such bounds add up from filter initiation to current time, and the overall mean estimation error can grow quickly and without bound.

Solution Separation Approach

Instead, a solution separation (SS) approach is used for detection of UO [2, 31-33]. In this context, multiple-hypothesis SS is interesting because it guarantees that the estimation error is zero-mean for the appropriate subset solution. SS exploits the redundancy of associations to detect UO. We must assume that only a subset of landmarks can be confused with UO. In this paper, we assume that a single landmark may be confused with an UO, but it can be any landmark in the map.

SS may seem cumbersome at first because it requires a bank of EKFs: one EKF for the full-set solution and one for each subset solution. Fortunately, there are methods to group fault hypotheses that limit the number of required subset solutions to two or three in this application. This point is further illustrated with performance evaluations in the next section.

To quantify the risk caused by potentially undetected UO, the initial integrity risk definition in (1) is modified: HMI is the joint event of the car being out of lane and of none of the solution separation tests triggering an alarm. The integrity risk is defined as:

$$P(HMI_j) = P\left(|\hat{\varepsilon}_j| > \ell \ \cap \ \left[\bigcap_{h=0}^{n_{H,j}} |\hat{\varepsilon}_j - \hat{\varepsilon}_h| \le T_h \right] \right)$$

where

$\hat{\varepsilon}_j$ is the 'full-set' solution, i.e., the state estimation error for the selected landmark feature measurement subset j

$\hat{\varepsilon}_h$ is 'subset' solution h, which is also the fault-free solution of the h^{th} UO fault hypothesis

h is an index that indicates UO hypotheses; it also designates the corresponding fault-free subset solution; $h = 0$ is the fault-free 'no UO' hypothesis

$n_{H,j}$ is the number of UO fault hypotheses; we have n_S single UO hypotheses for the selected subset j; we can reduce this number by grouping UO hypotheses; there are many ways to group hypotheses, as long as we consider a set of mutually exclusive, exhaustive hypotheses.

T_h is the detection threshold for the h^{th} solution separation: $|\hat{\varepsilon}_j - \hat{\varepsilon}_h|$.

The detection threshold T_h is set according to a continuity risk requirement C_{REQ} to limit the risk of false alarm. T_h is given by [31-33]:

$$T_h = Q^{-1}\left\{ 1 - C_{REQ}/n_H \right\} \sqrt{\sigma_h^2 - \sigma_j^2}$$

where

σ_j^2 is the variance of 'full-set' solution's estimation error for subset j (after landmark data selection)

σ_h^2 is the variance for subset solution h

The integrity risk bound accounting for the risk of not detecting an UO when one might be present is derived in Appendix. The derivation follows steps similar to the ones used in general SS methods [33], and uses the following key inequality:

$$|\hat{\varepsilon}_j| = |\hat{\varepsilon}_h + \hat{\varepsilon}_j - \hat{\varepsilon}_h| \le |\hat{\varepsilon}_h| + |\hat{\varepsilon}_j - \hat{\varepsilon}_h|$$

which can be further bounded by: $|\hat{\varepsilon}_j| \le |\hat{\varepsilon}_h| + T_h$ given that no alarm is triggered, i.e., that $|\hat{\varepsilon}_j - \hat{\varepsilon}_h| \le T_h$. The resulting integrity risk bound is (see the second section of Appendix):

$$P(HMI_j) \le \sum_{h=0}^{n_{H,j}} \left[1 - \left(1 - P(|\hat{\varepsilon}_h| + T_h > \ell \,|\, H_h,\ CA_{h,K}) \right) P(CA_{h,K} \,|\, H_h,\ d < L) \right] + I_{ALLOC} \tag{19}$$

where

H_h designates the h^{th} UO fault hypothesis; H_0 is the fault-free (no UO) hypothesis.

$CA_{h,K}$ is the correct association for the subset of measurements that is fault-free under hypothesis H_h

and where we used the notation: T_0 =0. Equation (19) can also be expressed as:

$$P(HMI_{j,k}) \le \sum_{h=0}^{n_{H,j}} \left(1 - \left[Q\left\{ \frac{-\ell + T_{h,k}}{\sigma_{h,k}} \right\} - Q\left\{ \frac{\ell + T_{h,k}}{\sigma_{h,k}} \right\} \right] \prod_{l=1}^{k} \left[1 - P\left(q_l^2 \ge \frac{L_l^2 \lambda_{j,l}^2}{4} \right) \right] \right) + I_{ALLOC} \tag{20}$$

PERFORMANCE ANALYSIS

In this section, example simulations used in [25-27, 34, 35] are modified to evaluate the above landmark data selection and UO detection methods.

Covariance Analysis of Landmark Data Selection for an Illustrative Three Landmark Scenario

The left charts on Figures 2 and 4 show a vehicle represented by an upward pointing triangle roving between three landmarks (black-shaded circles). The vehicle starts at an initial, known position at point (0, 0) in a local East-North reference frame, and uses measurements from a lidar to estimate its position. In this example, vehicle orientation is known. While roving along the North axis, the vehicle passes by three point-feature landmarks. Assuming that a precise map is available, landmark locations are known. Simulation parameters are listed in Table 1. Vehicle positioning errors at consecutive sample updates are represented by red covariance ellipses in Figures 2 and 4 (left charts). These ellipses assume consistently successful FE and DA. We focus on the lateral positioning error, perpendicular to vehicle's straight line trajectory.

The right chart in Figure 2 shows the integrity risk bound labeled $P(HMI_k)$, represented with a red curve, versus northward travel distance. The bound is always larger than $I_{ALLOC,k} = 10^{-9}$, which is our chosen integrity risk allocation to determine the landmark separation bound L in (13). This $P(HMI_k)$-bound is loose when $P(CA_K) \approx 1$ and $P(HMI_k \mid CA_K) << 10^{-9}$, which is the case for travel distances between 5 m and 32 m in Figure 2 (right). A loose bound is adequate when $P(HMI_k) << I_{REQ,k} = 10^{-7}$ because it does not impact availability of $P(HMI_k) < I_{REQ,k}$.

In parallel, the black curve in Figures 2 and 4 (right) represents the bound on $P(HMI_k \mid CA_K)$, which is fully determined by the alert limit ℓ and the vehicle positioning error covariance. This covariance-based vehicle navigation performance metric is often used to evaluate lidar-based navigation systems [18, 23]. The black curve converges with our integrity risk bound (red curve) for travel distances smaller than 5 m. However, for travel distances larger than 30 m, the $P(HMI_k)$-bound is orders of magnitude larger than the $P(HMI_k \mid CA_K)$ curve. This is because $P(HMI_k)$ accounts for the risk of incorrect associations (IA), whereas $P(HMI_k \mid CA_K)$ does not. The covariance-based metric is misleading in that regard. The reason for the sudden increase in integrity risk bound for travel distances of about 30 m is that landmark '1' gets hidden behind landmark '3', which causes confusion in DA.

In Figure 3, the $P(HMI_k)$ and $P(HMI_k \mid CA_K)$ curves are represented for the full-set and three candidate subsets. The selected subset is the one which satisfies (18): $\min_j[\max_k\{P(HMI_{j,k})\}]$. In this case, it is the subset comprising landmarks '1' and '2'. The $\max\{P(HMI_k)\}$ for the corresponding thick curve is significantly lower than for any of the other red curves.

Table 1. Simulation parameters

System Parameters	Values
Standard deviation on raw lidar ranging measurement	0.02 m
Standard deviation on raw lidar angular measurement	0.5 deg
Lidar range Limit	20 m
Lidar data sampling interval	0.5 s
Vehicle speed	1 m/s
Alert limit ℓ	0.5 m
Integrity risk allocation for FE, $I_{FE,ALLOC,k}$	10^{-9}

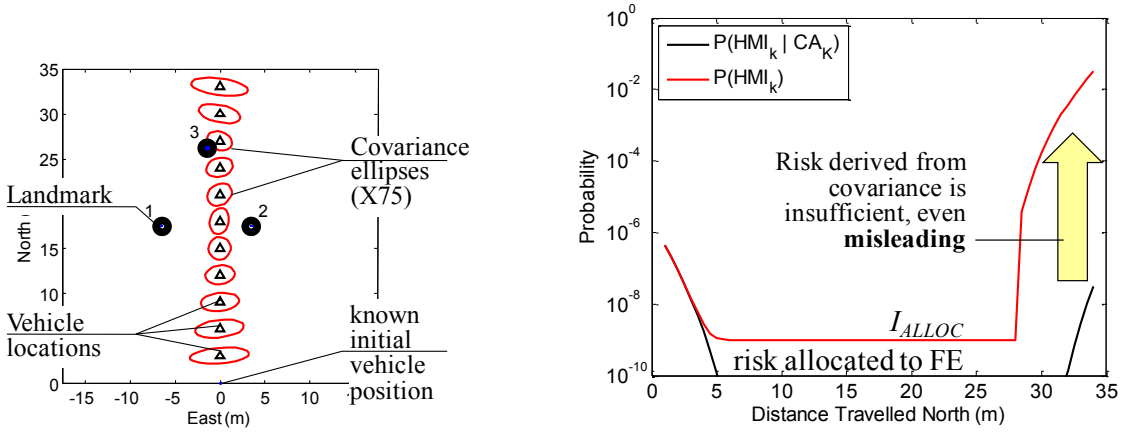

Figure 2. (LEFT) Covariance Analysis for the Illustrative Three Landmark Scenario; (RIGHT) Integrity Risk Bound for the Illustrative Three Landmark Scenario.

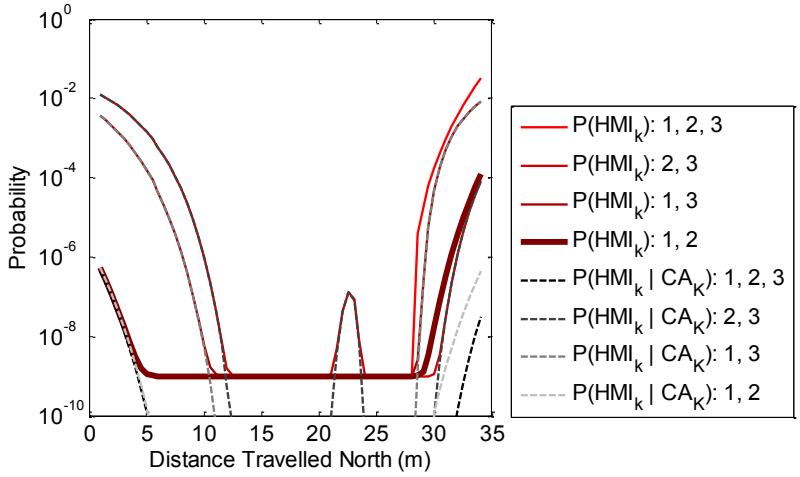

Figure 3. Integrity Risk Bounds for All Candidate Landmark Feature Measurement Subsets. All shades-of-red curves are considered in the data selection process, which uses a straightforward "brute-force" procedure. : the integrity-risk-minimizing measurement subset (thick line) includes landmarks '1' and '2', and excludes '3'.

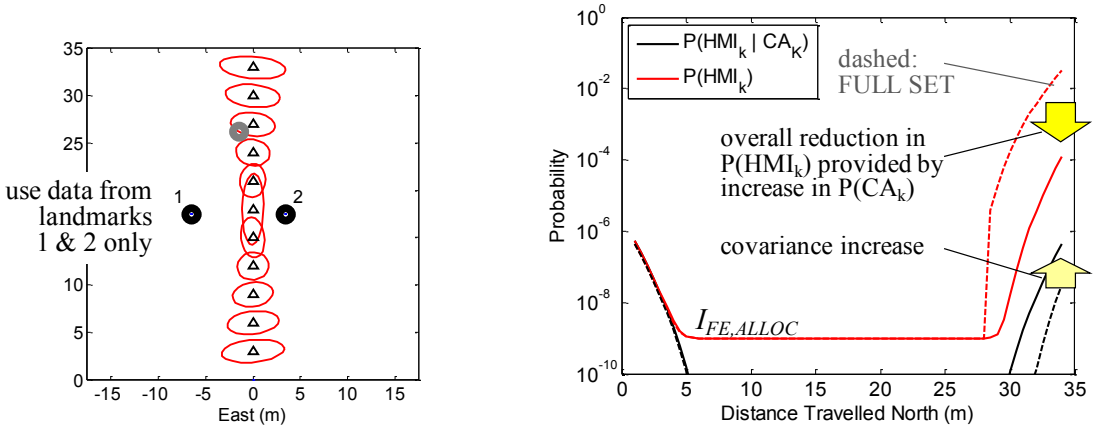

Figure 4. (LEFT) Three Landmark Scenario Used to Illustrate the Landmark Measurement Selection Process; (RIGHT) Integrity Risk Bound Comparison of the Full-Set Versus the Risk-Minimizing Subset of Landmark Data.

This point is further illustrated in Figure 4 (right chart), where the integrity risk curves for the full-set and for the optimal subset are compared. Although the black covariance-based curve is higher for the subset (solid) than for the full-set (dashed) —because fewer measurements are exploited using the subset—, the overall $P(HMI_k)$-bound curve is two orders of magnitude lower for the subset as compared to the full-set —because the risk of IA is reduced—.

Direct Simulation: Vehicle Roving Through a GNSS-Denied Area

This analysis investigates the safety performance of a GPS/lidar navigation system onboard a vehicle roving through a forest. GPS signals are blocked by the tree canopy, and low-elevation satellite signals do not penetrate under the trees. Tree trunks are used as landmarks by the lidar-based SLAM-type algorithm.

The measurement vector $\hat{\mathbf{z}}_k$ in (5) is augmented with GPS code and carrier measurements, and the state vector $\mathbf{x}_k$ is augmented to include an unknown GPS receiver clock bias and carrier phase cycle ambiguities. Time-correlated GPS signals and non-linear lidar data are processed in a unified time-differencing EKF derived in [34, 35]. The simulation parameter values are listed in Table 1, and a standard differential GPS measurement error model is used, which is described in [35]. In this scenario, GPS and lidars essentially relay each other with seamless transitions from open-sky through GPS-denied areas where landmarks modeled as poles with non-zero radii are visible.

Figures 5 to 7 and 9 consistently employ the following yellow-green-blue color code: the mission starts with the vehicle operating in a GPS available area (yellow-shaded). Satellite signals available during initialization enable accurate estimation of cycle ambiguities, so that vehicle positioning uncertainty does not exceed a few centimeters. Then, as the vehicle moves and crosses the GPS-and-lidar available area (green-shaded), and the lidar-only area (blue-shaded), seamless variations in covariance are achieved. A detailed description of this simulation is given in [35].

In this scenario, the likelihood of IA is high. One indicator of IA is displayed on the upper left-hand-side (LHS) plot in Figure 5. It shows that the actual cross-track positioning error versus distance travelled exceeds the corresponding one-sigma covariance envelope. This suggests that errors are impacting positioning, which are not captured by the covariance.

This is confirmed on the lower LHS chart in Figure 5, where the black curve showing the $P(HMI_k \mid CA_K)$-bound stays below 10^{-8}. In contrast, the red $P(HMI_k)$-bound curve reaches a first plateau of $I_{ALLOC,k}$ as soon as two landmarks are visible (for reasons explained in previous sections). The $P(HMI_k)$-bound curve then suddenly increases to 1, at approximately 22 m of travel distance. At this point, there is not enough information to guarantee that any candidate associations is correct.

The right-hand-side (RHS) chart in Figure 5 shows that, at the travel distance corresponding to the large increase in predicted integrity risk, landmark '6' is hidden behind landmark '5'. Before that point, it was either out of lidar range, or hidden behind landmark '5'. It first becomes visible to the lidar at the next time step, which makes correct measurement association with either landmark '5' or '6' extremely challenging. The $P(HMI_k)$-bound accounts for the risk caused by such events.

Our proposed landmark data selection method was implemented in Figure 6. The algorithm determined that the candidate subset that minimizes $\max\{P(HMI_k)\}$ is the one that excludes data from landmarks '6' to '8'. This makes measurements more distinguishable and therefore reduces occurrences of incorrect associations. It is confirmed in the upper LHS graph of Figure 6 where the positioning error now fits the covariance envelope. Risk reduction is also observed in the lower LHS chart, where even though the black $P(HMI_k \mid CA_K)$-curve increases for subset (solid) versus full-set (dashed), the $P(HMI_k)$-bound does not exceed 10^{-6} (solid red), versus 1 for the full-set solution (dashed red). The RHS graph in Figure 6 shows the vehicle-landmark geometry when the increase in $P(HMI_k)$-bound occurs.

In addition, the LHS in Figure 7 shows the $P(HMI_k)$-bound (thick black curve) accounting for the risk of not detecting an unwanted obstacle (UO) even though one might be present. This bound is given in equation (20). The resulting landmark data selection indicates that the risk-minimizing subset is the same with and without detection, so that data from landmarks '6' to '8' are again being excluded.

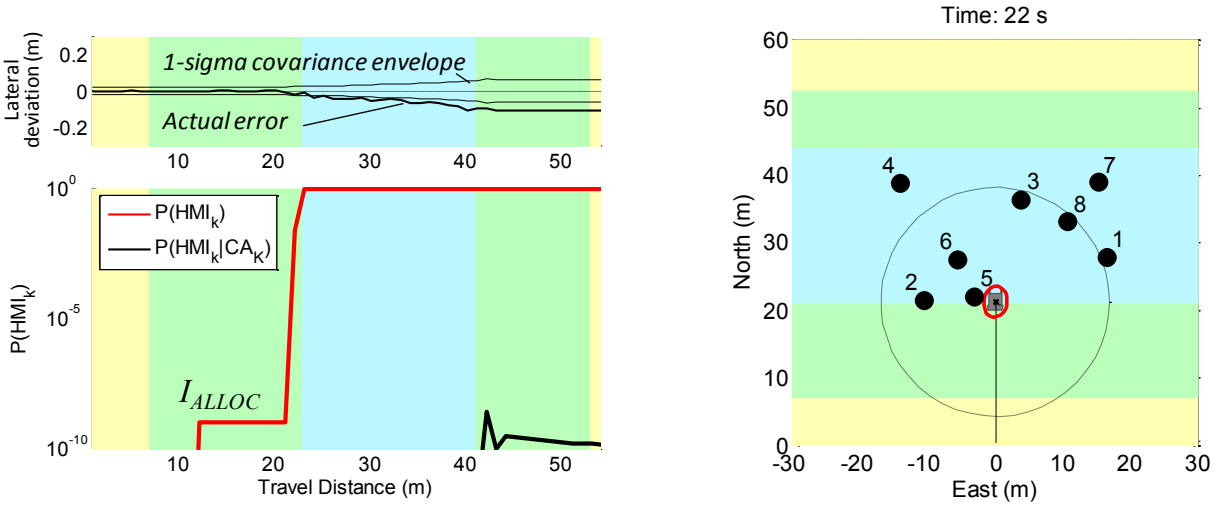

Figure 5. (LEFT) $P(HMI_k)$-Bound for the GPS-Denied-Area Crossing Scenario; (RIGHT) Snapshot Vehicle-Landmark Geometry at the Time Step Corresponding to the Large Increase in $P(HMI_k)$-Bound.

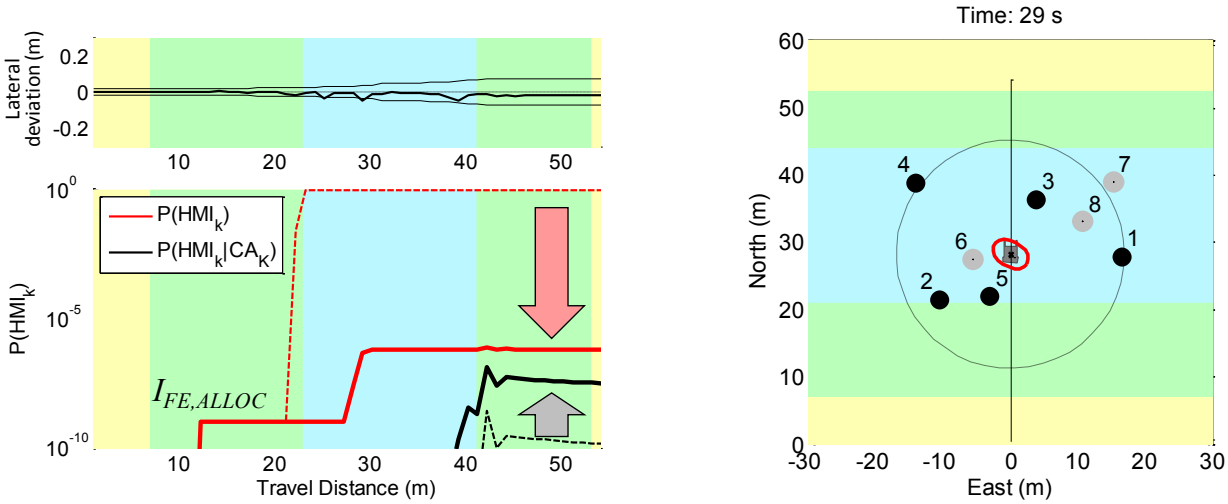

Figure 6. (LEFT) $P(HMI_k)$-Bound for the Case when Measurements from Landmarks '1' to '5' are Used, and Data from Landmarks '6' to '8' are Excluded (Solid Lines) as Compared to Using All Measurements (Dashed Lines). (RIGHT) Snapshot Vehicle-Landmark Geometry at the Time Step Corresponding to the Increase in $P(HMI_k)$-Bound.

$P(HMI_k)$-bounds were evaluated for the 'full data set' (after selection) and for the five subsets corresponding to individual single-UO hypothesis, following the proposed solution separation (SS) method. Grouping hypotheses reduces computation load. We found out that the subsets' accumulated risk contribution was not significantly different from the one obtained by grouping hypotheses as shown in the RHS of Figure 7. We considered two landmark subsets: {1, 2, 3} and {4, 5} in the solution separation approach. This guarantees that if a single UO is present, we always have a fault-free subset solution; e.g., if landmark '1' is confused with an UO, then subset solution {4, 5} is fault free. In this case, we need a bank of only three EKF for the full-set and subset solutions, instead of six when considering each individual single-UO hypothesis. Risk contributions for grouped subsets and full-set are color-coded and represented in Figure 7 (LHS).

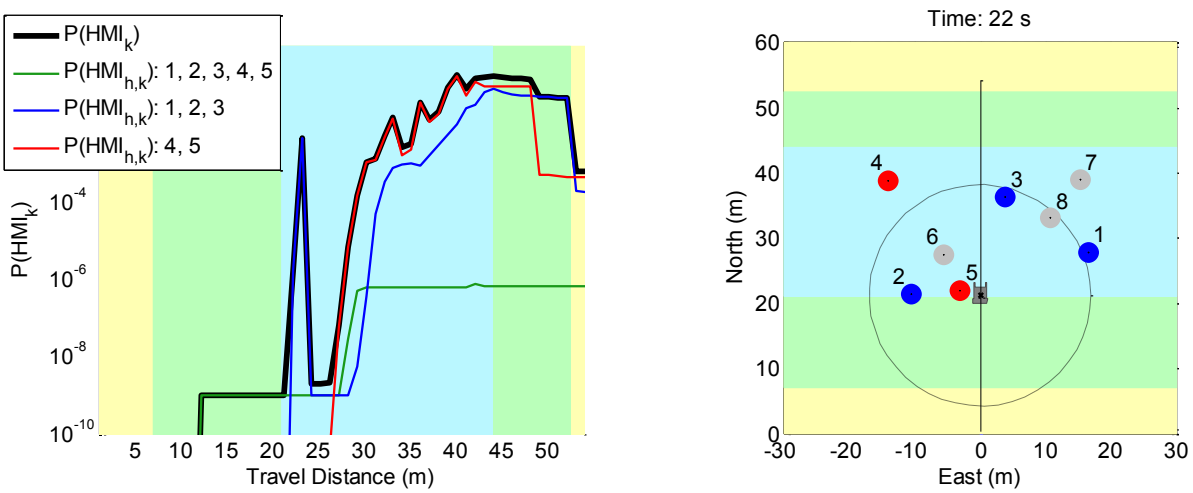

Figure 7. (LEFT) $P(HMI_k)$-Bound Contributions of the Full-Set and Subsets Using Solution Separation to Detect Unmapped Obstacles: the overall risk is the thick black line; (RIGHT) Color-Coded Subsets Used in Solution Separation: the 'full-set' after selection is {1, ..., 5}, the subset solutions are based on landmarks {1, 2, 3} and {4, 5}.

The LHS graph in Figure 7 shows that the $P(HMI_k)$-bound increases from 10^{-6} when assuming no UO to 10^{-2} when accounting for the potential presence and failed detection of a single UO. As expected in a SS implementation, the weakest landmark geometries are driving loss of integrity. In this example, the redundancy of distinguishable landmarks was too low to achieve low integrity risk. In response, one might have to incorporate other sensors such as inertial navigation systems, or one might exploit additional features (surface properties) and landmarks.

Preliminary Testing in an Incorrect-Association-Free Environment

Preliminary experimental testing is carried out using data collected in a structured environment shown in Figure 8. Static simple-shaped landmarks are located at locations sparse enough to ensure successful outcomes for FE and DA. Because the results presented here are free of incorrect associations, they describe the estimation process, and $P(HMI_k)$ is expected to match $P(HMI_k \mid CA_K)$. This test data is used to focus on the risk of UO misdetection.

Measurements from carrier phase differential GPS (CPDGPS) as well as lidar scanners are synchronized and recorded. In order to obtain a full 360 deg lidar scan, two 180 deg lidar scanners are assembled back-to-back. The lidar scanners have a specified 15-80 m range limit, a 0.5 deg angular resolution, a 5 Hz update rate and a ranging accuracy of 1-5 cm (1 sigma) [36]. The GPS antenna is mounted on top of the front lidar. The lever-arm distance between the two lidars is included in the measurement model. The two lidars and the GPS antenna are mounted on a rover also carrying the GPS receiver and data-link. An embedded computer onboard the vehicle records all measurements including the raw GPS data from the reference station transmitted via wireless spread-spectrum data-link. Truth trajectory is obtained using a fixed CPDGPS solution.

The upper LHS chart in Figure 9 confirms that this is an incorrect-association-free scenario because the actual error (thick line) fits within the covariance envelope (thin line) throughout the test. As expected, the landmark data selection method determined that using all available data would minimize $P(HMI_k)$.

In addition, the lower LHS graph in Figure 9 shows $P(HMI_k)$-bound contributions for each subset solution. In this case, we used three subset solutions, excluding landmarks as indicated by the color-code on the RHS chart. For example, the blue $P(HMI_k)$-curve on the LHS is for the subset that excludes data from landmarks '3' and '4'. If an UO is incorrectly associated with landmark '1', then the brown curve for the subset that excludes '1' and '6' is fault-free. The LHS graph shows again a substantial increase in $P(HMI_k)$ when accounting for undetected UO (thick black curve), as compared to ignoring their potential presence (dashed red line). This must be addressed using redundant information from other additional sensors, landmarks, or landmark features.

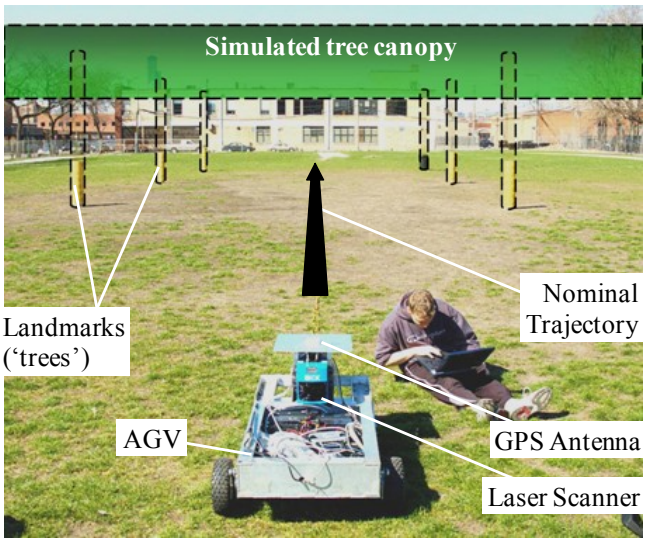

Figure 8. Experimental Setup of a Forest-Type Scenario, where a GPS/Lidar-Equipped Rover is driving by Six Landmarks (Cardboard Columns) in a GPS-Denied Area. GPS is Artificially Blocked by a Simulated Tree Canopy, and a Precise Differential GPS Solution is used for Truth Trajectory Determination.

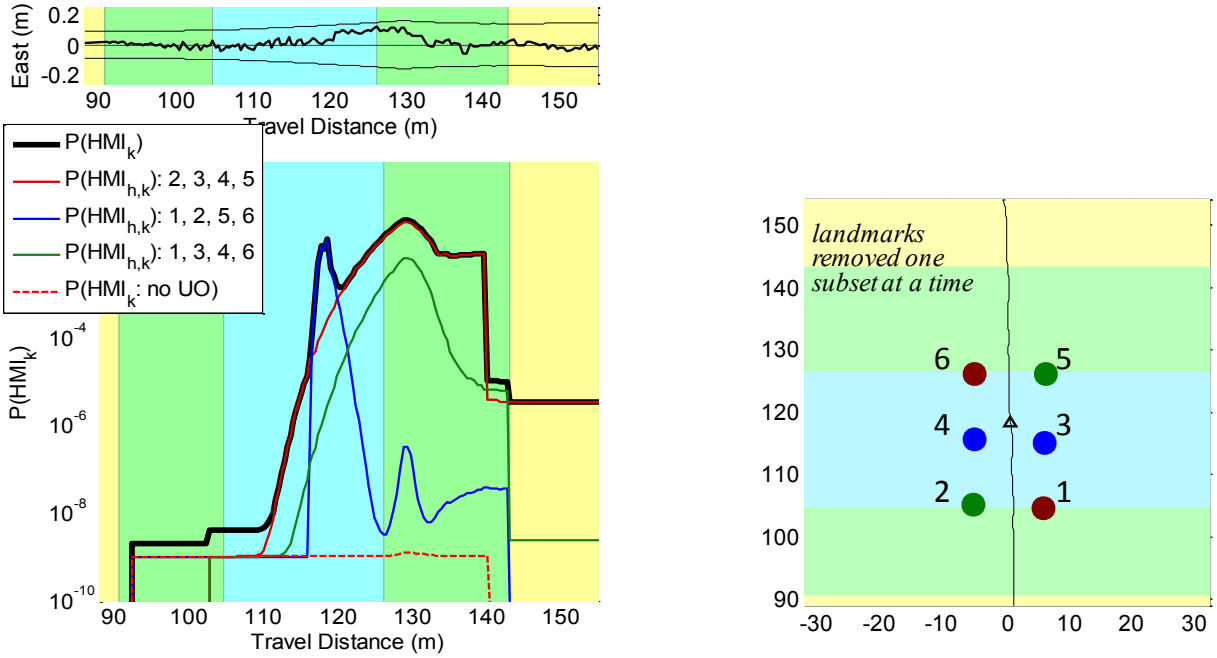

Figure 9. (LEFT) $P(HMI_k)$-Bound Contributions of the Full-Set and Subsets Using Solution Separation to Detect Unmapped Obstacles for the Preliminary Experimental Data Set: the overall risk is the thick black line; (RIGHT) Color-Coded Subsets Used in Solution Separation: the 'full-set' after selection is {1, ..., 6}, the subset solutions are based on landmarks {2, .., 5}, {1, 2, 5, 6} and {1, 3, 4, 6}.

CONCLUSION

This paper presents a new approach to quantifiably improve the safety of lidar-based navigation using feature extraction (FE) and data association (DA) by selecting landmark data and detecting unwanted obstacles (UO). First, we developed a new landmark data selection method that provides the means to choose a subset of extracted feature measurements while accounting for all mapped landmarks. 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. Then, we designed a multiple-hypothesis solution separation detector, and we established an analytical expression for the impact of undetected UO on integrity risk. Covariance analysis, direct simulation, and preliminary testing in a structured environment showed that feature selection can dramatically reduce integrity risk. Performance evaluations also suggested that additional, redundant information from other sensors would probably be needed to safely detect UO in the lidar's surroundings. Future work includes the evaluation of practical feature extractors to refine the single UO assumption, and the experimental validation of these new methods in a more realistic ground vehicle environment.

APPENDIX

Integrity Risk Derivation to Incorporate the Landmark Separation Bound

In order to incorporate the probabilistic landmark separation bound L into the integrity risk expression, we start from equation (1) and we use the law of total probability for mutually exclusive exhaustive hypotheses of events ($d \geq L$) and ($d < L$). Recalling that, by definition of L, $P(d < L) \leq I_{ALLC}$, and that $P(d \geq L) + P(d < L) = 1$, we consider the following inequalities:

$$
\begin{aligned}
P(HMI_k) &\leq 1 - [1 - P(HMI_k \mid CA_K)]P(CA_K) \\
&\leq 1 - [1 - P(HMI_k \mid CA_K)]\left[P(CA_K, d \geq L) + P(CA_K, d < L)\right] \\
&\leq P(d \geq L) + P(d < L) - [1 - P(HMI_k \mid CA_K)]\left[P(CA_K \mid d \geq L)P(d \geq L) + P(CA_K \mid d < L)P(d < L)\right] \\
&\leq P(d \geq L)\left\{ 1 - [1 - P(HMI_k \mid CA_K)]\ P(CA_K \mid d \geq L) \right\} \\
&\qquad + P(d < L)\left\{ 1 - [1 - P(HMI_k \mid CA_K)]\ P(CA_K \mid d < L) \right\} \\
&\leq 1 - [1 - P(HMI_k \mid CA_K)]\ P(CA_K \mid d \geq L) + I_{ALLOC}
\end{aligned}
$$

Integrity Risk Derivation to Incorporate the Risk of No Detection Using Solution Separation

The integrity risk bound using solution separation is derived by going back to the initial $P(HMI_{j,k})$ equation in (1), and considering the following inequalities:

$$
\begin{aligned}
P(HMI_j) &= P\left(\mid \hat{\varepsilon}_j \mid > \ell \ \cap \ \left[\bigcap_{h=1}^{n_H} \mid \hat{\varepsilon}_j - \hat{\varepsilon}_h \mid \leq T_h \right] \right) \\
&\leq \sum_{h=0}^{n_{H,j}} P(\mid \hat{\varepsilon}_j \mid > \ell \ \cap \ \mid \hat{\varepsilon}_j - \hat{\varepsilon}_h \mid \leq T_h \ \cap \ H_h) \\
&\leq \sum_{h=0}^{n_{H,j}} P(\mid \hat{\varepsilon}_j + \hat{\varepsilon}_h - \hat{\varepsilon}_h \mid > \ell \mid \ \cap \ \mid \hat{\varepsilon}_j - \hat{\varepsilon}_h \mid \leq T_h \ \cap \ H_h) \\
&\leq \sum_{h=0}^{n_{H,j}} P(\mid \hat{\varepsilon}_h \mid + \mid \hat{\varepsilon}_j - \hat{\varepsilon}_h \mid > \ell \mid \ \Big| \ \mid \hat{\varepsilon}_j - \hat{\varepsilon}_h \mid \leq T_h \ \cap \ H_h) \\
&\leq \sum_{h=0}^{n_{H,j}} P(\mid \hat{\varepsilon}_h \mid + T_h \ > \ell \mid \ H_h)
\end{aligned}
$$

$$P(HMI_j) \le \sum_{h=0}^{n_{H,j}} \left[P(|\hat{\varepsilon}_h| + T_h > \ell \cap CA_{h,K} \mid H_h) + P(|\hat{\varepsilon}_h| + T_h > \ell \cap IA_{h,K} \mid H_h) \right]$$

$$\le \sum_{h=0}^{n_{H,j}} \left(1 - [1 - P(|\hat{\varepsilon}_h| + T_h > \ell \mid CA_{h,K} \cap H_h)] P(CA_{h,K} \mid d \ge L \cap H_h) \right) + I_{ALLOC}$$

where we used the notation: $T_0 = 0$, and where $CA_{h,K}$ designates correct associations for the subset of measurements that is fault-free under hypothesis H_h. Also, in this derivation, for consistency with the '$\bigcap_{h=1}^{n_H}$' operator in the first line, we used the notation '$\cap$' to designate an intersection of events, but we use the more compact symbol '$,$' in the rest of the paper.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the National Science Foundation for supporting this research (NSF award #1637899). However, the opinions expressed in this paper do not necessarily represent those of any other organization or person.

REFERENCES

[1] RTCA Special Committee 159, "Minimum Aviation System Performance Standards for the Local Area Augmentation System (LAAS)," *RTCA/DO-245*, 2004, Appendix D.

[2] Y.C. Lee, "Analysis of Range and Position Comparison Methods as a Means to Provide GPS Integrity in the User Receiver," *Proc. of the 42nd Annual Meeting of The Institute of Navigation*, Seattle, WA, 1986, pp. 1-4.

[3] B.W. Parkinson, and P. Axelrad, "Autonomous GPS Integrity Monitoring Using the Pseudorange Residual," *NAVIGATION*, Vol. 35, No. 2, 1988, pp. 225-274.

[4] RTCA Special Committee 159, "Minimum Operational Performance Standards for Global Positioning System/Wide Area Augmentation System Airborne Equipment," *RTCA/DO-229C*, 2001, pp. 1-21.

[5] Y. Bar-Shalom, and T. E. Fortmann, "Tracking and Data Association," *Mathematics in Science and Engineering*, Vol. 179, Academic Press, 1988.

[6] T. Bailey, "Mobile Robot Localisation and Mapping in Extensive Outdoor Environments," PhD Dissertation, The University of Sydney, 2002.

[7] T. Bailey and J. Nieto, "Scan-SLAM: Recursive Mapping and Localisation with Arbitrary-Shaped Landmarks," *Workshop at IEEE RSS 2008*, Zurich, Switzerland, 2008.

[8] A.J. Cooper, "A Comparison of Data Association Techniques for Simultaneous Localization and Mapping," M.S. Thesis, Massachussetts Institute of Technology, 2005.

[9] S. T. Pfister, K. L. Kriechbaum, S. I. Roumeliotis, J. W. Burdick, "Weighted Range Sensor Matching Algorithms for Mobile Robot Displacement Estimation," *Proc IEEE ICRA*, 2002.

[10] S.T. Pfister, S.I. Roumeliotis, J.W. Burdick, "Weighted Line Fitting Algorithms for Mobile Robot Map Building and Efficient Data Representation Robotics and Automation," *Proc. IEEE ICRA*, 2003.

[11] I. Tena Ruiz, Y. Petillot, D. M. Lane, and C. Salson, "Feature Extraction and Data Association for AUV Concurrent Mapping and Localisation," *Proc. IEEE-ICRA*, 2001.

[12] S. Thrun, "Robotic Mapping: A Survey," Exploring Artificial Intelligence in the New Millenium, G. Lakemeyer and B. Nebel, February 2002.

[13] S. Thrun, W. Burgard, and D. Fox, "A Probabilistic Approach to Concurrent Mapping and Localization for Mobile Robots," *Machine Learning and Autonomous Robots*, Vol. 31, No. 5, 1998, pp. 1-25.

[14] S. Thrun, W. Burgard, and D. Fox, "A Real-Time Algorithm forMobile Robot MappingWith Applications to Multi-Robot and 3D Mapping," *Proc. IEEE ICRA 2000*, San Francisco, CA, 2000.

[15] V. Nguyen, A. Martinelli, N. Tomatis, and R. Siegwart, "A Comparison of Line Extraction Algorithms using 2D Laser Rangefinder for Indoor Mobile Robotics," *Proc. IEEE/RSJ IROS*, 2005.

[16] P. Nunez, R. Vazquez-Martin, J.C. del Toro, and A. Bandera, "Feature Extraction from Laser Scan Data based on Curvature Estimation for Mobile Robotics," *Proc. IEEE ICRA*, 2006.

[17] Y. Li, and E.B. Olson, "A general purpose feature extractor for light detection and ranging data," *Sensors*, Vol. 10, No. 11, 2010.

[18] R. Madhavan, H. Durrant-Whyte, and G. Dissanayake, "Natural Landmark-based Autonomous Navigation using Curvature Scale Space," *Proc. IEEE-ICRA*, 2002.

[19] D. Maksarov, and H. Durrant-Whyte, "Mobile Vehicle Navigation in Unknown environments: a multiple hypothesis approach," *Proc. IEEE Control Theory Applications*, Vol. 142 No. 4, 1995, pp. 385-400.

[20] J. Leonard, and H. Durrant-Whyte, "Directed Sonar Sensing for Mobile Robot Navigation," Kluwer Academic Publishers, Cambridge, MA, 1992, pp. 129-138.

[21] S.B. Williams, G. Dissanayake, and H. Durrant-Whyte, "An efficient Approach to the Simultaneous Localisation and Mapping Problem," *Proc. IEEE-ICRA*, 2002.

[22] Y. Bar-Shalom, F. Daum, and J. Huang, "The Probabilistic Data Association Filter," *IEEE Control Systems Magazine*, 2009, pp. 82-100.

[23] G. Dissanayake, P. Newman, S. Clark, H. Durrant-Whyte, and M. Csorba. "A Solution to the Simultaneous Localization and Map Building (SLAM) Problem." IEEE Transactions on Robotics Automation. 17.3 (2001): 229-241

[24] J. Areta, Y. Bar-Shalom, and R. Rothrock, "Misassociation Probability in M2TA and T2TA," *J. of Advances in Information Fusion*, Vol. 2, No. 2, 2007, pp. 113-127.

[25] M. Joerger, M. Jamoom, M. Spenko, and B. Pervan, "Integrity of Laser-Based Feature Extraction and Data Association," *Proc. IEEE/ION PLANS 2016*, Savannah, GA, 2016, pp. 557-571.

[26] M. Joerger, and B. Pervan, "Continuity Risk of Feature Extraction for Laser-Based Navigation," *Proc. ION ITM 2017*, Monterey, California, January 2017, pp. 839-855.

[27] M Joerger, and B. Pervan, "Quantifying Safety of Laser-Based Navigation," submitted to *IEEE Transactions on Aerospace and Electronic Systems*, 2017.

[28] B. DeCleene, "Defining Pseudorange Integrity – Overbounding," Proc. of ION GPS 2000, Salt Lake City, UT., 2000, pp. 1916-1924.

[29] J. Rife, S. Pullen, P. Enge, and B. Pervan, "Paired Overbounding for Nonideal LAAS and WAAS Error Distributions." IEEE TAES, Vol. 42, No. 4, 2006 pp. 1386-1395.

[30] G. Duenas Arana, M. Joerger, and M. Spenko, "Minimizing Integrity Risk via Landmark Selection in Mobile Robot Localization," in preparation for *IEEE Transactions on Robotics*, 2017.

[31] Brenner, Mats, "lntegrated GPS/lnertial Fault Detection Availability," *Proceedings of the 8th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1995)*, Palm Springs, CA, September 1995, pp. 1949-1958.

[32] Blanch, J., Ene, A., Walter, T., and Enge, P., "An Optimized Multiple Hypothesis RAIM Algorithm for Vertical Guidance," *Proceedings of the 20th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2007)*, Fort Worth, TX, September 2007, pp. 2924-2933.

[33] Joerger, M., Chan, F.-C., and Pervan, B., "Solution Separation Versus Residual-Based RAIM", *NAVIGATION, Journal of The Institute of Navigation*, Vol. 61, No. 4, Winter 2014, pp. 273-291.

[34] M. Joerger, and B. Pervan, "Measurement-Level Integration of Carrier-Phase GPS and Laser-Scanner for Outdoor Ground Vehicle Navigation," ASME Journal of Dynamic Systems, Measurement, and Control, Vol. 131, 2009.

[35] M. Joerger, "Carrier Phase GPS Augmentation Using Laser Scanners and Using Low Earth Orbiting Satellites," Ph.D. Dissertation, Illinois Institute of Technology, 2009.

[36] C. Ye, and J. Borenstein, "Characterization of a 2-D Laser Scanner for Mobile Robot Obstacle Negotiation," Proc. IEEE-ICRA, May 2002.