Robust Error Estimation Based on Factor-Graph Models for Non-Line-of-Sight Localization

O. Arda Vanli (oavanli@eng.fsu.edu)
Florida State University https://orcid.org/0000-0002-0478-4004
Clark N. Taylor
Air Force Institute of Technology

Research

Keywords: Bayesian Estimation, Robust Estimation, Multipath, Factor Graphs.

Posted Date: September 27th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-917431/v1

Version of Record: A version of this preprint was published at EURASIP Journal on Advances in Signal Processing on January 25th, 2022. See the published version at https://doi.org/10.1186/s13634-022-00837-8.

Robust Error Estimation Based on Factor-Graph Models for Non-line-of-sight Localization

O. Arda Vanli
Department of Industrial and Manufacturing Engineering
Florida A&M University, Florida State University
Tallahassee, FL, USA
oavanli@eng.fsu.edu

Clark N. Taylor
Autonomy and Navigation Technology Center
Air Force Institute of Technology
WPAFB, OH, USA
clark.taylor@afit.edu

September 20, 2021

Abstract

This paper presents a method to estimate the covariances of the inputs in a factor-graph formulation for localization under non-line-of-sight conditions. A general solution, based on covariance estimation and M estimators in linear regression problems, is presented that is shown to give unbiased estimators of multiple variances and are robust against outliers. An iteratively re-weighted least squares (IRLS) algorithm is proposed to jointly compute the proposed variance estimators and the state estimates for the non-linear factor graph optimization. The efficacy of the method is illustrated in a simulation study using a robot localization problem under various process and measurement models and measurement outlier scenarios. A case study involving a Global Positioning System (GPS) based localization in an urban environment and data containing multipath problems demonstrates the application of the proposed technique.

Keywords: Bayesian Estimation, Robust Estimation, Multipath, Factor Graphs.

1 Introduction

An important challenge in Global Positioning System (GPS) - based localization is the non-line-of-sight or multipath problem, commonly encountered in urban environments. When the direct line of sight to a satellite is blocked by a building, its signal may reach the receiver

on the ground via reflections from buildings, resulting in large errors or outliers in the pseudorange measurements [1]. Least squares based localization methods that typically rely on Gaussian model errors can be severely distorted due to outliers, resulting in poor localization accurracy. Some approaches in the recent literature to remedy this issue include, robust estimators that relies on downweighting of outlying observations [2], mixture distributions to explicitly model the outlying observations in the sensor model [3], switchable constraints [4] and Receiver Autonomous Integrity Monitoring (RAIM) to monitor the integrity of satellites [5]. However, no studies focused on developing unbiased estimators of the noise variances.

Many data fusion applications require combining measurements from heteregenous sensors or matching features of multiple measurements, both tasks requiring an accurate characterization of the noise covariance matrices of measurements. In particular, GPS-based localization in urban or contested environments relies on fusion of measurements from inertial measurement unit (IMU) or light detection and ranging (LiDAR) systems in the navigation solution [6] to mitigate multipath problems. In vision-based localization, combining sequential measurements requires landmarks identified in these measurements to be matched, which assumes accurate knowledge of covariances [7],[8]. One of the common assumptions made in Bayesian state estimation of dynamic processes is that the covariance matrices for the noise sources are known or some reliable estimates exist *a priori*. Unfortunately, accurate *a priori* knowledge or estimates of the covariance matrices in practice may not be available and inaccurate covariance estimates can lead to a significant degradation in the estimation quality of the system states. In addition, as the applications of Bayesian estimation move from high-quality, expensive systems (e.g., the Apollo mission) to lower-cost systems with lower-quality sensors (e.g., cell phones), the ability to accurately characterize the uncertainty of all inputs to the system become both less repeatable across a class of sensors and less economically feasible.

In this paper, we propose new unbiased estimators of the noise covariances in a factor graph formulation when sensor data is contaminated with outliers and Gaussianity of data is not satisfied, particularly applicable to the GPS-based localization problem under non-line-of-sight conditions. Unbiased variance estimators for factor-graph problem proposed in [9] was extended to handle the multipath problem, study both linear and nonlinear vehicle motion models and demonstrate the performance and benefits in GPS-based localization problems under non-line-of-sight conditions. A general solution, based on non-linear regression and robust estimation is proposed that is shown to give unbiased estimators of the multiple variances in factor graph formulation. To jointly compute the proposed variance estimators and the state estimates an iteratively re-weighted least squares (IRLS) algorithm is presented. By contrast to the existing approaches, which use sample variances of the residuals to estimate noise variances, the primary contribution of the paper is to incorporate unbiased estimate of the noise variances, which, as will be illustrated in both simulation and case studies, can achieve significant improvement in localization accurracy over the existing approaches [1].

[1]The software developed for this study can be obtained at https://github.com/ardaVfamufsu/Factor-Graph-Covar

2 Review of Relevant Literature

Factor graphs have become a popular smoothing based alternative to sequential estimators of dynamic systems. Figure 1 illustrates the factor graph representation of a robot localization problem, in which states $x_1, ..., x_n$ are the sequence of n unknown robot locations. The factor graph provides a convenient way to represent the multiplicative relation of the "factors" (shown as boxes with different colors), or the probabilistic representations of states and measurements based on prior beliefs, process dynamics and measurement equations, and construct the optimization problem from which to solve for the robot locations.

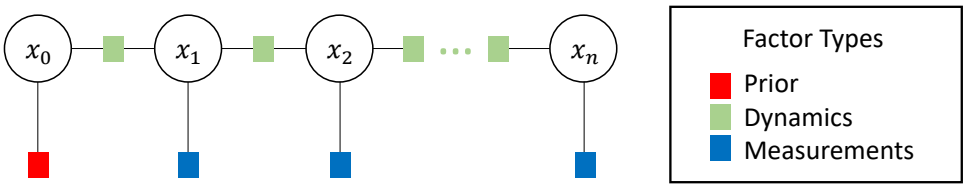


Figure 1: A simple factor graph expression of Equation (4).

In the simultaneous localization and mapping (SLAM) problem, where the factor graph formulation is commonly applied, it was shown that sequential state estimators may provide inconsistent solutions due to the inherent model nonlinearities and the accumulating effects of the linearization choices [10]. There has been a recent interest in smoothing as a fast alternative to sequential estimation, with their advantage of retaining the entire robot trajectory that significantly helps the performance [11]. Based off prior smoothing work [12, 13], the factor graph formulation of the estimation problem enables batch optimization of the system state estimates leading to better performance with non-linear systems. In addition, the factor graph formulation easily handles situations where the type, frequency, or quantity of measurements may change over time. In the SLAM problem, the subset of landmarks being tracked changes over time, meaning that the portions of the state that affect the measurement differ. This complex interaction of the state with the measurements is something easily handled in the factor graph formulation. Despite factor graph implementations being batch approaches, the sparsity of the matrix used to represent the factor graph leads to highly efficient implementations, including many that can be performed in real-time (e.g., [14, 15]).

To estimate the states of the factor graph formulation, several Bayesian estimation algorithms have been introduced, including the sequential approaches of (extended, unscented, ensemble) Kalman filter and particle filter, and smoothing based techniques [16]. For non-line-of-sight localization [17, 18] investigated weighted least squares and robust estimation methods to mitigate multipath effects. Most prior work on estimation methods, however, provided solutions assuming that noise covariances are known. Recently, methods based on the Expectation-Maximization (EM) framework have been proposed to estimate both the states and noise covariances in a factor graph framework [19, 20, 21, 22]. While the

objective of these papers was on making the factor graph or the estimation method robust to outliers, similar to our focus in the present paper, these approaches relied on sample covariances of the residuals (which maximizes the likelihood given the state estimates) as the direct estimators of the noise covariance matrices. As we will show in the paper these estimators are biased ([23, p. 103]) and may lead to a significant under-estimation of the covariance matrices, adversely impacting the localization accurracy. Many prior efforts also focused on estimating the *state covariance* for factor graphs (see e.g., [7]), however, these works all assume accurate input covariances are known a-priori. Our focus in this paper, by contrast, is on the estimation of *input noise* (measurement and process) covariance [24] for the factor graph.

3 Review of Factor Graph Formulation

In this section we first review the existing factor graph formulation of the localization, or state estimation, and the noise variance estimation problems. Consider a discrete-time dynamic system expressed as:

$$x_t = f(x_{t-1}, u_t) + v_t, \qquad v_t \sim N(0, Q) \tag{1}$$

$$z_t = g(x_t) + w_t, \qquad w_t \sim N(0, R) \tag{2}$$

where $x_t \in \mathbb{R}^{d_x}$ and $u_t \in \mathbb{R}^{d_u}$ are the state and control input vectors of the system at time t, respectively, $z_t \in \mathbb{R}^{d_z}$ is the sensor measurement vector, with d_x, d_u and d_z representing the dimensions of the corresponding vectors, $f(x, u)$ is the process dynamics model predicting what the state is at the next time step given the current state and the inputs, $g(x)$ is the measurement model that gives what the sensor measurement is at the current time given the current state, and v_t and w_t are white, zero-mean Gaussian noise sources with covariance matrices Q and R, respectively. For the GPS-based localization problem the states x_t may comprise the location coordinates of the vehicle and the measurements z_t may comprise vehicle odometry data or pseudorange measurements acquired from visible satellites. The measurements are conditionally independent. Localization problem consist of estimating the states, that is, solving for the sequence of vehicle locations $x_1, ..., x_n$ on the basis of a sequence of measurements $z_1, ..., z_n$ where n is the number of measurements available.

3.1 State Estimation

To find the maximum likelihood estimators of the system states the following optimization problem is solved [14, 15]:

$$X^* = \arg\max_X \; p(x_0) \prod_{t=1}^{n} p(x_t|x_{t-1}, u_t) \prod_{t=1}^{n} p(z_t|x_t), \tag{3}$$

where x_0 is the initial pose, and $X = \{x_t\}_{t=0}^{n}$ is the sequence of system states to be solved. This optimization problem can also be expressed as a factor graph as shown in Fig.1, where the factors (shown as squares) connect only with the state (or portion of the state) that are required for that measurement/dynamic equation to be evaluated.

Based on the Gaussian process dynamics and measurement models, the probability densities shown in the factorization are $p(x_t|x_{t-1}, u_t) \propto \exp -\frac{1}{2}\|f(x_{t-1}, u_t) - x_t\|^2_Q$ and $p(z_t|x_t) \propto \exp -\frac{1}{2}\|g(x_t) - z_t\|^2_R$, where the sign $\propto$ means "proportional to" and implies the density functions are given up to a proportionality constant. Note that because Q and R are assumed constant across time, the proportionality constants can be safely ignored. Because $\log(x)$ is a strictly monotonic function in x, we can rewrite Equation (3) as:

$$\begin{aligned}\arg\min_X \|x_0 - \mu_0\|^2_{\Sigma_0} + \sum_{t=1}^{n} \|f(x_{t-1}, u_t) - x_t\|^2_Q \\ + \sum_{t=1}^{n} \|g(x_t) - z_t\|^2_R \end{aligned} \tag{4}$$

assuming $p(x_0) \sim N(\mu_0, \Sigma_0)$, where μ_0 and Σ_0 are the mean and covariance of the initial pose.

Since the model equations can be nonlinear in general, the states are solved for by linearizing the process and measurement equations using a first order Taylor series expansion around an operating point $\{x^0_t\}^n_{t=0}$ as

$$\begin{aligned} f(x_{t-1}, u_t) - x_t &\cong f(x^0_{t-1}, u_t) \\ &+ F(x^0_{t-1})\Delta x_{t-1} - x^0_t - \Delta x_t \qquad (5) \\ g(x_t) - z_t &\cong g(x^0_t) + G(x^0_t)\Delta x_t - z_t \qquad (6) \end{aligned}$$

where $F(x_{t-1}) = \partial f(x_{t-1}, u_t)/\partial x_{t-1}$ and $G(x_t) = \partial g(x_t)/\partial x_t$ are the Jacobian matrices of the process and measurement equations, respectively, and $\Delta x_t = x_t - x^0_t$, $\Delta x_{t-1} = x_{t-1} - x^0_{t-1}$. Defining the vectors $a^0_t \triangleq x^0_t - f(x^0_{t-1}, u_t)$ and $c^0_t \triangleq z_t - g(x^0_t)$ and letting $\Delta x_0 = x_0 - x^0_0$, a linear least squares representation of the problem in (4) is obtained as:

$$\begin{aligned}\arg\min_{\Delta X} \|\Delta x_0 + x^0_0 - \mu_0\|^2_{\Sigma_0} \\ + \sum_{t=1}^{n} \|F_{t-1}\Delta x_{t-1} - a^0_t - \Delta x_t\|^2_Q \\ + \sum_{t=1}^{n} \|G_t\Delta x_t - c^0_t\|^2_R. \end{aligned} \tag{7}$$

where $\Delta X = \{\Delta x_t\}^n_{t=0}, F_{t-1} \triangleq F(x^0_{t-1})$, and $G_t \triangleq G(x^0_t)$. Each iteration of the state estimation problem can be written as that of finding the solution of the following system of linear equations:

$$b = A\Delta X + \epsilon \tag{8}$$

where $A \in \mathbb{R}^{m\times p}$ is the coefficient matrix, $b \in \mathbb{R}^m$ is the standardized measurement vector, $\epsilon \in \mathbb{R}^m$ is the vector of least squares solution errors, $m = (n+1)d_x + nd_z$ is the number of observations and $p = (n+1)d_x$ is the number of states. If the initial pose x_0 is known and not to be included in the estimation problem then $X = \{x_t\}^n_{t=1}$ are the states to be solved for, $m = nd_x + nd_z$ and $p = nd_x$. In the iterative scheme, ΔX is solved for using the Moore-

Penrose pseudo-inverse (typically $(A^TA)^{-1}A^T$), the state estimates for the next iteration is found as $X \leftarrow X + \Delta X$, and the new matrices A and b are found at the linearization point of X. These iterations are continued until convergence.

As an example, the factor-graph model given in Figure 1 for data set with $n = 4$ poses is written as

$$A = \begin{bmatrix} -\Sigma_0^{-\frac{1}{2}} & 0 & 0 & 0 & 0 \\ Q^{-\frac{1}{2}}F_1 & -Q^{-\frac{1}{2}} & 0 & 0 & 0 \\ 0 & Q^{-\frac{1}{2}}F_2 & -Q^{-\frac{1}{2}} & 0 & 0 \\ 0 & 0 & Q^{-\frac{1}{2}}F_3 & -Q^{-\frac{1}{2}} & 0 \\ 0 & 0 & 0 & Q^{-\frac{1}{2}}F_4 & -Q^{-\frac{1}{2}} \\ 0 & R^{-\frac{1}{2}}G_1 & 0 & 0 & 0 \\ 0 & 0 & R^{-\frac{1}{2}}G_2 & 0 & 0 \\ 0 & 0 & 0 & R^{-\frac{1}{2}}G_3 & 0 \\ 0 & 0 & 0 & 0 & R^{-\frac{1}{2}}G_4 \end{bmatrix}, \quad b = \begin{bmatrix} \Sigma_0^{-\frac{1}{2}}(x_0 - \mu_0) \\ Q^{-\frac{1}{2}}a_1 \\ Q^{-\frac{1}{2}}a_2 \\ Q^{-\frac{1}{2}}a_3 \\ Q^{-\frac{1}{2}}a_4 \\ R^{-\frac{1}{2}}c_1 \\ R^{-\frac{1}{2}}c_2 \\ R^{-\frac{1}{2}}c_3 \\ R^{-\frac{1}{2}}c_4 \end{bmatrix}. \tag{9}$$

3.2 Noise Variance Estimation

In the existing smoothing based factor graph methods [19, 20, 21, 22] the noise covariance matrices Σ_0, Q and R are estimated iteratively in conjunction with the state estimates, following an expectation and maximization (EM) approach. The maximization step in EM maximizes the likelihood, which minimizes the negative log likelihood (4), given current state estimates, to find the variances. This Maximum Likelihood (ML) estimator of variances, estimates the variances as the sample variance of state estimation errors $e_t = x_t - \hat{x}_t$ [25, p. 79], based on the best state estimates $\hat{x}_t$ and the true value of the state vectors x_t, as

$$\sigma^2 = \frac{1}{nd_x} \sum_{t=1}^{n} e_t^T e_t. \tag{10}$$

A separate maximum likelihood estimate with the residuals of the initial pose, process dynamics and measurements are calculated, respectively, and these estimates form the constant diagonal elements of the covariance matrices Σ_0, Q and R. To make the estimates robust against outliers, Sunderhauf et al., [1] discussed an extension to the EM method using the concept of switching constraints (SC), however, this approach is based on maximum likelihood to estimate the noise variances.

A drawback of the maximum likelihood (ML) estimators, such as (10), that are based on the measured estimation error sequences, in the context of localization is that they are asymptotically (as the data size grows) unbiased estimates of the true variances [24]. However, as we show in our derivations and the numerical results below, when data size is relatively small compared to the size of the state vector, a common situtaion in factor graph models where the number of states grow with data size, the ML estimators are biased [25] and loose their effectiveness. In the present paper we develop unbiased estimators of the multiple noise variances of the factor graph representation of dynamic systems, with a particular focus on GPS-based localization under non-line of sight conditions, and study the benefits

of these estimators in localization. In the simulation and case studies, we will compare the performance of the proposed unbiased covariance estimators, given by Equation (22) to the existing ML based and SC based existing methods. Both the proposed unbiased and the existing ML variance estimatiors will be implemented in an IRLS iterative scheme (explained in the next section) which is a form of EM approach.

4 Methods

In this section we present the proposed unbiased estimators for the noise variances of the process and measurement equations of the factor graph problem and an iteratively reweighted least squares algorithm that incorporates M estimators for jointly finding the variance and robust state estimates.

4.1 Unbiased noise variance estimation in factor graph problem

In a linear regression problem, a set of parameters Θ for the model $y = B\Theta + \epsilon$ are estimated, where $y \in \mathbb{R}^m$ is the measurement vector containing m observations, $B \in \mathbb{R}^{m \times p}$ is the regressor matrix for p regressor variables and $\epsilon \in \mathbb{R}^m$ is the vector of Gaussian distributed errors with 0 mean and a common variance of σ^2. It is clear that, the linearized factor graph formulation (8) is a linear regression problem, where b, A and ΔX correspond to y, B and Θ, respectively.

In linear regression problems, the variance of the measurements y is estimated based on the residuals of the fitted regression model, expressed as [25]

$$r = y - B\hat{\Theta} = (I - \tilde{H})y \tag{11}$$

where $\hat{\Theta}$ is the estimated parameter vector and $\tilde{H} = B(B^T B)^{-1} B^T$ is an orthogonal projection matrix and the superscript T denotes the matrix transpose. From this equation the sum of squared residuals $r^T r$ is

$$r^T r = y^T (I - \tilde{H}) y. \tag{12}$$

The expectation of the random quantity $r^T r$ with respect to the Gaussian probability distribution of the measurements can be found as [25, pp. 554-555]

$$E[r^T r] = \sigma^2 (m - p)$$

where $p = rank(\tilde{H})$. Following the "method of moments approach" for estimation [26], an unbiased estimate of the variance is [25, p. 77]

$$\sigma^2 = \frac{r^T r}{m - p}.$$

There are two assumptions in linear regression work, however, that differ from the factor graph formulation. First, linear regression assumes that a there is a single covariance for all inputs. The factor graph formulation, however, has at least two different covariances (R and Q), and often many more. Second, the residuals computed in a factor graph are

pre-weighted by the current covariance estimates. In our proposed method, we extend the linear regression based method for finding variances to the factor graph formulation of the system state estimation problem by incorporating these two important aspects.

Assume a reordering of the factor graph linear system (8) and divide the equations in to s partitions as

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_s \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_s \end{bmatrix} \Delta X + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_s \end{bmatrix} \tag{13}$$

where $b_1, ..., b_s$ correspond to the standardized measurements of the partitions containing $n_1, ..., n_s$ elements such that $\sum_{i=1}^{s} n_i = m$ and $\epsilon_1, ..., \epsilon_s$ are mutually un-correlated random errors of the partitions with 0 mean and unit variances. Each partition has its own covariance matrix before standardization. For example, the system described in Equations (9) can be split into $s = 3$ partitions corresponding to unstandardized measurement noise covariances Σ_0, Q and R, where σ_1^2, σ_1^2 and σ_3^2 are the constant diagonal elements of the respective 3 matrices.

The proposed iterative method will provide unbiased estimators of the (unstandardized) measurement variances $\sigma_1^2, ..., \sigma_s^2$. The standardized measurement vector b is a random variable with an identity covariance matrix, that is, $\Sigma_b = diag(I_{n_1}, ..., I_{n_s})$, where I_{n_i} is an n_i dimensional identity matrix. The unit variances of the standardized partitions are satisfied if the estimated noise variances $\sigma_1^2, ..., \sigma_s^2$ are correct. If the noise variances are incorrect, then the standardized measurements should have the covariance matrix

$$\Sigma_b = diag(k_1 I_{n_1}, ..., k_s I_{n_s}). \tag{14}$$

Therefore, if $k_i, i = 1, ..., s$ converge to 1 then the standardized noise covariances converge to unity and the unstandardized noise variances converge to the true values. In our approach, to accomplish convergence, the variance for the next iteration is found by scaling the variance in the current iteration with k_i as

$$\sigma_i^2 \leftarrow k_i \sigma_i^2 \text{ for } i = 1, ..., s. \tag{15}$$

To do this scaling, we must estimate k_i in each iteration. Let us define $H = I - \tilde{H}$ where $H \in \mathbb{R}^{m \times m}$ and $\tilde{H} = A(A^T A)^{-1} A^T$ is the projection matrix found by the coefficient matrix A obtained with the standardized measurements and the covariance estimates in the current iteration. Partition H as

$$H = \begin{bmatrix} H_{11} & H_{12} & \dots & H_{1s} \\ H_{21} & H_{22} & \dots & H_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ H_{s1} & H_{s2} & \dots & H_{ss} \end{bmatrix} = \begin{bmatrix} H_1 \\ H_2 \\ \vdots \\ H_s \end{bmatrix}. \tag{16}$$

where $H_{ij} \in \mathbb{R}^{n_i \times n_j}$ and $i, j = 1, 2, ..., s$. Note that because H is symmetric, $H_{ji} = H_{ij}^T$. The

residual vector for the i-th partition, $r_i \in \mathbb{R}^{n_i \times 1}$, is obtained as

$$\begin{aligned} r_i &= H_i b \\ &= \begin{bmatrix} H_{i1} & H_{i2} & \dots & H_{is} \end{bmatrix} b \end{aligned}$$

leading to the sum of squared residuals of the i-th partition:

$$r_i^T r_i = b^T H_i^T H_i b. \tag{17}$$

Because this is a scalar, we can take the "trace" of the right-hand side, leading to:

$$\begin{aligned} r_i^T r_i &= trace(b^T H_i^T H_i b) \\ &= trace(b b^T H_i^T H_i). \end{aligned}$$

Applying expectation operation, $E[.]$, we obtain:

$$\begin{aligned} E[r_i^T r_i] &= trace(E[bb^T] H_i^T H_i) && (18) \\ &= trace(\Sigma_b H_i^T H_i) && (19) \\ &= \sum_{j=1}^{s} trace(H_{ji} H_{ij}) k_j && (20) \\ &= \sum_{j=1}^{s} T_{ij} k_j. && (21) \end{aligned}$$

Equation (20) is obtained by inserting Equation (14) for Σ_b in Equation (19) and Equation (21) is obtained by defining $T_{ij} = trace(H_{ji} H_{ij})$. Equation (21) is what we needed in order to develop unbiased estimators of the scaling factors $k_1, ..., k_s$ based on the method of moments approach. The scaling factors can be easily solved by writing out a matrix equation setting equal the observed values $r_i^T r_i$ to their *e*xpected values, i.e., the right hand side of (21), for all partitions $i = 1, 2, ..., s$ as

$$\begin{bmatrix} r_1^T r_1 \\ r_2^T r_2 \\ \vdots \\ r_s^T r_s \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1s} \\ T_{21} & T_{22} & \cdots & T_{2s} \\ \vdots & & \ddots & \vdots \\ T_{s1} & T_{s2} & \cdots & T_{ss} \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_s \end{bmatrix}. \tag{22}$$

The solution $k_1, ..., k_s$ of the system (22) is a set of unbiased variance estimators of the standardized partitions $b_1, ..., b_s$. The unbiased variance estimators for the unstandardized partitions $\sigma_1^2, ..., \sigma_s^2$ are obtained by using Equation (15) given the state estimates ("maximization" step). The states are then estimated by solving (7) given the variance estimates ("expectation" step). The next subsection discusses the state estimation ("expectation") procedure. The following subsection discusses an iteratively reweighted least squares (IRLS) procedure that implements the expectation and maximization steps to estimate the noise variances and the states.

4.2 Robust state estimation

To make the state estimates robust to outliers that arise in non-line-of sight conditions, we present an approach to incorporate the proposed factor graph unbiased variance estimation method with M estimators. M estimators [25] is a class of popular robust estimators in linear regression models that make use of a weight matrix W that assigns a weight of 1 to inlying measurements that are consistent with the assumed distribution and smaller weights to outlying observations. In our factor graph formulation, we implement the M estimator by incorporating a weight matrix W within the coefficient matrix A and the measurement vector b of the system of equations (8) as follows:

$$A \triangleq W^{1/2}A, \text{ and } b \triangleq W^{1/2}b \tag{23}$$

in which the weight matrix is defined as

$$W = diag(\tau_1, ..., \tau_m) \tag{24}$$

where τ_i is the weight found as a function of the i-th scaled measurement defined as $z_i = b_i/\gamma$ and using the definitions given in Table 1 (for some of the popular M estimators), where γ is a nonparametric estimate of the standard deviation, defined as:

$$\gamma = \underset{1\leq i\leq n}{\text{median}}|b_i - \underset{1\leq i\leq n}{\text{median}}(b_i)|/0.6745 \tag{25}$$

where the constant 0.6745 is chosen so that γ is an unbiased estimator based on Gaussian error distribution [25, pp. 373-374].

To define the weights, various functions are used in the literature. For illustration purposes, we consider two of the more popular ones, Huber's t and Cauchy functions weight and loss functions, which are plotted in Figure 2 and defined in Table 1, in comparison to the ordinary least squares (OLS) estimator (3), which is a quadratic function of the standardized residuals and does not provide any robustness. For a given loss function, the constant (i.e., a_C or a_H) to be specified by the user, determines the strength of downweighting applied to outliers. For the factor graph formulation, the weights are only applied to measurement residuals and weights for process residuals are unity, that is: τ_i for $i = nd_x+1, ..., m$ is found from Table 1 and $\tau_i = 1$ for $i = 1, ..., nd_x$.

Table 1: Huber's t, Cauchy and OLS loss and weight functions

Method	Huber's t	Cauchy	OLS
Weight	$\tau_i = \begin{cases} 1.0 & \text{if } \lvert z_i\rvert \leq a_H \\ a_H/\lvert z_i\rvert & \text{if } \lvert z_i\rvert > a_H \end{cases}$	$\tau_i = 1/(1 + (z_i/a_C)^2)$	$\tau_i = 1$
Loss fnc	$\rho_i = \begin{cases} 1/2z_i^2 & \text{if } \lvert z_i\rvert \leq a_H \\ \lvert z_i\rvert a_H - 1/2a_H^2 & \text{if } \lvert z_i\rvert > a_H \end{cases}$	$(a_C^2/2)\log(1 + (z_i/a_C)^2)$	$1/2z_i^2$
Parameter	a_H	a_C	NA

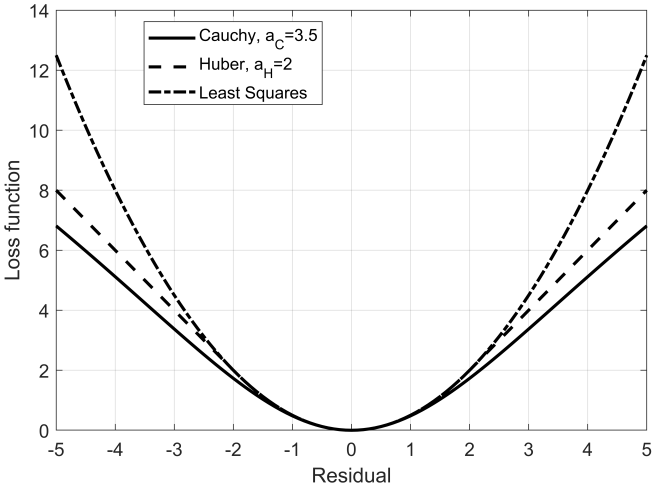


Figure 2: Huber's t, Cauchy and OLS loss functions

4.3 Iteratively reweighted least squares

In implementing the robust estimators in regression [27],[28] an Iteratively Reweighted Least Squares (IRLS) algorithm is often used to solve iteratively, for the best state estimate for a given set of noise variance values and the best variance values for a given state solution until convergence. These steps are analogous to the expectation and maximization steps of the methods in [19, 20, 21, 22] where the residuals are used to compute a covariance (the M-step), and the estimated covariances are used to compute new residuals (the E-step). The proposed IRLS algorithm to implement the proposed unbiased variance estimators of the factor graph formulation and the M estimators for robust state estimation following this iterative framework is described in Algorithm 1.

Note that, the measurement vector b and matrix A used in the algorithm already incorporates the M estimator weight matrix W, as shown in Equation (23). The iterative approach solves for the best state estimate in the factor graph framework using the current unbiased variance estimates $\sigma_1^2, \sigma_2^2, ..., \sigma_s^2$ and the scaling factors $k_1, k_2, ..., k_s$ of the partitions. This state estimate is then used to scale the variance estimates for the next iteration and the process is repeated until convergence.

5 Experiments

The efficacy of the proposed method has been illustrated using a simulation experiment involving a mobile robot and a case study experiment for localization based on real GPS data.

5.1 Simulation Experiment

The mobile robot simulation was based on a two-dimensional constant velocity model as shown in Figure 3. Figures 3a and 3b show simulations without and with outliers, respec-

Algorithm 1: IRLS algorithm for unbiased estimation of noise variances and robust estimation of states

Result: X, σ_1^2, σ_2^2, ... ,σ_s^2
Input: Z, U, Σ_0, μ_0, a_C (or a_H)
Initialization: X, σ_1^2, σ_2^2, ... ,σ_s^2
repeat
  Least Squares (E) step:
    repeat
      Based on $\sigma_1^2, \sigma_2^2, ..., \sigma_s^2$ compute W and $A(X, W, \sigma_i^2)$ and $b(X, W, \sigma_i^2)$;
      Solve Equation (8) for ΔX;
      $X \leftarrow X + \Delta X$
    until X *converges*;
  Reweighting (M) step:
    Use last $A(X, W, \sigma_i^2)$ to compute H;
    Solve Equation (22) for multiplicative scale factors $\{k_1, k_2 ... k_s\}$;
    for $i = 1...s$ **do**
      Update variance estimates using $\sigma_i^2 \leftarrow k_i \sigma_i^2$ (Equation (15))
    end
until X, σ_1^2, σ_2^2, ... ,σ_s^2 *converge*;

tively.

Two different motion models are considered. *(i)* A linear motion model where the state vector $x_t = (x_t, \dot{x}_t, y_t, \dot{y}_t)^T$ gives the position and velocity of the robot, the measurement vector $z_t = (x_t, y_t)$ gives the position of the robot in two dimensions and the control vector $u_t = (u_{1t}, u_{2t})^T$ affects the velocities $\dot{x}_t$ and $\dot{y}_t$. *(ii)* A nonlinear motion model, with state vector $x_t = (x_t, y_t, \theta_t, \lambda_t)^T$ where the θ_t is rotational velocity and λ_t is forward velocity; measurement vector $z_t = (x_t, y_t)$ is the position of the robot, and control vector $u_t = (u_{1t}, u_{2t})^T$ affects the velocities θ_t and λ_t. The number of timesteps for the robot motion simulation is n, from each, a measurement is obtained and used in the localization and noise variance estimation. Figure 3 shows simulations with $n = 20$ timesteps. In the analyses the cases of $n = 20$ and $n = 100$ timesteps are considered to study the influence of larger data sets on the quality of the noise variance estimates. The performance of robust estimators is studied by considering scenarios involving outliers in the measurements of the linear robot motion model.

Table 2 shows the definitions of the individual terms of the process dynamics equation (1) and measurement equation (2) of the linear and nonlinear motion models. The sampling period was taken as $T = 1$ and the initial pose was $x_0 = (0, 2, 0, 0)^T$. According to these models, variances correspond to process dynamics and measurements as follows. For linear motion σ_{Q1}^2 is for the dynamics of $(x_t, \dot{x}_t)$; σ_{Q2}^2 is for the dynamics of $(y_t, \dot{y}_t)$ and σ_R^2 is for the measurements of (x_t, y_t). For nonlinear motion σ_{Q1}^2 is for the dynamics of (x_t, y_t); σ_{Q2}^2 is for the dynamics of (θ_t, λ_t) and σ_R^2 is for the measurements of (x_t, y_t).

The proposed variance estimation method was applied by performing the partitioning shown in Equation (13) on the factor graph formulation of both motion models with $s = 3$ partitions. The partitioned matrices A_1, A_2, A_3 are formed such that they corresponding

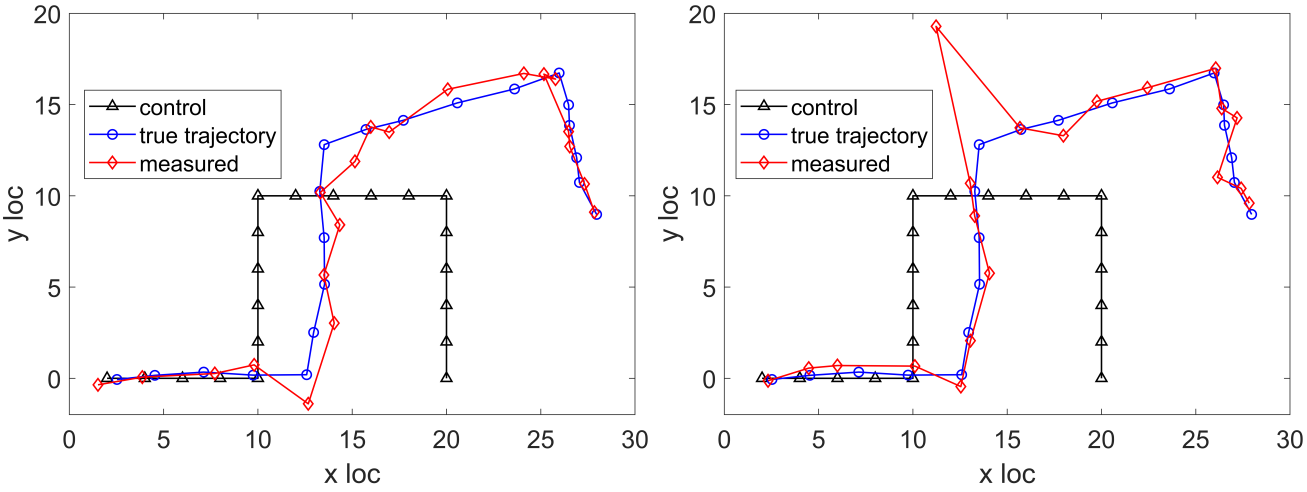


Figure 3: Mobile robot simulation for $n = 20$ poses showing control ($v_t = 0, w_t = 0$), true trajectories ($w_t = 0$) and measured trajectories (a) with no outliers (b) with outliers

to the rows of A containing $\sigma_{Q1}^2, \sigma_{Q2}^2$ and σ_R^2, respectively, and the partitions and b_1, b_2, b_3 are similarly obtained from the rows of b . See Table 2 for the definitions of Q and R and Equation (9) for how A and b are defined for an example with $n = 4$. In the simulation examples, the initial pose x_0 is assumed to be known and therefore, it treated as a constant in the factorization (7) and x_0 and Σ_0 are not estimated.

The control input for the linear motion case was defined as:

$$u_{1t} = \begin{cases} -2 & t = 5, 15, 25, ... \\ +2 & t = 10, 20, 30, ... \\ 0 & \text{otherwise} \end{cases} \tag{26}$$

$$u_{2t} = \begin{cases} +2 & t = 5, 20, 25, 40, 45, ... \\ -2 & t = 10, 15, 30, 35, 50, 55, ... \\ 0 & \text{otherwise} \end{cases} \tag{27}$$

which has the effect of (without process and measurement noises, i.e., $v_t = 0, w_t = 0$) making the robot move forward 10 units, then up 10 units, forward 10 units, then back down 10 units and as this pattern repeats, a "square wave" would be created. Figure 3a shows an example of running this system for 20 timesteps. The control input for the nonlinear motion case was defined as:

$$u_{1t} = \begin{cases} \pi/2 & t = 5, 15, 25, ... \\ -\pi/2 & t = 10, 20, 30, ... \\ 0 & \text{otherwise} \end{cases} \tag{28}$$

$$u_{2t} = 0 \text{ for all } t. \tag{29}$$

Similar to the linear motion model, this control input will make the robot to move forward 10 units, then up 10 units, forward 10 units, then back down 10 units. The control trajectory of the nonlinear motion is identical to that of linear motion shown in Figure 3a. The motion and measurement simulations were generated by using $\sigma_{Q1}^2 = 0.5, \sigma_{Q2}^2 = 0.2$ and $\sigma_R^2 = 1.5$ as the true variances of the models.

In order to more realistically represent the non-line-of-sight conditions, we used a two component mixture process for the measurements of the linear motion model in Table 2

$$w_t \sim (1-\alpha)N(0, \sigma_R^2 I_2) + \alpha N(0, 10^2) \tag{30}$$

where α is the mixture proportion and the second mixture component with variance 10^2 represents outliers with large variance (compared to the variance $\sigma_R^2 = 1.5^2$ of the inlying measurements). The mixture model injects unusually large or small deviations in randomly selected measurements, mimicking the effect of reflections of satellite signal from buildings. Figure 3b shows one simulation with the mixture parameter $\alpha = 0.1$ and $n = 20$ poses that resulted only one outlier.

5.2 Case Study Experiment: GPS based localization

In this section, we illustrate the application of the proposed method in a GPS-based localization under non-line-of-sight conditions using the real data set presented in Pfiefer and Protzel [19]. The data set was collected in the city center of Chemnitz, Germany, by driving an instrumented vehicle several times over a road network, as shown in Figure 4. The urban setting and the road layout consists of several tall buildings that cause a large number of outliers due to satellite blockages or reflections from the buildings. The data set contains high-precision inertial measurements of the vehicle position with a precision of 2 cm used as ground truth and GPS pseudorange measurements that will be used in the estimation procedure. The data set contained pseudorange measurements at 8,570 timesteps. At each timestep, between 7 and 12 satellites were visible to the vehicle and provided pseudorange measurements.

The proposed unbiased variance estimation method with M estimators was applied to the data set to find the location of the vehicle while making the solution robust to outliers. As an existing approach to compare our results, we consider the results of Sunderhauf et al. [1], who used the same data set and followed a "switching constraints" (SC) approach as an attempt to make their localization less sensitive to outliers, and the Maximum Likelihood (ML) variance estimators.

To develop the factor graph model we define the state vector as $x_t = (x_t, y_t, z_t, b_t, d_t)$ where b_t is the satellite clock error and $d_t = \dot{b}_t$ is clock error drift. The proces dynamics model, incorporating only the satellite clock error and drift, is $(b_t, d_t)^T = f(b_{t-1}, d_{t-1}) + v_t$ where

$$f(b_{t-1}, d_{t-1}) = \begin{bmatrix} b_{t-1} + T d_{t-1} \\ d_t \end{bmatrix} \tag{31}$$

where the sampling period in the data set was $T = 0.25$ seconds and $v_t \sim N(0, Q)$ is the

Table 2: Model terms of the mobile robot examples

Linear motion	Nonlinear motion
$x_t = f(x_{t-1}, u_t) + v_t$ $= Fx_{t-1} + Bu_t + v_t$ $v_t \sim N(0, Q)$	$x_t = f(x_{t-1}, u_t) + v_t$ $= x_{t-1} + \begin{bmatrix} \lambda_{t-1} T \cos\theta_{t-1} \\ \lambda_{t-1} T \sin\theta_{t-1} \\ 0 \\ 0 \end{bmatrix} + Bu_t + v_t$ $v_t \sim N(0, Q)$
$x_t = (x_t, \dot{x}_t, y_t, \dot{y}_t)$	$x_t = (x_t, y_t, \theta_t, \lambda_t)$
$z_t = g(x_t) + w_t$ $w_t \sim N(0, R)$	$z_t = g(x_t) + w_t$ $w_t \sim N(0, R)$
$F = \partial f / \partial x_{t-1}$ $= \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$F_{t-1} = \partial f / \partial x_{t-1}$ $= \begin{bmatrix} 1 & 0 & -\lambda_{t-1} T \sin\theta_{t-1} & T\cos\theta_{t-1} \\ 0 & 1 & \lambda_{t-1} T \cos\theta_{t-1} & T \sin\theta_{t-1} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
$B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$	$B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$
$G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$	$G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$
$Q = \begin{bmatrix} \sigma_{Q1}^2 & 0 & 0 & 0 \\ 0 & \sigma_{Q1}^2 & 0 & 0 \\ 0 & 0 & \sigma_{Q2}^2 & 0 \\ 0 & 0 & 0 & \sigma_{Q2}^2 \end{bmatrix}$, $R = \sigma_R^2 I_2$	$Q = \begin{bmatrix} \sigma_{Q1}^2 & 0 & 0 & 0 \\ 0 & \sigma_{Q1}^2 & 0 & 0 \\ 0 & 0 & \sigma_{Q2}^2 & 0 \\ 0 & 0 & 0 & \sigma_{Q2}^2 \end{bmatrix}$, $R = \sigma_R^2 I_2$

process error, with covariance matrix

$$Q = \begin{bmatrix} \sigma_b^2 & 0 \\ 0 & \sigma_d^2 \end{bmatrix} \tag{32}$$

where σ_b^2 and σ_d^2 are the error variances of clock offset and clock drift. The Jacobian matrix $F = \partial f / \partial x_{t-1}$ of the process equation is

$$F = \begin{bmatrix} 0 & 0 & 0 & 1 & T \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

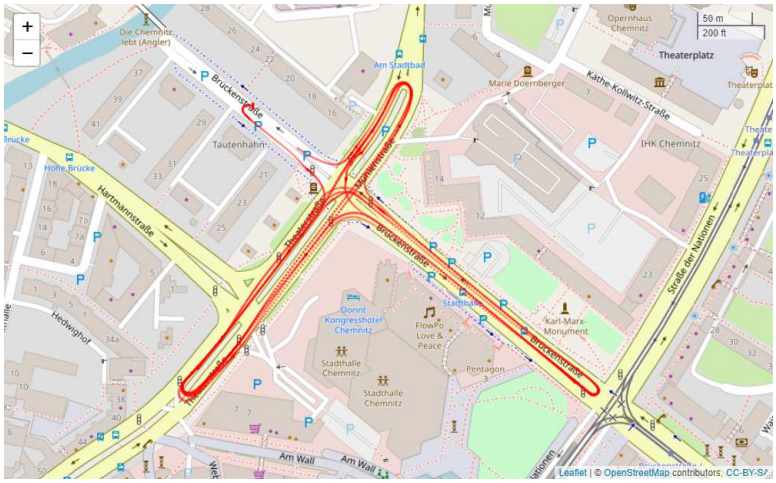


Figure 4: Chemnitz City urban environment and ground truth

The measurement equation consists of the pseudorange from the vehicle to the i-th satellite at time t

$$z_i(x_t) = g_i(x_t) + w_{it}, \quad i = 1, 2, ..., ns_t \tag{33}$$

where ns_t is the number of satellites visible at time t, w_{it} is the measurement error of the i-th satellite. The measurement errors of all satellites are Gaussian distributed $w_t \sim N(0, R)$ with covariance matrix is $R = \sigma_R^2 I_{ns_t}$. The noiseless measurement equation is (omitting the time index t in the state variables and satellite coordinates):

$$\begin{aligned} g_i(x) &= \sqrt{(S_{xi} - x)^2 + (S_{yi} - y)^2 + (S_{zi} - z)^2} \\ &+ c \times b + (\gamma/c)(S_{xi}y - S_{yi}x) \end{aligned} \tag{34}$$

where (S_{xi}, S_{yi}, S_{zi}) are the coordinates of the i-th satellite, $c = 3 \times 10^8 m/s$ is the speed of light, and $\gamma = 7.3 \times 10^{-5} rad/s$ is the earth's rotation speed. The measurement equation for the complete set of satellites is

$$z_t = g(x_t) + w_t \tag{35}$$

where $g = (g_1, ..., g_{ns_t})^T$. Sunderhauf et al. [1] use $\sigma_R^2 = 10^2$. The Jacobian matrix $G_t = \partial g / \partial x_t$ of the measurement equations is

$$G_t = \begin{bmatrix} -(S_{x1,t} - x_t)/\tilde{g}_1 - (\gamma/c)S_{y1,t} & -(S_{y1,t} - y_t)/\tilde{g}_1 + (\gamma/c)S_{x1,t} & -(S_{z1,t} - z_t)/\tilde{g}_1 & c & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -(S_{x_{nst},t} - x_t)/\tilde{g}_{ns_t} - (\gamma/c)S_{y_{nst},t} & -(S_{y_{ns_t},t} - y_t)/\tilde{g}_{ns_t} + (\gamma/c)S_{x_{ns_t},t} & -(S_{z_{ns_t},t} - z_t)/\tilde{g}_{ns_t} & c & 0 \end{bmatrix}$$

where $(S_{xi,t}, S_{yi,t}, S_{zi,t})$ are the coordinates of the i-th satellite at time t and $\tilde{g}_i = \sqrt{(S_{xi,t} - x_t)^2 + (S_{yi,t} - y_t)^2 + (S_{zi,t} - z_t)^2}$.

6 Results and Discussion

6.1 Mobile Robot Experiments

Using the example mobile robot systems, we evaluate the performance of the factor graph noise variance and localization estimates. 1,000 Monte Carlo replications of the robot simulation were generated for $n = 20$ timesteps and in each replication the location sequence of the robot and the noise variances are estimated. The proposed unbiased noise variance estimator (specified by Equations (22) and (15)) is compared to the to the existing maximum likelihood (ML) estimator (specified by Equation (10)). Both estimators are implemented within the IRLS scheme given in Algorithm 1.

We compared the performance of the estimators based on (i) bias in the variance estimates and (ii) Mahalanobis distance of the location solution. Bias of the estimates (of all 3 variances) measures the difference between the computed variances and the true variances. Since we have a variance estimate in each simulation, the average bias from N simulations is calculated as

$$C = 1/(3N) \sum_{i=1}^{3} \sum_{j=1}^{N} (\sigma_{ij}^2 - \sigma_{Ti}^2)^2 \tag{36}$$

where σ_{Ti}^2 is the true variance, σ_{ij}^2 is the estimated variance of ith component in jth simulation, $i = 1, 2, 3$ is the index of the variance component, $j = 1, ..., N$ is the index of the simulation run and $N = 1,000$.

Total squared Mahalonabis distance of the estimated locations from the truth over all timesteps is used to quantify the localization error. The total Mahalanobis localization error is found (in each simulation) as the sum of the errors over n poses as

$$G = \sum_{t=1}^{n} (x_t - \hat{x}_t)^T P_t^{-1} (x_t - \hat{x}_t) \quad , t = 1, 2, ..., n \tag{37}$$

in which x_t is the state and $\hat{x}_t$ is the state estimate of t-th pose, $P_t \in \mathbb{R}^{d_x \times d_x}$ is the covariance matrix of the state estimate $\hat{x}_t$ that can be obtained from the corresponding entries of the matrix $(A^T A)^{-1}$, which is an $m \times m$ matrix (See Section 3 for definition of m and d_x). If the predicted uncertainty on the state estimate is correct, then G should follow a chi-squared distribution with degree of freedom nd_x, with an expected value nd_x [23, 29].

The estimated poses and the 95% confidence ellipses using the two variance estimation approaches superimposed with the true poses from a single simulation of the linear motion problem is shown in Figure 5. The localization with the proposed unbiased variance estimates has the true location within its confidence interval at every timestep, while the existing ML method for estimating variances frequently has the true location either outside or very close to the confidence ellipse. This demonstrates some of the practical concerns of underestimation with the existing covariance estimation techniques.

Graphically, the distributions of the variance estimates $\sigma_R^2, \sigma_{Q1}^2$ and σ_{Q2}^2 obtained across the Monte Carlo simulations are shown in Figure 6 (for linear and nonlinear motion models) and the distributions of the Mahalanobis distances (for linear motion model) are shown in Figure 7. The distributions of the variance estimates obtained using the proposed method

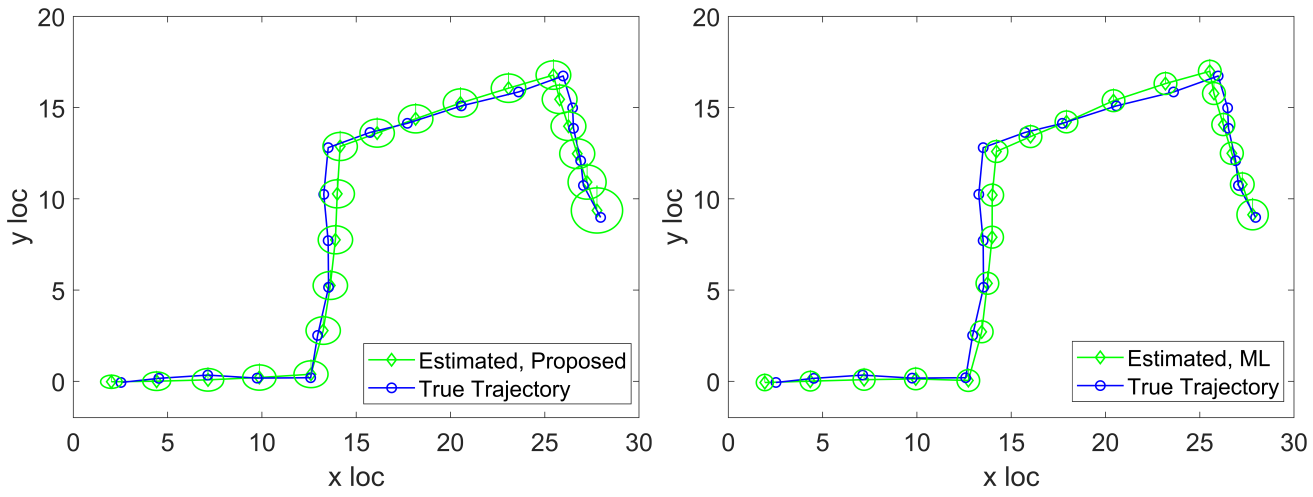


Figure 5: Estimated poses and 95% confidence ellipses for mobile robot using (a) Proposed covariance estimates and (b) Existing ML covariance estimates.

with $n = 20$ and 100 timesteps are shown in Figure 8. The summaries of the distributions are given in Table 3 which reports the average variance estimate, the bias of the variance estimates C and the average Mahalanobis localization error G, defined in Equations (36) and (37), respectively, computed in the simulations of both linear and nonlinear robot models with $n = 20$ timesteps.

The ML estimates are always significantly less than the true variance, while the proposed variance estimates are centered around the true variance (Figure 6), illustrating the ML estimators are underestimating the variances while the proposed approach is unbiased. The smaller ML variance estimates results in Mahalanobis errors being too large (Figure 7). Further, the proposed variance estimates have improved precision with less variability around the true value if larger number of data n is available while remaining unbiased (Figure 8). For both motion models, the existing ML estimates have far more bias and the resulting average Mahalanobis error is significantly larger than it should be, which is $nd_x = 20 \times 2 = 40$ (Table 3). In comparison, the proposed variance estimates agree much more closely with true values (significantly smaller bias) resulting in about 50% smaller Mahalanobis localization errors.

Table 3: Average variance estimates and localization errors, from simulations with $n = 20$

Motion model	Estimate	σ_R^2	σ_{Q1}^2	σ_{Q2}^2	C	G
Linear	Existing	0.81	0.13	0.06	0.628	103.93
	Proposed	1.46	0.52	0.17	0.003	51.13
	Truth	1.50	0.50	0.20	NA	NA
Nonlinear	Existing	0.49	0.77	0.18	1.097	95.81
	Proposed	1.20	1.22	0.38	0.638	59.57
	Truth	1.50	0.50	0.20	NA	NA

Performance under non-line-of-sight conditions: Monte Carlo simulations of the linear

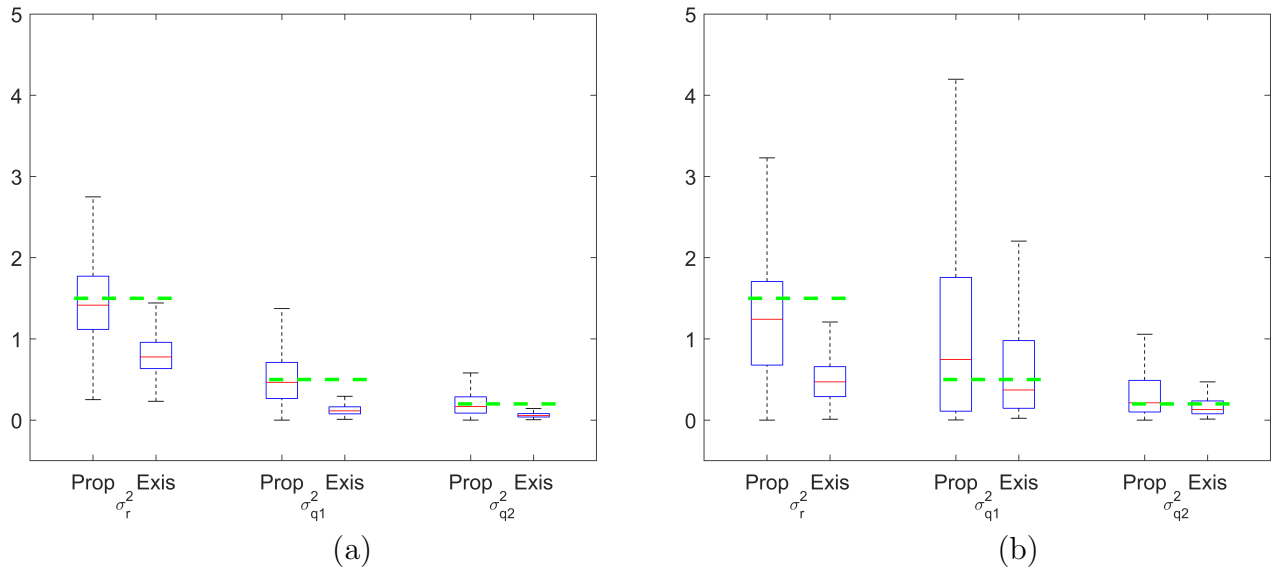


Figure 6: Variance estimates from simulations with Proposed and Existing ML methods. Horizontal dashed lines show the true variances. (a) linear and (b) nonlinear vehicle motion.

robot motion model with the mixture meausurement noise distribution Equation (30) and $n = 20$ poses were conducted to study the performance of the proposed variance estimator that incorporate M estimators and the proposed variance estimator without an M estimator (non-robust). Outlier mixture proportions $\alpha = 0.10$ and 0.25 were considered, in which the latter case contains more outliers than the former. The proposed method was implemented with a Cauchy M estimator using the parameter $a_C = 1.645$. Figure 9 shows the estimates of robot location for one realization with $\alpha = 0.10$. It can be seen that the localization estimates are similar with both methods, however, due to the presence of outliers, the confidence ellipses are significantly different depending on whether or not an M estimator is utilized. The confidence ellipses of the position solition with the M estimator (Figure 9a), by virtue of downweighting the unusual observations, are smaller than those of the non-robust method (Figure 9b) which are much larger (more uncertain).

Table 4 summarizes the means and biases of the variance estimates and the Mahalanobis distances from 1,000 Monte Carlo simulations. The benefit of using an M estimator in conjunction with the unbiased variance estimators on the quality of the estimates is evident, in particular when the outliers are more frequent (i.e., $\alpha = 0.25$ versus 0.10). With the presence of more outliers the variance estimates deviate more drastically from the true values if M estimator is not employed (bias increases from 94.3 to 570.3 due to outliers). By contrast, when an M estimator is used, in conjunction with the unbiased variance estimators, the bias in the estimates is much smaller (bias increases only from 0.05 to 4.31 due to outliers). The average Mahalanobis localization error is not drastically different from the expected value of $2n = 40$ due to the outliers, regardless of whether or not an M estimator is used, since unbiased variance estimators are used in both results. However, when M estimator is used the localization error is smaller, due to the downweighting of the unusual observations.

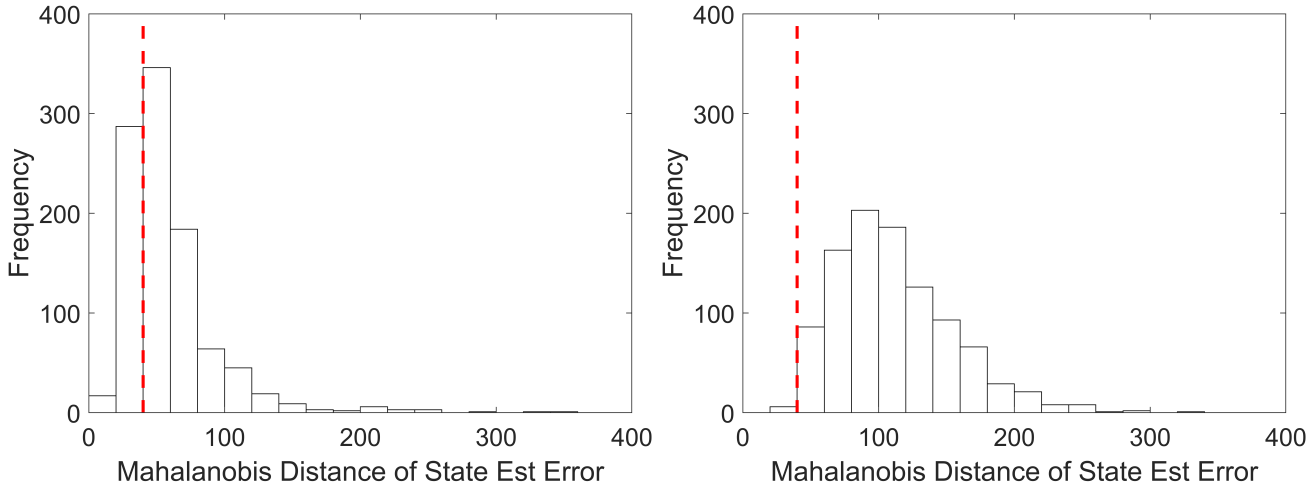


Figure 7: Mahalanobis localization errors from simulations of linear vehicle motion. Vertical line shows the true mean $2n = 40$. (a) Using proposed method variance estimators and (b) Using ML based variance estimators.

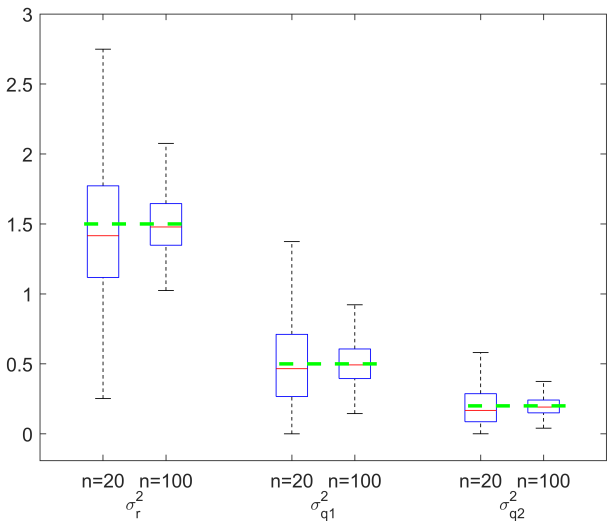
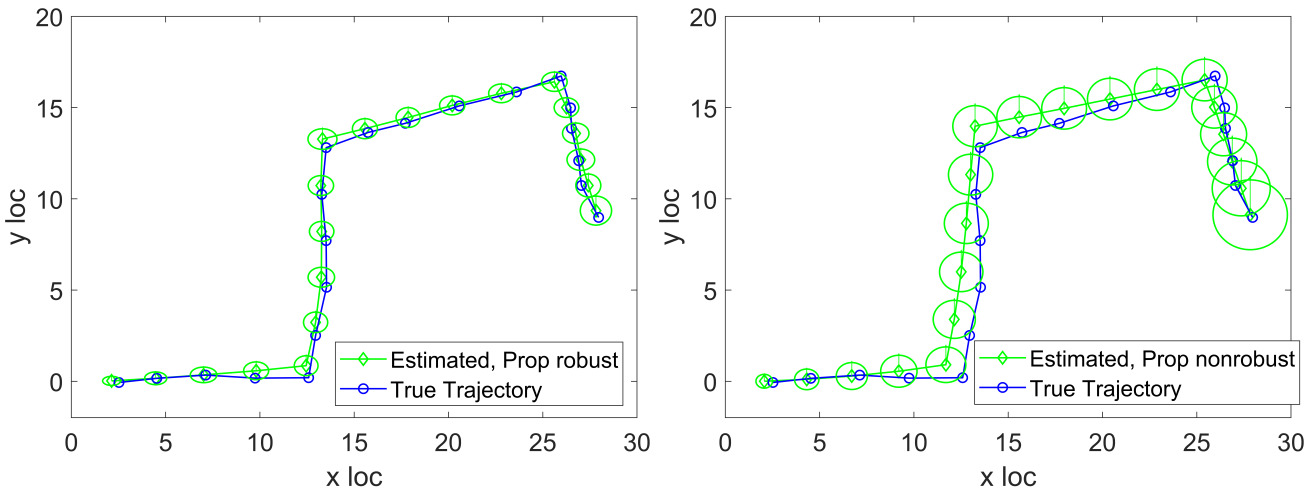


Figure 8: Effect of data size n used on the quality of the proposed variance estimates with the linear vehicle motion model. Horizontal dashed lines show the true variances.

6.2 GPS experiments

The proposed unbiased variance estimation and the existing ML variance estimation were applied to the GPS dataset of the vehicle driving experiment shown in Figure 4. The methods were used to estimate the variances σ_b^2, σ_d^2 and σ_R^2 of the clock error offset, drift and measurement error, respectively, and the vehicle trajectory by incorporating M estimators following the procedure described in Section 4. The Cauchy M estimator with parameter $a_C = 3.5$ was used to achieve robustness against outliers with both methods. As a further



(a) (b)

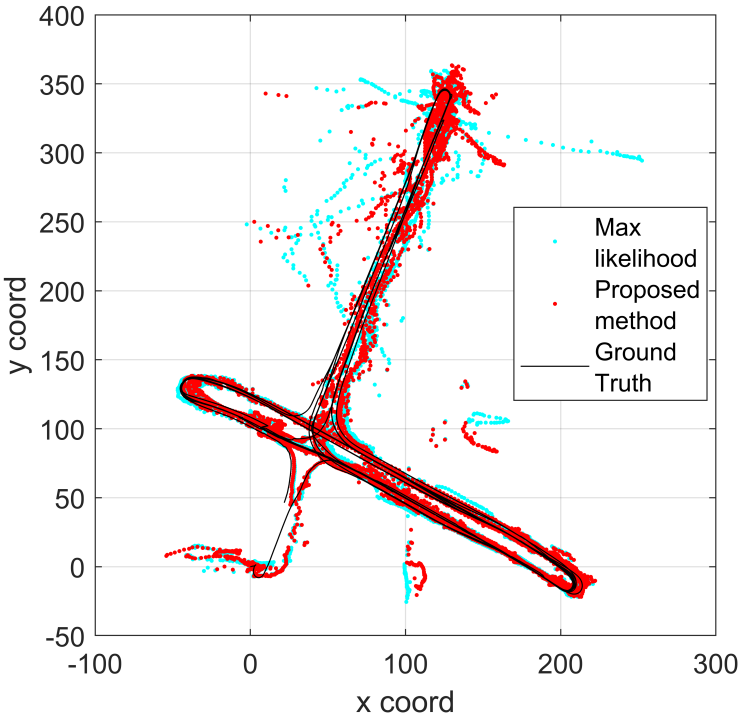
Figure 9: Estimated poses and 95% confidence ellipses using proposed method for linear motion with outliers with $\alpha = 0.1$. (a) With M estimator and (b) Without M estimator.

Table 4: Average variance estimates and localization errors- with outliers, linear robot model and $n = 20$

Method	α	σ_R^2	σ_{Q1}^2	σ_{Q2}^2	C	G
Proposed method	0.10	1.69	0.39	0.15	0.05	55.87
with M estimator	0.25	3.57	0.39	0.15	4.31	47.98
	Truth	1.50	0.50	0.20	NA	NA
Proposed method	0.10	11.19	1.04	0.59	94.28	56.00
without M estimator	0.25	25.37	0.96	0.73	570.28	59.37
	Truth	1.50	0.50	0.20	NA	NA

comparison, the Mahalanobis localization errors computed with the proposed and ML methods are compared to those reported by [1] that used the SC approach and the same data set.

Figure 10 shows the localization results with respect to the ground truth and Table 5 summarizes the variance estimates. It is clear that the localization with the proposed (unbiased) variance estimator is much closer to the ground truth than that obtained with the ML (biased) estimator. From Table 5 it can be seen that, similar to the simulation study, the ML variance estimates are much smaller than the proposed variance estimates, which indicates that it is likely that the ML estimates are underestimating the actual variances. Table 6 summarizes the localization errors, the average Mahalanabis distance from the truth over all poses, computed using Equation (37) with the localization results of the proposed and ML variance estimators and the results reported for the SC approach of [1] for the same data set. The proposed variance estimators, in conjunction with the M estimator, achieves significantly smaller mean and maximum localization error than the existing SC approach, while the median errors from both approaches are similar.



(a)

Figure 10: GPS based localization results using the proposed variance estimator and the ML variance estimator, by incorporating the Cauchy M estimator to achive robustness

Table 5: Process and measurement variance estimates obtained with the GPS data

Method	σ_R^2	σ_b^2	σ_d^2
Proposed method	19.24	5.498	0.704
ML method	18.35	4.84E-9	9.35e-04

7 Conclusions

This paper proposed a new methodology to estimate multiple noise variances in a factor-graph based formulation of the GPS based localization problem under non-line-of-sight conditions. Unbiased variance estimators of the factor graph formulation are developed following the method of moments approach in linear regression theory. An iteratively reweighted least squares (IRLS) approach was presented for jointly estimating the system state and the noise variances. The existing smoothing based state estimation approaches have largely relied on the maximum likelihood (or residual sample variance) estimators of variances in an expectation-maximization framework. While the ML estimators are easy to compute and have good performance in many applications, they can result in significant underestimation of the true variances which may in turn impact the localization accuracy.

A simulated mobile robot motion analysis showed that using the proposed variance estimators, the localization errors can be significantly improved over the ML estimators, considering model nonlinearities and measurement outliers. In addition, a case study involving

Table 6: Localization errors (m) of the proposed method and the SC approach [1] with the GPS data

Method	Mean	Median	Max
Proposed method	6.06	3.48	121.08
ML method	7.06	4.05	131.13
SC method	17.91	3.66	171.61

real GPS data compared the performance of the method to an existing switching constraints method for a localization problem in an urban environment with significant multipath problems. A major assumption in the current study was that the noise components are mutually uncorrelated and covariance matrix we estimate was diagonal. A future work of practical interest is to extend the presented approach to unbiased estimators of non-diagonal covariance matrices.

8 Abbreviations

GPS: Global Positioning System
RAIM: Receiver Autonomous Integrity Monitoring
LiDAR: Light Detection and Ranging
IMU: Inertial Measurement Unit
SLAM: Simultaneous Localization and Mapping
IRLS: iteratively Reweighted Least Squares
EM: Expectation Maximization
ML: Maximum Likelihood
SC: Switching Constraints
CPU: Central Processing Unit

9 Declarations

Ethics Approval and Consent to Participate

Not applicable.

Consent For Publication

Not applicable.

Availability of Data and Material

Data and computer software supporting our findings can be accessed at https://github.com/ardaVfamufsu/Factor-Graph-Covar.

Competing Interests

O. Arda Vanli and Clark N. Taylor declare that they have no conflict of interest.

Funding

This research was supported in part by the Air Force Research Laboratory, Air Force Institute of Technology (AFIT), through the Air Force Office of Scientific Research (AFOSR), Summer Faculty Fellowship Program (SFFP) ®, Contract Numbers FA8750-15-3-6003 and FA9550-15-0001.

Authors' Contributions

OAV and CNT developed the core concepts of the methodology presented within this manuscript, OAV implemented the algorithm for the simulation and case studies, CNT provided debugging and refinements on the algorithm, OAV developed and refined the manuscript, CNT provided proofreadings, refinements and revisions of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to thank the anonymous reviewers for their valuable suggestions. The first author (OAV) would like to acknowledge the support for this work from the AFOSR-SFFP program and the AFIT.

Authors' Information

Affiliations
Department of Industrial and Manufacturing Engineering, Florida A&M University, Florida State University, FL, USA
O. Arda Vanli
Air Force Institute of Technology, Autonomy and Navigation Technology Center, WPAFB, OH, USA
Clark N. Taylor

References

[1] N. Sünderhauf, M. Obst, G. Wanielik, and P. Protzel, "Multipath mitigation in gnss-based localization using robust optimization," in *2012 IEEE Intelligent Vehicles Symposium*. IEEE, 2012, pp. 784–789.

[2] G. Agamennoni, P. Furgale, and R. Siegwart, "Self-tuning m-estimators," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 4628–4635.

[3] E. Olson and P. Agarwal, "Inference on networks of mixtures for robust robot mapping," *The International Journal of Robotics Research*, vol. 32, no. 7, pp. 826–840, 2013.

[4] N. Sünderhauf and P. Protzel, "Switchable constraints for robust pose graph slam," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2012, pp. 1879–1884.

[5] S. Hewitson and J. Wang, "Gnss receiver autonomous integrity monitoring (raim) performance analysis," *Gps Solutions*, vol. 10, no. 3, pp. 155–170, 2006.

[6] A. Shetty and G. Xingxin Gao, "Adaptive covariance estimation of lidar-based positioning errors for uavs," *Navigation*, vol. 66, no. 2, pp. 463–476, 2019.

[7] M. Kaess and F. Dellaert, "Covariance recovery from a square root information matrix for data association," *Robotics and autonomous systems*, vol. 57, no. 12, pp. 1198–1210, 2009.

[8] J. Behley and C. Stachniss, "Efficient surfel-based slam using 3d laser range data in urban environments." in *Robotics: Science and Systems*, 2018.

[9] O. A. Vanli and C. N. Taylor, "Covariance estimation for factor graph based bayesian estimation," in *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*. IEEE, 2020, pp. 1–8.

[10] S. J. Julier and J. K. Uhlmann, "Simultaneous localisation and map building using split covariance intersection," in *Proceedings 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems. Expanding the Societal Role of Robotics in the the Next Millennium (Cat. No. 01CH37180)*, vol. 3. IEEE, 2001, pp. 1257–1262.

[11] F. Dellaert and M. Kaess, "Square Root SAM: Simultaneous localization and mapping via square root information smoothing," *The International Journal of Robotics Research*, vol. 25, no. 12, pp. 1181–1203, 2006.

[12] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear dynamic systems," *AIAA journal*, vol. 3, no. 8, pp. 1445–1450, 1965.

[13] G. J. Bierman, "Factorization methods for discrete sequential estimation," 1977.

[14] M. Kaess, A. Ranganathan, and F. Dellaert, "iSAM: Incremental smoothing and mapping," *IEEE Transactions on Robotics*, vol. 24, no. 6, pp. 1365–1378, 2008.

[15] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. J. Leonard, and F. Dellaert, "isam2: Incremental smoothing and mapping using the bayes tree," *The International Journal of Robotics Research*, vol. 31, no. 2, pp. 216–235, 2012.

[16] Z. Chen, "Bayesian filtering: From kalman filters to particle filters, and beyond," *Statistics*, vol. 182, no. 1, pp. 1–69, 2003.

[17] R. Casas, A. Marco, J. Guerrero, and J. Falco, "Robust estimator for non-line-of-sight error mitigation in indoor localization," *EURASIP Journal on Advances in Signal Processing*, vol. 2006, pp. 1–8, 2006.

[18] İ. Güvenç, C.-C. Chong, F. Watanabe, and H. Inamura, "Nlos identification and weighted least-squares localization for uwb systems using multipath channel statistics," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, pp. 1–14, 2007.

[19] T. Pfeifer and P. Protzel, "Robust sensor fusion with self-tuning mixture models," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 3678–3685.

[20] ——, "Expectation-maximization for adaptive mixture models in graph optimization," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 3151–3157.

[21] R. M. Watson, J. N. Gross, C. N. Taylor, and R. C. Leishman, "Enabling Robust State Estimation through Measurement Error Covariance Adaptation," *IEEE Transactions on Aerospace and Electronic Systems*, 2019.

[22] ——, "Robust incremental state estimation through covariance adaptation," *IEEE Robotics and Automation Letters*, 2020.

[23] Y. Bar-Shalom and X.-R. Li, "Estimation and tracking- principles, techniques, and software," *Norwood, MA: Artech House, Inc, 1993.*, 1993.

[24] J. Duník, O. Straka, O. Kost, and J. Havlík, "Noise covariance matrices in state-space models: A survey and comparison of estimation methodsPart I," *International Journal of Adaptive Control and Signal Processing*, 2017.

[25] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to linear regression analysis*. John Wiley & Sons, 2012, vol. 821.

[26] A. C. Rencher and G. B. Schaalje, *Linear models in statistics*. John Wiley & Sons, 2008.

[27] G. Grisetti, T. Guadagnino, I. Aloise, M. Colosi, B. Della Corte, and D. Schlegel, "Least squares optimization: from theory to practice," *arXiv preprint arXiv:2002.11051*, 2020.

[28] C. Zach, "Robust bundle adjustment revisited," in *European Conference on Computer Vision*. Springer, 2014, pp. 772–787.

[29] C. N. Taylor and S. Lubold, "Verifying the predicted uncertainty of bayesian estimators," in *Geospatial Informatics, Motion Imagery, and Network Analytics VIII*, vol. 10645. International Society for Optics and Photonics, 2018, p. 106450E.