# Mobile Localization in Non-Line-of-Sight Using Constrained Square-Root Unscented Kalman Filter

Siamak Yousefi, *Student Member, IEEE*, Xiao-Wen Chang, and Benoit Champagne, *Senior Member, IEEE*

***Abstract*—Localization and tracking of a mobile node (MN) in non-line-of-sight (NLOS) scenarios, based on time-of-arrival (TOA) measurements, is considered in this paper. We develop a constrained form of a square-root unscented Kalman filter (SRUKF), where the sigma points of the unscented transformation are projected onto the feasible region by solving constrained optimization problems. The feasible region is the intersection of several disks formed by the NLOS measurements. We show how we can reduce the size of the optimization problem and formulate it as a convex quadratically constrained quadratic program, which depends on the Cholesky factor of the *a posteriori* error covariance matrix of the SRUKF. As a result of these modifications, the proposed constrained SRUKF (CSRUKF) is more efficient and has better numerical stability compared to the constrained unscented Kalman filter (UKF). Through simulations, we also show that the CSRUKF achieves a smaller localization error compared to other techniques and that its performance is robust under different NLOS conditions.**

***Index Terms*—Constrained Kalman filter (KF), convex optimization, localization, non-line-of-sight (NLOS).**

## I. Introduction

NETWORK-BASED radio localization has received great attention in recent years due to limitations of the Global Positioning System in indoor places and dense urban areas and now finds numerous applications in surveillance, security, etc. [1]. In this technology, radio signals exchanged between a mobile node (MN) and fixed reference nodes (RNs) with known positions,[1] are exploited to determine the unknown location of the MN. Several different types of measurement can be used for localization, e.g., time of arrival (TOA), time difference of arrival, received signal strength (RSS), angle of arrival (AOA), and a hybrid of these. Among the different localization techniques, TOA-based methods in which the MN and RNs are synchronized are usually preferred, particularly in the context of IEEE 802.15.4a, which exploits ultrawideband (UWB) technology [2]. Indeed, measurement of TOA can be done accurately with UWB signaling due to its fine timing resolution and robustness against multipath and fading.

One of the main challenges in radio localization is the non-line-of-sight (NLOS) problem, which occurs due to the blockage of the direct sight between the MN and RNs. In an NLOS situation, due to either reflection of the radio waves by scatterers or penetration through blocking objects, the travel time of the received signals increases [3]–[5]. Consequently, the NLOS error of each measured TOA needs to be modeled as a random variable with a positive bias, which can be quite large [6]. The first step in dealing with the NLOS problem is to detect the NLOS measurements and, if necessary, discard them. To this end, some techniques estimate the variance of each measurement, and if it is above a given threshold, the corresponding link is identified as NLOS [7]–[9]. NLOS identification techniques using signal features have been also proposed for UWB applications [4], [5], [10]. If, however, the NLOS measurements cannot be discarded due to an insufficient number of LOS measurements for unambiguous localization, the next step is to mitigate their effect through further processing.

There are numerous works focusing on NLOS mitigation for the localization of stationary nodes, which are mostly based on (memoryless) constrained optimization techniques, e.g., [11] and [12]. In these approaches, the position of the MN is constrained to be within the convex hull formed by the intersection of multiple disks, each disk being centered at one of the NLOS RNs and with a radius equal to the corresponding measured range. By restricting the MN position in this way and by employing the LOS measurements in the cost function to be minimized, the unknown location can be found through solving a constrained optimization problem. For a survey on TOA-based memoryless localization in NLOS scenarios, see [6] and the references therein.

For an MN with available dynamic model, filtering techniques are preferred compared to memoryless methods. This is particularly the case when data from inertial measurements units (IMUs) are used in parallel with range information for tracking purposes [13], [14]. Some methods apply Kalman filter (KF) preprocessing on measured TOAs to smooth out the effect of the variances of the NLOS biases while scaling the covariance matrix in an extended KF (EKF) to further mitigate the effect of their means [7]–[9]. However, these approaches

Manuscript received December 8, 2013; revised May 10, 2014 and July 2, 2014; accepted July 4, 2014. Date of publication July 16, 2014; date of current version May 12, 2015. This work was supported by the Natural Sciences and Engineering Research Council of Canada. The review of this paper was coordinated by Prof. G. Mao.

S. Yousefi and B. Champagne are with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada (e-mail: siamak.yousefi@mail.mcgill.ca; benoit.champagne@mcgill.ca).

X.-W. Chang is with the School of Computer Science, McGill University, Montreal, QC H3A 0E9, Canada (e-mail: chang@cs.mcgill.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2014.2339734

[1]In wireless cellular networks, the RN are identified with the base stations, whereas in wireless sensor networks, they are called anchors.

can only achieve a moderate performance for large NLOS biases. In [15] and [16], it is assumed that the mean and variance of the NLOS biases are known; in practice, however, this information is not available accurately beforehand unless prior field measurements are obtained.

Some other approaches regard the NLOS bias as a nuisance parameter and try to estimate its distribution using kernel density estimation (KDE) techniques. In [17], a robust semiparametric EKF is proposed for NLOS mitigation of an MN. The performance of this technique is improved by the interacting multiple model algorithm in [18]. Although considered for TOA measurements, these techniques are also suitable when AOA, RSS, or a hybrid of these is employed. However, in addition to high computational cost, the performance of KDE still depends on how well it can model the distribution of the NLOS biases. It is claimed that for cellular applications, the performance is only satisfactory when the ratio of NLOS to LOS measurements is less than a half, and a higher ratio might result in divergence of KDE algorithms [18].

In some other techniques, the random NLOS biases are considered as parameters in the state vector to be jointly estimated with other state parameters [19]–[22], whereas the NLOS bias variation over time is modeled as a random walk. The technique in [19] uses EKF, whereas [20] and [21] use particle filters (PFs) that generally have a high computational cost. In [22], an improved EKF is used where bound constraints on the NLOS biases are enforced for improving the localization accuracy. Although the aforementioned techniques can mitigate the effect of NLOS biases to some extent, their performance might not be good due to the mismatch between the random walk model and the physical reality, which is unavoidable considering the unpredictable nature of the biases. Furthermore, by including the biases in the state vector, the computational cost of the filter grows noticeably [17].

In this paper, we propose an efficient square-root unscented KF (SRUKF) with convex inequality constraints for localization of an MN in NLOS situations. The proposed constrained SRUKF (CSRUKF) is based on a combination of the SRUKF in [23] for unconstrained problems and the constrained unscented KF (UKF) in [24]. In our proposed algorithm, similar to some memoryless approaches, the NLOS measurements are removed from the observation vector and are employed instead to form a closed convex constraint region [6]. At each time step, we use a SRUKF to estimate the state vector and compute the Cholesky factor of the error covariance matrix. To impose the constraints onto the estimated quantities, as proposed in [24], the sigma points of the unscented transformation (UT) may need to be projected onto the feasible region by solving a convex quadratically constrained quadratic program (QCQP). However, we show that the projection can be done in a more efficient and numerically stable way by solving a QCQP with reduced size, in which the cost function depends on the Cholesky factor of the *a posteriori* error covariance matrix, which is readily obtained from the SRUKF. Through simulations, our proposed algorithm is shown to achieve a good localization performance under different NLOS scenarios. In particular, in severe NLOS conditions and with small measurement noises, our method achieves a superior performance compared to other benchmark approaches. Another salient advantage is its robustness to false alarm (FA) errors[2] in NLOS identification, which makes it suitable for practical applications where such errors may be inevitable.

The organization of this paper is as follows: In Section II, the system model is described, and the problem formulation is presented. The proposed CSRUKF algorithm is developed in Section III, along with a discussion of computational complexity. The simulation results and comparisons with different algorithms are given in Section IV. Finally, Section V concludes this paper.

*Notation:* Small and capital bold letters represent vectors and matrices, respectively. The vector 2-norm operation is denoted by $\|\cdot\|$, and $(\cdot)^T$ and $(\cdot)^{-1}$ stand for matrix transpose and inverse operations, respectively. A diagonal matrix with entries $x_1, \ldots, x_M$ on the main diagonal is denoted by $\mathrm{diag}(x_1, \ldots, x_M)$. For $i \leq j$, $\boldsymbol{q}(i:j)$ denotes a vector of size $j-i+1$ obtained by extracting the $i$th to $j$th entries of vector $\boldsymbol{q}$, inclusively. The symbol $\boldsymbol{I}$ denotes an identity matrix of appropriate dimension. For a positive semidefinite Hermitian matrix $\boldsymbol{R}$, $\boldsymbol{R}^{1/2}$ denotes its unique positive semidefinite square-root matrix, i.e., such that $\boldsymbol{R}^{1/2}\boldsymbol{R}^{1/2} = \boldsymbol{R}$ [25].

## II. System Description and Problem Statement

### A. System Model

Consider a network of $M$ fixed RNs and one MN, distributed on a 2-D plane and exchanging timing signals via wireless links. With reference to a Cartesian coordinate system in this plane, let $\boldsymbol{a}^i \in \mathbb{R}^2$ denote the known position vector of the $i$th RN, where $i \in \{1, \ldots, M\}$, whereas $\boldsymbol{x}_k \in \mathbb{R}^2$ and $\boldsymbol{v}_k \in \mathbb{R}^2$ denote the unknown position and velocity vectors of the MN at discrete time instant $k$, respectively. Let the state vector be $\boldsymbol{s}_k = [\boldsymbol{x}_k^T, \boldsymbol{v}_k^T]^T \in \mathbb{R}^4$, which includes the position and velocity components of the MN. The motion model is assumed to be a nearly constant velocity model as

$$\boldsymbol{s}_k = \boldsymbol{F}\boldsymbol{s}_{k-1} + \boldsymbol{G}\boldsymbol{w}_{k-1} \tag{1}$$

where the matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ are

$$\boldsymbol{F} = \begin{bmatrix} 1 & 0 & \delta t & 0 \\ 0 & 1 & 0 & \delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{G} = \begin{bmatrix} \frac{\delta t^2}{2} & 0 \\ 0 & \frac{\delta t^2}{2} \\ \delta t & 0 \\ 0 & \delta t \end{bmatrix} \tag{2}$$

and $\delta t$ is the time step duration. The vector $\boldsymbol{w}_{k-1} \in \mathbb{R}^2$ in (1) is a zero-mean white Gaussian noise process (acceleration) with diagonal covariance matrix $\boldsymbol{Q} = \sigma_w^2 \boldsymbol{I}$.

In this paper, we consider TOA-based localization, in which the range between the MN and each RN is obtained by multiplying the time of flight of the radio wave by the speed of light. If the MN and RNs are accurately synchronized, then a one-way ranging scheme can be used; otherwise, a two-way ranging protocol may be employed where the relative clock offsets are removed from the TOA measurements [26]. Let $\mathcal{L}_k$ and $\mathcal{N}_k$ denote the index sets of the RNs that are identified as

[2]In this paper, an FA refers to the erroneous identification of an LOS link as being NLOS, whereas a missed detection (MD) refers to the opposite situation.

LOS and NLOS nodes at time instant $k$, respectively. The range measurements can thus be represented by vector $\boldsymbol{r}_k \in \mathbb{R}^M$ with components

$$r_k^i = \begin{cases} h^i(\boldsymbol{s}_k) + n_k^i, & i \in \mathcal{L}_k \\ h^i(\boldsymbol{s}_k) + b_k^i + n_k^i, & i \in \mathcal{N}_k \end{cases} \tag{3}$$

where $h^i(\boldsymbol{s}_k) = \|\boldsymbol{x}_k - \boldsymbol{a}^i\|$, $n_k^i$ is the measurement noise, and $b_k^i$ is a positive random NLOS bias, which is usually considered independent from $n_k^i$. The noise terms $n_k^i$, for $i \in \{1, \ldots, M\}$, are modeled as independent white Gaussian processes, with zero mean and known variance $\sigma_n^2 > 0$. The probability distributions of the biases $b_k^i$ are time varying due to the movement of the MN and other objects in the area. In the literature, different distributions have been considered for the biases, for instance: exponential [27], [28], shifted Gaussian [16], and uniform [4] are widely employed. However, having *a priori* knowledge about the distributions of the NLOS biases requires preliminary field measurements, which may not be possible in practical applications. Therefore, in this paper, we make no specific assumption about the distributions of the NLOS biases, although we suppose that the NLOS links are accurately identified at every time instant.[3]

The processing of range measurements for NLOS identification and mitigation can be done either at the MN or at a fusion center connected to the RNs. The former is used in the MN self-localization applications, whereas the latter is of interest to target tracking applications.

### B. Problem Formulation

The state vector $\boldsymbol{s}_k$ and the NLOS biases $b_k^i$ for $i \in \mathcal{N}_k$ are the unknown parameters in the aforementioned model. Representing the NLOS biases by a simple dynamic model such as a random walk might be justified for certain environments as considered in [20] and [21], but in general environments, this may only be considered an approximation. The optimal choice for the variance of the random walk increment is also intractable as discussed in [29]. Including the biases $b_k^i$ in the state vector also increases the computational complexity of the KF; therefore, it may not be computationally efficient as well.

Since the random walk model may not be an accurate approximation for the evolution of $b_k^i$ over time, we avoid using this model and estimating the biases. To simplify the problem and reduce the number of unknowns, we eliminate the NLOS measurements from the observation vector $\boldsymbol{r}_k$ and instead use the information carried out by the biases to restrict the position of the MN within a certain range. For instance, in many applications, it can be assumed that the TOA measurement noise $n_k^i$ is small compared to $b_k^i$ (particularly in UWB ranging), which implies that $b_k^i + n_k^i \geq 0$ [6]. In light of (3), this assumption is equivalent to

$$\|\boldsymbol{x}_k - \boldsymbol{a}^i\| \leq r_k^i, \qquad i \in \mathcal{N}_k \tag{4}$$

[3]We assume that for every time instant, an NLOS identification technique has been applied on the measured ranges before employing our proposed filter. There are numerous techniques that identify the NLOS link using the variance test [7]–[9]. For UWB applications, the features of the received TOA signal can also be employed for NLOS identifications, as proposed in [4], [5], and [10].

which is obviously a convex constraint, as in [30]. If the small noise assumption cannot be made, e.g., in narrow-band systems where TOA-based ranging measurement errors are relatively large, the constraints in (4) may not be satisfied. To avoid this limitation, we can generalize the latter as

$$\|\boldsymbol{x}_k - \boldsymbol{a}^i\| \leq r_k^i + \epsilon\sigma_n, \qquad i \in \mathcal{N}_k \tag{5}$$

where $\epsilon \geq 0$ is a small number to ensure that the MN is located inside a disk with radius $r_k^i + \epsilon\sigma_n$. Note that even if the bias is zero for a given link (i.e., LOS situation), it is more likely that the MN satisfies the constraint in (5) as compared to (4). Therefore, we propose to use the constraint in (5) throughout this paper due to its robustness against measurement noise and FA error in NLOS identification. In the sequel, the *feasible region*, which is denoted by $\mathcal{D}_k$, refers to the convex set formed by the intersection of the disks in (5); hence

$$\mathcal{D}_k = \left\{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{a}^i\| \leq r_k^i + \epsilon\sigma_n \quad \forall\, i \in \mathcal{N}_k\right\}. \tag{6}$$

At every time instant $k$, let us remove the NLOS measurements from the observations in (3) and only keep the LOS measurements, i.e., $r_k^i$ for all $i \in \mathcal{L}_k$. The remaining LOS range measurements can be represented by the vector $\boldsymbol{z}_k \in \mathbb{R}^{|\mathcal{L}_k|}$. Note that in the worst case, where all the measurements are identified as NLOS, the vector $\boldsymbol{z}_k$ is empty. The state space model and constraints can thus be expressed as

$$\boldsymbol{z}_k = \boldsymbol{h}(\boldsymbol{s}_k) + \boldsymbol{n}_k \tag{7a}$$

$$\boldsymbol{s}_k = \boldsymbol{F}\boldsymbol{s}_{k-1} + \boldsymbol{G}\boldsymbol{w}_{k-1} \tag{7b}$$

$$\|\boldsymbol{x}_k - \boldsymbol{a}^i\| \leq r_k^i + \epsilon\sigma_n, \qquad i \in \mathcal{N}_k \tag{7c}$$

where $\boldsymbol{h}(\boldsymbol{s}_k)$ and $\boldsymbol{n}_k$ are vectors whose entries are $h^i(\boldsymbol{s}_k)$ and $n_k^i$ for $i \in \mathcal{L}_k$, respectively. Under our previous assumptions on the measurement noise $n_k^i$ in (3), the covariance matrix of $\boldsymbol{n}_k$ is positive-definite diagonal, i.e., $\boldsymbol{R} = \mathbb{E}[\boldsymbol{n}_k\boldsymbol{n}_k^T] = \sigma_n^2\boldsymbol{I} \in \mathbb{R}^{|\mathcal{L}_k|\times|\mathcal{L}_k|}$. The constraints in (7c) are only on the first two elements of the state vector, i.e., $\boldsymbol{x}_k$, as we have a 2-D positioning scenario herein. Note that if the constraints in (7c) are removed from the state model, then an ordinary nonlinear filtering technique such as EKF can be used. This approach is also known as EKF with outlier rejection [20] since the NLOS measurements are regarded as outliers and therefore discarded.

In minimum mean square error (MMSE) estimation, e.g., Kalman-type filters, one tries to find the conditional mean and covariance matrix of the state vector $\boldsymbol{s}_k$ given the measurements up to current time instant $k$, as characterized by the conditional probability density function (pdf) $f(\boldsymbol{s}_k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k)$. However, when extra information about the state vector is available in the form of inequality constraints, the probability that the MN is outside the feasible region should be zero. Hence, a truncated or *constrained* conditional pdf, i.e., $f_c(.|.)$, can be defined as

$$f_c(\boldsymbol{s}_k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k) = \begin{cases} \frac{1}{\beta} f(\boldsymbol{s}_k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k), & \text{if } \boldsymbol{x}_k \in \mathcal{D}_k \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

where $\beta \triangleq \int_{\boldsymbol{x}_k \in \mathcal{D}_k} f(\boldsymbol{s}_k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k) d\boldsymbol{s}_k$ is a normalization constant. Therefore, one can estimate the state vector by finding

the conditional mean of $\boldsymbol{s}_k$ with truncated pdf as

$$\hat{\boldsymbol{s}}_k = \int\limits_{\boldsymbol{x}_k \in \mathcal{D}_k} \boldsymbol{s}_k f_c(\boldsymbol{s}_k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k)\, d\boldsymbol{s}_k \tag{9}$$

and the covariance matrix of the constrained state estimate can be found through

$$\hat{\boldsymbol{\Sigma}}_k = \int\limits_{\boldsymbol{x}_k \in \mathcal{D}_k} (\boldsymbol{s}_k - \hat{\boldsymbol{s}}_k)(\boldsymbol{s}_k - \hat{\boldsymbol{s}}_k)^T f_c(\boldsymbol{s}_k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k)\, d\boldsymbol{s}_k. \tag{10}$$

This idea is known as pdf truncation, where the distribution of the state vector given the measurements is forced to be zero outside the feasible region [31]. For a linear dynamic model with zero-mean Gaussian measurement and process noises, where the state vector is subject to linear inequality constraints, closed-form expressions for $\hat{\boldsymbol{s}}_k$ and $\hat{\boldsymbol{\Sigma}}_k$ in (9) and (10) have been obtained using pdf truncation along with the Gaussian assumption [32]. For nonlinear inequality constraints, it is proposed in [32] to do a Taylor series linearization of the constraints around the current state estimate and then apply the aforementioned method; however, this approach may not be accurate [33]. In general cases with nonlinear inequality constraints, pdf truncation requires multidimensional Monte Carlo (MC) integration, which becomes computationally expensive as the size of the state vector grows. Therefore, these computationally demanding techniques may not be suitable to solve our problem.

In the following section, we show how we can efficiently approximate $\hat{\boldsymbol{s}}_k$ and $\hat{\boldsymbol{\Sigma}}_k$ using an alternative approach that combines the SRUKF [21] for unconstrained problems with the projection-based constrained UKF in [22].

## III. Constrained Nonlinear Filtering With Sigma Point Projection

Another family of methods for imposing inequality constraints on the state vector are the projection-based techniques, in which the unconstrained state estimate, obtained through a Kalman-type filter, is projected onto the feasible region by solving an optimization problem [31]. However, by this approach, one cannot estimate the constrained error covariance matrix of the state, i.e., $\hat{\boldsymbol{\Sigma}}_k$, accurately. Therefore, in addition to the unconstrained state estimate, some representative sample points of the conditional pdf $f(\boldsymbol{s}_k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k)$ need to be projected onto the feasible region. For instance, the sigma points of the UT can give good statistical information about the mean and the error covariance matrix of the state estimate [34]. Based on this idea, in [24], a constrained UKF technique has been proposed in which the sigma points of the UKF violating the constraints are projected onto the feasible region. However, due to the dependence of the projection function on the inverse of the *a posteriori* error covariance matrix, the method in [24] may become numerically unstable [33]. In the following sections, we first describe a variation of the SRUKF that is better suited to our specific problem; then, to overcome the aforementioned numerical issue, we design a more efficient and numerically reliable method for projecting the sigma points generated from the *a posteriori* estimates onto the feasible region; finally, we summarize our algorithm and comment on its numerical complexity.

### A. Unconstrained SRUKF Algorithm

The proposed algorithm in this part is based on the SRUKF presented in [23] with slight modification such that the algorithm is more efficient and numerically reliable. Let $\boldsymbol{s}_{k-1|k-1}$ be the estimated state and $\boldsymbol{\Sigma}_{k-1|k-1}$ be the estimated error covariance matrix of the state, based on the available measurements up to current time instant $k-1$. Let $\boldsymbol{U}_{k-1|k-1}$ be the upper triangular Cholesky factor of $\boldsymbol{\Sigma}_{k-1|k-1}$, i.e., $\boldsymbol{\Sigma}_{k-1|k-1} = \boldsymbol{U}_{k-1|k-1}^T \boldsymbol{U}_{k-1|k-1}$. Then, for the next time instant, the *a priori* estimate of the state vector and the corresponding error covariance matrix, which are denoted by $\boldsymbol{s}_{k|k-1}$ and $\boldsymbol{\Sigma}_{k|k-1}$, respectively, can be obtained through prediction as

$$\boldsymbol{s}_{k|k-1} = \boldsymbol{F}\boldsymbol{s}_{k-1|k-1} \tag{11}$$

$$\boldsymbol{\Sigma}_{k|k-1} = \boldsymbol{F}\boldsymbol{\Sigma}_{k-1|k-1}\boldsymbol{F}^T + \boldsymbol{G}\boldsymbol{Q}\boldsymbol{G}^T. \tag{12}$$

Alternatively, the computation of (12) can be avoided as only the Cholesky factor of the *a priori* covariance matrix, which is denoted by $\boldsymbol{U}_{k|k-1}$, is required [23]. To this aim, let us rewrite (12) as

$$\boldsymbol{\Sigma}_{k|k-1} = \begin{bmatrix} \boldsymbol{F}\boldsymbol{U}_{k-1|k-1}^T & \boldsymbol{G}\boldsymbol{Q}^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_{k-1|k-1}\boldsymbol{F}^T \\ \boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{G}^T \end{bmatrix}. \tag{13}$$

If we compute the QR factorization of the second matrix on the right-hand side of (13), we obtain $\boldsymbol{U}_{k|k-1}$, i.e.,

$$\boldsymbol{U}_{k|k-1} = \mathrm{qr}\left\{ \begin{bmatrix} \boldsymbol{U}_{k-1|k-1}\boldsymbol{F}^T \\ \boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{G}^T \end{bmatrix} \right\} \tag{14}$$

where, by definition, the function $\mathrm{qr}\{.\}$ returns the upper triangular factor of the QR factorization of its matrix argument.

With the help of $U_{k|k-1}$, the sigma points of the SRUKF are generated as proposed in [23], i.e.,

$$\boldsymbol{s}_{k|k-1}^{(j)} = \begin{cases} \boldsymbol{s}_{k|k-1}, & j = 0 \\ \boldsymbol{s}_{k|k-1} + \sqrt{\eta_\alpha}\left(\boldsymbol{U}_{k|k-1}^T\right)_j, & j = 1, \ldots, N \\ \boldsymbol{s}_{k|k-1} - \sqrt{\eta_\alpha}\left(\boldsymbol{U}_{k|k-1}^T\right)_{j-N}, & j = N+1, \ldots, 2N \end{cases} \tag{15}$$

where $N$ is the dimension of the state vector (in this paper, $N = 4$), $(\boldsymbol{U}_{k|k-1}^T)_j$ denotes the $j$th column of matrix $\boldsymbol{U}_{k|k-1}^T$, and $\eta_\alpha$ is a tuning parameter that controls the spread of the sigma points. To better understand the geometric meaning of parameter $\eta_\alpha$, we can assume that $\boldsymbol{s}_{k|k-1}$ and $\boldsymbol{\Sigma}_{k|k-1}$ obtained through the proposed filter are approximately equal to the mean and covariance matrix of the conditional pdf $f(\boldsymbol{s}_k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{k-1})$. Define random variable $\eta_k = (\boldsymbol{s}_k - \boldsymbol{s}_{k|k-1})^T \boldsymbol{\Sigma}_{k|k-1}^{-1} (\boldsymbol{s}_k - \boldsymbol{s}_{k|k-1})$, which is the weighted squared distance between $\boldsymbol{s}_k$ and $\boldsymbol{s}_{k|k-1}$. Suppose that the parameter $\eta_\alpha$ in (15) is chosen

such that $\Pr(\eta_k \leq \eta_\alpha) = \alpha$, where $0 < \alpha < 1$ represents a desired confidence level. Then, the region of $\mathbb{R}^N$ defined by $\eta_k \leq \eta_\alpha$ represents a confidence ellipsoid, on the boundary of which the sigma points in (15) (except $\boldsymbol{s}_{k|k-1}^{(0)}$) fall. For example, if $\alpha = 0.9$, the probability for $\boldsymbol{s}_k$ to lie inside the ellipsoid delimited by the sigma points with the corresponding $\eta_\alpha$ is 90%. If we assume that $f(\boldsymbol{s}_k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{k-1})$ is approximately Gaussian, then the random variable $\eta$ has a chi-square distribution with $N$ degrees of freedom, and it becomes easy to find a value for $\eta_\alpha$ corresponding to a certain ellipsoid with confidence level $\alpha$.[4]

The generated sigma points are transformed through the nonlinear measurement function as

$$\boldsymbol{z}_{k|k-1}^{(j)} = \boldsymbol{h}(\boldsymbol{s}_{k|k-1}^{(j)}), \qquad j = 0, \ldots, 2N. \tag{16}$$

Then, the mean, cross-covariance matrix, and error covariance matrix of the transformed sigma points can be estimated by means of weighted sums as in [35]

$$\hat{\boldsymbol{z}}_{k|k-1} = \sum_{j=0}^{2N} w^{(j)} \boldsymbol{z}_{k|k-1}^{(j)} \tag{17}$$

$$\boldsymbol{\Sigma}_{k|k-1}^{\boldsymbol{s},\boldsymbol{z}} = \sum_{j=0}^{2N} w^{(j)} \left(\boldsymbol{s}_{k|k-1}^{(j)} - \boldsymbol{s}_{k|k-1}\right)\left(\boldsymbol{z}_{k|k-1}^{(j)} - \hat{\boldsymbol{z}}_{k|k-1}\right)^T \tag{18}$$

$$\boldsymbol{P}_{k|k-1}^{z} = \sum_{j=0}^{2N} w^{(j)} \left(\boldsymbol{z}_{k|k-1}^{(j)} - \hat{\boldsymbol{z}}_{k|k-1}\right)\left(\boldsymbol{z}_{k|k-1}^{(j)} - \hat{\boldsymbol{z}}_{k|k-1}\right)^T + \boldsymbol{R} \tag{19}$$

where $\boldsymbol{R}$ is the covariance matrix of the measurement noise $\boldsymbol{n}_k$ in (7a), and the weights $w^{(j)}$ appearing in these expressions are defined in a similar way as in [36]

$$w^{(j)} = \begin{cases} 1 - \frac{N}{\eta_\alpha}, & j = 0 \\ \frac{1}{2\eta_\alpha}, & j = 1, \ldots, 2N \end{cases} \tag{20}$$

and therefore satisfy $\sum_{j=0}^{2N} w^{(j)} = 1$.

If the weight $w^{(0)}$ in (20) is negative, it is possible that the covariance matrix obtained through (19) becomes indefinite (i.e., with negative eigenvalues). However, by choosing a sufficiently large value of $\alpha$, we can guarantee that $\eta_\alpha \geq N$; in turn, this implies that $w^{(0)} \geq 0$, and the covariance matrix (19) then becomes positive definite. In this paper, we are interested in projecting the sigma points that are far away from the mean, and it is therefore legitimate to consider ellipsoids with larger confidence levels so that the above issue can be naturally avoided.[5] In our dynamic model, with $N = 4$ and based on the chi-square assumption for $\eta_k$, it follows that if $\alpha > 0.6$, then $\eta_\alpha > N$, and the positive definiteness of (19) is guaranteed.

For numerical stability, instead of forming $\boldsymbol{P}_{k|k-1}^{z}$ explicitly, its Cholesky factor is calculated. Specifically, if we let

$$\mathbf{e}_z^{(j)} = \sqrt{w^{(j)}}\left(\boldsymbol{z}_{k|k-1}^{(j)} - \hat{\boldsymbol{z}}_{k|k-1}\right), \qquad j = 0, \ldots, 2N \tag{21}$$

then the upper triangular Cholesky factor of $\boldsymbol{P}_{k|k-1}^{z}$, which is denoted by $\boldsymbol{U}_{\boldsymbol{z}_k}$, is obtained through

$$\boldsymbol{U}_{\boldsymbol{z}_k} = \mathrm{qr}\left\{\left[\mathbf{e}_z^{(0)}, \mathbf{e}_z^{(1)}, \ldots, \mathbf{e}_z^{(2N)}, \boldsymbol{R}^{\frac{1}{2}}\right]^T\right\}. \tag{22}$$

It is proposed in [23] to first compute the Kalman gain

$$\boldsymbol{K}_k = \boldsymbol{\Sigma}_{k|k-1}^{\boldsymbol{s},\boldsymbol{z}}\left(\boldsymbol{P}_{k|k-1}^{z}\right)^{-1} = \boldsymbol{\Sigma}_{k|k-1}^{\boldsymbol{s},\boldsymbol{z}} \boldsymbol{U}_{z_k}^{-1} \boldsymbol{U}_{z_k}^{-T} \tag{23}$$

and then, the *a posteriori* state estimate and the Cholesky factor of the error covariance matrix can be updated through

$$\boldsymbol{s}_{k|k} = \boldsymbol{s}_{k|k-1} + \boldsymbol{K}_k(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_{k|k-1}) \tag{24}$$

$$\boldsymbol{U}_{k|k} = \mathrm{cholupdate}\left\{\boldsymbol{U}_{k|k-1}, \boldsymbol{K}_k \boldsymbol{U}_{z_k}^T, -1\right\} \tag{25}$$

where $\mathrm{cholupdate}\{\boldsymbol{U}_{k|k-1}, \boldsymbol{K}_k\boldsymbol{U}_{z_k}^T, -1\}$ is the consecutive downdates[6] of the Cholesky factor of $\boldsymbol{U}_{k|k-1}^T\boldsymbol{U}_{k|k-1}$ using the columns of $\boldsymbol{K}_k\boldsymbol{U}_{z_k}^T$. Note that (25) follows from the covariance matrix update

$$\boldsymbol{\Sigma}_{k|k} \triangleq \boldsymbol{U}_{k|k}^T\boldsymbol{U}_{k|k} = \boldsymbol{U}_{k|k-1}^T\boldsymbol{U}_{k|k-1} - \boldsymbol{K}_k\boldsymbol{U}_{z_k}^T\boldsymbol{U}_{z_k}\boldsymbol{K}_k^T. \tag{26}$$

Herein, however, we propose a more efficient and numerically reliable way to compute $\boldsymbol{s}_{k|k}$ and $\boldsymbol{U}_{k|k}$. Instead of the Kalman gain $\boldsymbol{K}_k$, we compute

$$\boldsymbol{T}_k = \boldsymbol{\Sigma}_{k|k-1}^{\boldsymbol{s},\boldsymbol{z}}\boldsymbol{U}_{z_k}^{-1} \tag{27}$$

which can be obtained by solving multiple triangular linear systems $\boldsymbol{T}_k\boldsymbol{U}_{z,k} = \boldsymbol{\Sigma}_{k|k-1}^{\boldsymbol{s},\boldsymbol{z}}$. Then, it follows from (23) that $\boldsymbol{K}_k = \boldsymbol{T}_k\boldsymbol{U}_{z_k}^{-T}$. Substituting this expression into (24), we obtain

$$\boldsymbol{s}_{k|k} = \boldsymbol{s}_{k|k-1} + \boldsymbol{T}_k\boldsymbol{U}_{z_k}^{-T}(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_{k|k-1}) \tag{28}$$

where the vector $\boldsymbol{y}_k \triangleq \boldsymbol{U}_{z_k}^{-T}(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_{k|k-1})$ can be obtained by solving the triangular linear system

$$\boldsymbol{U}_{z_k}^T\boldsymbol{y}_k = \boldsymbol{z}_k - \hat{\boldsymbol{z}}_{k|k-1}. \tag{29}$$

From (25) and (27), it follows that the covariance matrix can be updated as

$$\boldsymbol{\Sigma}_{k|k} = \boldsymbol{U}_{k|k-1}^T\boldsymbol{U}_{k|k-1} - \boldsymbol{T}_k\boldsymbol{T}_k^T. \tag{30}$$

Hence, the Cholesky factor of $\boldsymbol{\Sigma}_{k|k}$ can be computed as

$$\boldsymbol{U}_{k|k} = \mathrm{cholupdate}\{\boldsymbol{U}_{k|k-1}, \boldsymbol{T}_k, -1\}. \tag{31}$$

[4]The MATLAB function `chi2inv`$(\alpha, N)$ can be used for this purpose.

[5]In [34], a scaled version of the UT has been proposed to capture higher moments of the nonlinear measurement function, where the generated sigma points are located in the vicinity of each other. This method also guarantees positive definiteness of the covariance matrix. However, our problem is not highly nonlinear, and we are interested to generate sigma points that might be far away from one another; therefore, our parameter selection is different from [23] and [34].

[6]In MATLAB, the built-in function `Cholupdate` can be employed to do rank-1 Cholesky update or downdate, indicated by the third argument of the function.

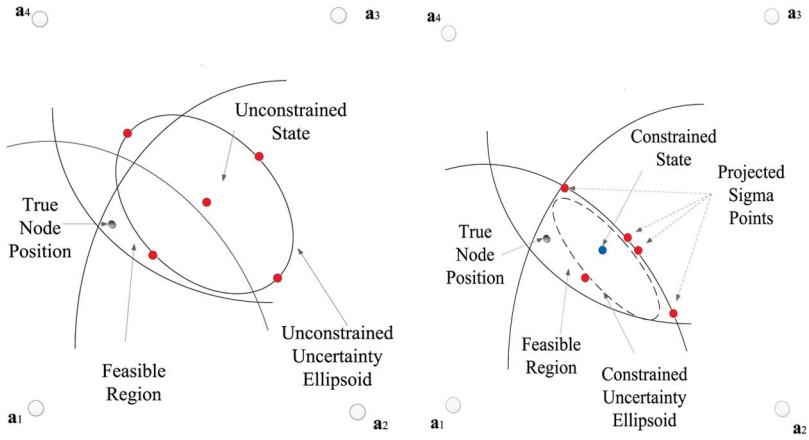


Fig. 1. Proposed projection technique. (a) Unconstrained state estimate and the uncertainty ellipsoid of sigma points, among which, some are outside the feasible region. (b) The projected sigma points fall inside the feasible region, and the uncertainty ellipsoid is shrunk.

Compared with the algorithm in [23], this modified algorithm for the estimation of $\boldsymbol{s}_{k|k}$ and $\boldsymbol{U}_{k|k}$ saves about $2N|\mathcal{N}_k|^2$ flops at each time step $k$. It is also more numerically reliable as it avoids solving some linear systems, which could be ill-conditioned, and computing some matrix–matrix multiplications.

Note that if all the measurements at time instant $k$ are in NLOS, then the measurement vector $\boldsymbol{z}_k$ is empty. Hence, we will use the predicted state in (11) and the Cholesky factor of the predicted covariance matrix in (14) to replace the *a posteriori* state vector in (28) and Cholesky factor of the error covariance matrix in (31), respectively.

### B. Imposing the Constraints on the Estimates

Up to this point, the *a posteriori* state estimate and the Cholesky factor of the *a posteriori* error covariance matrix have been obtained using a SRUKF without taking the constraints (7c) into account. To impose the constraints on the estimated state and error covariance matrix, similar to [24], a new set of sigma points is generated according to

$$\boldsymbol{s}_{k|k}^{(j)} = \begin{cases} \boldsymbol{s}_{k|k}, & j = 0 \\ \boldsymbol{s}_{k|k} + \sqrt{\eta_\alpha}(\boldsymbol{U}_{k|k}^T)_j, & j = 1, \ldots, N \\ \boldsymbol{s}_{k|k} - \sqrt{\eta_\alpha}(\boldsymbol{U}_{k|k}^T)_{j-N}, & j = N+1, \ldots, 2N. \end{cases} \tag{32}$$

The generated sigma points (except $\boldsymbol{s}_{k|k}^{(0)}$) form an uncertainty ellipsoid with $\boldsymbol{s}_{k|k}$ at its center, as shown in Fig. 1, for $N = 2$. After the generation of sigma points $\boldsymbol{s}_{k|k}^{(j)}$ with desired confidence ellipsoid, those that violate the constraints are projected onto the convex feasible region through

$$\mathcal{P}\left(\boldsymbol{s}_{k|k}^{(j)}\right) = \arg\min_{\boldsymbol{q}} \left\{ \left(\boldsymbol{q} - \boldsymbol{s}_{k|k}^{(j)}\right)^T \boldsymbol{W}_k \left(\boldsymbol{q} - \boldsymbol{s}_{k|k}^{(j)}\right) \right\}$$
$$\text{s.t.} \quad \left\|\boldsymbol{q}(1:2) - \boldsymbol{a}^i\right\| \le r_k^i + \epsilon\sigma_n, \quad i \in \mathcal{N}_k \tag{33}$$

where $\boldsymbol{W}_k$ is a symmetric positive-definite (SPD) weighting matrix [33], [35]. One reasonable choice is $\boldsymbol{W}_k = \boldsymbol{\Sigma}_{k|k}^{-1}$, which gives the smallest estimation error covariance matrix when a linear KF is applied to a system with linear dynamic equations and with zero-mean Gaussian observation and excitation noises [37].

The optimization problem in (33) is a QCQP, which is convex since $\boldsymbol{W}_k$ is SPD and the constraints are convex [38, p. 153]. As the constraints are only on the first two elements of the state vector, it is possible to reduce the size of the QCQP problem. A conventional way to do so is as follows. If we suppose that $\boldsymbol{q}(1:2)$ is fixed, then we can find the optimal $\boldsymbol{q}(3:N)$, which is a function of $\boldsymbol{q}(1:2)$. By substituting the optimal $\boldsymbol{q}(3:N)$ into the cost function, we obtain a QCQP, which only involves the unknown $\boldsymbol{q}(1:2)$.

However, in the aforementioned approach, we first need to find the matrix $\boldsymbol{W}_k$ through an inverse operation, which is both unnecessarily costly and numerically unstable if the covariance matrix $\boldsymbol{\Sigma}_{k|k}$ is ill-conditioned. To avoid these shortcomings, we propose to use an idea from [39] to reformulate and reduce the size of the convex QCQP problem in (33) such that it can be solved in a more numerically reliable way. Recalling that $\boldsymbol{\Sigma}_{k|k} = \boldsymbol{U}_{k|k}^T \boldsymbol{U}_{k|k}$, the objective function in (33) can be expressed as $\left(\boldsymbol{q} - \boldsymbol{s}_{k|k}^{(j)}\right)^T \boldsymbol{U}_{k|k}^{-1} \boldsymbol{U}_{k|k}^{-T} \left(\boldsymbol{q} - \boldsymbol{s}_{k|k}^{(j)}\right)$. To get around the inverse operation, we define

$$\boldsymbol{u} = \boldsymbol{U}_{k|k}^{-T}\left(\boldsymbol{s}_{k|k}^{(j)} - \boldsymbol{q}\right) \tag{34}$$

from which it follows that:

$$\boldsymbol{q} = \boldsymbol{s}_{k|k}^{(j)} - \boldsymbol{U}_{k|k}^T \boldsymbol{u}. \tag{35}$$

It is convenient to partition the lower triangular matrix $\boldsymbol{U}_{k|k}^T$ as follows:

$$\boldsymbol{U}_{k|k}^T = \begin{bmatrix} \boldsymbol{L}_{11} & \boldsymbol{0} \\ \boldsymbol{L}_{21} & \boldsymbol{L}_{22} \end{bmatrix} \tag{36}$$

where $\boldsymbol{L}_{11} \in \mathbb{R}^{2\times 2}$ and $\boldsymbol{L}_{22} \in \mathbb{R}^{(N-2)\times(N-2)}$ are lower triangular. Then, it follows from (35) that

$$\boldsymbol{q}(1:2) = \boldsymbol{s}_{k|k}^{(j)}(1:2) - \boldsymbol{L}_{11}\boldsymbol{u}(1:2). \tag{37}$$

Using (34) and (37), we can reformulate the QCQP problem (33) as

$$\begin{aligned} \min_{\boldsymbol{u}} \quad & \left\{\boldsymbol{u}^T(1:2)\boldsymbol{u}(1:2) + \boldsymbol{u}^T(3:N)\boldsymbol{u}(3:N)\right\} \\ \text{s.t.} \quad & \left\|\boldsymbol{L}_{11}\boldsymbol{u}(1:2) - \left(\boldsymbol{s}_{k|k}^{(j)}(1:2) - \boldsymbol{a}^i\right)\right\| \\ & \leq r_k^i + \epsilon\sigma_n, \quad i \in \mathcal{N}_k. \end{aligned} \tag{38}$$

Since the constraints do not include $\boldsymbol{u}(3:N)$, the optimal choice is obviously $\boldsymbol{u}(3:N) = \boldsymbol{0}$, and the optimization problem (38) becomes

$$\begin{aligned} \min_{\boldsymbol{u}(1:2)} \quad & \left\{\boldsymbol{u}^T(1:2)\boldsymbol{u}(1:2)\right\} \\ \text{s.t.} \quad & \left\|\boldsymbol{L}_{11}\boldsymbol{u}(1:2) - \left(\boldsymbol{s}_{k|k}^{(j)}(1:2) - \boldsymbol{a}^i\right)\right\| \\ & \leq r_k^i + \epsilon\sigma_n, \quad i \in \mathcal{N}_k. \end{aligned} \tag{39}$$

This 2-D convex QCQP problem can now be efficiently solved using iterative techniques [38].

After finding the optimal $\boldsymbol{u}(1:2)$, we can compute the optimal $\boldsymbol{q}$ using (35) and the fact that the optimal $\boldsymbol{u}(3:N) = \boldsymbol{0}$ as follows:

$$\mathcal{P}\left(\boldsymbol{s}_{k|k}^{(j)}\right) \triangleq \boldsymbol{q} = \boldsymbol{s}_{k|k}^{(j)} - \begin{bmatrix} \boldsymbol{L}_{11} \\ \boldsymbol{L}_{21} \end{bmatrix} \boldsymbol{u}(1:2). \tag{40}$$

The aforementioned approach for reducing the size of the QCQP problem (33) not only avoids a matrix inverse computation, which may cause numerical instability (see [39]), but it is also computationally efficient. This approach is even more suitable when a SRUKF is employed since the Cholesky factor $\boldsymbol{U}_{k|k}$ of $\boldsymbol{\Sigma}_{k|k}$ is readily provided in (31). We note that in some particular scenarios, particularly under FA in NLOS identification, it is possible that the feasible region $\mathcal{D}_k$ in (6) becomes empty, and consequently, (39) has no solution. In this case, we simply propose to increase $\epsilon$ until $\mathcal{D}_k$ becomes nonempty.

After finding the projected sigma points through (40), the mean and covariance matrix may be estimated through weighted averaging

$$\boldsymbol{s}_{k|k}^{\mathcal{P}} = \sum_{j=0}^{2N} w^{(j)} \mathcal{P}\left(\boldsymbol{s}_{k|k}^{(j)}\right) \tag{41}$$

$$\boldsymbol{\Sigma}_{k|k}^{\mathcal{P}} = \sum_{j=0}^{2N} w^{(j)} \left(\mathcal{P}\left(\boldsymbol{s}_{k|k}^{(j)}\right) - \boldsymbol{s}_{k|k}^{\mathcal{P}}\right)\left(\mathcal{P}\left(\boldsymbol{s}_{k|k}^{(j)}\right) - \boldsymbol{s}_{k|k}^{\mathcal{P}}\right)^T. \tag{42}$$

As before, instead of (42), we compute the Cholesky factor $\boldsymbol{U}_{k|k}^{\mathcal{P}}$ of $\boldsymbol{\Sigma}_{k|k}^{\mathcal{P}}$, i.e.,

$$\begin{aligned} \mathbf{e}_{\mathcal{P}}^{(j)} &= \sqrt{w^{(j)}}\left(\mathcal{P}\left(\boldsymbol{s}_{k|k}^{(j)}\right) - \boldsymbol{s}_{k|k}^{\mathcal{P}}\right), \qquad j = 0, \ldots, 2N \\ \boldsymbol{U}_{k|k}^{\mathcal{P}} &= \text{qr}\left\{\left[\mathbf{e}_{\mathcal{P}}^{(0)}, \mathbf{e}_{\mathcal{P}}^{(1)}, \ldots, \mathbf{e}_{\mathcal{P}}^{(2N)}\right]^T\right\}. \end{aligned} \tag{43}$$

As shown in Fig. 1, we note that the projected sigma points generally have a different mean and covariance matrix. The weighted average of the sigma points achieved through this technique lies inside the feasible region since the average of selected points in a convex feasible region must lie in it [40]. Furthermore, the covariance matrix of the error is generally reduced as the sigma points have moved closer to each other.

Finally, in the next iteration of the unconstrained SRUKF, the constrained *a posteriori* state estimate $\boldsymbol{s}_{k|k}^{\mathcal{P}}$ and the Cholesky factor of the corresponding error covariance matrix $\boldsymbol{U}_{k|k}^{\mathcal{P}}$ replace $\boldsymbol{s}_{k|k}$ and $\boldsymbol{U}_{k|k}$, respectively, as

$$\boldsymbol{s}_{k|k} = \boldsymbol{s}_{k|k}^{\mathcal{P}} \tag{44}$$

$$\boldsymbol{U}_{k|k} = \boldsymbol{U}_{k|k}^{\mathcal{P}}. \tag{45}$$

### C. Algorithm Summary and Computational Analysis

The proposed CSRUKF algorithm, which is summarized in Algorithm 1, consists of two main stages: modified version of the SRUKF and projection of sigma points, which are discussed in more detail below.

**Algorithm 1** CSRUKF

1: Initialize $\boldsymbol{s}_{0|0}$ and set $\boldsymbol{\Sigma}_{0|0}$ to a large SPD diagonal matrix.
2: Set $\eta_\alpha$ and $\epsilon$
3: **for** $k = 1, \ldots, K$ **do**
4: Prediction of $\boldsymbol{s}_{k|k-1}$ using (11), and $\boldsymbol{U}_{k|k-1}$ using (14).
5: **if** $|\mathcal{L}_k| = 0$ **then**
6: Set $\boldsymbol{s}_{k|k} = \boldsymbol{s}_{k|k-1}$ and $\boldsymbol{U}_{k|k} = \boldsymbol{U}_{k|k-1}$.
7: **else**
8: Find the predicted measurement through (16).
9: Calculate the predicted mean (17) and implement qr$\{.\}$ in (22).
10: Estimate the cross-covariance in (18).
11: Solve (27) to find $\boldsymbol{T}_k$.
12: Estimate the *a posteriori* mean $\boldsymbol{s}_{k|k}$ using (28) and Cholesky factor of *a posteriori* covariance matrix $\boldsymbol{U}_{k|k}$ using (31).
13: **end if**
14: Generate the sigma points using (32).
15: For every sigma point whose first two elements fall outside $\mathcal{D}_k$ solve (39) and find the projected point (40).
16: Estimate $\boldsymbol{s}_{k|k}^{\mathcal{P}}$ using (41) and $\boldsymbol{U}_{k|k}^{\mathcal{P}}$ using (43).
17: Replace $\boldsymbol{s}_{k|k}^{\mathcal{P}}$ and $\boldsymbol{U}_{k|k}^{\mathcal{P}}$ as the *a posteriori* estimates, i.e., (44) and (45).
18: **end for**

The SRUKF is more efficient and numerically stable than UKF, and the computational complexity analysis has also been presented in [23], in which it is shown that this algorithm requires $O(N^3)$, where $N$ is the size of the state vector. However, the cost of the first stage of our algorithm is generally small compared to the cost of the projection operations in the second stage.

The QCQP in (39) is a convex optimization problem, which is not NP-hard [38, p.153], and can be solved in polynomial time using an extended optimization package in MATLAB such as SeDuMi [41]. Since $\boldsymbol{u}(1:2) \in \mathbb{R}^2$, the optimization

problem can be solved with moderate cost for $2N+1$ sigma points at most. However, these calculations can be performed in parallel and independently of each other; hence, our technique is suitable for parallel processing. The computational cost of the algorithm depends on the number of sigma points in (32) that fall inside the feasible region, as the projection operation needs not to be applied on them. By tuning the parameter $\alpha$, we can achieve a tradeoff between accuracy and computational cost. On the one hand, if $\alpha$ is small, then it is more likely that many sigma points will fall inside the feasible region, resulting in a lower computational cost. However, selecting a small $\alpha$ may degrade the localization performance as the estimated quantities remain unchanged after applying the constraints. On the other hand, selecting a large $\alpha$ increases the computational cost but, at the same time, may result in sampling many of the nonlocal points, and thus, the linearization of $\boldsymbol{h}(\boldsymbol{s}_k)$ might be inaccurate [34]. In our simulations, it is observed that selecting $0.65 \leq \alpha \leq 0.85$ can offer a reasonable tradeoff in terms of accuracy and computational cost.

## IV. Simulation Results

The simulations are implemented in MATLAB 2010b on a 64-bit computer with Intel i7-2600 3.4-GHz processor and 12 GB of RAM. We consider a 2-D area with $M=5$ fixed RNs located at known positions $\boldsymbol{a}^1=[0,0]$, $\boldsymbol{a}^2=[2000,0]^T$, $\boldsymbol{a}^3=[0,2000]^T$, $\boldsymbol{a}^4=[-2000,0]^T$, and $\boldsymbol{a}^5=[0,-2000]^T$, where the units are in meters. A mobile agent moves on this 2-D plane according to the motion model considered earlier in (1) with noise covariance matrix $\boldsymbol{Q}=0.04\boldsymbol{I}_2$, and the time step duration is set to $\delta t=0.2$ s for $K=1000$ time instants. The initial MN state vector, including the position and velocity components, is normally distributed with zero mean and covariance matrix $\mathrm{diag}([10^4,10^4,10^2,10^2])$.

To model the range measurement, the true distance between each RN and MN is perturbed with an additive zero-mean Gaussian noise. We consider two different measurement noise scenarios: large noise with standard deviation $\sigma_n=150$ m and small noise with $\sigma_n=15$ m, where in our algorithm, we assume that these values are known.[7] The large noise assumption can model general applications such as narrow-band cellular mobile positioning as considered in [17], [42], whereas the small noise assumption is suitable for localization applications with accurate ranging, e.g., IEEE 802.15.4.a. Note that the accuracy of UWB ranging can be improved by increasing the bandwidth of the system [43]. We also perturb some of the measurements by NLOS biases that are modeled as exponential random variables with parameter $\gamma=500$ m [18].[8]

To consider the possible transition of an RN from LOS to NLOS and *vice versa*, we assume that the status of each RN can change with a certain probability after every 250 time instants. This assumption is reasonable and in line with [9] and [44] as the channel conditions might not change drastically for an MN moving with moderate speed. We consider three different scenarios as follows where the transition from LOS to NLOS and *vice versa* is done with probability of 0.5.

- Scenario I: There are four NLOS RNs all the time, whereas the other RN (the one in the center of the plane) can change between LOS and NLOS.
- Scenario II: There are three NLOS RNs all the time, whereas the other two RNs can transit between LOS and NLOS.
- Scenario III: There are two NLOS RNs all the time, whereas the other three RNs can change between LOS and NLOS.

For the proposed CSRUKF, we consider $\epsilon=3$ and $\alpha=70\%$, which corresponds to $\eta_\alpha=4.8784$ under the Gaussian posterior pdf assumption. Note that for the CSRUKF, all the sigma points violating the constraints are projected onto the feasible region. For solving the QCQP problem, we use the optimization toolbox `YALMIP` [45] and `SeDuMi` solver [41].

To see if projecting all the sigma points is necessary to achieve a good result in NLOS scenarios, we first consider the common projection technique where only the *a posteriori* state estimate of a KF is projected onto the feasible region [37]. Therefore, $\boldsymbol{s}_{k|k}$ obtained through the SRUKF is projected onto the feasible region; thus, the new *a posteriori* state estimate satisfies the constraints. However, the *a posteriori* covariance matrix is not changed and remains the same as the unconstrained case. This approach has, in general, a lower computational cost compared to the proposed CSRUKF algorithm since at most one projection operation needs to be done at each iteration. We denote this approach by the projection KF (PKF), and for solving the optimization problem, we follow the similar procedure as for the CSRUKF.

For comparison purposes, we consider the conventional techniques proposed in [8], [44], and [46], in which the range measurements are processed using a KF, and then, the smoothed range measurements are used in an EKF, where the diagonal elements of the covariance matrix corresponding to the NLOS measurements are scaled for further mitigation of NLOS bias. While these techniques differ slightly in terms of preprocessing and variance calculation, we consider the simple one in [44] denoted by the smooth EKF (SEKF) with scaling factor 1.5 and assume that the NLOS identification and variance calculation are done without error.

The Cramer–Rao lower bound analysis in NLOS shows that if no prior statistics about the distribution of the NLOS bias are available, then the optimal strategy is to discard the NLOS measurements and only use LOS ones [47]. If prior statistics are available, then the NLOS measurements should also be used to achieve a lower mean square error. However, this bound can only be practical if there are enough LOS measurements for unambiguous localization; hence, for a small number of LOS links, it may not be useful. Although the posterior Cramer–Rao bound (PCRB) on positioning root mean square error (RMSE) has been derived approximately in [48] and [49], these derivations are based on the assumption that the NLOS bias has a Gaussian distribution with known mean and variance.

[7] In practice, knowledge about $\sigma_n$ can be obtained by means of preliminary calibration experiments in a given environment or through online calculation based on the path-loss model for radio propagation.

[8] Our algorithm can still work well with a range-dependent NLOS bias model as considered in [28]; however, we avoid considering this case due to the lack of space.

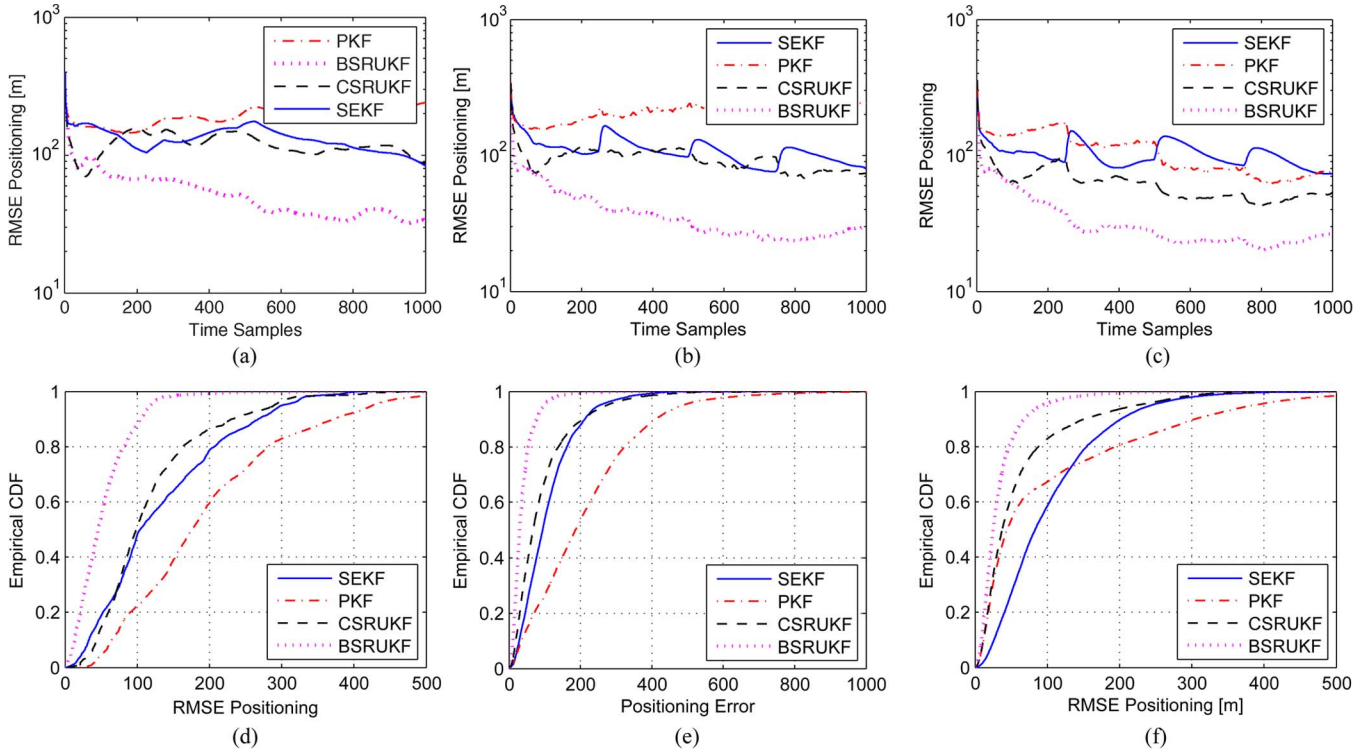


Fig. 2. Comparison of different techniques for large measurement noise with $\sigma_n = 150$ m and exponentially distributed NLOS bias with parameter $\gamma = 500$ m. (a) RMSE for scenario I. (b) RMSE for scenario II. (c) RMSE for scenario III. (d) CDF for scenario I. (e) CDF for scenario II. (f) CDF for scenario III.

Evaluating the PCRB for other NLOS distributions such as exponential is even more challenging. Since in this paper there is no information about the distribution of the NLOS biases, except that they are positive, the mentioned lower bound is still loose and cannot accurately show the lowest possible error in estimating the state vector.

Due to these limitations in finding a lower bound on the positioning RMSE, we consider a semi-ideal situation where the mean and variance of the NLOS bias of each link are known. To apply a KF to this case, the mean of the bias is subtracted from each NLOS measurement, and the error covariance matrix $\boldsymbol{R}_r$ of the measurement vector $\boldsymbol{r}_k = [r_k^1, r_k^2, \ldots, r_k^M]^T$ is scaled according to the variance of the corresponding NLOS bias. Then we apply an unconstrained SRUKF to a dynamic system with the same state motion model as in (1) and with an unbiased set of measurements. For instance, if $i \in \mathcal{N}_k$, then $\boldsymbol{R}_r(i,i) = \sigma_n^2 + \sigma_b^2$, where $\sigma_b^2$ is the variance of the NLOS bias. Note that after subtracting the mean of the NLOS bias from each NLOS range measurement, the remaining error is a combination of a shifted exponentially distributed variable with zero mean and a zero-mean Gaussian noise. Therefore, if the error is dominated by the measurement noise, i.e., $\sigma_n^2 \gg \sigma_b^2$, then nonlinear KFs give nearly MMSE estimation performance for moderately nonlinear systems. However, if the error is dominated by the NLOS bias, i.e., $\sigma_b^2 \gg \sigma_n^2$, these filters are unlikely to give nearly optimal performance in the MMSE sense. Although this approach, which is denoted by bias-aware SRUKF (BSRUKF), is not optimal in the MMSE sense and may not be even a performance lower bound for our technique when the mean and variance of the NLOS bias are known, it can be regarded as a useful benchmark for comparison with our method.

To evaluate the performance of the algorithms in different scenarios, we perform $T = 100$ MC trials for each scenario and consider different trajectories at each trial. Let $\boldsymbol{x}_k^t$ and $\boldsymbol{x}_{k|k}^t$ denote the true MN position and its estimated vectors at the $k$th time step of the trajectory over the $t$th MC trial, respectively. The performance metrics are the cumulative distribution function (cdf) of the positioning error $e_k$, which are expressed as

$$\text{cdf}(e_k) = \mathbb{P}\left[\left\|\boldsymbol{x}_k^T - \boldsymbol{x}_{k|k}^T\right\| \leq e_k\right] \tag{46}$$

and the RMSE of the position estimates at time step $k$, which is defined as

$$\bar{e}_k = \sqrt{\mathbb{E}\left[\left(\boldsymbol{x}_k^T - \boldsymbol{x}_{k|k}^t\right)^T \left(\boldsymbol{x}_k^T - \boldsymbol{x}_{k|k}^t\right)\right]} \tag{47}$$

where $\mathbb{P}$ and $\mathbb{E}$, which are the probability function and expectation operator, respectively, are evaluated approximately using MC trials.

In the following, we compare the effect of measurement noise and NLOS bias on the performance of different techniques in each considered scenario. We assume that the initial estimate $\boldsymbol{s}_{0|0}$ is normally distributed with mean equal to the true state $\boldsymbol{s}_0$ and covariance matrix $\boldsymbol{\Sigma}_{0|0} = \text{diag}([9 \times 10^4, 9 \times 10^4, 10^3, 10^3])$.

### A. Large Measurement Noise

In the first scenario, we consider the case of a narrow-band ranging application where the noise variance is relatively high, i.e., $\sigma_n = 150$ m is considered. The RMSE versus time step is

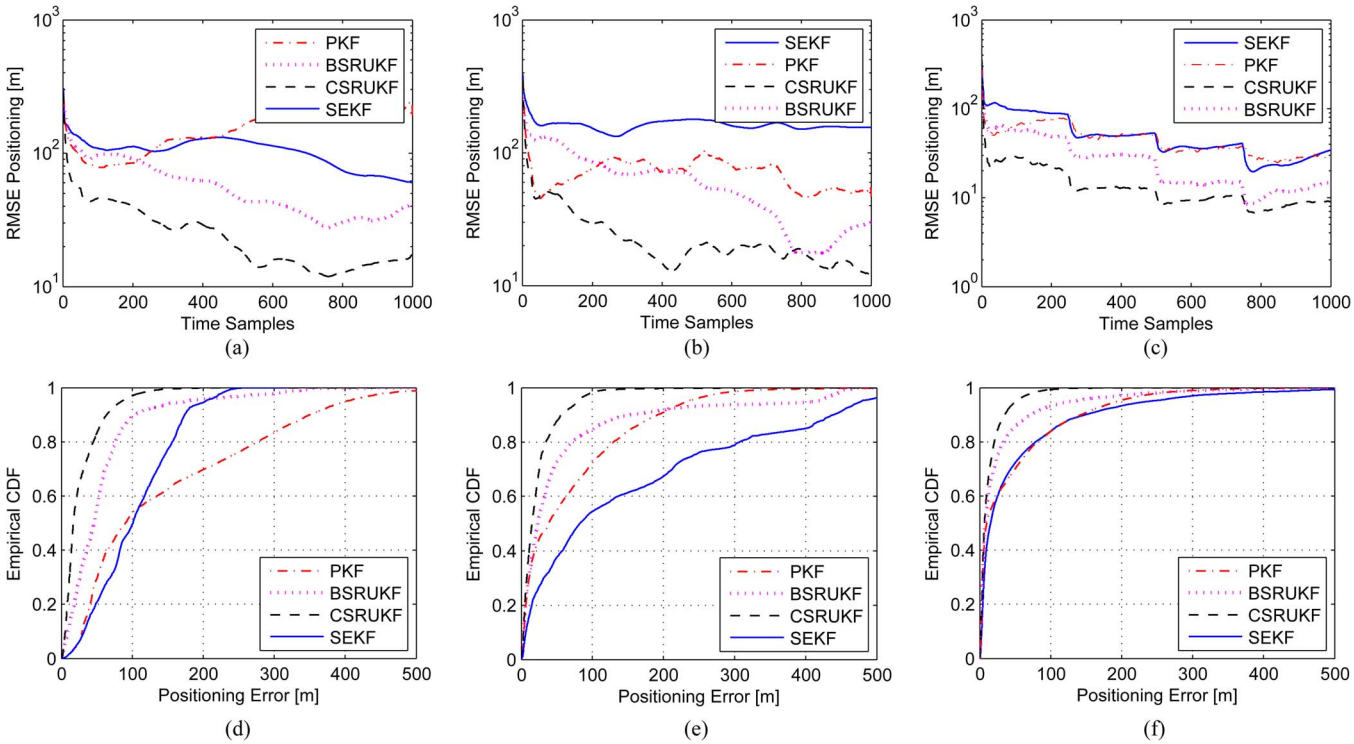


Fig. 3. Comparison of different techniques for small measurement noise $\sigma_n = 15$ m and with exponentially distributed NLOS bias with parameter $\gamma = 500$ m. (a) RMSE for scenario I. (b) RMSE for scenario II. (c) RMSE for scenario III. (d) CDF for scenario I. (e) CDF for scenario II. (f) CDF for scenario III.

shown in Fig. 2 for scenarios I, II, and III. The corresponding cdf of the positioning error is also plotted for each scenario.

We can observe that for scenarios I and II, the CSRUKF performs almost similar to the SEKF, whereas the RMSE of the PKF is relatively high. This shows that to obtain a decent localization performance, the projection of all the sigma points in the CSRUKF is necessary as compared to projecting only the mean as done in the PKF. The RMSE of all the techniques are lower bounded by the RMSE of the BSRUKF, which uses more prior information about the NLOS biases. The RMSE and cdf of scenario III indicate that the performance of the CSRUKF is better than those of the PKF and SEKF noticeably.

### B. Small Measurement Noise

For further verification of our algorithm, we consider a case where the noise variance is relatively small, i.e., $\sigma_n = 15$ m, which can model errors in UWB ranging applications. The RMSE and cdf of the estimation error are shown in Fig. 3 for scenarios I, II, and III.

As observed in Fig. 3, in all the scenarios, the proposed CSRUKF performs better than all the other methods, particularly the BSRUKF. There are several reasons why the CSRUKF can outperform the BSRUKF to this extent for small measurement noise: First, the BSRUKF cannot necessarily provide a performance lower bound, since after removing the mean of the bias from the NLOS range measurements, the remaining error term does not follow a Gaussian distribution; hence, applying a KF to this problem is not the optimum MMSE estimation technique. In small noise scenarios, the NLOS bias dominates over the measurement noise, and therefore, the distribution of the error in the BSRUKF is far from the Gaussian distribution. For large measurement noise scenarios in Fig. 2, where the NLOS bias is not significantly larger than the measurement noise, the error distribution was closer to a Gaussian one, which was one of the reasons that the BSRUKF was performing better than the CSRUKF. Second, when $\sigma_n$ is large, the feasible region $\mathcal{D}_k$ becomes larger compared to the case that $\sigma_n$ is small. Therefore, it is more likely that most of the sigma points lie inside $\mathcal{D}_k$ and no projection is done; thus, the second stage of our algorithm does not improve the *a posteriori* estimate. Note that by restricting the sigma points to be within a smaller feasible region, a better location estimate may be obtained.

### C. Robustness to Errors in NLOS Identification

In this part, we analyze the performance of our proposed technique in the presence of NLOS identification errors, i.e., FA and MD, which are inevitable in some applications.

To see the effect of FA on the proposed CSRUKF, we assume that we have one LOS and four NLOS RNs. However, due to the FA, the LOS link is also wrongly detected as being NLOS. Therefore, the CSRUKF and PKF wrongly remove the LOS measurements from the measurement vector and employ the wrongly detected measurement to impose a constraint on the state vector. Since the use of a larger $\epsilon$ can increase the chance that a LOS measurement also satisfies the constraint in (5), it is expected that FA does not severely degrade the performance of our proposed technique. The simulation results are shown in Fig. 4, where it is observed that the CSRUKF is robust against FA error in NLOS identification and outperforms the SEKF. Note that the BSRUKF algorithm is evaluated under perfect NLOS identification, whereas its performance is still worse than our proposed technique.

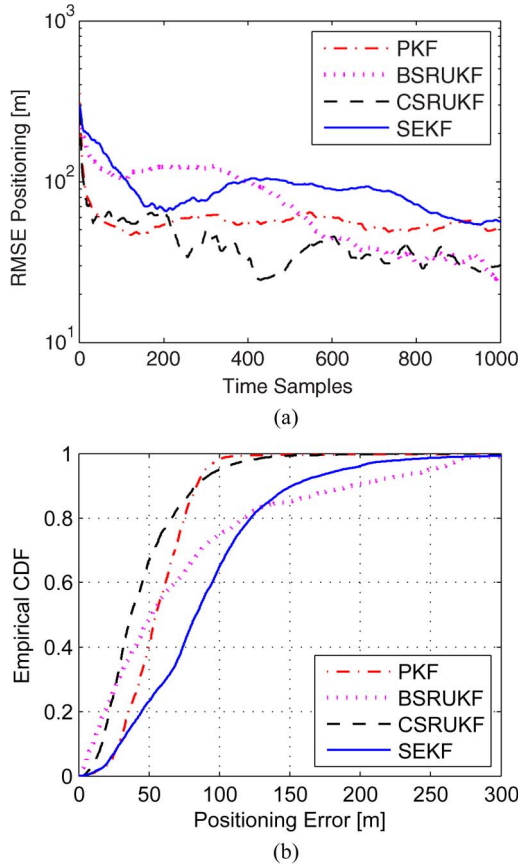


Fig. 4. Comparison of different techniques for $\sigma_n = 15$ m and exponentially distributed NLOS bias with parameter $\gamma = 500$ m and FA in identification of an LOS RN. (a) RMSE. (b) CDF.

If the NLOS links are regarded as LOS ones, i.e., in the presence of NLOS MD error, all the Kalman-type filters have to use a biased measurement in their observation vector, and thus, it is not surprising that their performances are degraded. Therefore, for our algorithm to perform well in most of the times, the threshold used for NLOS identification should change such that the probability of MD becomes very small.

### *D. Computation Time*

The average computation time of each algorithm is calculated for each scenario and is shown in Table I. The SEKF has a very small computation time, because it is essentially an ordinary KF. The computation time of the PKF, where only the state vector needs to be projected onto the feasible region, is obviously lower than the CSRUKF because only one QCQP problem might need to be solved at every time step. The most computationally demanding part of the CSRUKF is the projection of the sigma points, which might be implemented in parallel form. However, in the simulation, we have done these projections sequentially, and therefore, the total elapsed time is larger for the CSRUKF. Still, the highest computation time of the CSRUKF is lower than the total elapsed time for the entire trajectory (200 s), meaning that with the computer used here, the algorithm can be applied for online tracking. Note that the computational cost of many other popular methods, such as the KDE or PFs, is also much higher than the ordinary EKF or SRUKF, and therefore, our algorithm remains competitive in terms of computation time.

TABLE I
AVERAGE RUNNING TIME OF EACH ALGORITHM IN EACH SCENARIO EVALUATED FOR THE ENTIRE TRAJECTORY (200 s) IN SECONDS

| | Scenario I | Scenario II | Scenario III |
|---|---|---|---|
| **SEKF** | 0.38 | 0.41 | 0.45 |
| **PKF** | 25.72 | 10.5935 | 1.9141 |
| **CSRUKF** | 64.03 | 31.4485 | 8.3804 |

## V. CONCLUSION AND FUTURE WORK

A CSRUKF with projection technique was proposed in this paper for the aim of TOA-based localization of an MN in NLOS scenarios. The NLOS measurements were removed from the measurement vector, and instead, they were employed to impose quadratic constraints onto the position coordinates of the MN. The sigma points of the UKF that violated the constraints were projected on the feasible region by solving a convex QCQP. As compared to other constrained UKF techniques, we considered a square-root filter and avoided computing the inverse of the state covariance matrix both in the KF and in the optimization steps; consequently, our approach has better numerical stability and lower computational cost. In the simulation experiments, the proposed CSRUKF performed better than other approaches in different NLOS scenarios. In particular, CSRUKF performance was excellent when a small measurement noise variance was considered, suggesting that it is particularly suitable for high-resolution TOA-based UWB localization. Another advantage of our technique is its robustness to FA errors in NLOS identification. The proposed CSRUKF can be extended to the situations where the information of an IMU is fused with range measurements for more accurate mobile localization. The proposed CSRUKF can also be extended to cooperative tracking of multiple MNs. However, several challenges remain in this respect such as the communication overload and the efficient implementation of a distributed version of the CSRUKF. We leave the investigation of such cooperative localization scenarios to future work.

## ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their helpful comments and suggestions, which improved the quality of this paper.

## REFERENCES

[1] A. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location: Challenges faced in developing techniques for accurate wireless location information," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 24–40, Jul. 2005.

[2] *IEEE Standard for Local and Metropolitan Area Networks-Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs) Amendment 1: MAC Sublayer*, IEEE Std. 802.15.4e-2012, 2012, pp. 1–225.

[3] K. Pahlavan *et al.*, "Indoor geolocation in the absence of direct path," *IEEE Wireless Commun.*, vol. 13, no. 6, pp. 50–58, Dec. 2006.

[4] S. Venkatesh and R. Buehrer, "Non-line-of-sight identification in ultra-wideband systems based on received signal statistics," *IET Microw., Antennas Propag.*, vol. 1, no. 6, pp. 1120–1130, Dec. 2007.

[5] S. Maranò, W. M. Gifford, H. Wymeersch, and M. Z. Win, "NLOS identification and mitigation for localization based on UWB experimental data," *IEEE J. Sel. Area Commun.*, vol. 28, no. 7, pp. 1026–1035, Sep. 2010.

[6] I. Guvenc and C.-C. Chong, "A survey on TOA based wireless localization and NLOS mitigation techniques," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 3, pp. 107–124, 2009.
[7] N. Thomas, D. Cruickshank, and D. Laurenson, "Performance of a TDOA-AOA hybrid mobile location system," in *Proc. 2nd Int. Conf. 3G Mobile Commun. Technol.*, 2001, pp. 216–220.
[8] C.-D. Wann, Y.-J. Yeh, and C.-S. Hsueh, "Hybrid TDOA/AOA indoor positioning and tracking using extended Kalman filters," in *Proc. IEEE Veh. Technol. Conf.-Spring*, May 2006, vol. 3, pp. 1058–1062.
[9] K. Yu and E. Dutkiewicz, "NLOS identification and mitigation for mobile tracking," *IEEE Trans. Aeros. Electron. Syst.*, vol. 49, no. 3, pp. 1438–1452, Jul. 2013.
[10] I. Guvenc, C.-C. Chong, and F. Watanabe, "NLOS identification and mitigation for UWB localization systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2007, pp. 1571–1576.
[11] S. Venkatesh and R. Buehrer, "NLOS mitigation using linear programming in ultrawideband location-aware networks," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 3182–3198, Sep. 2007.
[12] C.-L. Chen and K.-T. Feng, "An efficient geometry-constrained location estimation algorithm for NLOS environments," in *Proc. Int. Conf. Wireless Netw., Commun. Mobile Comput.*, Jun. 2005, vol. 1, pp. 244–249.
[13] J. Hol, F. Dijkstra, H. Luinge, and T. Schon, "Tightly coupled UWB/IMU pose estimation," in *Proc. IEEE Int. Conf. Ultra-Wideband*, Sep. 2009, pp. 688–692.
[14] J. Youssef, B. Denis, C. Godin, and S. Lesecq, "Loosely-coupled IR-UWB handset and ankle-mounted inertial unit for indoor navigation," in *Proc. IEEE Int. Conf. Ultra-Wideband*, Sep. 2011, pp. 160–164.
[15] B.-S. Chen, C.-Y. Yang, F.-K. Liao, and J.-F. Liao, "Mobile location estimator in a rough wireless environment using extended Kalman-based IMM and data fusion," *IEEE Trans. Veh. Technol.*, vol. 58, no. 3, pp. 1157–1169, Mar. 2009.
[16] C. Fritsche, U. Hammes, A. Klein, and A. Zoubir, "Robust mobile terminal tracking in NLOS environments using interacting multiple model algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2009, pp. 3049–3052.
[17] U. Hammes, E. Wolsztynski, and A. Zoubir, "Robust tracking and geolocation for wireless networks in NLOS environments," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 5, pp. 889–901, Oct. 2009.
[18] U. Hammes and A. Zoubir, "Robust MT tracking based on M-estimation and interacting multiple model algorithm," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3398–3409, Jul. 2011.
[19] M. Najar, J. Huerta, J. Vidal, and J. Castro, "Mobile location with bias tracking in non-line-of-sight," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2004, vol. 3, pp. 956–959.
[20] D. Jourdan, J. J. Deyst, M. Win, and N. Roy, "Monte Carlo localization in dense multipath environments using UWB ranging," in *Proc. IEEE Int. Conf. Ultra-Wideband*, Sep. 2005, pp. 314–319.
[21] J. González *et al.*, "Mobile robot localization based on ultra-wide-band ranging: A particle filter approach," *Robot. Auton. Syst.*, vol. 57, no. 5, pp. 496–507, May 2009.
[22] S. Yousefi, X. Chang, and B. Champagne, "Improved extended Kalman filter for mobile localization with NLOS anchors," in *Proc. Int. Conf. Wireless Mobile Commun.*, Jul. 2013, pp. 25–30.
[23] R. Van der Merwe and E. Wan, "The square-root unscented Kalman filter for state and parameter-estimation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2001, vol. 6, pp. 3461–3464.
[24] R. Kandepu, B. Foss, and L. Imsland, "Applying the unscented Kalman filter for nonlinear state estimation," *J. Process Control*, vol. 18, no. 7/8, pp. 753–768, Aug./Sep. 2008.
[25] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.
[26] D. Dardari, R. D'Errico, C. Roblin, A. Sibille, and M. Win, "Ultrawide bandwidth RFID: The next generation?" *Proc. IEEE*, vol. 98, no. 9, pp. 1570–1582, Sep. 2010.
[27] A. Molisch, "Ultrawideband propagation channels—Theory, measurement, modeling," *IEEE Trans. Veh. Technol.*, vol. 54, no. 5, pp. 1528–1545, Sep. 2005.
[28] K. Yu and Y. Guo, "Improved positioning algorithms for nonline-of-sight environments," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2342–2353, Jul. 2008.
[29] R. M. Vaghefi and R. M. Buehrer, "Target tracking in NLOS environments using semidefinite programming," in *Proc. IEEE Mil. Commun. Conf.*, Nov. 2013, pp. 169–174.
[30] L. Doherty, K. Pister, and L. El Ghaoui, "Convex position estimation in wireless sensor networks," in *Proc. IEEE 20th Annu. Joint Conf. Comput. Commun. Soc.*, Apr. 2001, vol. 3, pp. 1655–1663.
[31] D. Simon, "Kalman filtering with state constraints: A survey of linear and nonlinear algorithms," *IET Control Theory Appl.*, vol. 4, no. 8, pp. 1303–1318, Aug. 2010.
[32] D. Simon and D. L. Simon, "Constrained Kalman filtering via density function truncation for turbofan engine health estimation," *Int. J. Syst. Sci.*, vol. 41, no. 2, pp. 159–171, Feb. 2010.
[33] J. Lan and X. Li, "State estimation with nonlinear inequality constraints based on unscented transformation," in *Proc. 14th Int. Conf. Inf. Fusion*, Jul. 2011, pp. 1–8.
[34] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.
[35] D. Simon, *Optimal State Estimation: Kalman, H-Infinity, Nonlinear Approaches*. Hoboken, NJ, USA: Wiley-Interscience, Aug. 2006.
[36] D. Zachariah, I. Skog, M. Jansson, and P. Handel, "Bayesian estimation with distance bounds," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 880–883, Dec. 2012.
[37] D. Simon and D. L. Simon, "Kalman filtering with inequality constraints for turbofan engine health estimation," *Proc. Inst. Elect. Eng. Control Theory Appl.*, vol. 153, no. 3, pp. 371–378, May 2006.
[38] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.
[39] C. Paige, "Computer solution and perturbation analysis of generalized linear least squares problems," *J. Math. Comput.*, vol. 33, no. 145, pp. 171–183, Jan. 1979.
[40] D. Luenberger and Y. Ye, *Linear and Nonlinear Programming*, 3rd ed. New York, NY, USA: Springer-Verlag, 2008.
[41] J. F. Strum, Using SeDuMi 1.02, a MATLAB Toolbox for Optimization Over Symmetric Cones 1998.
[42] U. Hammes and A. Zoubir, "Robust mobile terminal tracking in NLOS environments based on data association," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5872–5882, Nov. 2010.
[43] S. Gezici *et al.*, "Localization via ultra-wideband radios: A look at positioning aspects for future sensor networks," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 70–84, Jul. 2005.
[44] B. L. Le, K. Ahmed, and H. Tsuji, "Mobile location estimator with NLOS mitigation using Kalman filtering," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2003, vol. 3, pp. 1969–1973.
[45] J. Lofberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proc. IEEE Int. Symp. Comput. Aided Control Syst. Design*, Sep. 2004, pp. 284–289.
[46] K. Yu and E. Dutkiewicz, "Improved Kalman filtering algorithms for mobile tracking in NLOS scenarios," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2012, pp. 2390–2394.
[47] Y. Qi, H. Kobayashi, and H. Suda, "Analysis of wireless geolocation in a non-line-of-sight environment," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 672–681, Mar. 2006.
[48] L. Chen, S. Ali-Löytty, R. Piché, and L. Wu, "Mobile tracking in mixed line-of-sight/non-line-of-sight conditions: Algorithm and theoretical lower bound," *Wireless Pers. Commun.*, vol. 65, no. 4, pp. 753–771, Aug. 2012.
[49] C. Fritsche, A. Klein, and F. Gustafsson, "Bayesian Cramer–Rao bound for mobile terminal tracking in mixed LOS/NLOS environments," *IEEE Wireless Commun. Lett.*, vol. 2, no. 99, pp. 335–338, Jun. 2013.

**Siamak Yousefi** (S'12) received the B.Sc. degree in electrical engineering from Iran University of Science and Technology, Tehran, Iran, in 2007 and the M.Sc. degree in communication engineering from Chalmers University of Technology, Gothenburg, Sweden, in 2010.

From March to August 2010, he was a Research Engineer with the School of Electrical Engineering, Royal Institute of Technology (KTH), Stockholm, Sweden. His current research includes cooperative localization of wireless mobile nodes in harsh propagation environments. He is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada.

Mr. Yousefi has received several grants and awards, including the McGill Graduate Research Enhancement and Travel Award, the McGill International Doctoral Award, the McGill Graduate Research Mobility Award, the Top 10 Student Paper Award at the IEEE Radio and Wireless Symposium in 2010, and the Best Paper Award at the IARIA International Conference on Wireless and Mobile Communications in 2013.

**Xiao-Wen Chang** received the B.S. and M.S. degrees in computational mathematics from Nanjing University, Nanjing, China, in 1986 and 1989, respectively, and the Ph.D. degree (with Dean's honors list) in computer science from McGill University, Montreal, QC, Canada, in 1997.

He is currently an Associate Professor with the School of Computer Science, McGill University. He has published about 50 journal papers and 15 conference papers. His current research interests are in the area of scientific computing, with particular emphasis on various least squares methods and their applications in communications, localization in wireless sensor networks, and Global Navigation Satellite System-based positioning. His research has been funded by the Natural Sciences and Engineering Research Council of Canada and the "Fonds de Recherche sur la Nature et les Technologies" from the Government of Quebec.

Dr. Chang is a member of the Editorial Board of *Numerical Algebra, Control and Optimization*.

**Benoit Champagne** (S'87–M'89–SM'03) received the B.Ing. degree in engineering physics from the Ecole Polytechnique de Montréal, Montreal, QC, Canada, in 1983, the M.Sc. degree in physics from the Université de Montréal in 1985, and the Ph.D. degree in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 1990.

From 1990 to 1999, he was an Assistant and then Associate Professor with INRS-Telecommunications, Université du Quebec, Montreal. In 1999, he joined McGill University, Montreal, where he is now a Full Professor within the Department of Electrical and Computer Engineering. He served as Associate Chairman of Graduate Studies of the Department from 2004 to 2007. His research focuses on the study of advanced algorithms for the processing of information bearing signals by digital means. His interests span many areas of statistical signal processing, including detection and estimation, sensor array processing, adaptive filtering, and applications thereof to broadband communications and audio processing, where he has coauthored nearly 200 referred publications. His research has been funded by the Natural Sciences and Engineering Research Council of Canada, the "Fonds de Recherche sur la Nature et les Technologies" from the Government of Quebec, as well as some major industrial sponsors, including Nortel Networks, Bell Canada, InterDigital, and Microsemi.

Dr. Champagne has served on the Technical Committees of several international conferences in the fields of communications and signal processing. In particular, he was Co-Chair, Wide Area Cellular Communications Track, for the IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (Toronto, September 2011); Co-Chair, Antenna and Propagation Track, for the IEEE Vehicular Technology Conference Fall (Los Angeles, CA, USA, September 2004); and Registration Chair for the IEEE International Conference on Acoustics, Speech and Signal Processing (Montreal, May 2004). He has been an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the *EURASIP Journal on Applied Signal Processing*.