Network-based RF Localization in Adverse Propagation Environments

Siamak Yousefi

Department of Electrical & Computer Engineering
McGill University
Montreal, Canada

August 2015

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

2015/08/14

Abstract

Network-based radio frequency (RF) localization has gained considerable attention in the past decade due to the tremendous number of applications where location information is required while conventional localization systems, such as the global positioning system (GPS), can not be employed. Positioning in harsh propagation environments such as dense urban areas, indoor places and underground areas, where the GPS satellites are not visible to the receiver and the GPS signal is attenuated, are among such applications. Network-based RF localization, where instead of satellites, a set of fixed reference nodes transmit and/or receive signal to/from a wireless device (target) can overcome the limitations of the GPS. These reference nodes could be base stations (BSs) in a cellular network, or anchors in a wireless sensor network (WSN). Since the number of reference nodes is limited and might not be sufficient for unambiguous localization, it becomes necessary and beneficial for the targets to make pairwise measurements and exchange information with their neighbours; consequently cooperative localization has gained much attention. Nevertheless, localization in harsh propagation environments is challenging and can lead to large errors due to the multipaths and non-line of sight (NLOS) propagation. The problem of multipaths can be overcome using high resolution ultra wide-band (UWB) timing pulses so that accurate time of arrival (TOA) measurements can be obtained. However, the NLOS problem, in which the range measurements become positively biased, still remains the main challenge.

The first contribution of this thesis focuses on mobile localization in NLOS using TOA measurements, where a constrained square-root unscented Kalman filter (CSRUKF) is developed. While the proposed filter is based on a constrained unscented Kalman filter (UKF) with sigma point projection, its efficiency and numerical stability are improved by using the idea of square-root unscented Kalman filter (SRUKF). The proposed filter is also extended to a cooperative localization scenario where a centralized CSRUKF is employed for multi-target tracking in NLOS scenarios. The simulation results illustrate that the proposed filter can yield a good localization accuracy in severe NLOS situations and is also robust against false alarm (FA) in NLOS identification.

The centralized techniques have limitation in computations and are not scalable with the size of the network. Therefore, distributed localization techniques are preferred in many applications and this is the second contribution of this thesis. To this end, a two-stage robust distributed algorithm is proposed for cooperative sensor network localization

using TOA data without prior NLOS identification. In the first stage, to overcome the effect of outliers, a convex relaxation of the Huber loss function is applied so that by using iterative optimization techniques, coarse estimates of the true sensor locations can be obtained. In the second stage, the original (non-relaxed) Huber cost function is further minimized to obtain refined location estimates based on the estimated positions obtained in the first stage. Through simulations and real data analysis, it is shown that the proposed algorithm can achieve a lower root mean squared error (RMSE) compared to other existing algorithms, and can achieve a performance close to that of an idealized approach which assumes *a priori* knowledge of the NLOS links.

The third contribution of this thesis is to study a geometric problem in WSNs, where the TOA measurements are positively biased, thus every target is restricted to lie inside the intersection of several disks, forming a convex hull. To quantify the size of these convex hulls, iterative techniques based on the sum-product algorithm over wireless network (SPAWN) have been developed in the literature, where a crucial step in the algorithm is to find a tight ellipsoidal outer approximation (OA) of the intersection of ellipses. There exist sub-optimal convex optimization techniques in the literature to solve this problem, however they are not always tight. To overcome this limitation, two novel techniques are developed and studied to find tighter OAs in 2-dimensional space. Through simulations, it is shown that SPAWN using the proposed ellipsoidal OA converges to some limit rapidly and offers a tighter estimate of the convex hulls as compared to the case when conventional ellipsoidal OA is used.

The new localization algorithms developed in this thesis are well suited to the next generation of cellular and WSN positioning systems, where efficient distributed techniques that are robust against NLOS propagation will be needed to support various applications.

Sommaire

La localisation par fréquence radio (RF) dans les réseaux sans fil a attiré une attention considérable dans la dernière décennie en raison de grand nombre d'applications qui ont besoin d'information de localisation, tandis que les systèmes de localisation classiques tels que le système mondial de localisation (GPS) ne peuvent pas être utilisés. Parmi ces applications, nous notons le positionnement dans les environnements avec conditions de propagation radio d'afordables, tels que les zones urbaines denses ainsi que les lieux intérieurs et souterrains où les satellites GPS ne sont pas visibles au récepteur et où le signal GPS est considérablement atténué. La localisation RF basée sur les réseau terrestriels, qui au lieu de stellites, utilise un ensemble de points de référence radio fixes pour transmettre et/ou recevoir des signaux vers/à partir d'un appareil sans fil (cible) peut surmonter les limitations du GPS. Les points de référence pouvent être des stations de base (BSs) dans un réseau cellulaire, ou des ancres dans un réseau de capteurs sans fil (WSN). Puisque le nombre de points de référence est limité et pourrait ne pas être suffisant pour la localisation sans ambigüité, il serait nécessaire et bénéfique que les cibles de localisation prennent des mesures par paires et échangent de l'information avec leurs voisins; par conséquent, la localisation coopérative a attiré beaucoup d'attention au cours des dernières année. Néanmoins, la localisation dans des environnements avec propagation difficile est problématique et peut mener à des erreurs importantes en raison de propagation multi-trajet et sans ligne de vue (NLOS). Le problème de propagation multi-trajet peut être surmonté en utilisant des impulsions de synchronisation haute résolution à ultra large bande (UWB), de telle sorte que des mesures exactes des temps d'arrivé (TOA) peuvent être obtenues. Cependant, le problème de locailsation NLOS, dans lequel les mesures de la distance deviennent positivement biasées, reste toujours le défi principal.

La première contribution de cette thèse met l'accent sur la localisation mobile NLOS en utilisant des mesures TOA, où un nouveau filtre de Kalman, de type racine-carrè sans parfum (unscented) avec contrainte (CSRUKF) est développé. Alors que le filtre proposé est basé sur le filtre contraint de Kalman sans parfum (UKF) avec projection des points sigma, son efficacité et sa stabilité numérique sont améliorées en utilisant l'idée de filtre de Kalman sans parfum à racine carrée (SRUKF). Le filtre proposé est ensuite utilisé dans un scénario de localisation coopérative où un CSRUKF centralisé est utilisé pour le suivi multi-cibles dans les scénarios NLOS. Les résultats des simulations démontrent que le filtre

proposé peut donner une bonne précision de localisation dans les situations NLOS sévères et est également robuste contre les fausses alarmes (FA) aux identifications NLOS.

Les techniques centralisées souffrent de la complexité calculs et ne sont pas extensibles en proportion à la taille du réseau. Par conséquent, les techniques de localisation distribuées sont préférables dans de nombreuses applications, et elle constituent la deuxième contribution de cette thèse. Plus particulièrement, un algorithme robuste distribué en deux étapes est proposé pour la localisation coopérative avec réseau de capteurs en utilisant les données TOA sans identification préalable NLOS. Dans la premiére étape, pour surmonter l'effet des donnés extrèmes, une relaxation convexe de la fonction de perte Huber est appliquée de telle sorte que, en utilisant des techniques d'optimisation itérative, des estimés approximatifs des vrais emplacements de capteurs soient obtenues. Dans la deuxiéme étape, la fonction d'origine de coût Huber (non relaxée) est minimisée afin d'obtenir des meilleurs estimés de localisation par rapport à ceux obtenus dans la premiére étape. Par des simulations et l'analyse de données réelle, il est démontré que l'algorithme proposé peut atteindre une racine de l'erreur quadratique moyenne (RMSE) plus faible que celle des algorithmes existants, et ainsi atteindre une performance se rapprochant du cas idéal dans lequel la connaissance *a priori* des liens NLOS est disponible.

La troisième contribution de cette thèse porte sur l'analyse d'un problème géométrique dans les réseaux de capteurs où les mesures de TOA sont biaisées positivement en raison d'une condition NLOS. Ainsi, la position de chacune des cibles est contrainte de se trouver à l'intérieur de l'intersection de plusieurs disques, formant une coque convexe. Pour quantifier la taille de ces enveloppes convexes, des techniques itératives basées sur l'algorithme somme-produit sur le réseau sans fil (SPAWN) ont été développées dans la littérature. Un élément crucial de cet algorithme est de trouver un rapprochement extérieur (outer approximation ou OA) ellipsoïdal serré de l'intersection de plusieurs ellipses. Il existe des techniques sous-optimales d'optimisation convexe dans la littérature pour résoudre ce problème, mais les solutions ne sont pas toujours serrées. Pour surmonter cette limitation, deux nouvelles techniques sont développées et étudiées pour trouver des OAs plus serrées dans lespace à deux dimensions. Les simulations démontrent que l'algorithme SPAWN qui utilise ces nouvelles techniques OA ellipsoïdales converge rapidement vers une limite et offre une estimation plus serrée des enveloppes convexes par rapport au cas où OA ellipsoïdale classique est utilisée.

Les nouveaux algorithmes de localisation développés dans cette thèse sont bien adaptés

à la prochaine génération de systèmes de positionnement cellulaires et WSN, où des techniques distribuées efficaces qui sont robustes contre la propagation NLOS seront nécessaires afin de soutenir diverses applications.

Acknowledgments

I would like to express my gratitude and appreciation towards my supervisors, Professor Xiao-Wen Chang and Professor Benoit Champagne for helping me travel through this journey. Without the useful discussions with Professor Chang, the progress of my research would not have been possible, while the feedbacks of Professor Champagne on the papers have definitely played an important role in improving the quality of my work. I would like to thank them for supporting me financially during these years and paying the expenses for travelling to conference venues, which has been of great help in my research. Their flexibility towards collaboration with other laboratories has also allowed me to take advantage of the knowledge of some other experts in my research area and thus progress effectively.

I am grateful to know Professor Henk Wymeersch of Chalmers University of Technology in Sweden with whom I had the pleasure to collaborate and write papers. His expertise in the area of distributed localization has helped me significantly in developing my ideas. I would also like to thank Professor Mark Coates and Professor Ioannis Psaromiligkos for their role as my PhD supervisory committee. I would also like to thank other faculty members of the Department of Electrical and Computer Engineering (ECE), especially Professor Michael Rabbat with whom I had useful discussions.

I would like to thank all the students of the Statistical Signal Processing Laboratory in the Department of ECE, as well as those of the Scientific Computing Laboratory in the School of Computer Science, with whom I have had such good interactions.

And last but not least, I would like to thank my family who has been always supportive of me to finish this journey: My father, Dr Ahmadali Yousefi who has always encouraged me to obtain high education and has supported me financially; My mother, Dr Mina Jamzad who always gave me hope to keep on trying hard and not being disappointed by the defeats. Without their encouragement I might have quit the program at some point due to the difficulties I was encountering. My brother has also been a good advisor for me as he went through a similar path during his PhD studies.

In the end, I would like to dedicate this PhD thesis to my family.

Preface and Contribution of the Authors

The research presented in this dissertation was carried out in the Department of Electrical and Computer Engineering (ECE) of McGill University from September 2010 to March 2015. This dissertation is the result of my original work and created under the supervisions of Prof. Benoit Champagne from the same department, and Prof. Xiao-Wen Chang from the School of Computer Science at McGill University. The research results obtained during the first four years of my PhD thesis were published in a journal (J-1), and several well-known conferences (C-1), (C-2), (C-3), (C-4). In the last year of My PhD studies I started collaboration with Prof. Henk Wymeersch of Chalmers University of Technology in Sweden, one of the leading researchers in the area of distributed sensor network localization using Bayesian inference techniques, who also provided insightful and useful feedback about my research. Together with him and my supervisors, we submitted a letter (L-1) and are preparing another journal paper (J-2). I also collaborated with Reza Monir Vaghefi and Prof. Michael Buehrer of Virginia Tech in USA on joint localization and synchronization techniques in wireless sensor networks under non-line of sight situations. We wrote a joint paper, which has been recently presented in the conference (C-5). The full list of my publications during my PhD studies is given below:

- **Journal Papers:**

  (J-1) S. Yousefi, X.-W. Chang, and B. Champagne, "Mobile Localization in Non-Line-of-Sight Using a Constrained Square Root Unscented Kalman Filter", in *IEEE Trans. on Vehicular Tech.*, vol. 64, pp. 2071-2083, May 2015.

  (J-2) S. Yousefi, X.-W. Chang, and B. Champagne, H. Wymeersch, "Ellipse Outer-Approximation of the Intersection Region of Multiple Ellipses", in preparation for submission to *Journal of Computational Geometry*.

- **Letters:**

  (L-1) S. Yousefi, H. Wymeersch, X.-W. Chang, and B. Champagne, "Tight Outer-approximation of the Uncertainty Regions in Wireless Sensor Network", *IEEE Communication Letters*, to be submitted.

- **Conference Papers:**

(C-1) S. Yousefi, X.-W. Chang, and B. Champagne, "An Improved Extended Kalman Filter for Localization of a Mobile Node with NLOS Anchors", in *Proc. Int. Conf. on Wireless Mobile Radio Communication*, pp. 25-30, Jul. 2013, Nice, France, (**best paper award**).

(C-2) S. Yousefi, X.-W. Chang, and B. Champagne, "A Joint Localization and Synchronization Technique Using Time of Arrival at Multiple Antenna Receivers", in *Proc.* 44^{th} *Asilomar Conf. on Signals, systems, and Computers*, pp. 2017-2021, Nov. 2013, Pacific Grove, USA.

(C-3) S. Yousefi, X.-W. Chang, and B. Champagne, "Distributed Cooperative Localization without NLOS Identification", in *Proc. IEEE Wireless Positioning and Networking Conf. (WPNC)*, pp. 1-6, Mar. 2014, Dresden, Germany.

(C-4) S. Yousefi, X.-W. Chang, and B. Champagne, "Cooperative Localization of Mobile Nodes in NLOS", in *Proc.* 25^{th} *Annual Int. Symp. on Personal Indoor and Mobile Radio Communication (PIMRC)*, Sep. 2014, Washington D.C., USA.

(C-5) S. Yousefi, R. M. Vaghefi, X.-W. Chang, B. Champagne, and M. Buehrer, "Sensor Localization in NLOS Environments with Anchor Uncertainty and Unknown Clock Parameters", in *Proc.* 3^{rd} *ICC Workshop on Advances in Network Localization and Navigation*, Jun. 2015, London, UK.

Contents

List of Figures

List of Tables

List of Acronyms

AOA	Angle of Arrival
AOD	Angle of Departure
AP	Access Point
BP	Belief Propagation
BS	Base Station
CRLB	Cramer Rao Lower Bound
BSRUKF	Biased Square Root Unscented Kalman Filter
CSRUKF	Constrained Square Root Unscented Kalman Filter
DOA	Direction of Arrival
DOD	Direction of Departure
EKF	Extended Kalman Filter
FA	False Alarm
FCC	Federal Communications Commission
GPS	Global Positioning System
IMM	Interacting Multiple Model
IMU	Inertial Measurement Unit
IPPM	Iterative Parallel Projection Method
KDE	Kernel Density Estimate
LOS	Line of Sight
MAI	Multiple Access Interference
MC	Monte Carlo Simulation
MD	Missed Detection
MIMO	Multiple Input Multiple Output
ML	Maximum Likelihood

MMSE	Minimum Mean Square Error
MN	Mobile Node
MS	Mobile Station
MSE	Mean Square Error
NLOS	Non-Line-of-Sight
NLS	Nonlinear Least Squares
OA	Outer Approximation
PAN	Personal Area Network
PDF	Probability Density Function
PKF	Projected Kalman Filter
POCS	Projection Onto Convex Sets
PF	Particle Filter
QCQP	Quadratically Constrained Quadratic Programming
RFID	Radio Frequency Identification
RN	Reference Node
RSS	Received Signal Strength
SDP	Semi Definite Programming
SKF	Smooth Kalman Filter
SIUKF	Sigma-point Interval Unscented Kalman Filter
SNR	Signal to Noise Ratio
SOCP	Second Order Cone Programming
SPAWN	Sum Product Algorithm over a Wireless Network
SPD	Symmetric Positive Definite
SRUKF	Square Root Unscented Kalman Filter
TDOA	Time Difference of Arrival
TOA	Time of Arrival
TWR	Two Way Ranging
UKF	Unscented Kalman Filter
UWB	Ultra Wide-band
WLS	Weighted Lest Squares
WNLS	Weighted Nonlinear Lest Squares
WSN	Wireless Sensor Network

Chapter 1

Introduction

In this chapter, a general overview of the network-based radio frequency (RF) localization problem and a literature review of the available techniques are given in Section 1.1 and Section 1.2, respectively. We then explain the objectives and contributions of this thesis in Section 1.3 and finally summarize its organization in Section 1.4.

1.1 Wireless Positioning

The global-positioning-system (GPS) is a conventional satellite based system, for RF localization or positioning. In spite of being extensively employed in different applications, the GPS system fails to operate efficiently in indoor places or dense urban areas due to the weak received GPS signal and the multipath fading. Moreover, the size and the high energy consumption of the GPS-based devices are other disadvantages. Therefore, a network-based positioning system is preferred as a replacement for GPS in indoor places and dense urban areas.

Network-based Radio Frequency (RF) positioning or localization refers to finding the coordinates of a target in a network, based on radio signals received or sent by some reference nodes (usually fixed) with known positions. In cellular network positioning, the base stations (BS) are used as reference nodes while the mobile terminal (MT) is the target whose location needs to be estimated. The localization of mobile phones has become important since, due to an order by the Federal Communications Commission (FCC) in 1996, wireless positioning is supposed to be a mandatory public safety feature of all cellular systems by 2020 [3]. It is therefore important for service providers to be able to deliver an

accurate position of a mobile user, who calls 911, to the public safety access points [3]. In addition to the cellular application, localization can be done for a specific wireless sensor network (WSN) system in which the positions of anchors, which are the reference nodes, are known, while the positions of the sensors are to be estimated [4]. Network-based localization has many applications in healthcare, surveillance, military systems, and public safety.

The scheme of a range-based cellular network positioning system is illustrated in Fig. 1.1, where the circles represent the measured ranges between the MT (target) and BSs. Under the assumption of error-free range measurements, by intersecting three circles, the position of the MT can be obtained in 2-dimensional (2-D) space. In the presence of measurement noise and other sources of error in the measured RF signals, such as shadowing, multipath fading and non-line-of-sight (NLOS), the localization problem becomes more difficult. Improved and more robust algorithms therefore need to be developed for this task, which is the main topic of this thesis.

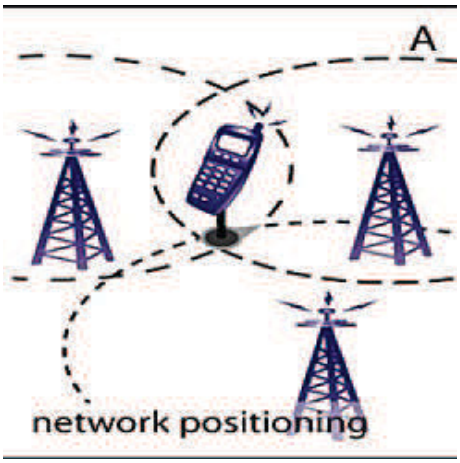

Fig. 1.1 Network-based localization using a cellular network. Figure from: http://www.e-cartouche.ch

1.2 Literature Review on Network-based RF Localization

Network-based RF localization can be done by making RF measurements between the target nodes and the fixed reference nodes, based on which the unknown positions of the targets can be estimated using different techniques [3]. In order to find the position of

the targets in a network, different types of RF measurements can be employed, including: received-signal-strength (RSS), time-of-arrival (TOA), and angle-of-arrival (AOA). In many works, it has been assumed that the radio measurement among the nodes are obtained under a line of sight (LOS) condition, in which the direct view between the transmitter and the receiver is not blocked and hence the obtained signal comes from a direct path. A great number of localization techniques and algorithms have been developed under this assumption. However, having access to LOS propagation at all times is unrealistic for indoor places and dense urban areas and therefore, in the past decade, many techniques have been developed by considering the NLOS measurements as well. The NLOS refers to a situation where the direct sight between the transmitter and the receiver is blocked, which typically occurs in two different cases. In the first case, the direct path (i.e., LOS) signal is attenuated such that the receiver can not distinguish it from the background noise; instead the copies of the signal, which are bounced off through the wall and other objects surrounding the transmitter and the receiver, are detected first. Thus, the receiver, which tries to detect the first distinguishable arriving path, wrongly selects one of the reflected signals as the first detectable path. This is known as a hard NLOS situation [26]. In the second case, known as soft NLOS, the signal passes through the walls or other objects blocking the way. The speed of the wave propagation inside different materials is lower than that in the air, hence an extra delay is observed in the detected signal. Therefore, for any of the aforementioned cases, the modelling of the received NLOS signal has to be different that that of the LOS one.

On the above basis, we can separate the literature survey into two main parts, corresponding to localization techniques developed for LOS and NLOS scenarios, described below.

1.2.1 Network-Based RF Localization under LOS Condition

In traditional localization techniques, it is assumed that the measurements are obtained in LOS situation. In RSS-based methods, the distance between the transmitter and the receiver is estimated by considering a mathematical model for the received signal power at the receiver. In a 2-D plane, three range measurements are needed to locate the target by the trilateration method without ambiguity. However, the RSS methods are not accurate in a dense multipath environment, due to a lack of an accurate model to relate the measured

RSS to the position of the target. In AOA-based methods, the direction of arrival of the received signal at each reference node is estimated by employing smart antenna arrays along with AOA estimation methods such as MUSIC [5] or ESPRIT [6]. In a 2-D space, the target position can then be estimated by having access to the angles at two receivers. Using antenna arrays increases the cost of the system and therefore the AOA-based methods need a more expensive infrastructure. In TOA-based methods, the range between the transmitter and the receiver is found by measuring the travel time of the signal between them. The position is then estimated using the trilateration or nonlinear least squares (NLS) techniques. To estimate the TOA at the receiver side, the transmitter and the receiver need to be accurately synchronized in time, which is a major challenge for TOA-based localization. The synchronization is usually done at the network layer, however the need for a precise synchronization can be avoided by using a two-way ranging (TWR) protocol, in which the signal is received and sent back to the transmitter and by adding the two measurements and having the knowledge of the internal delays of the receiver, most of the clock error terms will be removed [7]. In several applications, it may be difficult and costly to maintain the synchronization between the target and the anchors, and the target may not be capable of obtaining the TWR measurements. In this case, a time-difference-of-arrival (TDOA) measurement can be found by computing the difference between the TOA measurements obtained at a receiver and a reference receiver (usually the home BS in cellular networks). If all the receivers are synchronized with one another, then the clock error of the transmitting target will be omitted and then a hyperbolic localization technique can be applied to estimate the target position if at least four receivers exist [8]. However, the TDOA-based methods require one more anchor for localization compared to the trilateration method, as well as a reliable LOS reference node which is not guaranteed all the times. It is worth mentioning that under certain conditions, the joint localization and synchronization can be performed using the TOA data [9–12]. If the synchronization can be done accurately, either jointly with localization or separately, then the TOA-based methods are generally preferred to the TDOA-based ones in many applications.

There are several RF technologies in the market that can be exploited for sensor network localization, e.g., infra-red, Bluetooth, Wi-Fi, and ultra-wideband (UWB) [4]. Among these, the UWB technology is proven to be the best option especially for indoor localization due to its fine timing resolution and robustness against multipath and fading [4]. The Cramer-Rao lower bound (CRLB) analysis of the TOA measurements in UWB systems

shows that the variance of the estimation error can be reduced by increasing the bandwidth and the SNR of the system. Therefore, the TOA estimates in a UWB system have a good resolution due to the high bandwidth of the short transmitted pulses. Since the resolution of the TOA data is not very high for narrowband systems, AOA information can help in determining the position. On the contrary, in wideband systems like UWB, the timing resolution is high while the angle measurement is challenging and inaccurate due to the large number of rays. In a dense and scattered environment, the number of propagation paths might be very large [13]. The UWB channel can not be considered as a flat fading channel any more and there are many different models that can be considered instead. For a detailed survey on UWB channel modelling see in [14]. In addition to the narrow-band assumption which is missing in UWB array processing, in indoor applications, the far field approximation may not hold true such that the rays arriving at each antenna cannot be regarded as parallel planar wave-fronts. Therefore, due to the challenges mentioned above, the angle measurement in UWB system has to be studied with more care. To obtain the angular measurements, different methods have been proposed in the literature for the joint estimation of the TOA and AOA of UWB signals with an antenna array [15–19]. The assumption of having access to both TOA and AOA is therefore reasonable and using both measurements can improve the localization accuracy in LOS scenarios.

While in traditional localization the aim is to localize a single wireless device, in cooperative localization the sensors (targets) exchange measurements and information with each other and they can be localized either through a centralized system or in a distributed manner throughout the network. This can be very useful when the number of reference nodes is limited and the targets can not be localized by only relying on these nodes. For a general survey on cooperative localization, see also [20–22]. The cooperative localization techniques can be classified into two main categories, namely: probabilistic and deterministic [22]. In the deterministic case, the localization is done by solving an NLS problem, which can often be relaxed to a convex cost function (with constraints) and is thus easier to solve with good accuracy. The probabilistic methods make use of the probability distribution of the measurement error and then by estimating the posterior PDF of the sensors positions given the measurements, the localization can be done.

In the deterministic approaches, the aim is to solve an optimization problem to minimize a selected performance criterion such as the mean squared error, which typically leads to an NLS problem. Cooperative localization problem, formulated as an NLS problem,

which amounts to minimizing the 2-norm of the error between the true ranges and measured ranges, is NP-hard and may not be solved efficiently. Therefore, relaxation techniques have been proposed to formulate the problem as a convex optimization problem, which is easier to solve. For instance, in [23], the localization problem is relaxed to a semi-definite programming (SDP). However, the formulated SDP may only be solved in a centralized manner and therefore, it may not be scalable with the size of the WSN, and faces limitations in practice. In [24] the localization problem is formulated as a second order cone programming (SOCP) which is a looser form of relaxation compared to SDP in [23], and hence its localization performance is in general worse. There have been some efforts to implement the proposed SOCP technique in a distributed manner over the network in [25], however, the proposed SOCP can not be implemented optimally in a distributed manner and the estimates are not in general the same as the ones obtained by the centralized SOCP technique. The SOCP relaxation is similar to the idea of projection onto convex sets (POCS) where the estimate of sensor's position is projected onto the disks obtained from the range measurements whenever the current estimate is outside the corresponding disk, while otherwise no projection is done. The cooperative version of POCS is considered in [2]. While this method converges to a local minimum, since it is based on an approximation, its performance may not be good at all times. Another relaxation which is conceptually similar to POCS and SOCP is considered in [26] where by using iterative gradient techniques, it is observed that the solution to the relaxed minimization problem can be obtained in a distributed manner over the network. Since the proposed technique might not offer a very accurate estimate, after converging to some stationary points, the original NLS problem is minimized iteratively using gradient-descent technique in a distributed manner. If the initialization is good, then this proposed 2-stage technique can offer a solution close to the maximum likelihood (ML) estimate. In fact, the iterative parallel projection method (IPPM) considered in [27], where the projection of the estimates onto the disk boundaries is done at all times, is based on minimizing the original NLS problem. The estimate obtained by IPPM may converge to a local minimum, which is not necessarily near the global minimum. However, if properly initialized by techniques such as cooperative POCS, then IPPM can yield a better result than POCS. Since these iterative techniques are based on the gradient descent optimization approach, the convergence speed is low and hence faster iterative techniques may need to be developed.

In probabilistic approaches, information about the probability distribution of measure-

ment error is employed instead of directly solving a deterministic optimization problem. One of the popular techniques is the so-called belief propagation (BP) which is a distributed algorithm for factorizing the posterior distribution of the sensor positions given the measurements, using a graphical model. In BP, each sensor keeps the knowledge about its own position in terms of a probability distribution, which is known as *belief*. Then it can update its own belief by using the *messages* obtained from the neighbouring sensors. Calculation of both messages and belief at each node requires multi-dimensional integrations which in general are implemented using Monte Carlo techniques. In non-parametric versions of BP [28], kernel density estimates (KDE) are used and thus the robustness of BP algorithm against probability distribution mismatch and outliers is improved. The BP algorithm may not be efficient enough for the particular problem of network-based localization as the information that is transmitted over the network is very large, which makes it almost intractable for low-cost and low-power sensor network localization. By assuming symmetric range measurements between each pair of neighbouring nodes, in [21], by introducing the so-called sum-product algorithm over wireless network (SPAWN), the authors improve the efficiency of BP algorithm such that every sensor needs only to transmit the belief information in terms of the mean and covariance matrix of the positioning error, and there is no need to transmit the message information as the messages are computed at the receiving sensor; therefore, the data traffic over the network is reduced significantly. Still, the convergence of the BP algorithm or SPAWN for localization applications is not in general proved analytically and remains an open problem.

1.2.2 Network-based RF Localization in NLOS

There are several challenges for localization in dense environments and indoor places including multipath propagation, multiple access interference (MAI), and NLOS [13]. The multipath propagation problem can be resolved by exploiting a high resolution timing system, e.g., UWB. The MAI problem can be overcome by time multiplexing as implemented for example in IEEE 802.15.3 PAN [13], but in general, this remains an open issue. One of the most serious impairments in indoor geo-location is known as the NLOS problem, in which the direct sight between the transmitter and receiver is blocked by an object. As explained earlier, due to the blockage of the direct view, either soft or hard NLOS cases might occur. While the distribution of range measurement errors in hard and soft NLOS

situations are different, in both cases, the errors would be positively biased and the position estimate would be severely degraded. Therefore, the NLOS problem has to be studied with great care in order to provide a good location estimate for a mobile node. In dealing with the NLOS problem, there are two consecutive steps that have to be taken in order to achieve accurate localization, namely: NLOS identification and NLOS mitigation [8], where the latter is dependent on the result of the former.

In the NLOS identification techniques, the aim is to detect the fixed anchors which are in NLOS situation. It has been shown in [29] that the CRLB on the positioning error depends only on LOS anchors if no information about the statistics of the NLOS error is available. This means that it is usually preferred to discard the NLOS measurements as it degrades the performance of the other reference nodes that are in LOS condition. There are different techniques for NLOS identification using TOA data, e.g., based on the residual test algorithm [30], or based on features of the UWB channel, such as root mean squared (RMS) delay spread or mean delay spread [31–34]. Other approaches that exploit TDOA data [35] or TDOA-AOA measurements [36] have also been proposed. For a more detailed description of the NLOS identification techniques, see [8] and the references therein.

In some scenarios, the number of LOS links is less than the minimum required number for unambiguous localization, or in the worst case, all of the links are facing a NLOS situation. In this case, the system needs to use the NLOS measurements, while mitigating their effect to obtain a more accurate location estimate. If the LOS/NLOS identification can be done accurately, then the NLOS measurements are generally given less weight compared to the LOS ones [37]. However, there are always probabilities of false alarm (FA) and missed-detection (MD) in NLOS identification. Motivated by such considerations, significant research has been recently undertaken in the area of NLOS mitigation techniques, as described below.

Several localization methods are available in the literature which use TOA measurements of synchronized nodes and consider mixed LOS/NLOS conditions. In these works, it is assumed that due to the NLOS, the estimated range is always greater than the exact range, and thus, the target has to be located inside a disk (or a ball in 3-D space) with the NLOS anchor at its center. The uncertainty region around each anchor forms a closed convex feasible region, in which the target is located with a high likelihood. In [38] the uncertainty region formed by the intersection of biased NLOS measurements is formed and then the position is estimated through a quadratic programming approach by solving a

constrained least squares problem with nonlinear quadratic constraints. Since the nonlinear constraints increase the computational cost, a relaxation of the constraints into linear form along with a linear programming approach has been proposed in [39, 40]. Another geometry-constrained localization approach is given in [41] which uses a two step ML algorithm. In [42], an interior-point minimization method for solving a constrained non-linear least squares problem is proposed for treating the bias. The accuracy of the method depends on a good selection of the upper bound for the bias of the range measurement of each receiver in NLOS. In [43], a constrained minimization technique using the sequential quadratic programming (SQP) is proposed to estimate the bias and the position simultaneously. It is shown that the method in [43] outperforms the method in [42]. A survey on these methods has been given in [44].

In some other approaches, additional measurements, especially the angular information, are exploited along with the TOA data. In [45], it is assumed that the angles at both the target and the anchors, i.e., the angle of departure (AOD) and AOA, respectively, are available. Then, a geometric scatterer-based approach is proposed to estimate the position of the scatterers and the location of the target using the AOA and AOD information. The joint estimation of the angles at both sides using a multiple-input multiple-output (MIMO) system is proposed in [46]. However, it is usually impractical to have antenna arrays or directional antennas on the MT, e.g., the hand-held devices, due to the cost, size, or effect of the human body on the radio propagation [47]. In [47], the authors assume that the AOD measurement at the mobile node is unknown and instead, consider that it is moving with a certain speed for which the Doppler frequency can be measured. Using the measured Doppler frequency, the TOA, and the AOA at the anchors, the effect of the NLOS can be reduced. A method for joint estimation of TOA, AOA, AOD, and Doppler frequency is proposed in [48]. However, this method might not be useful for slowly moving targets or targets moving along an unpredictable trajectory, since in such cases, it becomes intractable to measure a well-defined Doppler frequency. In [49], a constrained minimization technique is proposed for the case that three BSs measure the TOAs and at least one of them (the home BS) measures the AOA while all of them are in NLOS situation. The method is based on a constrained optimization technique, which due to being nonlinear and non-convex might not be solvable efficiently. Some other approaches have also been proposed that exploit the hybrid TDOA-AOA measurements to mitigate the NLOS effect [50, 51].

While the above techniques make a one-shot estimate of the position of the target, for

a mobile target with available dynamic model, filtering techniques are preferred and can result in a smoother trajectory estimate. This is especially the case when data from inertial measurements units (IMU) are used in parallel with range information for tracking purposes [52], [53]. Some methods apply Kalman filter preprocessing on measured TOAs to smooth out the effect of the variances of the NLOS biases, while scaling the covariance matrix in an extended Kalman filter (EKF) to further mitigate the effect of their means [50], [51], [54]. However, these approaches can only achieve a moderate performance for large NLOS biases. In [55, 56], 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 KDE techniques. In [57], a robust semi-parametric EKF is proposed for NLOS mitigation of a mobile node. The performance of this technique is improved by employing the interacting multiple model (IMM) algorithm in [58]. These semi-parametric techniques are also suitable when AOA, RSS or a mixture of these measurements is employed. However, in addition to a high computational cost, the performance of KDE still depends on how well it can model the PDF 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 [58]. In some other techniques, the random NLOS biases are considered as parameters in the state vector, to be jointly estimated with other state parameters [59–62], where the NLOS bias variation over time is modelled as a random walk. The technique in [59] uses EKF, while [60] and [61] use particle filters (PFs) that generally have a high computational cost. Although the above 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 [57]. Therefore, there is still a potential for improving tracking of a mobile node with dynamic equation in NLOS scenarios.

In the above works, it has been assumed that only fixed reference points help in estimating the location of a single sensor. However, when there are multiple sensors and there is a possibility of communication among them, the localization can be done in a cooperative manner. In cooperative localization, if a sensor has enough LOS anchors around

itself, it is able to help localizing another sensor which does not have access to enough LOS anchors. In this case, first the former sensor is regarded as a pseudo-anchor and together with the other anchors, a location estimate is made for the second sensor. Although many works have addressed cooperative localization in LOS, the works that consider the NLOS problem are relatively recent. Basically, the NLOS errors result in positively biased range measurements and this information has to be incorporated into the cooperative localization framework. In [63] and [64], centralized techniques based on SDP are proposed for localization in NLOS. In [2], [27], distributed iterative techniques for localization in NLOS are considered by using the idea of projection. Probabilistic techniques such as BP or SPAWN can not be easily applied to NLOS localization problem as the distribution of NLOS bias is generally unknown.

However, by assuming that the bias is positive, each sensor is restricted to the intersection of multiple balls (a convex set) corresponding to the range measurements of neighbouring nodes. By outer-approximation of each convex set, an estimation of each sensor's position uncertainty can be obtained. Since the position of sensors are unknown, the balls corresponding to them have unknown locations, thus finding an outer-approximation of each convex set is not straightforward. To overcome this limitation, a distributed algorithm, which is an approximation of SPAWN algorithm was proposed in [65], to find tight ellipsoidal outer-approximation of each convex set. In this distributed algorithm, each sensor calculates the messages of its neighbouring anchors or sensors, which are approximated to be uniformly distributed on a ball or an extended ellipsoid, respectively. Then the belief will be uniformly distributed on the intersection of several balls or ellipsoid, where conventional techniques for finding the tightest ellipsoid containing such an intersection region can be used [66]. However, these techniques may not always provide a tight ellipsoidal outer-approximation of the intersection of several balls and ellipsoids, and thus tighter results need to be developed, especially for 2-D space, in order to improve the performance of the proposed distributed algorithm for localization purposes. The proposed algorithm can be used as a pre-processing step in a SPAWN algorithm which employs the LOS measurement only thus computational cost of SPAWN can be reduced significantly.

1.3 Thesis Objectives and Contributions

This Ph.D. thesis addresses several key problems related to the practical application of network-based RF localization in harsh propagation environments, such as dense urban areas, indoor places and underground, where the range measurements become positively biased due to the NLOS condition. Our contributions cover considerable ground, including the development of new Bayesian tracking schemes, robust cooperative localization methods and geometrical techniques for bounding of uncertainty regions in WSNs. The main objectives of the research work underlying this Ph.D. thesis can be stated as follows:

1. To develop an efficient filtering technique for tracking a single or multiple mobile targets under NLOS conditions.

2. To develop a robust distributed technique for estimation of sensors locations in a WSN in which the NLOS links are not identified.

3. To propose a novel technique in 2-D for tight outer-approximation of the intersection of multiple ellipses, and then apply it to a distributed algorithm based on SPAWN for outer-approximation of convex sets in WSNs under NLOS conditions.

Below, we summarize the main contributions of the thesis as they relate to these objectives:

The first objective is focused on localization in NLOS scenarios using efficient Kalman filter-type techniques. To address this problem, an efficient square root unscented Kalman filter (SRUKF) with convex inequality constraints for mobile localization is proposed. First, a non-cooperative scenario is considered for localization of a single mobile node in NLOS scenarios. The proposed constrained SRUKF (CSRUKF) is based on a combination of the SRUKF in [67] for unconstrained problems and the constrained UKF in [68]. 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 [44]. 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 [68], the sigma points of the unscented transformation 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, 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 in NLOS identification, which makes it suitable for practical applications where such errors may be inevitable. In this work, FA refers to the erroneous identification of an LOS link as being NLOS, while a missed detection (MD) refers to the opposite situation. These findings have been reported and published in (J-1), as listed in the section "Preface and Contribution of the Authors". Subsequently, the proposed centralized CSRUKF is extended to a cooperative localization scenario where multiple mobile nodes are tracked by combining the information at a fusion center. Although the proposed filter is centralized, due to the independence of the QCQP optimization problems, the computations may be done in parallel at several processors. Through simulations, it is shown that the proposed CSRUKF can perform well even in severe NLOS situations. The results of this study are published in (C-4).

Towards the second objective, a robust distributed cooperative localization technique is proposed for static networks to overcome the scalability issue faced with the centralized technique described above as well as the issue of making an error in identification of LOS links from NLOS ones. This technique consists of two-stages based on Huber M-estimation for distributed cooperative localization in the presence of unidentified NLOS links. In the first stage, a similar convex relaxation, as considered in [26], is applied to the Huber cost function, so that relatively decent sensor locations are iteratively estimated. Since the performance may not necessarily be good under a situation with low ratio of NLOS to LOS links, in the second stage, the original Huber cost function is minimized iteratively with a suitable choice of tuning parameter, and using the estimates obtained in the previous stage as initial values. For iterative optimization in both stages, we use a simple gradient descent technique since it can be easily implemented in a distributed manner. Through simulations, we first show that the proposed convex relaxation gives a reliable estimate in different NLOS scenarios. Furthermore, we show that the position estimates are generally improved in the second stage as we minimize the original Huber cost function. The robustness of our algorithm to outliers is also evaluated by using real sensor measurement sets, as obtained by measurement campaign in [69]. The results appear in the conference paper (C-3).

Towards the final objective of this thesis, we consider a specific problem in WSNs under NLOS conditions, where the sensor positions are restricted to be within the intersection of several disks, which form a closed convex set. The goal is to outer-approximate each convex set by an ellipse. These outer-approximations can be used as constraints of the sensor locations in conventional cooperative localization algorithms. The considered algorithm for this purpose is an approximation of sum-product algorithm for wireless network (SPAWN), considered as well in [65]. In this algorithm, each node finds an estimate of its convex set in the form of an ellipse. An intermediate step in this algorithm is to find the tightest ellipse which contains the intersection of several (a finite number of) ellipses. While different algorithms exist in the literature for this problem [66], herein, we develop a novel method to find tighter outer-approximating ellipses in 2-dimensional (2-D) space. Through numerical analysis we show that the proposed outer-approximation is tighter than the ones obtained by state-of-the-art algorithms. By applying the proposed technique along with SPAWN for the distributed cooperative outer-approximation of convex sets in WSNs, more accurate results can be obtained. The results of this study appear in (J-1), (J-2), and (L-1).

1.4 Thesis Overview and Notations

A general overview of network-based RF localization techniques, with consideration of non-cooperative and cooperative approaches in both LOS and NLOS scenarios, is presented in Chapter 2. The proposed CSRUKF algorithm for both non-cooperative and cooperative Bayesian tracking is presented in Chapter 3. In Chapter 4, a deterministic and distributed cooperative localization technique based on Huber M-estimation is proposed for robust positioning without NLOS identification. In Chapter 5, distributed outer-approximation of uncertainty regions in WSNs under NLOS conditions is considered where a novel method for finding a tight ellipse containing the intersection of multiple ellipses is proposed and evaluated. The conclusion of our studies and potential avenues for future works are presented in Chapter 6.

Notation: Lower-case and upper-case bold letters represent vectors and matrices, respectively. The vector 2-norm operation is denoted by $\|\cdot\|$, while $(\cdot)^T$ 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 $\boldsymbol{q}$, inclu-

sively. The symbol $\boldsymbol{I}$ denotes an identity matrix of appropriate dimension. For a positive semi-definite symmetric matrix $\boldsymbol{A}$, $\boldsymbol{A}^{1/2}$ denotes its unique positive semi-definite square root matrix, i.e. such that $\boldsymbol{A}^{1/2}\boldsymbol{A}^{1/2} = \boldsymbol{A}$ [70]. The inner product of two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ is denoted by $\boldsymbol{A} \bullet \boldsymbol{B}$ which is equivalent to Trace($\boldsymbol{A}^T\boldsymbol{B}$). We use $\mathcal{N}(\mu, \sigma^2)$ to denote the normal probability distribution with mean μ and variance σ^2.

Chapter 2

Background on Network-based Localization

In this chapter, we present background material needed for understanding the new localization techniques presented in this thesis. We first consider the case of LOS scenarios in Section 2.1 where different approaches are reviewed. We then consider the case of NLOS scenarios in Section 2.2 and how to identify and mitigate the NLOS effect. For each one of the LOS and NLOS scenarios, we present both non-cooperative and cooperative localization techniques.

2.1 Localization in LOS Scenarios

In this section, different radio localization approaches based on TOA, TDOA, AOA, RSS, and hybrid of the aforementioned measurements, under LOS conditions are reviewed. In the end, the cooperative localization scenarios and related popular techniques are presented.

2.1.1 Time of Arrival

In TOA-based methods, the travel times of a signal between a sensor (target) and multiple anchors (reference nodes) have to be measured, from which the corresponding range can be computed. The target sends a timing signal to the anchors, followed by another signal which carries the information of the time-stamp observed on its internal clock. Each anchor detects the signal and records the reading of its own clock and then subtracts the difference

between the two time-stamps to get the TOA measurement.

Let us assume that the time shown by the clock of the target is given by $C(t)$:

$$C(t) = \alpha^{(t)}t + \theta^{(0)} \tag{2.1}$$

where $\theta^{(0)}$ is the clock offset (bias) of the target at time $t = 0$, and $\alpha^{(t)}$ is the clock skew of the target at time t. The measurements of the clock skew show that it should be regarded as a random process rather than a constant [71]. However, for a short period of time during which the localization is performed and under constant temperature conditions, it can be assumed that the clock skew remains constant. Therefore, throughout this section, the superscript is removed for the sake of simplicity and the skews are regarded as constant. The clock reading of the i-th anchor at absolute time t can also be modelled as $C_i(t)$, where

$$C_i(t) = \alpha_i t + \theta_i, \tag{2.2}$$

where α_i and θ_i are the clock skew and offset of the i-th anchor, respectively. Using the above model, in the absence of noise, the TOA measured at the i-th anchor, is therefore

$$\tau_i^{(1)} = \alpha_i t_i + \theta_i - (\alpha t_s + \theta) \tag{2.3}$$

where t_s is the absolute time at which the signal is sent and t_i is the absolute time at which the signal is received by the i-th anchor. Equivalently, (2.3) can be expressed as

$$\tau_i^{(1)} = \alpha_i(t_i - t_s) + (\alpha_i - \alpha)t_s + (\theta_i - \theta) \tag{2.4}$$

where the term $t_i - t_s$ is the exact travel time of the signal from the target to the i-th anchor, which is related to the exact range d_i as

$$t_i - t_s = \frac{d_i}{c} \tag{2.5}$$

where c is the speed of the radio waves in the air.

TOA with Prior Synchronization

For the aim of accurate TOA estimation, the target and the anchors have to be precisely synchronized, i.e., the clock parameters need to be equal, or accurately estimated and their effects mitigated. If the target and anchors are synchronized then the clock skews and offsets are equal, i.e., $\alpha_i = \alpha$ and $\theta_i = \theta$ for all measurements, and thus the second and third terms in (2.4) can be omitted.

In general, if the target has the hardware capability of receiving signals as well, then a two-way ranging (TWR) protocol can be exploited, as will be described below. Let us assume that the anchors send a timing signal back to the target, δt seconds after receiving the signal from the latter, where δt is small enough such that clock skew remains constant and the target movement is negligible. So, assume that the signal is received by the target at absolute time t_r^i, which is read by its clock as $\alpha t_r^i + \theta$. By subtracting the received time-stamp and the sent time-stamp, the time of flight of the signal from the anchors to the target will be expressed as

$$\tau_i^{(2)} = \alpha t_r^i + \theta - (\alpha_i(t_i + \delta t) + \theta_i). \tag{2.6}$$

In this case, by adding the two measurements from the target to i-th anchor in (2.3) with the one from the latter to the former in (2.6), some of the clock terms in (2.4) can be cancelled from the measurements as

$$\tau_i^{(1)} + \tau_i^{(2)} = \alpha(t_r^i - t_s) - \alpha_i \delta t = \alpha(2(t_i - t_s) + \delta t) - \alpha_i \delta t \tag{2.7}$$

where in the above we have replaced the true round trip time (RTT), i.e., $t_r^i - t_s$ with its equivalent $2(t_i - t_s) + \delta t$. Finally, (2.7) can be expressed as

$$\tau_i^{(1)} + \tau_i^{(2)} = 2\alpha(t_i - t_s) + (\alpha - \alpha_i)\delta t. \tag{2.8}$$

Since it is easy to keep the anchors synchronized through a network synchronization protocol [7] and by means of wires, we can assume that they are synchronized, which amounts to setting $\alpha_i = 1$ for all i. By assuming that the error in clock skew of the target is small, i.e., $\alpha \approx 1$, the second term in (2.8) will be small compared to the first term since δt is set to be a small value. Also, with $\alpha \approx 1$ the first term is approximately equal to $2(t_i - t_s)$.

Therefore, by means of TWR the effect of clock parameters will be negligible and thus accurate ranging may be obtained. However, if the target clock is very erroneous, the first term will be noticeably different than $2(t_i - t_s)$ and the second term may not be simply neglected, hence their effects can cause large localization errors [7]. In this case, α should also be estimated together with position of the target. Note that we can assume that the parameter δt is known by the target, which can be done by transmitting this information from the anchors to the target through the communication links.

Suppose that the mobile target, at time step k, is located at position $\boldsymbol{x}[k] = [x[k], y[k]]^T \in \mathbb{R}^2$ and the known position of the i-th anchor is $\boldsymbol{p}_i = [X_i, Y_i]^T \in \mathbb{R}^2$ for $i = 1, ..., M$. By assuming that the target and the anchors are synchronized, the clock skews and biases are known and can be cancelled out from the TOA measurements, thus the range measurements can be expressed as

$$z_i[k] = (t_i - t_s)c + n_i[k] = \|\boldsymbol{p}_i - \boldsymbol{x}[k]\| + n_i[k] \tag{2.9}$$

where the relation (2.5) between the true TOA and the true range is used in the second equality, and $n_i[k]$ is the measurement noise. In practice, this error is often modelled as a white Gaussian noise with zero-mean as $n_i[k] \sim \mathcal{N}(0, \sigma_i^2)$, where σ_i^2 is the corresponding variance. Although the above measurements are nonlinear with respect to $\boldsymbol{x}[k]$, they can be linearised by a technique known as trilateration [8], which is explained below. Note that in a 2-D space at least three range measurements are needed for localization using the trilateration approach. Alternatively one may also use Taylor series expansion to linearise the equations and then, by having two linear equations, the target location can be estimated. However, using Taylor series requires an accurate initial position estimate to guarantee a good localization performance therefore, trilatertion approach may be preferred in many applications. In trilateration approach, after squaring the range measurements in (2.9) and by using the first anchor as the reference for subtraction, the measurement equations can be restated as (see, e.g., in [8])

$$\boldsymbol{H}\boldsymbol{x}[k] = \boldsymbol{q} + \boldsymbol{\psi}[k] \tag{2.10}$$

where

$$\boldsymbol{H} = \begin{bmatrix} X_2 - X_1 & Y_2 - Y_1 \\ \vdots & \vdots \\ X_M - X_1 & Y_M - Y_1 \end{bmatrix}, \tag{2.11}$$

$$\boldsymbol{q} = \frac{1}{2} \begin{bmatrix} z_1^2[k] - z_2^2[k] + X_2^2 + Y_2^2 - (X_1^2 + Y_1^2) \\ \vdots \\ z_1^2[k] - z_M^2[k] + X_M^2 + Y_M^2 - (X_1^2 + Y_1^2) \end{bmatrix} \tag{2.12}$$

and $\boldsymbol{\psi}[k]$ can be assumed to be zero-mean; however, it does not follow a Gaussian distribution [72]. Then, the least squares (LS) estimate of $\boldsymbol{x}[k]$ (which may not be the optimum estimator in the mean square sense since the noise no longer follows a Gaussian distribution) can be expressed as

$$\hat{\boldsymbol{x}}[k] = (\boldsymbol{H}^T \boldsymbol{H})^{-1} \boldsymbol{H}^T \boldsymbol{q} \tag{2.13}$$

where it is assumed that $\boldsymbol{H}$ has full column rank, which requires that at least three anchors are not co-linear on the plane. Note that if the noise variances of the different measurements are not equal, then a weighted least squares (WLS) method should be applied.

Joint TOA-based Synchronization and Localization

In the above approach, it is assumed that the target and the anchors are precisely synchronized, or that the fixed clock parameters are cancelled out by means of the TWR algorithm. Although the TWR protocol can provide accurate range estimates in LOS scenarios, in several applications the target is only a small wireless device that is only capable of transmitting a signal, thus TWR can not be done.

One solution is to use techniques that model the clock parameters using a random process and track its changes by a multi-model EKF, e.g., [71], [73]. However, these techniques can not be employed by a small target due to the power and space limitations. In the literature, there are alternative approaches that combine the localization and synchronization together [9, 11, 12]. In these methods, the clock parameters and the location of the target are unknown and they have to be estimated jointly for each time instant. Furthermore, different alternatives to the TWR protocol have been proposed in [10]. In the above joint synchronization and localization methods, only the LOS scenario is considered. The reason is that by including NLOS bias parameters into the joint localization and synchronization,

the estimation becomes difficult to handle.

2.1.2 Time Difference of Arrival

Another localization technique is based on using the differences between the range measurements corresponding to different anchors. If the anchors make the range measurements as in (2.6), then each of the measurements should be subtracted from a similar measurement of a reference anchor. While if the target makes the range measurements, then (2.3) should be subtracted from a similar measurement corresponding to a reference anchor. Usually TDOA-based localization is used for the case that the target transmit the signal and the anchor nodes aim at tracking it. Therefore, in this case, by subtracting the range measurements in (2.6) from the one corresponding to a reference anchor, the term $\alpha t_s + \theta$ will be cancelled and by assuming perfect synchronization among the anchors, there will be no clock error term in the resulting equations. Hence, the location of the target can be found through the hyperbolic localization where one of the anchors, let us say the first one is considered as the reference node. By defining the extended location vector $\boldsymbol{x}_r[k] = [x[k], y[k], d_1[k]]^T$, where $d_1[k] = \|\boldsymbol{x}[k] - \boldsymbol{p}_1\|$, it can be verified that

$$\boldsymbol{H}\boldsymbol{x}_r[k] = \boldsymbol{q} + \boldsymbol{\psi}[k] \tag{2.14}$$

where

$$\begin{aligned} \boldsymbol{H} &= \begin{bmatrix} X_2 - X_1 & Y_2 - Y_1 & z_2[k] - z_1[k] \\ \vdots & \vdots & \\ X_M - X_1 & Y_M - Y_1 & z_M[k] - z_1[k] \end{bmatrix}, \\ \boldsymbol{q} &= \frac{1}{2} \begin{bmatrix} X_2^2 + Y_2^2 - (X_1^2 + Y_1^2) - (z_2[k] - z_1[k])^2 \\ \vdots \\ X_M^2 + Y_M^2 - (X_1^2 + Y_1^2) - (z_M[k] - z_1[k])^2 \end{bmatrix}. \end{aligned} \tag{2.15}$$

A solution can be obtained using a weighted least squares (WLS) given by

$$\hat{\boldsymbol{x}}_r[k] = (\boldsymbol{H}^T\boldsymbol{Q}\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{Q}\boldsymbol{q} \tag{2.16}$$

where $\boldsymbol{Q}$ is the covariance matrix of the noises in the linearised equations as given by

$$\boldsymbol{Q} = \begin{bmatrix} \sigma_{n_1}^2 + \sigma_{n_2}^2 & \dots & \sigma_{n_1}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{n_1}^2 & \dots & \sigma_{n_1}^2 + \sigma_{n_M}^2 \end{bmatrix}. \tag{2.17}$$

and $\sigma_{n_i}^2$ is the variance of the measurement error of the i-th anchor. Due to the linearisation of the equations in (2.14), the solution in (2.16) is still biased; to overcome this problem, a bias reduction approach has been given in [74]. For a better performance the relation between $d_1[k]$ and $\boldsymbol{x}[k]$ can be considered as an equality constraint and then a constrained LS problem can be solved instead of (2.16), as done in [75, 76].

2.1.3 Angle of Arrival

In LOS situation, location of a target can also be determined if the angle of arrival (AOA) of the signal emitted from the target can be measured at two or more receivers, using smart antenna arrays or directional antennas. For narrowband signals, the AOA measurements can be achieved using the high resolution techniques like MUSIC [5] and ESPIRIT [77]. However, for signals with a wide bandwidth like in UWB systems, the narrowband assumption is no longer satisfied. Therefore, different methods have been proposed in the literature for the joint estimation of the TOA and AOA with the UWB antenna arrays [15–19]. A disadvantage of the AOA-based methods is the extra cost due to the antenna arrays, while their main advantages are the reduced number of anchors and removing the need for synchronization.

With only two angle measurements, one can find the location of a target in a 2D area using a triangulation technique as follows. Assuming that the AOA of the signal at each anchor is measured with respect to a common reference axis, thus we can write

$$\theta_i[k] = \arctan\left(\frac{x[k] - X_i}{y[k] - Y_i}\right) + n_{\theta,i}[k] \tag{2.18}$$

where $n_{\theta,i}$ is the measurement error modelled as zero-mean Gaussian noise for LOS measurements. Then it is easy to see that

$$(X_i - x[k])\sin(\theta_i[k]) = (Y_i - y[k])\cos(\theta_i[k]) + \psi_i[k], \tag{2.19}$$

where $\psi_i[k]$ has a nonlinear relation with $n_\theta[k]$, and in matrix form it follows that

$$\boldsymbol{H}\boldsymbol{x}[k] = \boldsymbol{q} + \boldsymbol{\psi}[k] \tag{2.20}$$

where $\boldsymbol{\psi}[k]$ is the vector including the $\psi_i[k]$s and

$$\boldsymbol{H} = \begin{bmatrix} -\sin(\theta_1[k]) & \cos(\theta_1[k]) \\ \vdots & \vdots \\ -\sin(\theta_M[k]) & \cos(\theta_M[k]) \end{bmatrix}, \tag{2.21}$$

$$\boldsymbol{q} = \begin{bmatrix} Y_1 \sin(\theta_1[k]) - X_1 \cos(\theta_1[k]) \\ \vdots \\ Y_M \sin(\theta_M[k]) - X_M \cos(\theta_M[k]) \end{bmatrix}. \tag{2.22}$$

One way to find an estimate of $\boldsymbol{x}[k]$ is to use LS method as

$$\hat{\boldsymbol{x}}[k] = (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{q} \tag{2.23}$$

although it is not necessarily the optimum in the mean square sense since $\boldsymbol{\psi}[k]$ is not a zero-mean Gaussian process. One may use the obtained estimate as initialization in a nonlinear least-squares (NLS) problem which is formed based on (2.18), however, there is no guarantee for convergence and improvements in performance due to non-convexity.

2.1.4 Received Signal Strength

The RSS is a measure of the power of the received signal at the receiver. The power of a signal is attenuated by different factors. A constant attenuation is due the propagation path loss which is dependent on the distance between the transmitter and the receiver. Other propagation effects such as shadowing, multipath, scattering and diffraction will also affect the signal attenuation [37]. RSS measurements can be employed in two different ways for localization purposes as will be discussed in the sequel.

RSS with Known Anchor Position

By measuring the power of the received signal, and considering a path-loss model, the range between the target and each anchor is estimated and the location of the former can be computed based on the trilateration method discussed in Section 2.1.1. In many situations, the transmit power may not be known to the receivers and needs to be estimated. In [78], an approach is considered by approximating the underlying ML estimation problem by a convex optimization problem formulated in turn as a standard SDP. In [79], a minimax SDP is employed which is shown to achieve the CRLB for sufficiently large SNRs. In many applications, such as dense urban areas or indoor places, the path-loss exponent which relates the received and transmit powers of the signal, is not fixed and its exact value is unknown. Therefore, in [80], a technique is proposed for RSS-based localization under unknown channel parameters such as path-loss exponent and transmit power.

In general, the RSS-based methods exhibit a poor performance when the RSS data is converted to range information. This is because there is no accurate model that can relate the location of the target to the RSS, especially due to changes in the environment. Also, the received RSS measurements are very noisy and accurate range information may not be easily extracted from them. Besides, in certain applications, e.g., Wi-Fi based localization, the exact location of the wifi access points (APs) might not be available. However, there is another way to exploit the RSS measurements as will be described below.

RSS with Fingerprinting

To overcome the aforementioned limitations, the fingerprinting approach has been proposed for the aim of localization where a database of signal patterns is stored in the system beforehand [81,82]. This is known as the training phase, where the RSS measurements are made at N different locations distributed over the area under consideration. This training data, known as a radio map, is saved in a database and represented as

$$\mathcal{R} = \{(\tilde{\boldsymbol{p}}_i, \boldsymbol{F}(\tilde{\boldsymbol{p}}_i)) | i = 1 \dots N\} \tag{2.24}$$

where $\tilde{\boldsymbol{p}}_i$ is the coordinate of the i-th point in the area, $\boldsymbol{F}(\tilde{\boldsymbol{p}}_i) = [\boldsymbol{z}_i(1), \dots, \boldsymbol{z}_i(\tilde{n})]$ is a fingerprint matrix and $\tilde{n}$ is the number of training samples that could be obtained at

location $\tilde{\boldsymbol{p}}_i$.[1] The vector $\boldsymbol{z}_i(t) = [z_i^1(t), \ldots, z_i^M(t)]^T$ consists of the RSS measurements obtained at M different APs [83]. This database has to be updated later on as the channel conditions change over time due to the change of environment. In the second phase, known as the positioning phase, the RSS measurements are made by the target and stored in another vector. Then by minimizing the euclidean distance between the measured and already stored vectors in the database, the location of the target is estimated. Among the advantages of fingerprinting, we note that it does not require hardware modification and synchronization among the stations. It is quite useful for applications like Wi-Fi where the location of the APs may not be known and only the fingerprint data is essential. However, the main disadvantage of fingerprinting is the need for a database, which also needs to be updated due to changes in the environment. In addition, different target positions may result in similar power profiles across the APs, so that the mapping from location to RSS measurement is not invertible. Finally, the movement of people surrounding the target Wi-Fi device, change in the location of objects such as walls, corridors and partitions in the environment, and the possibility of facing unpredictable blockages make the RSS characteristics different from those obtained in the training phase [83]. Therefore, Wi-Fi fingerprinting might result in large estimation errors. To overcome this issue, the estimate obtained by fingerprinting is used as a measurement vector in a Kalman or particle filter. The measurement model will thus be linear but modelling the distribution of the error is challenging [83]. With the use of long tailed distributions in the PF, the estimated positions will be more robust against model mismatch. Therefore, with the help of PF, and using the indoor map as well, a smoother track and more accurate position can be estimated for the target compared to the case that Wi-Fi fingerprinting is used alone. Since Wi-Fi is already commercialized and widely used, it seems to be a leading technology for assisted GPS (AGPS) or non-GPS aided localization systems. Therefore, the fingerprinting techniques are promising for mobile positioning in Wi-Fi networks.

2.1.5 Hybrid Approaches

By combining two or more of the above techniques, a hybrid localization approach is obtained where the localization performance can be improved [8]. For instance, in an LOS scenario, TDOA and AOA measurements can be used jointly, where a hybrid form of hy-

[1] To simplify the presentation, we assume that this number is the same for every location.

perbolic localization and trilateration can be achieved by combining the equations in (2.14) with (2.20). This can be formulated as

$$\boldsymbol{H}\boldsymbol{x}_a[k] = \boldsymbol{q} + \boldsymbol{\psi}[k]. \tag{2.25}$$

where the extended vector $\boldsymbol{x}_a[k] = [x[k], y[k], d_1[k]]^T$, with $d_1[k] = \|\boldsymbol{x}[k] - \boldsymbol{p}_i\|$, and

$$\boldsymbol{H} = \begin{bmatrix} X_2 - X_1 & Y_2 - Y_1 & z_2[k] - z_1[k] \\ \vdots & \vdots & \vdots \\ X_M - X_1 & Y_M - Y_1 & z_M[k] - z_1[k] \\ -\sin(\theta_1[k]) & -\cos(\theta_1[k]) & 0 \\ \vdots & \vdots & \vdots \\ -\sin(\theta_M[k]) & -\cos(\theta_M[k]) & 0 \end{bmatrix}, \tag{2.26}$$

$$\boldsymbol{q} = \frac{1}{2}\begin{bmatrix} K_2 - K_1 - (z_2[k] - z_1[k])^2 \\ \vdots \\ K_M - K_1 - (z_M[k] - z_1[k])^2 \\ 2Y_1\sin(\theta_1[k]) - 2X_1\cos(\theta_1[k]) \\ \vdots \\ 2Y_M\sin(\theta_M[k]) - 2X_M\cos(\theta_M[k]) \end{bmatrix}, \tag{2.27}$$

with $K_i = X_i^2 + Y_i^2$ and $\boldsymbol{\psi}[k]$ is the vector including the measurement errors. In the following we show an almost optimum solution for the localization problem as described in [84]. Before doing that, the vector $\boldsymbol{\psi}[k]$ and its covariance matrix should be derived as will be described below. By replacing the equivalent value of the range differences and AOA measurements in $\boldsymbol{H}$ and $\boldsymbol{q}$, it follows that

$$\boldsymbol{H} = \begin{bmatrix} X_2 - X_1 & Y_2 - Y_1 & d_2[k] - d_1[k] + n_{21}[k] \\ \vdots & \vdots & \vdots \\ X_M - X_1 & Y_M - Y_1 & d_M[k] - d_1[k] + n_{M1}[k] \\ -\sin(\theta_1^0[k] + n_{\theta,1}[k]) & -\cos(\theta_1^0[k] + n_{\theta,1}[k]) & 0 \\ \vdots & \vdots & \vdots \\ -\sin(\theta_M^0[k] + n_{\theta,M}[k]) & -\cos(\theta_M^0[k] + n_{\theta,M}[k]) & 0 \end{bmatrix},$$

$$\boldsymbol{q} = \frac{1}{2}\begin{bmatrix} K_2 - K_1 - (d_2[k] - d_1[k] + n_{21}[k])^2 \\ \vdots \\ K_M - K_1 - (d_M[k] - d_1[k] + n_{M1}[k])^2 \\ 2Y_1 \sin(\theta_1^0[k] + n_{\theta,1}[k]) - 2X_1 \cos(\theta_1^0[k] + n_{\theta,1}[k]) \\ \vdots \\ 2Y_M \sin(\theta_M^0 + n_{\theta,M}[k]) - 2X_M \cos(\theta_M^0 + n_{\theta,M}[k]) \end{bmatrix}$$

and where $\theta_i^0[k]$ is the true angle and $n_{i,1}$ is the TDOA measurement noise for i-th anchor. The terms that do not include the noise terms cancel each other and by using the assumption that the noise of the AOA measurement is small, it follows that $\sin(n_\theta[k]) \approx n_{\theta,i}[k]$ and $\cos(n_{\theta,i}[k]) \approx 1$. Therefore, with a good approximation it follows that

$$\begin{aligned} \boldsymbol{\psi}[k] &\approx \begin{bmatrix} (d_2[k] - d_1[k])n_{21}[k] + \frac{1}{2}n_{21}^2[k] + d_1[k]n_{21}[k] \\ \vdots \\ (d_M[k] - d_1[k])n_{L1}[k] + \frac{1}{2}n_{M1}^2[k] + d_1[k]n_{M1}[k] \\ n_{\theta,1}[k][(x[k] - X_1)\cos(\theta_1^0[k]) + (y[k] - Y_1)\sin(\theta_1^0[k])] \\ \vdots \\ n_{\theta,M}[k][(x[k] - X_M)\cos(\theta_M^0[k]) + (y[k] - Y_M)\sin(\theta_M^0[k])] \end{bmatrix} \\ &= \begin{bmatrix} d_2[k]n_{21}[k] + \frac{1}{2}n_{21}^2[k] \\ \vdots \\ d_M[k]n_{M1}[k] + \frac{1}{2}n_{M1}^2[k] \\ n_{\theta,1}[k]d_1[k] \\ \vdots \\ n_{\theta,1}[k]d_M[k] \end{bmatrix}. \end{aligned} \tag{2.28}$$

Since in practice $|n_{i1}[k]| \ll d_i[k]$ is satisfied, by ignoring the terms $n_{i1}^2[k]$ in (2.28) it follows approximately that

$$\boldsymbol{\psi} \approx \boldsymbol{B}\boldsymbol{n} \tag{2.29}$$

where

$$\boldsymbol{B} = \text{diag}\{d_2[k], \ldots, d_M[k], d_1[k], \ldots, d_M[k]\} \in \mathbb{R}^{2M-1 \times 2M-1}$$
$$\boldsymbol{n} = [n_{21}[k], n_{31}[k], \ldots, n_{M1}[k], n_{\theta,1}[k], \ldots, n_{\theta,M}[k]]^T.$$

and we have omitted the time dependence in k to simplify the notations.

Thus, $\boldsymbol{\psi}$ can be assumed to be approximately Gaussian [72] with covariance matrix

$$\boldsymbol{\Psi} = E[\boldsymbol{\psi}\boldsymbol{\psi}^T] = \boldsymbol{B}\boldsymbol{Q}\boldsymbol{B}^T \tag{2.30}$$

and $\boldsymbol{Q}$, the covariance matrix of the measurement equation, is given by

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{Q}_r & \boldsymbol{Q}_{r\theta} \\ \boldsymbol{Q}_{r\theta}^T & \boldsymbol{Q}_\theta \end{bmatrix} \tag{2.31}$$

where

$$\boldsymbol{Q}_r = \begin{bmatrix} \sigma_{n_1}^2 + \sigma_{n_2}^2 & \ldots & \sigma_{n_1}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{n_1}^2 & \ldots & \sigma_{n_1}^2 + \sigma_{n_M}^2 \end{bmatrix}, \boldsymbol{Q}_\theta = \begin{bmatrix} \sigma_{\theta,1}^2 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \sigma_{\theta,M}^2 \end{bmatrix} \tag{2.32}$$

$$\boldsymbol{Q}_{r\theta} = \begin{bmatrix} \alpha_1\sigma_1\sigma_{\theta,1} & 0 & 0 & \ldots & 0 \\ 0 & \alpha_2\sigma_2\sigma_{\theta,2} & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \alpha_M\sigma_M\sigma_{\theta,M} \end{bmatrix} \tag{2.33}$$

with α_i being a parameter (known experimentally) showing the amount of cross-correlation between the angular and range measurements at the i-th anchor. The variance of each link is calculated as discussed in the previous section. In [84], the correlation between the AOA and TDOA measurements is not considered; however, we consider them herein for more generality.

In order to calculate $\boldsymbol{B}$, the exact value of ranges $d_i[k]$ are needed, however, one can first find a rough position estimate by TDOA or AOA method and then have an estimate of $d_1[k]$, say $\tilde{d}[k]$. Then as $d_i[k] \approx z_i[k] - z_1[k] + d_1[k]$, replace $d_i[k]$ in matrix $\boldsymbol{B}$ with its approximate value $\tilde{d}_i[k] = z_i[k] - z_1[k] + \tilde{d}_1[k]$ to get an approximation as $\tilde{\boldsymbol{B}}$. Then an

approximate value of $\boldsymbol{\Psi}$ is obtained as

$$\tilde{\boldsymbol{\Psi}} = \tilde{\boldsymbol{B}}\boldsymbol{Q}\tilde{\boldsymbol{B}}^T. \tag{2.34}$$

The solution to the unknown vector $\boldsymbol{x}_a$ is then given by solving the following constrained optimization problem

$$\begin{aligned} &\min_{\boldsymbol{x}_a[k]} (\boldsymbol{H}\boldsymbol{x}_a[k] - \boldsymbol{q})^T \tilde{\boldsymbol{\Psi}}^{-1} (\boldsymbol{H}\boldsymbol{x}_a[k] - \boldsymbol{q}) \\ &\text{s.t.} \quad (\boldsymbol{x}_a[k] - \boldsymbol{p}_1)^T \boldsymbol{\Sigma}_a (\boldsymbol{x}_a[k] - \boldsymbol{p}_1) = 0 \end{aligned} \tag{2.35}$$

where $\boldsymbol{\Sigma}_a = \text{diag}(1, 1, -1)$. Note that the equality constraint in (2.35) is the matrix formulation of $(x[k] - X_1)^2 + (y[k] - Y_1)^2 = d_1^2[k]$. Once an estimate of $\boldsymbol{x}_a[k]$ is obtained, then a more accurate $d_1[k]$ can be estimated from which the matrices $\tilde{\boldsymbol{B}}$ and then $\tilde{\boldsymbol{\Psi}}$ can be updated and the optimization problem in (2.35) can be solved again. This process can continue iteratively until convergence.

Other hybrid combinations of the measurements may also be considered, for which the readers are referred to [84].

2.1.6 Localization with Available Dynamic Equation

In the above techniques, only the measurements which correspond to the geometric relation between the sensor and anchors are used for the aim of localization. However, for mobile targets, these memoryless techniques may not offer a smooth estimate of the trajectory and sometimes large spikes are observed in the positioning error. In case the motion of the target can be accurately modelled by a dynamic equation, filtering techniques are usually preferred. For general nonlinear dynamic systems, state vector at the $(k+1)$-th time instant $\boldsymbol{s}[k+1]$ is related to the state at the k-th time instant $\boldsymbol{s}[k]$ as

$$\boldsymbol{s}[k+1] = \boldsymbol{f}(\boldsymbol{s}[k]) + \boldsymbol{w}[k], \tag{2.36}$$

where $\boldsymbol{f}$ is a nonlinear function of the state vector and $\boldsymbol{w}[k]$ is the process noise with covariance matrix $\boldsymbol{Q}$. The measurements at the k-th time instant, $\boldsymbol{z}[k]$ can be expressed in general form as

$$\boldsymbol{z}[k] = \boldsymbol{h}(\boldsymbol{s}[k]) + \boldsymbol{n}[k] \tag{2.37}$$

where $\boldsymbol{h}$ is a nonlinear function of the state vector and $\boldsymbol{n}[k]$ is the measurement noise with covariance matrix $\boldsymbol{R}$.

One of the most popular filtering techniques to estimate the state vector and its uncertainty in the form of a covariance matrix is the extended Kalman filter (EKF), which is based on the first order Taylor series approximation. Due to the poor performance of EKF for highly nonlinear problems, unscented Kalman filter (UKF) which can offer accuracy of the linearisation at least up to the second order Taylor series term is proposed [85]. The performances of these Kalman-based filters are only good if the measurement and process noises are normally distributed. For the case of non-Gaussian errors, the particle filters (PFs), which are Monte Carlo techniques for estimation of the posterior PDF using particles and weights assigned to them, may be employed [86]. In general, PFs suffer from high computational cost for moderately large state vectors, thus their computational cost is a limiting factor in low-power and low-cost sensor networks.

Considering nonlinear filtering techniques for tracking has been done during the last decades and there are numerous works available. The readers are referred to [8] and the references therein for further explanations.

2.1.7 Cooperative Localization Techniques in LOS

The cooperative techniques have received great attention in ad-hoc and sensor networks where each sensor can communicate with its neighbours in order to do self-localization. In a wireless sensor network, the localization can be done in a cooperative fashion so that the neighbouring targets help each other in finding their coordinates. The cooperative techniques can be classified into different categories. For example, the problem formulation can be categorised in two ways: probabilistic and non-probabilistic. In probabilistic methods, the localization is referred to as a probabilistic inference problem where the belief in the position of each target is computed and sent to other targets. Alternatively, the non-probabilistic approaches provide a deterministic estimate of the location of each target based on the measurements. The probabilistic methods provide a reliable result and are suitable for distributed implementations but with the downside that they generally have a higher computational cost compared to deterministic approaches.

Another classification of the cooperative techniques can be done based on the computation center as centralized or distributed techniques. The centralized approaches suffer from

the high computational cost when the network grows in size, therefore, these approaches are not scalable. On the other hand, in distributed sensor network localization, the computations are distributed over the nodes in the network and the nodes only need the local information of their neighbours in order to estimate their own location. In distributed implementation, the localization can be done sequentially or in parallel for every sensor . In the sequential methods, first, all nodes estimate their range to all neighbouring nodes and anchors. Then all the nodes with three or more LOS anchors determine their position and are regarded as virtual anchors afterwards. The virtual anchors together with the anchors help the nodes that do not have enough LOS measurements to localize themselves. This process continues until the position of all the nodes are updated. In parallel techniques, every node updates its position at every time instant and sends this information to the neighbouring nodes. Based on the updated location information, the next iteration is done for every node until some convergence is achieved for every node.

Deterministic Approaches

In deterministic localization, the problem is formulated in order to minimize the mean squared error (MSE). For example, consider a cooperative network with N sensor nodes for which the position matrix $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N] \in \mathbb{R}^{2\times N}$ is unknown (the extension to 3-D positioning is straightforward), together with M anchor nodes with known position $\boldsymbol{p}_k$, for $k \in \{N+1, \ldots N+M\}$. All the neighbouring sensors and the anchors measure the range among each other as

$$\begin{aligned} z_{kj} &= \|\boldsymbol{p}_k - \boldsymbol{x}_j\| + n_{kj}, \quad \forall (k,j) \in \mathcal{N}_a \\ z_{ij} &= \|\boldsymbol{x}_i - \boldsymbol{x}_j\| + n_{ij}, \quad \forall (i,j) \in \mathcal{N}_x \end{aligned} \tag{2.38}$$

where $\mathcal{N}_a$ and $\mathcal{N}_x$ are the sets of pairwise indices of neighbour anchor-sensor and neighbour sensor-sensor nodes, respectively. The aim is to find the position matrix $\boldsymbol{X}$ by minimizing the mean squared errors as

$$\min_{\boldsymbol{X}} \Big(\sum_{(k,j)\in\mathcal{N}_a} (\|\boldsymbol{x}_j - \boldsymbol{p}_k\| - z_{kj})^2 + \sum_{(i,j)\in\mathcal{N}_x} (\|\boldsymbol{x}_i - \boldsymbol{x}_j\| - z_{ij})^2 \Big) \tag{2.39}$$

which is a non-linear and non-convex optimization problem and NP hard. There are several relaxation techniques to modify the cost function, such that the global optimum of the

modified function can be obtained, which might be close to the true global minimum.

Below we discuss two of the most popular techniques, one uses semi-definite programming (SDP) and the other is based on another relaxation which makes the computations easily implementable in a distributed manner.

SDP Approach: The optimization problem in (2.39) is in general non-convex, however, it can be relaxed and reformulated as an SDP problem, which is convex and can be solved more efficiently in polynomial time [23,87]. The SDP has received great attention in cooperative and large scale WSN and wireless ad-hoc networks, due to the great computational cost of finding the location of all the unknown positions with standard optimization techniques [23, 87, 88]. Herein, we consider the noise-free measurements and try to solve the problem of finding the location of sensors, given the true distances between neighbouring nodes. This problem is also NP-hard and finding the exact solution can not be done in polynomial time. First consider for all $(i,j) \in \mathcal{N}_x$, $i<j$:

$$\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2 = \boldsymbol{e}_{ij}^T \boldsymbol{X}^T \boldsymbol{X} \boldsymbol{e}_{ij} \tag{2.40}$$

and for all $(k,j) \in \mathcal{N}_a$:

$$\|\boldsymbol{p}_k - \boldsymbol{x}_j\|^2 = \begin{bmatrix} \boldsymbol{p}_k^T & -\boldsymbol{e}_j^T \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_d^T \\ \boldsymbol{X}^T \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_d & \boldsymbol{X} \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_k \\ -\boldsymbol{e}_j \end{bmatrix} \tag{2.41}$$

where $\boldsymbol{e}_j \in \mathbb{R}^n$ is a vector of entry 1 at the j-th row and all zeros elsewhere; $\boldsymbol{e}_{ij} \in \mathbb{R}^n$ is a vector which is 1 at the i-th row, -1 at the j-th row and zero elsewhere. Therefore, we should find a symmetric matrix $\boldsymbol{Y} \in \mathbb{R}^{n\times n}$ and a matrix $\boldsymbol{X} \in \mathbb{R}^{d\times n}$ such that

$$\boldsymbol{e}_{ij}^T \boldsymbol{Y} \boldsymbol{e}_{ij} = z_{ij}^2 \quad \forall (i,j) \in \mathcal{N}_x \tag{2.42}$$

$$\begin{bmatrix} \boldsymbol{p}_k^T & -\boldsymbol{e}_j^T \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_d & \boldsymbol{X} \\ \boldsymbol{X}^T & \boldsymbol{Y} \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_k \\ -\boldsymbol{e}_j \end{bmatrix} = z_{kj}^2 \quad \forall (k,j) \in \mathcal{N}_a \tag{2.43}$$

$$\boldsymbol{Y} = \boldsymbol{X}^T \boldsymbol{X} \tag{2.44}$$

where $\boldsymbol{I}_d \in \mathbb{R}^{d\times d}$ is an identity matrix. In the SDP relaxation tecnique, the constraint

$\boldsymbol{Y} = \boldsymbol{X}^T\boldsymbol{X}$ is relaxed to $\boldsymbol{Y} \succeq \boldsymbol{X}^T\boldsymbol{X}$ [88] [2], or equivalently the following matrix inequality

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{I}_d & \boldsymbol{X} \\ \boldsymbol{X}^T & \boldsymbol{Y} \end{bmatrix} \succeq \boldsymbol{0}, \tag{2.45}$$

where $\boldsymbol{W} \in \mathbb{R}^{(d+n)\times(d+n)}$. In the end, the problem can be formulated as an SDP problem (see [88]):

$$\begin{aligned}
&\max_{\boldsymbol{W}} \ 0 \\
&\text{s.t. } \boldsymbol{W}_{1:d,1:d} = \boldsymbol{I}_d && (2.46)\\
&\left(\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{e}_{ij} \end{bmatrix} \begin{bmatrix} \boldsymbol{0}^T & \boldsymbol{e}_{ij}^T \end{bmatrix} \right) \bullet \boldsymbol{W} = z_{ij}^2 \quad \forall (i,j) \in \mathcal{N}_x && (2.47)\\
&\left(\begin{bmatrix} \boldsymbol{p}_k \\ -\boldsymbol{e}_j \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_k^T & -\boldsymbol{e}_j^T \end{bmatrix} \right) \bullet \boldsymbol{W} = z_{kj}^2 \quad \forall (k,j) \in \mathcal{N}_a && (2.48)\\
&\boldsymbol{W} \succeq \boldsymbol{0} && (2.49)
\end{aligned}$$

where the operator $\bullet$ represents the inner product, as defined in Section 1.4. Once a matrix $\boldsymbol{W}$ is found, the matrix $\boldsymbol{X}$ can also be calculated, which contains the estimated positions of the sensors.

Extension of this SDP approach for the case with noisy range measurements has been considered in [23], in which the NLS problem for finding the location of the sensors by minimizing the 2-norm of the error is relaxed to an SDP. Although the proposed SPD technique performs relatively accurate localization in polynomial time, it is based on the fact that all the information about the sensors is collected at a fusion center, which makes it a centralized approach. The centralized techniques suffer from high computational cost at the fusion center and network overload. Furthermore, the centralized techniques may not be scalable with the size of the network, therefore, distributed techniques are preferred. There exist some sub-optimal implementations of the proposed SDP relaxation in a distributed manner, however, their performances may not be good and convergence to the centralized SDP solution may not be guaranteed, therefore, it is still an open problem.

In the sequel, other techniques are described where they are mostly suitable for dis-

[2] The notation $\boldsymbol{A} \succeq \boldsymbol{B}$ means $\boldsymbol{A} - \boldsymbol{B} \succeq 0$, i.e., $\boldsymbol{A} - \boldsymbol{B}$ is positive semidefinite.

tributed implementation over the network.

Other Relaxations: Another convex relaxation technique which can be useful for distributed localization has been proposed in [26]. This convex relaxation modifies the original cost function in (2.39) as

$$\sum_{(i,j)\in\mathcal{N}_x}\Big((\|\boldsymbol{x}_j-\boldsymbol{x}_i\|-z_{ij})_+\Big)^2+\sum_{(j,k)\in\mathcal{N}_a}\Big((\|\boldsymbol{x}_j-\boldsymbol{p}_k\|-z_{kj})_+\Big)^2 \tag{2.50}$$

where for a real value u,

$$u_+=\begin{cases} u & u\geq 0\\ 0 & \text{otherwise}\end{cases} \tag{2.51}$$

Further explanation about the convexity of this cost function is given in [89]. The concept of this relaxation is similar to POCS proposed first in [90] and considered for cooperative sensor network localization in [2]. By implementing the gradient descent algorithm on the convex cost function, it will be observed that the calculation of the gradients can be done locally at each sensor node. Therefore, each sensor iteratively minimizes the modified local cost function and then transmits the estimate of its own location to its neighbours to be used in the next iteration. Therefore, the global minimum of the convex cost function will be reached after enough iterations. This relaxation is weaker than the SDP relaxation mentioned earlier, in the sense that the sensor estimates are generally less accurate. However, the computational cost for solving the SOCP is less than that of the SDP problem therefore it is preferred in that sense.

Projection onto Convex Sets: The idea of POCS was first proposed in [90], and has been applied to target localization problem in [91]. For the j-th target, the convex set is $\cap_{i\in\mathcal{N}(j)}\mathcal{D}_{ij}$, where $\mathcal{N}(j)$ is the index set of neighbouring nodes of j-th sensor, and

$$\mathcal{D}_{ij}=\{\boldsymbol{x}\in\boldsymbol{R}^2:\|\boldsymbol{x}-\boldsymbol{x}_i\|\leq z_{ij}\} \tag{2.52}$$

is the disc with centre $\boldsymbol{x}_i$ and radius z_{ij}, which contains the j-th sensor. The POCS method is an iterative algorithm for projecting the parameters onto a convex set where at the l-th iteration, the position of a single target j is updated as

$$\boldsymbol{x}_j^{(l+1)}=\boldsymbol{x}_j^{(l)}+\lambda_j^{(l)}\sum_{i\in\mathcal{N}(j)}w_{ij}(\mathcal{P}_{\mathcal{D}_{ij}}(\boldsymbol{x}_j^{(l)})-\boldsymbol{x}_j^{(l)}) \tag{2.53}$$

where $\mathcal{P}_{\mathcal{D}_{ij}}(\boldsymbol{x}_j^{(l)})$ is a projection operation, $\epsilon_1 < \lambda_j^{(l)} < 2-\epsilon_2$ for arbitrary small and positive ϵ_1 and ϵ_2, and the weights w_i are such that

$$\sum_{i\in\mathcal{N}(j)} w_{ij} = 1, \quad w_{ij} > 0 \tag{2.54}$$

and the optimum value of the projection onto a disc is

$$\mathcal{P}_{\mathcal{D}_{ij}}(\boldsymbol{x}_j^{(l)}) = \begin{cases} \boldsymbol{x}_i + \frac{\boldsymbol{x}_j^{(l)}-\boldsymbol{x}_i}{\|\boldsymbol{x}_j^{(l)}-\boldsymbol{x}_i\|} z_{ij}, & \|\boldsymbol{x}_j^{(l)} - \boldsymbol{x}_i\| \geq z_{ij} \\ \boldsymbol{x}_j^{(l)}, & \|\boldsymbol{x}_j^{(l)} - \boldsymbol{x}_i\| \leq z_{ij} \end{cases} \tag{2.55}$$

with $\boldsymbol{x}_i$ being the centre of a disc $\mathcal{D}_{ij}$. If i-th nodes is a sensor and hence $\boldsymbol{x}_i$ is unknown then its estimate at the l-th iteration, i.e., $\boldsymbol{x}_i^{(l)}$ is used in (2.55).

The POCS can provide a relatively reliable solution with any initialization. However, the POCS method might face a situation that there exist disks such that $\cap_{i\in\mathcal{N}(j)}\mathcal{D}_{ij} = \emptyset$. Therefore the POCS will never converge to a feasible solution as there is no feasible solution in this scenario [2]. Furthermore, cooperative POCS is suitable when most of the measurements are positively biased, i.e., are in NLOS. In the presence of zero-mean measurement noises, a disk may not contain the position of the corresponding sensor, and the solution obtained by POCS is far from optimal. Therefore, a better solution can be obtained by minimizing the original cost function. A method known as iterative parallel projection method (IPPM) for distributed cooperative localization is proposed in [27] that will be described below.

Iterative Parallel Projection Method (IPPM): The idea behind the IPPM is developed from the modified parallel projection method (MPPM) as explained in [27] and the references therein. In MPPM, the feasibility problem is formulated as a weighted LS problem when there exists no solution satisfying all convex feasibility sets. The methods such as POCS are examples of the methods which suffer from the inconsistency in the intersection of feasibility sets. While the MPPM is developed for the non-collaborative localization, however, a modification is made in IPPM for the collaborative scenario. Basically, in the IPPM, each node uses the MPPM to update its position estimated using the range measurements of the neighbouring nodes. If the i-th and j-th nodes are neighbours, with estimated positions at l-th iteration $\hat{\boldsymbol{x}}_i^{(l)}$ and $\hat{\boldsymbol{x}}_j^{(l)}$, respectively, the projection of $\hat{\boldsymbol{x}}_i^{(l)}$ onto

the feasibility set given by the range measurement z_{ij} is

$$\mathcal{P}_{ij}^{\text{col}}(\hat{\boldsymbol{x}}_i^{(l)}) = \hat{\boldsymbol{x}}_j^{(l)} + z_{ij}\frac{\hat{\boldsymbol{x}}_i^{(l)} - \hat{\boldsymbol{x}}_j^{(l)}}{\|\hat{\boldsymbol{x}}_i^{(l)} - \hat{\boldsymbol{x}}_j^{(l)}\|} \tag{2.56}$$

where "col" stands for collaborative, and the residual is defined as

$$\Phi^{\text{col}}(\hat{\boldsymbol{x}}_i^{(l)}) = \frac{1}{|\mathcal{N}(i)|}\sum_{j\in\mathcal{N}(i)}(z_{ij} - \|\hat{\boldsymbol{x}}_i^{(l)} - \hat{\boldsymbol{x}}_j^{(l)}\|)^2 \tag{2.57}$$

where $|.|$ denotes the cardinality. Let the indexes $i = 1, \ldots, N$ be for the sensors and the indexes $i = N+1, \ldots, N+M$ be for the anchors. The IPPM algorithm is summarized in Algorithm 1 where Φ^{ncl} is the non-collaborative residual computed by averaging the residuals corresponding to the neighbouring anchors of node i only.

Algorithm 1 IPPM

Initialization:
Use tri-lateration of anchors to initialize $[\hat{\boldsymbol{x}}_1, \ldots, \hat{\boldsymbol{x}}_N]$.
Set $l = 0$, $\Phi_l = \Phi^{\text{ncl}}(\hat{\boldsymbol{x}})$, K as the maximum number of iterations, and δ as a small positive number. Let $F_i = 0$ and $W_i = 0$ for $i = 1, 2, \ldots, N$.
While (any $F_i = 0$) {
for $i = 1, 2, \ldots, N$ **do**
  if $F_i = 0$ **then**
    $\hat{\boldsymbol{x}}_i = \frac{1}{|\mathcal{N}(i)|}\sum_{j\in\mathcal{N}(i)} P_{ij}^{\text{col}}(\hat{\boldsymbol{x}}_i), \Phi_{i,l+1} = \Phi^{\text{ col}}(\hat{\boldsymbol{x}}_i)$
    if $|\Phi_{i,l} - \Phi_{i,l+1}| < \delta$ **then**
      Let $W_i = W_i + 1$; If $W_i \geq K$ set $F_i = 1$
    else
      Let $W_i = 0$;
    end if
  end if
end for
$l = l + 1$;
}

In Algorithm 1, when $F_i = 1$ the i-th node is considered as localized. The iteration for each node stops when the number of iterations passes the parameter K or the difference between the values of Φ is smaller than the considered threshold δ.

Probabilistic Approaches

Many problems in signal processing and coding might be solved using a so-called belief/probability propagation algorithm, which is based on marginalization of the posterior PDF using a facto-graph [21]. Specific examples of such algorithms include Kalman filtering and smoothing, the forward backward algorithm for hidden Markov models, probability propagation in Bayesian networks, decoding algorithms for error correcting codes such as the Viterbi algorithm, the BCJR algorithm, and the iterative decoding of turbo codes, low-density parity check codes, and similar codes [21].

The aim of belief propagation technique is to approximately marginalize the posterior PDF $f(\boldsymbol{x}|\boldsymbol{z})$, where in the localization context $\boldsymbol{x} = [\boldsymbol{x}_1^T, \ldots, \boldsymbol{x}_N^T]^T$ and $\boldsymbol{z}$ is defined by stacking all the range measurements z_{ij}. Therefore, the beliefs can be computed through iterative message passing on the facto graph corresponding to $f(\boldsymbol{x}|\boldsymbol{z})$. In a belief propagation (BP) algorithm, the belief (probabilistic knowledge) of node i about its location at the l-th iteration is

$$b_i^{(l)}(\boldsymbol{x}_i) \propto p(\boldsymbol{x}_i) \prod_{j \in \mathcal{N}(i)} \mu_{j \to i}^{(l)}(\boldsymbol{x}_i) \tag{2.58}$$

where $p(\boldsymbol{x}_i)$ is prior belief (PDF) of node i about $\boldsymbol{x}_i$ and for node j, which is a neighbouring node of node i, $\mu_{j \to i}^{(l)}(\boldsymbol{x}_i)$ is the so-called *message* of j-th node to i-th node, defined as

$$\mu_{j \to i}^{(l)}(\boldsymbol{x}_i) \propto \int p(z_{ij}|\boldsymbol{x}_i, \boldsymbol{x}_j) \frac{b_j^{(l-1)}(\boldsymbol{x}_j)}{\mu_{i \to j}^{(l-1)}(\boldsymbol{x}_j)} d\boldsymbol{x}_j. \tag{2.59}$$

The initial values are set to be $\mu_{i \to j}^{(0)}(\boldsymbol{x}_j) = 1$ and $b_j^{(0)}(\boldsymbol{x}_j) = p(\boldsymbol{x}_i)$. In BP algorithm, in order for node i to calculate the belief $b_i^{(l)}(\boldsymbol{x}_i)$, the messages of all neighbour nodes $j \in \mathcal{N}(i)$ needs to be transmitted to it. Since in general, the message is represented approximately by a large set of discrete samples, the communication load in the network will hence be high and thus this method will be difficult to implement for low-cost sensor networks. Another variation of BP known as non-parametric BP, which is suitable when the noise is multi-modal has been proposed in [28], however, it also suffers from high computational cost.

A more efficient implementation of BP, especially suitable for distributed localization, is known as sum-product algorithm over wireless network (SPAWN) [21]. In SPAWN, each node i computes the belief about its own variable $\boldsymbol{x}_i$ defined the same way as (2.58). The

message of node j to i is defined in SPAWN as

$$\mu_{j\to i}^{(l)}(\boldsymbol{x}_i) \propto \int p(z_{ij}|\boldsymbol{x}_i, \boldsymbol{x}_j) b_j^{(l-1)}(\boldsymbol{x}_j) d\boldsymbol{x}_j. \tag{2.60}$$

In SPAWN, each node only transmits its belief information (e.g., mean and covariance for Gaussian beliefs) to its neighbours and the messages in (2.60) are computed at the destination node analytically or through numerical integration. Calculating the belief in (2.58) requires multiplication of multiple messages calculated earlier, which can be done using a parametric representation for known error distributions. The probabilistic approaches are accurate for distributed localization, if the distribution of measurement error is known. However, the knowledge of the error distribution might be limited in some scenarios. Another important disadvantage of these techniques is the high computational cost that might not be tolerated for low cost sensor networks.

2.2 Localization in NLOS Scenarios

In this section, the NLOS problem is first introduced. To deal with NLOS situations, the NLOS links need to identified, and then their effects need to be mitigated, which will be described in this section, sequentially. Finally, cooperative localization in NLOS situations is introduced.

2.2.1 NLOS Problem

One of the main challenges in the localization is the NLOS problem, i.e., when the direct path between the transmitter and the receiver is blocked. In this case, the regular approaches like tri-lateration, triangulation, and other approaches designed for LOS situations yield poor estimate due to the additive bias term in the measurements. The NLOS bias of TOA measurements is a positive random variable with a large variance, thus making the measurements imprecise. The large variance of NLOS measurements compared to LOS ones have been observed through real experiments in [8,14,37]. Since the NLOS errors can significantly deteriorate the localization accuracy, the effect of the NLOS bias should be mitigated.

Let us assume that the TOA measurement is done through the common technique, and that the range between the transmitter and the receiver is computed. Let us assume that the nodes are synchronized with each other, so that the clock parameters do not appear in the range measurement equations. The mobile target, at time step k, is located at position $\boldsymbol{x}[k] = [x[k], y[k]]^T \in \mathbb{R}^2$. The known position of the i-th anchor is $\boldsymbol{p}_i = [X_i, Y_i]^T \in \mathbb{R}^2$ for $i = 1, ..., M$. Let us consider the first M_N anchors ($M_N \leq M$) to be the ones facing NLOS and the remaining $L = M - M_N$ anchors as the ones in LOS situation. Let the range measurements at time instant k be denoted by $z_i[k]$ where

$$z_i[k] = \begin{cases} d_i[k] + b_i[k] + n_i[k], & i = 1, \ldots, M_N \\ d_i[k] + n_i[k], & i = M_N + 1, \ldots, M \end{cases}$$

where $d_i[k] = \|\boldsymbol{x}[k] - \boldsymbol{p}_i\|$, $n_i[k]$ is a white Gaussian noise modelled as $n_i[k] \sim \mathcal{N}(0, \sigma_i^2)$, and $b_i[k]$ is the positive bias due to the NLOS with mean $\overline{b}_i[k]$ and variance $\sigma_{b_i}^2[k]$. Different distributions have been considered for $b_i[k]$, e.g., exponential, Gaussian, Gamma, etc. [14].

The CRLB derivation of the TOA-based localization based on the range measurements

shows that when the statistics of b_i are not available, the CRLB is dependent only on the LOS measurements [29]. This means that the best localization performance can be achieved if the NLOS anchors are identified perfectly and then the corresponding set of measurements are discarded from the total set of measurements. This has also been verified for localization methods based on TDOA, AOA, RSS, and hybrid thereof. Discarding can be done if there would be enough LOS fixed anchors for unambiguous localization. For a 2-D TOA-based localization, at least three anchors are required. Therefore, in many applications like in cellular systems, the discarding technique might make it almost impossible to estimate the mobile position without ambiguity. Moreover, in [29], it is shown that if prior NLOS statistical information is available then the generalized CRLB (G-CRLB) is also dependent on the NLOS anchors. In the case that the PDF of the NLOS bias $p_b(b)$ is available, the G-CRLB can be asymptotically achieved by the maximum *a-posteriori* (MAP) estimator without discarding any NLOS measurement. The performance can be guaranteed when σ_b^2 is small enough (although it is usually not satisfied) and when b is in the neighbourhood of a local maximum of $p_b(b)$.

Therefore, for a good position estimation technique, it is essential that the NLOS anchors be first identified and that some prior knowledge about the NLOS bias be also available. The prior knowledge could include only the variance, but in the best case the PDF is also desired. The distribution of the NLOS is location dependent and when the target moves inside a room, the distribution is varying. Therefore, the methods that consider a prior distribution are of limited practical interest. However, the variance of the range measurements could be estimated by statistical information about the range measurements over time. A smooth cubic regression technique for variance estimation is presented in [92]. The variance calculation might be accurate for localization of a fixed node, however, for mobile nodes, it might not be done with good accuracy.

Below, some popular NLOS identification techniques are first illustrated and then we move forward with the NLOS mitigation techniques.

2.2.2 NLOS Identification

The NLOS identification is the initial phase in dealing with an NLOS situation in order to improve the performance of the localization. If the identification of the NLOS anchors

is done with a high precision, then the localization error can be reduced to some extent. Below, we describe identification techniques using TOA and other sets of measurements.

Identification Using TOA Data

In this part, we summarize some of the popular identification techniques using TOA measurements.

Residual Test Algorithm: In this approach [30], one considers different combinations of at least three out of M total fixed anchors, the number of such combinations will be

$$S_0 = \sum_{i=3}^{M} \binom{M}{i} \tag{2.61}$$

For each combination, different location coordinates $\hat{x}_k$ and $\hat{y}_k$ are estimated using the approximate maximum likelihood algorithm, where $k \in \{1, 2, \ldots, S_0\}$. For each location estimate, the square of the normalized residuals are computed as

$$\chi_x^2(k) = \frac{(\hat{x}_k - \hat{x}_{S_0})^2}{I_x(k)}, \quad \chi_y^2(k) = \frac{(\hat{y}_k - \hat{y}_{S_0})^2}{I_y(k)} \tag{2.62}$$

where $\hat{x}_{S_0}$ and $\hat{y}_{S_0}$ are the coordinate estimates using all the anchors, and $I_x(k)$ and $I_y(k)$ are the approximation of CRLBs for the estimation errors in x and y coordinates, using the k-th combination, respectively [44]. If all the anchors in the k-th combination are LOS then the residuals $\chi_x^2(k)$ and $\chi_y^2(k)$ have Chi-square distributions with one degree of freedom, otherwise, they have non-centralized Chi-distributions with non-centrality parameter depending on the NLOS bias. In the case that the distribution happens to be non-central Chi-square, there is at least one NLOS anchor. Then the algorithm forms combinations with $M-1$ anchors in each set. If the distribution is a centralized Chi-square then those $M-1$ anchors are selected as LOS, otherwise the algorithm continues until there are at least three LOS anchors available for localization.

Statistical Methods in UWB NLOS Identification: In [32], the multipath channel statistics are used for NLOS identification of UWB links. It is shown that for TOA data, there is a correlation between the NLOS bias and different channel characteristics of the signal, e.g., energy and maximum amplitude of the received signal, root mean squared (RMS) delay spread, mean delay spread, kurtosis, and rise time. Therefore, using a likelihood ratio test,

the NLOS/LOS identification can be done.

In [33], a NLOS identification technique using real UWB data is proposed. By exploiting the aforementioned feature of the UWB channel and with the aid of machine learning techniques, the UWB data which are in a NLOS situation are identified. In [34], it is shown that the highest correlation between the TOA measurements and the UWB data features are in the RMS delay spread, mean delay spread, and maximum amplitude of the signal. After identification, it is further shown that, using an iterative technique the NLOS biases can be estimated and subtracted from the measurements. Other UWB measurement campaigns for NLOS identification are also given in the literature [31], [32].

The results obtained for UWB NLOS identification shows that a good performance may be achieved in general with a low chance of wrong classification of a LOS or NLOS link. Therefore, the assumption that the NLOS links are identified accurately may be justified.

Identification Using Other Sets of Data

In wireless networks, there are other types of measurements available such as TDOA, AOA, and RSS. In this part, some of the popular cases are briefly described.

Identification using TDOA: In [35], a localization technique using TDOA data is proposed. The TDOA residual is defined as the norm of the difference between the measured TDOA and calculated TDOA using the initial location estimate.

Identification Using TDOA/AOA: By having access to the AOA of the home BS, the authors in [35] have extended their work and improved it in [36]. The residual is defined in a different way to take into account the variance of each measurement and the sign of the difference between the measured TDOA and the calculated TDOA. The method is based on the conditional probability of the TDOA measurements assuming the LOS condition, which follows a Gaussian distribution. In this algorithm, first, all the TDOA measurements together with the AOA from the home BS are used to obtain an initial position. Using the initial point, the residual from the conditional PDF of the TDOA data, which equates to the PDF of the TDOA measurement noise, is computed. If the residual is above a threshold then the BS is detected as NLOS. Thus, the higher the difference between the measured TDOAs and the computed range differences using the estimated position, the more likely the BS is in NLOS. By this method, the BSs which are in more severe NLOS situation will be detected as they have larger residuals.

2.2.3 NLOS Mitigation Using Non-cooperative Techniques

After detecting which anchors are facing LOS and which ones are in NLOS, the next step is to reduce the effect of the biased NLOS measurements.

NLOS Mitigation Using TOA

As the UWB technology has great timing resolution, many localization networks are based on the use of the TOA data. When the direct view of the mobile terminal and a fixed station is blocked, the direct path signal is attenuated such that it can not be detected at the receiver. However, the earliest detectable signal arriving at the receiver is due to the reflection from the surrounding objects. The reflected signal has travelled a longer distance compared to the direct path, hence its travel time is positively biased. Moreover, due to the absorption and reflection from the objects, the noise variance of the received signal is generally higher than that in the LOS case.

ML Based Algorithm with Known NLOS Bias Distribution: In the ML-based approach proposed in [8], the distribution of the NLOS bias is assumed to be known. For example as considered in [8], the NLOS bias corresponding to the i-th anchor is exponentially distributed with parameter λ_i. Since the additive noise in the range measurements is assumed to be Gaussian, the exact distribution results from the convolution of a zero mean Gaussian distribution and a non-zero mean exponential distribution, with the resulting PDF given by

$$P(z) = \lambda_i \exp(-\lambda_i(z - \lambda_i\sigma_i^2/2))Q(\lambda_i\sigma - z/\sigma_i) \tag{2.63}$$

where $Q(.)$ is the Gaussian Q-function [8]. Thus, the exact ML solution is given by

$$\hat{\boldsymbol{x}}_{\mathrm{ML}} = \arg\min_{\boldsymbol{x}}\Big(\sum_{i\in\mathcal{N}_N} \lambda_i(z_i - \|\boldsymbol{x} - \boldsymbol{p}_i\| - \lambda_i\sigma_i^2/2) - \sum_{i\in\mathcal{N}_N} \log[Q(\lambda_i\sigma_i - \frac{z_i - \|\boldsymbol{x} - \boldsymbol{p}_i\|}{\sigma_i})] + \sum_{i\in\mathcal{N}_L} \frac{(z_i - \|\boldsymbol{x} - \boldsymbol{p}_i\|)^2}{2\sigma_i^2}\Big) \tag{2.64}$$

where z_i is the measured range corresponding to the i-th anchor with location $\boldsymbol{a}_i$, and $\mathcal{N}_N$

and $\mathcal{N}_L$ are the sets which include the indices of the NLOS and LOS links, respectively. Since the ML approach is computationally expensive, an approximate ML (AML) solution can be expressed by means of weighted nonlinear least squares (WNLS) with weights inversely proportional to the variance of the noise, as in [8]

$$\hat{\boldsymbol{x}}_{\mathrm{AML}} = \arg\min_{\boldsymbol{x}} \{ \sum_{i \in \mathcal{N}_N} \frac{(z_i - \|\boldsymbol{x} - \boldsymbol{p}_i\| - \lambda_i)^2}{\lambda_i^2} + \sum_{i \in \mathcal{N}_L} \frac{(z_i - \|\boldsymbol{x} - \boldsymbol{p}_i\|)^2}{\sigma_i^2} \} \tag{2.65}$$

Localization Using the Feasible Region Constraints: In this approach the NLOS anchors are not discarded but they are used together with the LOS measurements to construct a feasible region. If there are more NLOS anchors, then the feasible region becomes the intersection of more discs, thus it is generally smaller. If the NLOS bias is always larger than the zero-mean measurement noise, i.e., $b_i + n_i \geq 0$ with probability 1, then the noise can be neglected in the NLOS measurements. This is usually satisfied for NLOS measurements as the NLOS bias is a very large positive random variable. Under this assumption, for the i-th anchor which is identified to be in NLOS, we have

$$\|\boldsymbol{x} - \boldsymbol{p}_i\| \leq z_i. \tag{2.66}$$

In a network with fixed reference nodes facing NLOS situation a closed feasible set will be achieved as shown in Fig. 2.1.

Since the constraints in (2.66) are nonlinear, using them directly in an optimization problem increases the computation time noticeably. For reduced computational cost, the inequalities are relaxed to the rectangular constraints:

$$X_i - z_i \leq \quad x \quad \leq z_i + X_i \tag{2.67}$$

$$Y_i - z_i \leq \quad y \quad \leq z_i + Y_i. \tag{2.68}$$

After finding the approximate feasible region, only the LOS anchors are exploited to find the position. To this aim, the LOS measurement equations are linearised, therefore, the problem changes to a linear program which can be solved with low computational cost [39, 40].

A Constrained Bias Estimation Technique with Taylor Series Approximation: In this technique [42], the nonlinear equations are first linearised using the Taylor series, so the

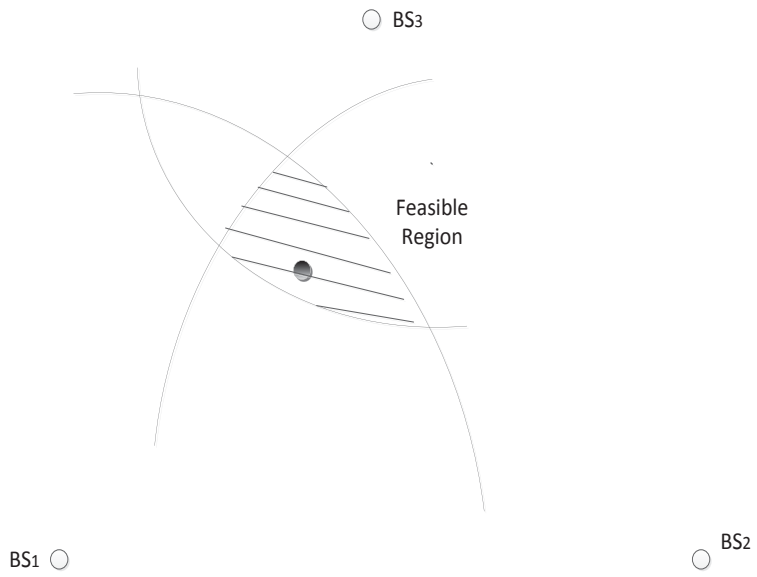

Fig. 2.1 The constraint region made by the disks defined in (2.66)

NLOS measurements can be expressed as

$$\boldsymbol{y} \approx \boldsymbol{H}_0\boldsymbol{x} + \boldsymbol{b} + \boldsymbol{n}, \tag{2.69}$$

where $\boldsymbol{H}_0$ is the Jacobian matrix of the measurement equations with respect to currently estimated location coordinates, $\boldsymbol{b}$ is the vector of NLOS biases, and $\boldsymbol{n}$ is the vector of measurement noises. If the bias vector is known then the bias-free position estimate is

$$\hat{\boldsymbol{x}} = \tilde{\boldsymbol{x}} + \boldsymbol{V}\boldsymbol{b}, \tag{2.70}$$

where

$$\tilde{\boldsymbol{x}} = -\boldsymbol{V}\boldsymbol{y}, \tag{2.71}$$

and $\boldsymbol{V}$ is the matrix:

$$\boldsymbol{V} = -(\boldsymbol{H}_0^T\boldsymbol{R}^{-1}\boldsymbol{H}_0)^{-1}\boldsymbol{H}_0^T\boldsymbol{R}^{-1}, \tag{2.72}$$

in which $\boldsymbol{R}$ is the covariance matrix of the measurement noise vector $\boldsymbol{n}$. Since in practice the bias is unknown, the following quantity can be defined

$$\boldsymbol{u} = \boldsymbol{y} - \boldsymbol{H}_0\tilde{\boldsymbol{x}} = (\boldsymbol{I} + \boldsymbol{H}_0\boldsymbol{V})\boldsymbol{b} + \boldsymbol{H}_0(\boldsymbol{x} - \hat{\boldsymbol{x}}) + \boldsymbol{n} = \boldsymbol{S}\boldsymbol{b} + \boldsymbol{w}, \tag{2.73}$$

where $\boldsymbol{w} = \boldsymbol{H}_0(\boldsymbol{x}-\hat{\boldsymbol{x}})+\boldsymbol{n}$ is an error with covariance matrix $\boldsymbol{Q}_w = \boldsymbol{H}_0(\boldsymbol{H}_0\boldsymbol{R}^{-1}\boldsymbol{H}_0^T)^{-1}\boldsymbol{H}_0^T + \boldsymbol{R}$. Then the bias is estimated by solving the following constrained optimization problem using interior point methods

$$\begin{aligned}\hat{\boldsymbol{b}} =& \arg\min_{\boldsymbol{b}}(\boldsymbol{u}-\boldsymbol{S}\boldsymbol{b})^T\boldsymbol{Q}_w^{-1}(\boldsymbol{u}-\boldsymbol{S}\boldsymbol{b}) \\ &\text{s.t. } l_i \leq b_i \leq u_i, \quad i \in \mathcal{N}_N\end{aligned} \tag{2.74}$$

where the lower bounds l_i are usually set to zero and the upper bounds u_i can be selected as in [42]. The tighter the bounds, the better the estimation of b_i. Using the estimated biases, the position can be computed from (2.70) and this process continues iteratively until convergence.

NLOS Mitigation using Hybrid Approaches

Unlike the above methods which only exploit the TOA data for localization, several techniques are proposed which exploit a hybrid of AOA, angle of departure (AOD), RSS, TOA, or Doppler spread.

AOA/AOD Measurements: In the hybrid scheme in [45], the anchors and target have to be equipped with directional antennas or smart antenna arrays. Using the fact that the scatterer, from which the strongest signal reaches the receiver, has to be located on the intersection of the line passing by the anchor with direction defined by the AOA, and the line passing by the target with direction defined by the AOD, one linear equation can be formed. By having another AOA and AOD measurements from another anchor the linear equations can be combined to give a location estimate.

TOA/AOA/Doppler Spread Measurements: The above approach requires the mobile node to be equipped with directional antennas. This is not practical and increases the cost and size of the sensor. Therefore, in [47], the Doppler spread of the moving node is also assumed to be estimated. The joint estimation of TOA/AOA/Doppler spread is given in [48]. Using the information of the Doppler spread, the AOD of the emitted wave can be estimated and then the position of the target can be found as in [45].

TOA/AOA Measurements: In many applications, the Doppler spread is not tractable and cannot be estimated. In this case only the TOA and AOA can be exploited as reliable data. If the TOA and AOA data are available then the location of the target can be

estimated using only one anchor in LOS scenario. Since there is a great chance that at least one LOS anchor is available in the network, having access to TOA and AOA data is preferred in several applications. An ESPRIT-based joint estimation of TOA and AOA using a narrowband signalling scheme is given in [93]. As the UWB has a large bandwidth, the regular techniques proposed earlier cannot yield good estimates. Therefore, several techniques have been proposed to provide the AOA information for large bandwidth signals. In [18, 94], a joint TOA-AOA estimation technique for IR-UWB is proposed. A more accurate frequency domain method is also given in [16]. In [49], a constrained optimization technique is proposed where the location of the scatterers and the target are jointly estimated. The equality and inequality constraints are non-linear and nonconvex and the approach may not be efficient. A grid-search-based technique, suited for cellular network, has been proposed in [95], where the constraints are modified. The technique in [95] outperforms the one in [49] in terms of localization performance.

TDOA/AOA Measurements: In several applications, it is difficult to synchronize the nodes and the two way communication might not be done. Therefore, the TDOA methods are preferred and exploiting the AOA information would further improve the performance in NLOS. A NLOS mitigating technique using the hybrid TDOA-AOA technique has been proposed in [36].

NLOS Localization with Known Dynamic Model

If the target is moving and its movement can be modelled by a dynamic equation, then filtering techniques are preferred as they can track the trajectory of the target more smoothly.

In addition to the aforementioned measurements, data from inertial measurements units (IMU) can be used in parallel with range information for tracking purposes [52], [53]. Some methods apply Kalman filter preprocessing on measured TOAs to smooth out the effect of the variances of the NLOS biases, while scaling the covariance matrix in an extended Kalman filter (EKF) to further mitigate the effect of their means [50], [51], [54]. However, these approaches can only achieve a moderate performance for large NLOS biases. In [55, 56], 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 esti-

mate its distribution using Kernel density estimation (KDE) techniques. In [57], a robust semi-parametric EKF is proposed for NLOS mitigation of a mobile target. The performance of this technique is improved by the interacting multiple model (IMM) algorithm in [58]. Although considered for TOA measurements, these techniques are also suitable when AOA, RSS or a hybrid of these are 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 [58].

In some other techniques, the random NLOS biases are considered as parameters in the state vector, that is, $\boldsymbol{s}[k]$ is augmented with all the biases. The biases are then jointly estimated with other state parameters [59–62], while the NLOS bias variation over time is modelled as a random walk as

$$b_i[k+1] = b_i[k] + w_b[k], \quad i \in \mathcal{N}_N \tag{2.75}$$

where $w_b[k]$ is a process noise considered to model the variation of the NLOS bias over time and $\mathcal{N}_N$ is the set including the indices of NLOS links. As an example let the unknown state vector, which includes the biases, be defined as $\boldsymbol{s}[k] = [x[k], y[k], v_x[k], v_y[k], b_1[k], \ldots, b_{|\mathcal{N}_N|}[k]]^T$. The state equation is given by

$$\boldsymbol{s}[k+1] = \boldsymbol{A}_a \boldsymbol{s}[k] + \boldsymbol{B}_a \boldsymbol{w}_a[k], \tag{2.76}$$

where

$$\boldsymbol{A}_a = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_N \end{bmatrix}, \ \boldsymbol{B}_a = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_N \end{bmatrix}, \ \boldsymbol{w}_a[k] = \begin{bmatrix} \boldsymbol{w}[k] \\ \boldsymbol{w}_b[k] \end{bmatrix},$$

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & \delta t & 0 \\ 0 & 1 & 0 & \delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 0.5\delta t^2 & 0 \\ 0 & 0.5\delta t^2 \\ \delta t & 0 \\ 0 & \delta t \end{bmatrix},$$

$\boldsymbol{w}[k]$ is normally distributed with zero mean and covariance matrix $\mathrm{diag}(\sigma_x^2, \sigma_y^2)$, $\boldsymbol{w}_b[k]$, which is uncorrelated with $\boldsymbol{w}[k]$, has a uniform distribution with zero mean and covariance matrix $\sigma_{w_b}^2 \boldsymbol{I}_N$, and δt is the time step duration. The aim is to find the position and

velocity vectors of the target at the k-th time instant, i.e., $\boldsymbol{x}[k] = [x[k], y[k]]^T$ and $\boldsymbol{v}[k] = [v_x[k], v_y[k]]^T$, and maybe the biases, based on all the past and current measurements $\boldsymbol{z}[j]$ for time instants $j \in \{1, \ldots, k\}$.

The technique in [59] uses EKF, while [60] and [61] use particle filters (PFs) that generally have a high computational cost. Although the above 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 [57].

Instead of the above models, one may impose constraints on the biases to be positive or with good approximation imposes geometrical constraints on the position coordinates, as done in (2.66). Then, there will be a need to use constrained Bayesian estimation techniques such as constrained EKF, constrained UKF, and constrained PFs. For a survey on constrained Bayesian techniques see in [96] and the references therein. In constrained EKF, the estimated state vector is projected onto the feasible region by solving an optimization problem but usually for applying the constraints on the error covariance, approximations need to be used. In [68], a constrained UKF technique has been proposed in which the sigma points of the UKF violating the constraints are projected onto the feasible region. In this way, both the *a-posteriori* state estimate and the corresponding error covariance matrix are modified according to the constraints. Constrained PFs have also been considered in the tracking community for instance assuming that a car is moving on a road or a robot moving inside a corridor in a building. There are different particle filters (PFs) that can take constraints on the state vector into account, e.g., either by rejecting every particle that falls outside the feasible region, or by continuing to sample until the particle satisfies the constraints. However, the former might result in lack of enough particles to model the posterior distribution while the latter might increase the computation time and makes the PF inefficient. Another technique is to give a zero weight to every particle that violates the constraints but in the NLOS localization problem there might be no particle left in the feasible region to be given noticeable weight. Although there are other PFs in the literature, based on solving an optimization problem for every particle violating the constraints, they still suffer from high computational cost and are thus intractable for practical applications. Therefore, the current PFs available in the literature which deal with constrained state space models may not be a suitable option for the NLOS localization problem.

2.2.4 NLOS Mitigation Using Cooperative Localization Techniques

When there is a possibility of collaboration between the sensor nodes and the number of LOS anchors is limited, the sensors can help each other in improving their location estimate. Although the main idea of cooperative localization is to overcome the problem of NLOS and outliers, still the NLOS is an issue and has to be considered with great care. The cooperative localization techniques can again be categorized in different ways. In the sequel we divide them into two categories as deterministic or probabilistic, then within each category the possibility of distributed implementation of each technique is discussed.

Deterministic Cooperative Localization in NLOS

Similar to LOS scenarios, in NLOS scenarios, the relaxation techniques can be employed by giving the NLOS measurements less weight or using them as constraints. In the sequel we discuss some of the techniques already proposed.

SDP Approaches in NLOS: In order to make the cooperative localization applicable to NLOS scenarios, in [64], the NLOS measurements are regarded as constraints as in (2.66). These constraints are then substituted into the SDP framework that was considered for LOS measurements. Another technique has also been proposed in [63], where the nodes to be localized are assumed to be inside two circular discs. Furthermore, more efficient versions of SDP relaxation, known as edge SDP (ESDP) are considered, which are shown to perform well with lower computational complexity.

Still the SDP relaxation can not be accurately implemented in a distributed manner. Below, we explain other techniques that are suitable for distributed implementation.

Cooperative POCS in NLOS: The cooperative POCS discussed earlier performs reasonably well in the presence of high NLOS contamination ratio as the NLOS measurements are positively biased. It also does not require prior identification of NLOS links and therefore it is advantageous in that sense. However, if the NLOS links can be identified accurately, then it will be often better to reject the NLOS measurements or use them as constraints and minimize the original cost function. In that case, the cooperative POCS can be used as a step to provide the sensor with reliable initial estimates [2].

IPPM in NLOS: The IPPM presented earlier works well when it is well initialized, e.g., using cooperative POCS, and when the range measurements have zero-mean errors, i.e., there exists no NLOS links. This is not practical since in indoor places there are often nodes

that have no direct sight of their neighbouring nodes. Therefore, the IPPM for NLOS has been considered in [27]. In this technique, the set of nodes around the un-localized i-th node is divided into two categories: $\mathcal{N}_L(i)$ and $\mathcal{N}_N(i)$ which correspond to the neighbouring nodes of i-th node which obtain pairwise LOS or NLOS measurement from it, respectively. Then the set of the nodes exploited for localization of the i-th node is defined as

$$\mathcal{N}^A_{\hat{\boldsymbol{x}}_i}(i) = \mathcal{N}_L(i) \cup \{j | j \in \mathcal{N}_N(i), \|\hat{\boldsymbol{x}}_i - \hat{\boldsymbol{x}}_j\| \geq z_{ij}\}. \tag{2.77}$$

where $\hat{\boldsymbol{x}}_i$ is the estimated position of i-th node in the previous iteration and z_{ij} is the pairwise range measurement between i-th and j-th nodes. The algorithm is almost the same as IPPM considered for LOS scenario described earlier, except that only the set $\mathcal{N}^A_{\hat{\boldsymbol{x}}_i}(i)$ is exploited for the aim of localization instead of $\mathcal{N}(i)$ used earlier.

Probabilistic Techniques for Cooperative Localization in NLOS

Although the belief propagation has been applied to cooperative sensor localization in LOS, in [97] by assuming *a-priori* known distribution of the NLOS, its extension to NLOS has been considered. Other works such as [98], only provide comparisons about which message passing algorithm works better in NLOS situation.

When the distribution of the NLOS is not known, applying probabilistic approaches seems irrelevant. However, by assuming that the NLOS measurements are identified and the measurement errors are positively biased, it can be assumed that the target is uniformly distributed inside a ball with its neighbour node at the centre and with radius equal to the pairwise range measurement. Therefore, in [65], by assuming that the measurements are positively biased and each sensor is inside the intersection of several balls, forming a convex set, the authors try to outer approximate each convex set by an ellipsoid. The main concept of the considered algorithm is based on SPAWN with the assumption that each sensor is uniformly distributed inside the intersection of the balls of its neighbour nodes. Therefore, each sensor iteratively finds an ellipsoid, and sends the parameters of its ellipse to its neighbours. In the end, each sensor obtains an ellipsoid which bounds the convex hull. In this algorithm the intersection of multiple ellipsoids needs to be outer approximated by the tightest ellipsoid, which is an NP-complete problem and there exist only sub-optimal solutions to solve it, to the best of our knowledge. In [65], the authors use one of the conventional techniques in [99] where first the maximum volume ellipsoid inscribed by the

intersection of ellipsoids is found and then, by expanding it with the dimension of the space, an ellipsoid containing the intersection region is obtained. However, this technique may not always yield a tight outer-approximation, therefore, there is a possibility of improving the tightness especially in 2-D by resorting to geometrical techniques.

2.3 Chapter Summary

In this chapter, we provided a brief survey of conventional localization methods for both LOS and NLOS situations. The cooperative localization scenarios were also explained in each section separately. The localization techniques in LOS scenarios have been considered for several decades and there is not much room left for investigation. Localization techniques in NLOS scenarios have also been considered in mobile positioning in urban areas and indoor places during the recent decades. Although several works have been proposed for localization in NLOS situations, there is still a potential for developing more novel methods with a better localization performance and lower computational cost. Therefore, in the following chapters, we focus on different aspects of localization techniques under NLOS scenarios.

Chapter 3

Constrained Kalman Filter for Mobile Localization in NLOS

In this chapter, a constrained square-root unscented Kalman filter (CSRUKF) is proposed for mobile localization in NLOS. In Section 3.2, a non-cooperative scenario is considered where a single target is tracked using the proposed filter. In Section 3.3, the proposed CSRUKF is extended to a cooperative scenario where multiple mobile targets are tracked. Numerical simulation results, characterizing the performance of the proposed filter for both non-cooperative and cooperative scenarios are presented in Section 3.4. Finally, Section 3.5 concludes this chapter.[1]

3.1 Introduction

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. [40], [41]. In these approaches, the position of the mobile node (MN) is constrained to be within the convex hull formed by the intersection of multiple discs, each disc 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

[1]Part of this chapter has been published in journal paper (J-1) and conference paper (C-4).

NLOS scenarios, see [44] and the references therein.

When the motion of the MN can be described by state equations, filtering techniques are preferred as they can provide smooth estimates of the target trajectory. As mentioned in Chapter 2, several techniques have been proposed for filtering techniques in NLOS scenarios. However, they are either not accurate and robust in every NLOS scenario, or computationally demanding and cannot be implemented in real time. Furthermore, using constrained Bayesian estimation techniques for tracking an MN with its location restricted to be inside the intersection of the disks corresponding to NLOS measurements has not been done before.

In this chapter, we propose an efficient square root unscented Kalman filter (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 [67] for unconstrained problems and the constrained UKF in [68]. 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 [44]. 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 [68], the sigma points of the unscented transformation 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, 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.

[2]In this work, a false alarm refers to the erroneous identification of an LOS link as being NLOS, while a missed detection (MD) refers to the opposite situation.

3.2 Centralized Non-cooperative Constrained Kalman Filter

The organization of this section is as follows: In Section 3.2.1, the system model is described and the problem formulation is presented. The proposed constrained SRUKF algorithm is developed in Section 3.2.2, along with a discussion of computational complexity.

3.2.1 Problem Formulation of Non-Cooperative Scenario

System Model

Consider a network of M fixed RNs and one MN, distributed on a 2-dimensional (2D) 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 $i \in \{1, \ldots, M\}$, while $\boldsymbol{x}_k \in \mathbb{R}^2$ 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 random acceleration model as

$$\boldsymbol{s}_k = \boldsymbol{F}\boldsymbol{s}_{k-1} + \boldsymbol{G}\boldsymbol{w}_{k-1}, \tag{3.1}$$

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

$$\boldsymbol{F} = \begin{bmatrix} 1 & 0 & \delta t & 0 \\ 0 & 1 & 0 & \delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{G} = \begin{bmatrix} \frac{\delta t^2}{2} & 0 \\ 0 & \frac{\delta t^2}{2} \\ \delta t & 0 \\ 0 & \delta t \end{bmatrix}, \tag{3.2}$$

and δt is the time step duration. The vector $\boldsymbol{w}_{k-1} \in \mathbb{R}^2$ in (3.1) is a zero-mean white Gaussian noise process (acceleration) with diagonal covariance matrix $\boldsymbol{Q} = \sigma_w^2 \boldsymbol{I}$.

In this work, 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 [7]. Let $\mathcal{L}_k$ and $\mathcal{N}_k$ denote the index sets of the RNs that are identified as 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

$$r_k^i = \begin{cases} h^i(\boldsymbol{s}_k) + n_k^i, & i \in \mathcal{L}_k, \\ h^i(\boldsymbol{s}_k) + b_k^i + n_k^i, & i \in \mathcal{N}_k, \end{cases} \tag{3.3}$$

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 modelled as independent white Gaussian processes, with zero-mean and known variance σ_n^2. 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 [14], [43], shifted Gaussian [56], and uniform [100] 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 work, we do not make any specific assumption about the distributions of the NLOS biases, although we suppose that the NLOS links are identified at every time instant.[3]

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

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 above model. Representing the NLOS biases by a simple dynamic model such as a random walk might be justified for certain environments as considered in [60], [61], 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 [101]. Including the biases b_k^i in the state vector also increases the computational complexity of the Kalman filter, therefore, it may not be computationally efficient as well.

Since the random walk model may not be an accurate approximation for the evolution

[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 which identify the NLOS link using the variance test [50], [51], [54]. For UWB applications, the features of the received TOA signal can also be employed for NLOS identifications as proposed in [32, 33, 100].

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 (especially in high SNRs), which implies that $b_k^i + n_k^i \geq 0$ [44]. In light of (3.3), this assumption is equivalent to

$$\|\boldsymbol{x}_k - \boldsymbol{a}^i\| \leq r_k^i, \quad i \in \mathcal{N}_k, \tag{3.4}$$

which is obviously a convex constraint as in [87]. If the small noise assumption cannot be made, e.g., in narrowband systems where TOA-based ranging measurement errors are relatively large, the constraints in (3.4) may not be satisfied. To avoid this limitation, we can generalize the latter inequality as

$$\|\boldsymbol{x}_k - \boldsymbol{a}^i\| \leq r_k^i + \epsilon\sigma_n, \quad i \in \mathcal{N}_k, \tag{3.5}$$

where $\epsilon \geq 0$ is a small number to ensure that the MN is located inside a disc 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 (3.5) as compared to (3.4). Therefore, we propose to use the constraint in (3.5) throughout this work due to its robustness against measurement noise and FA error in NLOS identification. In the sequel, the *feasible region*, denoted by $\mathcal{D}_k$ refers to the convex set formed by the intersection of the discs in (3.5); hence

$$\mathcal{D}_k = \Big\{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{a}^i\| \leq r_k^i + \epsilon\sigma_n, \forall i \in \mathcal{N}_k\Big\}. \tag{3.6}$$

At every time instant k, let us remove the NLOS measurements from the observations in (3.3) 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

$$\boldsymbol{z}_k = \boldsymbol{h}(\boldsymbol{s}_k) + \boldsymbol{n}_k, \tag{3.7a}$$

$$\boldsymbol{s}_k = \boldsymbol{F}\boldsymbol{s}_{k-1} + \boldsymbol{G}\boldsymbol{w}_{k-1}, \tag{3.7b}$$

$$\|\boldsymbol{x}_k - \boldsymbol{a}^i\| \leq r_k^i + \epsilon\sigma_n, \quad i \in \mathcal{N}_k, \tag{3.7c}$$

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 (3.3), 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 (3.7c) are only on the first two elements of the state vector, i.e., $\boldsymbol{x}_k$, as we have a 2D positioning scenario herein. Note that if the constraints in (3.7c) 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 [60] 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, $f_c(.|.)$, can be defined as

$$f_c(\boldsymbol{s}_k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k) = \begin{cases} \frac{1}{\beta} f(\boldsymbol{s}_k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k), & \text{if } \boldsymbol{x}_k \in \mathcal{D}_k, \\ 0, & \text{otherwise}, \end{cases} \tag{3.8}$$

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

$$\hat{\boldsymbol{s}}_k = \int_{\boldsymbol{x}_k \in \mathcal{D}_k} \boldsymbol{s}_k f_c(\boldsymbol{s}_k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k) d\boldsymbol{s}_k, \tag{3.9}$$

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

$$\hat{\boldsymbol{\Sigma}}_k = \int_{\boldsymbol{x}_k \in \mathcal{D}_k} (\boldsymbol{s}_k - \hat{\boldsymbol{s}}_k)(\boldsymbol{s}_k - \hat{\boldsymbol{s}}_k)^T f_c(\boldsymbol{s}_k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k) d\boldsymbol{s}_k. \tag{3.10}$$

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 [96]. 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 (3.9)-(3.10) have been obtained using PDF truncation along with the Gaussian assumption [102]. For

nonlinear inequality constraints, it is proposed in [102] 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 due to the linearization error [103]. In general cases with nonlinear inequality constraints, PDF truncation requires multidimensional Monte Carlo 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].

3.2.2 Non-cooperative Constrained Nonlinear Filter

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 [96]. 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 unscented transformation (UT) can give good statistical information about the mean and the error covariance matrix of the state estimate [85]. Based on this idea, in [68], 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 [68] may become numerically unstable [103]. In the following subsections, to improve the numerical stability and make the filter more efficient, we use a variation of the square root version of the UKF, known as SRUKF. The proposed variation of the SRUKF is better suited to our specific problem. Then, to overcome the above mentioned 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.

Unconstrained SRUKF Algorithm

The proposed algorithm in this part is based on the SRUKF presented in [67] 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., $\boldsymbol{\Sigma}_{k-1|k-1} = \boldsymbol{U}^T_{k-1|k-1}\boldsymbol{U}_{k-1|k-1}$. Then, for the next time instant, the *a priori* estimate of the state vector and the corresponding error covariance matrix, denoted as $\boldsymbol{s}_{k|k-1}$ and $\boldsymbol{\Sigma}_{k|k-1}$, respectively, can be obtained through prediction as

$$\boldsymbol{s}_{k|k-1} = \boldsymbol{F}\boldsymbol{s}_{k-1|k-1}, \tag{3.11}$$

$$\boldsymbol{\Sigma}_{k|k-1} = \boldsymbol{F}\boldsymbol{\Sigma}_{k-1|k-1}\boldsymbol{F}^T + \boldsymbol{G}\boldsymbol{Q}\boldsymbol{G}^T. \tag{3.12}$$

Alternatively, the computation of (3.12) can be avoided as only the Cholesky factor of the *a priori* covariance matrix, denoted by $\boldsymbol{U}_{k|k-1}$ is required [67]. To this aim, let us rewrite (3.12) as

$$\boldsymbol{\Sigma}_{k|k-1} = \begin{bmatrix} \boldsymbol{F}\boldsymbol{U}^T_{k-1|k-1} & \boldsymbol{G}\boldsymbol{Q}^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_{k-1|k-1}\boldsymbol{F}^T \\ \boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{G}^T \end{bmatrix}, \tag{3.13}$$

If we compute the QR factorization of the second matrix on the right hand side of (3.13), we obtain $\boldsymbol{U}_{k|k-1}$:

$$\boldsymbol{U}_{k|k-1} = \mathrm{qr}\left\{ \begin{bmatrix} \boldsymbol{U}_{k-1|k-1}\boldsymbol{F}^T \\ \boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{G}^T \end{bmatrix} \right\}, \tag{3.14}$$

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

With the help of $\boldsymbol{U}_{k|k-1}$, the sigma points of the SRUKF are generated as proposed in [67], i.e.:

$$\boldsymbol{s}^{(j)}_{k|k-1} = \begin{cases} \boldsymbol{s}_{k|k-1}, & j = 0, \\ \boldsymbol{s}_{k|k-1} + \sqrt{\eta_\alpha}(\boldsymbol{U}^T_{k|k-1})_j, & j = 1, \ldots, 4, \\ \boldsymbol{s}_{k|k-1} - \sqrt{\eta_\alpha}(\boldsymbol{U}^T_{k|k-1})_{j-4}, & j = 5, \ldots, 8, \end{cases} \tag{3.15}$$

$(\boldsymbol{U}^T_{k|k-1})_j$ denotes the j-th column of matrix $\boldsymbol{U}^T_{k|k-1}$, and η_α is a tuning parameter which controls the spread of the sigma points. To better understand the geometric meaning

of parameter η_α, 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 η_α in (3.15) is chosen such that $\Pr(\eta_k \leq \eta_\alpha) = \alpha$, where $0 < \alpha < 1$ represents a desired confidence level. Then, the region of $\mathbb{R}^4$ defined by $\eta_k \leq \eta_\alpha$ represents a confidence ellipsoid, on the boundary of which the sigma points in (3.15) (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 η_α 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 η has a Chi-square distribution with 4 degrees of freedom and it becomes easy to find a value for η_α corresponding to a certain ellipsoid with confidence level α.[4]

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

$$\boldsymbol{z}_{k|k-1}^{(j)} = \boldsymbol{h}(\boldsymbol{s}_{k|k-1}^{(j)}), \quad j = 0,\ldots,8. \tag{3.16}$$

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 [104]:

$$\hat{\boldsymbol{z}}_{k|k-1} = \sum_{j=0}^{8} w^{(j)}\boldsymbol{z}_{k|k-1}^{(j)}, \tag{3.17}$$

$$\boldsymbol{\Sigma}_{k|k-1}^{\boldsymbol{s},\boldsymbol{z}} = \sum_{j=0}^{8} w^{(j)}(\boldsymbol{s}_{k|k-1}^{(j)} - \boldsymbol{s}_{k|k-1})(\boldsymbol{z}_{k|k-1}^{(j)} - \hat{\boldsymbol{z}}_{k|k-1})^T, \tag{3.18}$$

$$\boldsymbol{P}_{k|k-1}^{z} = \sum_{j=0}^{8} w^{(j)}(\boldsymbol{z}_{k|k-1}^{(j)} - \hat{\boldsymbol{z}}_{k|k-1})(\boldsymbol{z}_{k|k-1}^{(j)} - \hat{\boldsymbol{z}}_{k|k-1})^T + \boldsymbol{R}, \tag{3.19}$$

where $\boldsymbol{R}$ is the covariance matrix of the measurement noise $\boldsymbol{n}_k$ in (3.7a) and the weights $w^{(j)}$ appearing in these expressions are defined in a similar way as in [105]:

$$w^{(j)} = \begin{cases} 1 - \frac{4}{\eta_\alpha}, & j = 0, \\ \frac{1}{2\eta_\alpha}, & j = 1,\ldots,8, \end{cases} \tag{3.20}$$

[4]The Matlab built-in function `chi2inv`$(\alpha, 4)$ can be used for this purpose.

and therefore satisfy $\sum_{j=0}^{8} w^{(j)} = 1$.

If the weight $w^{(0)}$ in (3.20) is negative, it is possible that the covariance matrix obtained through (3.19) becomes indefinite (i.e., with negative eigenvalues). However, by choosing a sufficiently large value of α, we can guarantee that $\eta_\alpha \geq 4$; in turn, this implies that $w^{(0)} \geq 0$ and the covariance matrix (3.19) then becomes positive definite. In this work, 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 state vector of dimension 4 and based on the Chi-square assumption for η_k, it follows that if $\alpha > 0.6$, then $\eta_\alpha > 4$ and the positive definiteness of (3.19) is guaranteed.

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

$$\mathbf{e}_z^{(j)} = \sqrt{w^{(j)}}(\boldsymbol{z}^{(j)}_{k|k-1} - \hat{\boldsymbol{z}}_{k|k-1}), \quad j = 0, \ldots, 8, \tag{3.21}$$

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

$$\boldsymbol{U}_{\boldsymbol{z}_k} = \mathrm{qr}\left\{\left[\mathbf{e}_z^{(0)}, \mathbf{e}_z^{(1)}, \ldots, \mathbf{e}_z^{(8)}, \boldsymbol{R}^{\frac{1}{2}}\right]^T\right\}. \tag{3.22}$$

It is proposed in [67] to first compute the Kalman gain

$$\boldsymbol{K}_k = \boldsymbol{\Sigma}^{\boldsymbol{s},\boldsymbol{z}}_{k|k-1}(\boldsymbol{P}^{z}_{k|k-1})^{-1} = \boldsymbol{\Sigma}^{\boldsymbol{s},\boldsymbol{z}}_{k|k-1}\boldsymbol{U}^{-1}_{z_k}\boldsymbol{U}^{-T}_{z_k}, \tag{3.23}$$

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

$$\boldsymbol{s}_{k|k} = \boldsymbol{s}_{k|k-1} + \boldsymbol{K}_k(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_{k|k-1}), \tag{3.24}$$

$$\boldsymbol{U}_{k|k} = \mathrm{cholupdate}\{\boldsymbol{U}_{k|k-1}, \boldsymbol{K}_k\boldsymbol{U}^{T}_{z_k}, -1\}, \tag{3.25}$$

[5] In [85], a scaled version of the unscented transformation 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 [85] and [67].

where 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 (3.25) follows from the covariance matrix update

$$\boldsymbol{\Sigma}_{k|k} \triangleq \boldsymbol{U}_{k|k}^T\boldsymbol{U}_{k|k} = \boldsymbol{U}_{k|k-1}^T\boldsymbol{U}_{k|k-1} - \boldsymbol{K}_k\boldsymbol{U}_{z_k}^T\boldsymbol{U}_{z_k}\boldsymbol{K}_k^T. \tag{3.26}$$

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

$$\boldsymbol{T}_k = \boldsymbol{\Sigma}_{k|k-1}^{\boldsymbol{s},\boldsymbol{z}}\boldsymbol{U}_{z_k}^{-1}, \tag{3.27}$$

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 (3.23) that $\boldsymbol{K}_k = \boldsymbol{T}_k\boldsymbol{U}_{z_k}^{-T}$. Substituting this expression into (3.24) we obtain

$$\boldsymbol{s}_{k|k} = \boldsymbol{s}_{k|k-1} + \boldsymbol{T}_k\boldsymbol{U}_{z_k}^{-T}(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_{k|k-1}), \tag{3.28}$$

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

$$\boldsymbol{U}_{z_k}^T\boldsymbol{y}_k = \boldsymbol{z}_k - \hat{\boldsymbol{z}}_{k|k-1}. \tag{3.29}$$

From (3.25) and (3.27), it follows that the covariance matrix can be updated as

$$\boldsymbol{\Sigma}_{k|k} = \boldsymbol{U}_{k|k-1}^T\boldsymbol{U}_{k|k-1} - \boldsymbol{T}_k\boldsymbol{T}_k^T, \tag{3.30}$$

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

$$\boldsymbol{U}_{k|k} = \text{cholupdate}\{\boldsymbol{U}_{k|k-1}, \boldsymbol{T}_k, -1\}. \tag{3.31}$$

Compared to the algorithm in [67], this modified algorithm for the estimation of $\boldsymbol{s}_{k|k}$ and $\boldsymbol{U}_{k|k}$ saves about $2L|\mathcal{N}_k|^2$ 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 (3.11) and the Cholesky factor of the predicted covariance matrix in (3.14) to replace the *a posteriori* state vector in (3.28)

[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.

and Cholesky factor of the error covariance matrix in (3.31), respectively.

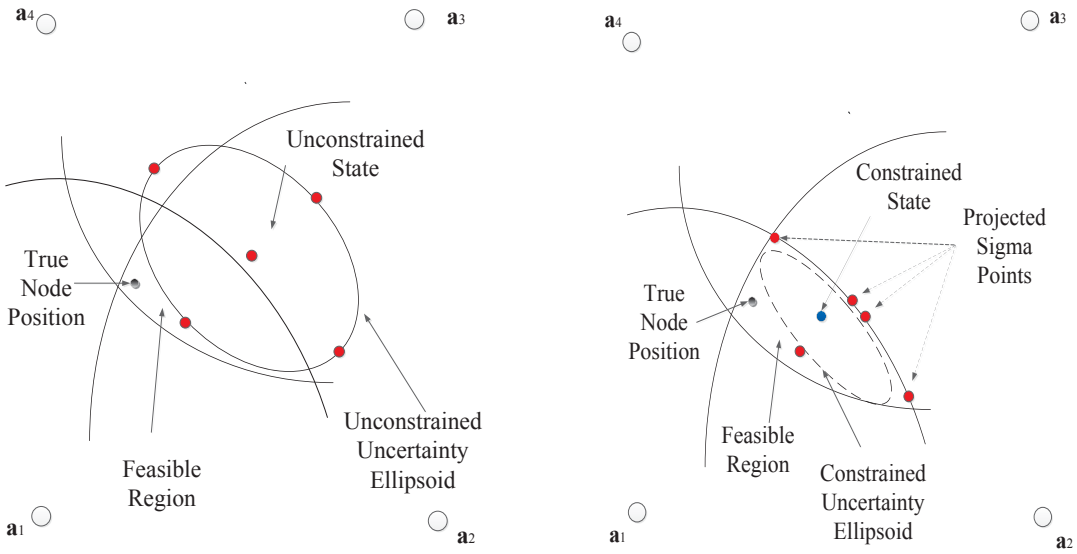

Fig. 3.1 Left: Unconstrained state estimate and the uncertainty ellipsoid of sigma points. Right: The projected sigma points and the shrunk uncertainty ellipsoid.

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 (3.7c) into account. To impose the constraints on the estimated state and error covariance matrix, similar to [68], a new set of sigma points are generated according to

$$\boldsymbol{s}_{k|k}^{(j)} = \begin{cases} \boldsymbol{s}_{k|k}, & j = 0, \\ \boldsymbol{s}_{k|k} + \sqrt{\eta_\alpha}(\boldsymbol{U}_{k|k}^T)_j, & j = 1, \ldots, 4, \\ \boldsymbol{s}_{k|k} - \sqrt{\eta_\alpha}(\boldsymbol{U}_{k|k}^T)_{j-4}, & j = 5, \ldots, 8. \end{cases} \tag{3.32}$$

The generated sigma points (except $\boldsymbol{s}_{k|k}^{(0)}$) form an uncertainty ellipsoid with $\boldsymbol{s}_{k|k}$ at its centre as illustrated in Fig. 3.1 for the case that the state vector is of size 2, i.e., no velocity components exist. After the generation of sigma points $\boldsymbol{s}_{k|k}^{(j)}$ with desired confidence ellipsoid, those which violate the constraints are projected onto the convex feasible region

through

$$\begin{aligned} \mathcal{P}(\boldsymbol{s}_{k|k}^{(j)}) = \arg\min_{\boldsymbol{q}} \ \Big\{ (\boldsymbol{q} - \boldsymbol{s}_{k|k}^{(j)})^T \boldsymbol{W}_k (\boldsymbol{q} - \boldsymbol{s}_{k|k}^{(j)}) \Big\}, \\ \text{s.t.} \quad \big\| \boldsymbol{q}(1:2) - \boldsymbol{a}^i \big\| \leq r_k^i + \epsilon\sigma_n, \quad i \in \mathcal{N}_k, \end{aligned} \tag{3.33}$$

where $\boldsymbol{W}_k$ is a symmetric positive definite (SPD) weighting matrix [103], [104]. One reasonable choice is $\boldsymbol{W}_k = \boldsymbol{\Sigma}_{k|k}^{-1}$ which takes into account the uncertainty in $\boldsymbol{s}_{k|k}$, and 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 [106]. In the state space model under consideration, since the non-linearity is not too high, and the noises are Gaussian, using the inverse of the covariance matrix as $\boldsymbol{W}_k$ might still give a solution that is close to optimal. Therefore, this weighting matrix is used for the projection operation.

The optimization problem in (3.33) is a quadratically constrained quadratic program (QCQP), which is convex since $\boldsymbol{W}_k$ is SPD and the constraints are convex [99, 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. Suppose that $\boldsymbol{q}(1:2)$ is fixed, then we can find the optimal $\boldsymbol{q}(3:4)$, which is a function of $\boldsymbol{q}(1:2)$. By substituting the optimal $\boldsymbol{q}(3:4)$ into the cost function, we obtain a QCQP, which only involves the unknown $\boldsymbol{q}(1:2)$.

However, in the above 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 [107] in order to reformulate and reduce the size of the convex QCQP problem in (3.33) 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 (3.33) can be expressed as $(\boldsymbol{q} - \boldsymbol{s}_{k|k}^{(j)})^T \boldsymbol{U}_{k|k}^{-1} \boldsymbol{U}_{k|k}^{-T} (\boldsymbol{q} - \boldsymbol{s}_{k|k}^{(j)})$. To get around the inverse operation, we define

$$\boldsymbol{u} = \boldsymbol{U}_{k|k}^{-T} (\boldsymbol{s}_{k|k}^{(j)} - \boldsymbol{q}), \tag{3.34}$$

from which it follows that

$$\boldsymbol{q} = \boldsymbol{s}_{k|k}^{(j)} - \boldsymbol{U}_{k|k}^T \boldsymbol{u}. \tag{3.35}$$

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

$$\boldsymbol{U}_{k|k}^T = \begin{bmatrix} \boldsymbol{L}_{11} & \boldsymbol{0} \\ \boldsymbol{L}_{21} & \boldsymbol{L}_{22} \end{bmatrix}, \tag{3.36}$$

where $\boldsymbol{L}_{11} \in \mathbb{R}^{2\times 2}$ and $\boldsymbol{L}_{22} \in \mathbb{R}^{2\times 2}$ are lower triangular. Then it follows from (3.35) that

$$\boldsymbol{q}(1:2) = \boldsymbol{s}_{k|k}^{(j)}(1:2) - \boldsymbol{L}_{11}\boldsymbol{u}(1:2). \tag{3.37}$$

Using (3.34) and (3.37), we can reformulate the QCQP problem (3.33) as

$$\begin{aligned} &\min_{\boldsymbol{u}} \Big\{ \boldsymbol{u}^T(1:2)\boldsymbol{u}(1:2) + \boldsymbol{u}^T(3:4)\boldsymbol{u}(3:4) \Big\}, \\ &\text{s.t. } \big\| \boldsymbol{L}_{11}\boldsymbol{u}(1:2) - (\boldsymbol{s}_{k|k}^{(j)}(1:2) - \boldsymbol{a}^i) \big\| \leq r_k^i + \epsilon\sigma_n,\ i \in \mathcal{N}_k. \end{aligned} \tag{3.38}$$

Since the constraints do not include $\boldsymbol{u}(3:4)$, the optimal choice is obviously $\boldsymbol{u}(3:4) = \boldsymbol{0}$ and the optimization problem (3.38) becomes

$$\begin{aligned} &\min_{\boldsymbol{u}(1:2)} \Big\{ \boldsymbol{u}^T(1:2)\boldsymbol{u}(1:2) \Big\}, \\ &\text{s.t. } \big\| \boldsymbol{L}_{11}\boldsymbol{u}(1:2) - (\boldsymbol{s}_{k|k}^{(j)}(1:2) - \boldsymbol{a}^i) \big\| \leq r_k^i + \epsilon\sigma_n,\ i \in \mathcal{N}_k. \end{aligned} \tag{3.39}$$

This 2D convex QCQP problem can now be solved efficiently using iterative techniques [99].

After finding the optimal $\boldsymbol{u}(1:2)$, we can compute the optimal $\boldsymbol{q}$ using (3.35) and the fact that the optimal $\boldsymbol{u}(3:4) = \boldsymbol{0}$ as follows:

$$\mathcal{P}(\boldsymbol{s}_{k|k}^{(j)}) \triangleq \boldsymbol{q} = \boldsymbol{s}_{k|k}^{(j)} - \begin{bmatrix} \boldsymbol{L}_{11} \\ \boldsymbol{L}_{21} \end{bmatrix} \boldsymbol{u}(1:2). \tag{3.40}$$

The above approach for reducing the size of the QCQP problem (3.33) not only avoids a matrix inverse computation, which may cause numerical instability (see [107]), 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 (3.31). We note that in some particular scenarios, especially under FA in NLOS identification, it is possible that the feasible region $\mathcal{D}_k$ in (3.6) becomes empty and consequently, (3.39) has no solution. In this case, which might rarely happen, we simply propose to increase ϵ until $\mathcal{D}_k$ becomes

non-empty.

After finding the projected sigma points through (3.40), the mean and covariance matrix may be estimated through weighted averaging

$$\boldsymbol{s}_{k|k}^{\mathcal{P}} = \sum_{j=0}^{8} w^{(j)} \mathcal{P}(\boldsymbol{s}_{k|k}^{(j)}), \tag{3.41}$$

$$\boldsymbol{\Sigma}_{k|k}^{\mathcal{P}} = \sum_{j=0}^{8} w^{(j)} (\mathcal{P}(\boldsymbol{s}_{k|k}^{(j)}) - \boldsymbol{s}_{k|k}^{\mathcal{P}})(\mathcal{P}(\boldsymbol{s}_{k|k}^{(j)}) - \boldsymbol{s}_{k|k}^{\mathcal{P}})^T. \tag{3.42}$$

As before, instead of (3.42) we compute the Cholesky factor $\boldsymbol{U}_{k|k}^{\mathcal{P}}$ of $\boldsymbol{\Sigma}_{k|k}^{\mathcal{P}}$:

$$\begin{aligned} \mathbf{e}_{\mathcal{P}}^{(j)} &= \sqrt{w^{(j)}}(\mathcal{P}(\boldsymbol{s}_{k|k}^{(j)}) - \boldsymbol{s}_{k|k}^{\mathcal{P}})), \quad j = 0, \ldots, 4, \\ \boldsymbol{U}_{k|k}^{\mathcal{P}} &= \mathrm{qr}\left\{ [\mathbf{e}_{\mathcal{P}}^{(0)}, \mathbf{e}_{\mathcal{P}}^{(1)}, \ldots, \mathbf{e}_{\mathcal{P}}^{(8)}]^T \right\}. \end{aligned} \tag{3.43}$$

As illustrated in Fig. 3.1, the projected sigma points 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 [108]. 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

$$\boldsymbol{s}_{k|k} = \boldsymbol{s}_{k|k}^{\mathcal{P}}, \tag{3.44}$$

$$\boldsymbol{U}_{k|k} = \boldsymbol{U}_{k|k}^{\mathcal{P}}. \tag{3.45}$$

3.2.3 Non-cooperative CSRUKF Summary

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

The SRUKF is more efficient and numerically stable than UKF and the computational complexity analysis has also been presented in [67], in which it is shown that this algorithm

Algorithm 2 CSRUKF

Initialize $\boldsymbol{s}_{0|0}$ and set $\boldsymbol{\Sigma}_{0|0}$ to a large SPD diagonal matrix.
Set η_α and ϵ
for $k = 1, \ldots, K$ **do**
  Prediction of $\boldsymbol{s}_{k|k-1}$ using (3.11), and $\boldsymbol{U}_{k|k-1}$ using (3.14).
  if $|\mathcal{L}_k| = 0$ **then**
    Set $\boldsymbol{s}_{k|k} = \boldsymbol{s}_{k|k-1}$ and $\boldsymbol{U}_{k|k} = \boldsymbol{U}_{k|k-1}$.
  else
    Find the predicted measurement through (3.16).
    Calculate the predicted mean (3.17) and implement qr$\{.\}$ in (3.22).
    Estimate the cross-covariance in (3.18).
    Solve (3.27) to find $\boldsymbol{T}_k$.
    Estimate the *a posteriori* mean $\boldsymbol{s}_{k|k}$ using (3.28) and Cholesky factor of *a posteriori* covariance matrix $\boldsymbol{U}_{k|k}$ using (3.31).
  end if
  Generate the sigma points using (3.32).
  For every sigma point whose first two elements fall outside $\mathcal{D}_k$ solve (3.39) and find the projected point (3.40).
  Estimate $\boldsymbol{s}^{\mathcal{P}}_{k|k}$ using (3.41) and $\boldsymbol{U}^{\mathcal{P}}_{k|k}$ using (3.43).
  Replace $\boldsymbol{s}^{\mathcal{P}}_{k|k}$ and $\boldsymbol{U}^{\mathcal{P}}_{k|k}$ as the *a posteriori* estimates, i.e., (3.44) and (3.45).
end for

requires at each time step $\mathcal{O}(D_s^3)$, where D_s is the dimension of the state vector. Since the dimension of the state vector is small and fixed, the computational cost of the first stage of our algorithm is generally small compared to the cost of the second stage where the projection operations are done.

The QCQP in (3.39) is a convex optimization problem and can be solved in polynomial time using an extended optimization package in Matlab such as Sedumi [109]. Since $\boldsymbol{u}(1:2) \in \mathbb{R}^2$, the optimization problem can be solved with moderate cost for 9 sigma points at most. In addition, 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 (3.32) that fall inside the feasible region, as the projection operation needs not to be applied on them. By tuning the parameter α we can achieve a trade-off between accuracy and computational cost. On the one hand, if α 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 α may degrade the localization performance as the estimated quantities remain unchanged after applying the constraints. On the other hand, selecting a large α increases the computational cost but at the same time may result in sampling many of the non-local points, and thus the linearisation of $\boldsymbol{h}(\boldsymbol{s}_k)$ might be inaccurate [85]. In our simulations, it is observed that selecting $0.65 \leq \alpha \leq 0.85$ can offer a reasonable trade-off in terms of accuracy and computational cost.

3.3 Centralized Cooperative Constrained Kalman Filter

In this section, we extend the centralized CSRUKF proposed in the previous section to the cooperative localization scenario where the are N target to be tracked and there are pairwise sensor-sensor measurements in addition to sensor-anchor measurements. The variables used throughout this section are different from the ones used for non-cooperative case in Section 3.2, unless it is stated explicitly. The organization of the this section is as follows. The system model and problem formulation are presented in Subsection 3.3.1. The proposed algorithm is developed in Subsection 3.3.2 and summarized in Subsection 3.3.3.

3.3.1 Problem Statement of Cooperative Localization

We consider a 2D WSN (the extension to 3D is straightforward) comprised of N MNs with unknown locations $\boldsymbol{x}_k^i \in \mathbb{R}^2$ and unknown speeds $\boldsymbol{v}_k^i \in \mathbb{R}^2$ for $i = \{1, \ldots, N\}$ at discrete time instant k, and of M anchors with known and fixed locations $\boldsymbol{a}^i$ for $i = \{N+1, \ldots, N+M\}$. We assume that each MN moves independently according to a random acceleration model as

$$\boldsymbol{x}_{k+1}^i = \boldsymbol{x}_k^i + \boldsymbol{v}_k^i \delta t + \boldsymbol{w}_k^i \frac{\delta t^2}{2}, \quad i = 1, \ldots, N \tag{3.46}$$

where δt is the time step duration and $\boldsymbol{w}_k^i \in \mathbb{R}^2$ is a zero-mean white Gaussian random process.

Pairwise range measurements between neighbouring nodes, obtained either by one way ranging and prior network synchronization or through TWR, are modelled as

$$r_k^{ij} = \begin{cases} \|\boldsymbol{x}_k^i - \boldsymbol{a}^j\| + n_k^{ij}, & (i,j) \in \mathcal{L}_a \\ \|\boldsymbol{x}_k^i - \boldsymbol{x}_k^j\| + n_k^{ij}, & (i,j) \in \mathcal{L}_s \\ \|\boldsymbol{x}_k^i - \boldsymbol{a}^j\| + b_k^{ij} + n_k^{ij}, & (i,j) \in \mathcal{N}_a \\ \|\boldsymbol{x}_k^i - \boldsymbol{x}_k^j\| + b_k^{ij} + n_k^{ij}, & (i,j) \in \mathcal{N}_s \end{cases} \tag{3.47}$$

where

$$\begin{aligned} \mathcal{L}_a &: \{(i,j) : \text{LOS link between } i\text{-th sensor and } j\text{-th anchor}\} \\ \mathcal{L}_s &: \{(i,j) : \text{LOS link between } i\text{-th and } j\text{-th sensors}\} \\ \mathcal{N}_a &: \{(i,j) : \text{NLOS link between } i\text{-th sensor and } j\text{-th anchor}\} \\ \mathcal{N}_s &: \{(i,j) : \text{NLOS link between } i\text{-th and } j\text{-th sensors}\} \end{aligned}$$

In (3.47), n_k^{ij} is a zero-mean Gaussian noise with known variance σ_n^2, and b_k^{ij} is the NLOS bias which is a positive random variable.

The NLOS biases are large random variables and for some systems with high SNR it can be assumed that $b_k^{ij} + n_k^{ij} \geq 0$, which is equivalent to stating that for all the NLOS

measurements

$$\|\boldsymbol{x}_k^i - \boldsymbol{a}^j\| \leq r_k^{ij}, \quad (i,j) \in \mathcal{N}_a \tag{3.48}$$

$$\|\boldsymbol{x}_k^i - \boldsymbol{x}_k^j\| \leq r_k^{ij}, \quad (i,j) \in \mathcal{N}_s \tag{3.49}$$

In order to increase the robustness against large noise samples we use the following constraints instead

$$\|\boldsymbol{x}_k^i - \boldsymbol{a}^j\| \leq r_k^{ij} + \epsilon\sigma_n, \quad (i,j) \in \mathcal{N}_a \tag{3.50}$$

$$\|\boldsymbol{x}_k^i - \boldsymbol{x}_k^j\| \leq r_k^{ij} + \epsilon\sigma_n, \quad (i,j) \in \mathcal{N}_s \tag{3.51}$$

where $\epsilon \geq 0$ is a tuning parameter which increases the chance that the inequalities hold true. Note that the new inequalities in (3.50)-(3.51) might also hold true for LOS measurements; therefore, if a link is LOS but wrongly identified as being NLOS, the inequalities in (3.50)-(3.51) have a higher chance to be satisfied compared to the ones in (3.48)-(3.49).

In some works, the NLOS biases have been modelled by the random walk model, and therefore, they are included in the state vector and estimated together with other unknowns [60], [62]. However, random walk approximately models the relationship between b_k^{ij} and b_{k+1}^{ij}, and selecting a suitable variance for the increment of the random walk is not easy for dynamic environments. We therefore, avoid including them in the state vector and estimating them, however, we remove the NLOS measurements from the observation vector and instead use these measurements to impose the constraints in (3.50) and (3.51) on the positions of sensors. The measurement vector $\boldsymbol{z}_k$ is obtained by stacking together all the LOS measurements r_k^{ij} for $(i,j) \in \mathcal{L}_a \cup \mathcal{L}_s$. The state of all the sensors can also be expressed in a vector form as $\boldsymbol{s}_k = [\boldsymbol{x}_k^1, \boldsymbol{x}_k^2, \ldots, \boldsymbol{x}_k^N, \dot{\boldsymbol{x}}_k^1, \dot{\boldsymbol{x}}_k^2, \ldots, \dot{\boldsymbol{x}}_k^N]^T \in \mathbb{R}^{4N}$. We can finally formulate the constrained state space model as

$$\boldsymbol{z}_k = \boldsymbol{h}(\boldsymbol{s}_k) + \boldsymbol{n}_k \tag{3.52}$$

$$\boldsymbol{s}_k = \boldsymbol{F}\boldsymbol{s}_{k-1} + \boldsymbol{G}\boldsymbol{w}_k \tag{3.53}$$

$$\text{s.t.} \quad \|\boldsymbol{x}_k^i - \boldsymbol{a}^j\| \leq r_k^{ij} + \epsilon\sigma_n, \quad (i,j) \in \mathcal{N}_a \tag{3.54}$$

$$\|\boldsymbol{x}_k^i - \boldsymbol{x}_k^j\| \leq r_k^{ij} + \epsilon\sigma_n, \quad (i,j) \in \mathcal{N}_s \tag{3.55}$$

where $\boldsymbol{h}(\boldsymbol{s}_k)$ is the vector of true ranges, $\boldsymbol{n}_k$ is the vector of measurement errors with zero-

mean and covariance matrix $\mathbf{R} = \sigma_n^2\mathbf{I}$, $\boldsymbol{w}_k = [(\boldsymbol{w}_k^1)^T, (\boldsymbol{w}_k^2)^T, \ldots, (\boldsymbol{w}_k^N)^T]^T$ is a zero-mean Gaussian vector with covariance matrix $\boldsymbol{Q}$ and

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{I}_{2N} & \delta t \boldsymbol{I}_{2N} \\ \boldsymbol{0}_{2N} & \boldsymbol{I}_{2N} \end{bmatrix} \in \mathbb{R}^{4N\times 4N}, \quad \boldsymbol{G} = \begin{bmatrix} \frac{\delta t^2}{2}\boldsymbol{I}_{2N} \\ \delta t \boldsymbol{I}_{2N} \end{bmatrix} \in \mathbb{R}^{4N\times 2N} \tag{3.56}$$

In the following, we show how to estimate $\boldsymbol{s}_k$ based on the history of the range measurements and the constraints.

3.3.2 Centralized Cooperative CSRUKF

In this section, the proposed CSRUKF for non-cooperative scenario is extended to cooperative case. The first stage, i.e., an unconstrained SRUKF [67], is almost the same as the one presented for non-cooperative case. The second stage varies slightly as the optimization problem and the constraints have changed due to the extra sensor-sensor measurements and the increased dimension of the state vector.

Unconstrained SRUKF

If there are no constraints on the state vector, then a nonlinear Kalman filter can be used for the *a-posteriori* estimation of state and the Cholesky factor of the corresponding error covariance matrix, i.e., $\boldsymbol{s}_{k|k} \in \mathbb{R}^{4N}$ and $\boldsymbol{U}_{k|k} \in \mathbb{R}^{4N\times 4N}$, respectively, where $\boldsymbol{U}_{k|k}^T\boldsymbol{U}_{k|k} = \boldsymbol{\Sigma}_{k|k}$. We use a SRUKF for this aim, as proposed in [67], and find $\boldsymbol{s}_{k|k}$ and the $\boldsymbol{U}_{k|k}$, as described in Subsection 3.2.2, where the size of the state vector is now $4N$. A detailed explanation about selection of parameters η_α and ϵ and their relation with the weights $w^{(l)}$ can be found in Section 3.2.

Projection of Sigma Points

Assume that 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 into account. To impose the constraints on the estimated state and error covariance matrix,

similar to [68], a new set of sigma points are generated according to

$$\boldsymbol{s}_{k|k}^{(l)} = \begin{cases} \boldsymbol{s}_{k|k}, & l = 0, \\ \boldsymbol{s}_{k|k} + \sqrt{\eta_\alpha}(\boldsymbol{U}_{k|k}^T)_l, & l = 1, \ldots, 4N, \\ \boldsymbol{s}_{k|k} - \sqrt{\eta_\alpha}(\boldsymbol{U}_{k|k}^T)_{l-4N}, & l = 4N+1, \ldots, 8N. \end{cases} \tag{3.57}$$

After the generation of sigma points $\boldsymbol{s}_{k|k}^{(l)}$, those which violate the constraints are projected onto the convex feasible region through

$$\begin{aligned} \mathcal{P}(\boldsymbol{s}_{k|k}^{(l)}) = \arg\min_{\boldsymbol{q}} \ & \Big\{(\boldsymbol{q} - \boldsymbol{s}_{k|k}^{(l)})^T \boldsymbol{W}_k (\boldsymbol{q} - \boldsymbol{s}_{k|k}^{(l)})\Big\}, \\ \text{s.t. } & \big\|\boldsymbol{q}(2i-1{:}2i) - \boldsymbol{a}^j\big\| \le r_k^{ij} + \epsilon\sigma_n, \quad (i,j) \in \mathcal{N}_a \\ & \big\|\boldsymbol{q}(2i-1{:}2i) - \boldsymbol{q}(2j-1{:}2j)\big\| \le r_k^{ij} + \epsilon\sigma_n, \ (i,j) \in \mathcal{N}_s \end{aligned} \tag{3.58}$$

where $\boldsymbol{W}_k = \boldsymbol{\Sigma}_{k|k}^{-1}$ is chosen as done for non-cooperative case.

As the constraints are only on the first $2N$ elements of the state vector, i.e., the position coordinates, it is possible to reduce the size of the QCQP problem similar to before. Recalling that $\boldsymbol{\Sigma}_{k|k} = \boldsymbol{U}_{k|k}^T \boldsymbol{U}_{k|k}$, the objective function in (3.58) can be expressed as $(\boldsymbol{q} - \boldsymbol{s}_{k|k}^{(l)})^T \boldsymbol{U}_{k|k}^{-1} \boldsymbol{U}_{k|k}^{-T} (\boldsymbol{q} - \boldsymbol{s}_{k|k}^{(l)})$. To avoid the inverse operation, we define

$$\boldsymbol{u} = \boldsymbol{U}_{k|k}^{-T}(\boldsymbol{s}_{k|k}^{(l)} - \boldsymbol{q}). \tag{3.59}$$

Then it follows that

$$\boldsymbol{q} = \boldsymbol{s}_{k|k}^{(l)} - \boldsymbol{U}_{k|k}^T \boldsymbol{u}. \tag{3.60}$$

We partition the lower triangular matrix $\boldsymbol{U}_{k|k}^T$ as follows:

$$\boldsymbol{U}_{k|k}^T = \begin{bmatrix} \boldsymbol{L}_{11} & \boldsymbol{0} \\ \boldsymbol{L}_{21} & \boldsymbol{L}_{22} \end{bmatrix}, \tag{3.61}$$

where $\boldsymbol{L}_{11} \in \mathbb{R}^{2N \times 2N}$ and $\boldsymbol{L}_{22} \in \mathbb{R}^{2N \times 2N}$ are lower triangular. Then from (3.60) we have

$$\boldsymbol{q}(1{:}2N) = \boldsymbol{s}_{k|k}^{(l)}(1{:}2N) - \boldsymbol{L}_{11}\boldsymbol{u}(1{:}2N). \tag{3.62}$$

Using (3.60) and (3.62), and noting that the optimal $\boldsymbol{u}(2N+1:4N) = \boldsymbol{0}$ (since the

constraints only depend on $\boldsymbol{u}(1:2N)$) we can reformulate the QCQP problem (3.58) as (3.63).

$$\begin{aligned}
&\min_{\boldsymbol{u}(1:2N)} \left\{ \boldsymbol{u}^T(1{:}2N)\boldsymbol{u}(1{:}2N) \right\} \qquad (3.63)\\
&\text{s.t. } \left\| \boldsymbol{s}_{k|k}^{(l)}(2i-1{:}2i) - \boldsymbol{L}_{11}(2i-1:2i)\boldsymbol{u}(1:2N) - \boldsymbol{a}^j \right\| \leq r_k^{ij} + \epsilon\sigma_n,\ (i,j) \in \mathcal{N}_a\\
&\quad \left\| \boldsymbol{s}_{k|k}^{(l)}(2i-1{:}2i) - \boldsymbol{L}_{11}(2i-1:2i)\boldsymbol{u}(1:2N) \right.\\
&\quad \left. - \boldsymbol{s}_{k|k}^{(l)}(2j-1{:}2j) + \boldsymbol{L}_{11}(2j-1:2j)\boldsymbol{u}(1:2N) \right\| \leq r_k^{ij} + \epsilon\sigma_n,\ (i,j) \in \mathcal{N}_s
\end{aligned}$$

This convex QCQP problem can now be solved efficiently using iterative techniques [99]. After finding the optimal $\boldsymbol{u}(1:2N)$, we compute the optimal $\boldsymbol{q}$ using (3.60) and the fact that the optimal $\boldsymbol{u}(2N+1{:}4N) = \boldsymbol{0}$ as follows:

$$\mathcal{P}(\boldsymbol{s}_{k|k}^{(l)}) \triangleq \boldsymbol{q} = \boldsymbol{s}_{k|k}^{(l)} - \begin{bmatrix} \boldsymbol{L}_{11} \\ \boldsymbol{L}_{21} \end{bmatrix} \boldsymbol{u}(1{:}2N). \qquad (3.64)$$

After finding the projected sigma points through (3.64), the state and the Cholesky factor of the error covariance matrix may be estimated through weighted averaging and QR factorization, respectively as

$$\boldsymbol{s}_{k|k}^{\mathcal{P}} = \sum_{l=0}^{8N} w^{(l)} \mathcal{P}(\boldsymbol{s}_{k|k}^{(l)}), \qquad (3.65)$$

$$\boldsymbol{U}_{k|k}^{\mathcal{P}} = \mathrm{qr}\left\{ [\mathbf{e}_{\mathcal{P}}^{(0)}, \mathbf{e}_{\mathcal{P}}^{(1)}, \ldots, \mathbf{e}_{\mathcal{P}}^{(8N)}]^T \right\}, \qquad (3.66)$$

where

$$\mathbf{e}_{\mathcal{P}}^{(l)} = \sqrt{w^{(l)}}(\mathcal{P}(\boldsymbol{s}_{k|k}^{(l)}) - \boldsymbol{s}_{k|k}^{\mathcal{P}})), \quad l = 0, \ldots, 8N.$$

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

$$\boldsymbol{s}_{k|k} = \boldsymbol{s}_{k|k}^{\mathcal{P}}, \qquad (3.67)$$

$$\boldsymbol{U}_{k|k} = \boldsymbol{U}_{k|k}^{\mathcal{P}}. \qquad (3.68)$$

3.3.3 Centralized Algorithm Summary

The proposed two stage centralized CSRUKF for the cooperative scenario is summarized in Algorithm 3. By setting $N = 1$ in this algorithm and removing the sensor-sensor measurements, the filter changes to non-cooperative CSRUKF proposed in Section 3.2 for tracking a single target.

Algorithm 3 CSRUKF

Initialize $\boldsymbol{s}_{0|0}$ and set $\boldsymbol{\Sigma}_{0|0}$ to a large SPD diagonal matrix
Set η_α and ϵ
for $k = 1, \ldots, K$ **do**
  Estimate $\boldsymbol{s}_{k|k}$ and $\boldsymbol{U}_{k|k}$ using a conventional SRUKF as described in [67].
  Generate the sigma points using (3.57).
  For every sigma point whose first two elements fall outside the feasible region solve (3.63) and find the projected point (3.64).
  Estimate $\boldsymbol{s}^{\mathcal{P}}_{k|k}$ using (3.65) and $\boldsymbol{U}^{\mathcal{P}}_{k|k}$ using (3.66).
  Update the *a posteriori* estimates, i.e., (3.67) and (3.68).
end for

3.4 Simulation Results

3.4.1 Non-cooperative Case

The simulations are implemented in Matlab 2010b on a 64-bit computer with Intel i7-2600 3.4GHz processor and 12GB of RAM. We consider a 2-D area with $M = 5$ fixed RNs located at known positions $\boldsymbol{a}^1 = [0, 0]$, $\boldsymbol{a}^2 = [2000, 0]^T$, $\boldsymbol{a}^3 = [0, 2000]^T$, $\boldsymbol{a}^4 = [-2000, 0]^T$, $\boldsymbol{a}^5 = [0, -2000]^T$, where the units are in meters. An MN moves on this 2-D plane according to the motion model considered earlier in (3.1) 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 $\text{diag}([10^4, 10^4, 10^2, 10^2])$.

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

[7]In practice, knowledge about σ_n can be obtained by means of preliminary calibration experiments in

noise assumption can model general applications like narrowband cellular mobile positioning as considered in [57], [110], while the small noise assumption is suitable for localization applications with more accurate ranging and higher SNRs. We also perturb some of the measurements by NLOS biases which are modelled as exponential random variables with parameter $\gamma = 500$m [58][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 [111], [54] 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 4 NLOS RNs all the time, while the other RN (the one in the center of the plane) can change between LOS and NLOS.
- Scenario II: There are 3 NLOS RNs all the time, while the other 2 RNs can transit between LOS and NLOS.
- Scenario III: There are 2 NLOS RNs all the time, while the other 3 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` [112] and `Sedumi` solver [109].

In order 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 [106]. 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

a given environment, or through on-line calculation based on path-loss model for radio propagation.

[8]Our algorithm can still work well with a range dependent NLOS bias model as considered in [43].

one projection operation needs to be done at each iteration. We denote this approach by projection Kalman filter (PKF) and for solving the optimization problem we follow the similar procedure as done for CSRUKF.

For comparison purposes, we consider the conventional techniques proposed in [51], [111], [113], 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 pre-processing and variance calculation, we consider the simple one in [111] 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 (CRLB) analysis in NLOS shows that if no prior statistics about the distribution of the NLOS bias is available then the optimal strategy is to discard the NLOS measurements and only use LOS ones [29]. If prior statistics are available then the NLOS measurements should also be used to achieve a lower MSE. 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. Even though the posterior Cramer-Rao bound (PCRB) on positioning RMSE has been derived approximately in [114,115], these derivations are based on the assumption that the NLOS bias has a Gaussian distribution with known mean and variance. 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 is 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 (3.1) and with 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 σ_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 Kalman filters 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 $T = 100$ Monte Carlo (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 Monte Carlo trial, respectively. The performance metrics are the cumulative distribution function (CDF) of the positioning error e_k, expressed as

$$\mathrm{CDF}(e_k) = \mathbb{P}\Big[\|\boldsymbol{x}_k^t - \boldsymbol{x}_{k|k}^t\| \leq e_k\Big], \tag{3.69}$$

and the root mean square error (RMSE) of the position estimates at time step k, defined as

$$\bar{e}_k = \sqrt{\mathbb{E}\Big[(\boldsymbol{x}_k^t - \boldsymbol{x}_{k|k}^t)^T(\boldsymbol{x}_k^t - \boldsymbol{x}_{k|k}^t)\Big]}, \tag{3.70}$$

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 $\boldsymbol{\Sigma}_{0|0} = \mathrm{diag}([9 \times 10^4, 9 \times 10^4, 10^3, 10^3])$.

Large Measurement Noise

In the first scenario, we consider the case of a narrowband ranging application where the noise variance is relatively high, i.e., $\sigma_n = 150$m is considered. The RMSE versus time step is illustrated in Fig. 3.2 for scenarios I, II, and III. The corresponding CDF of the positioning error is also plotted for each scenario in Fig. 3.3.

We can observe that for scenario I and II, the CSRUKF performs almost similar to

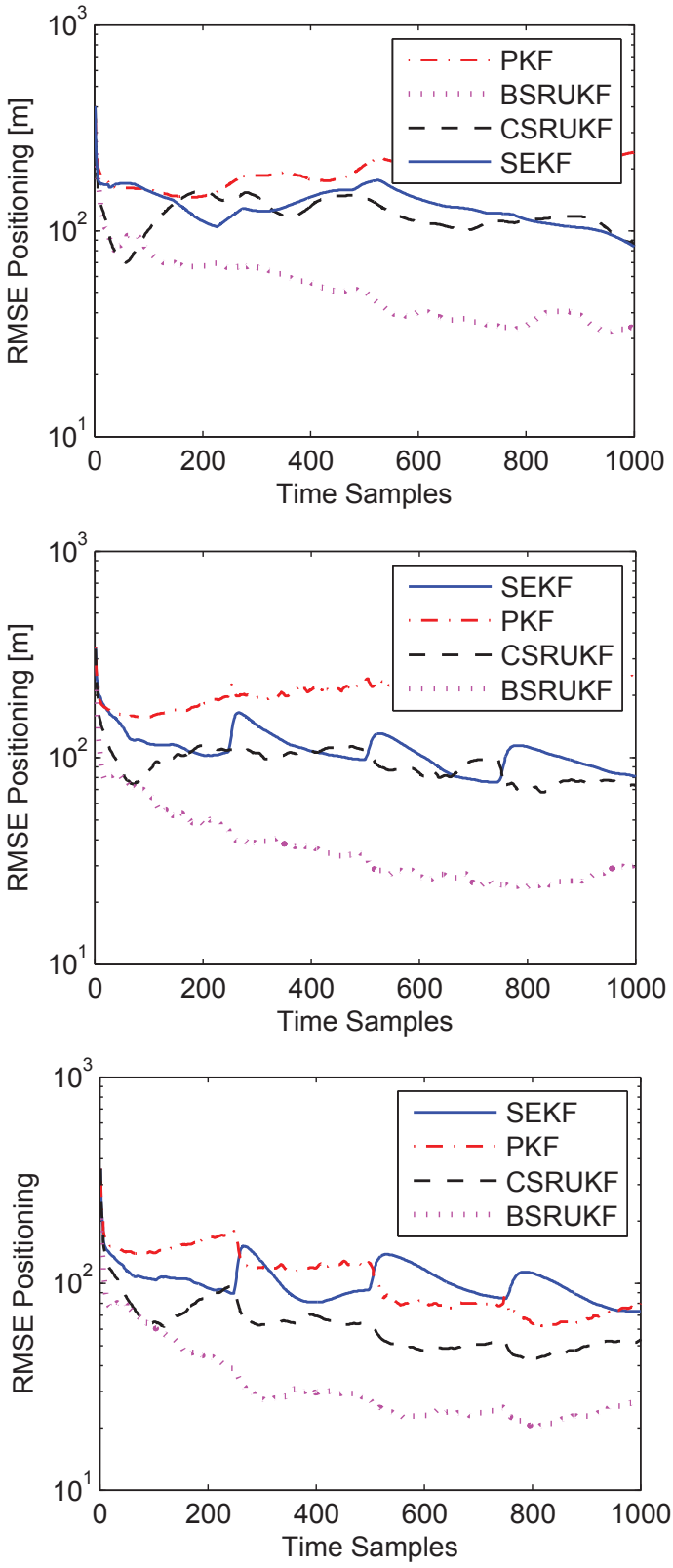

Fig. 3.2 From top to bottom: RMSE for scenarios I, II, and III.

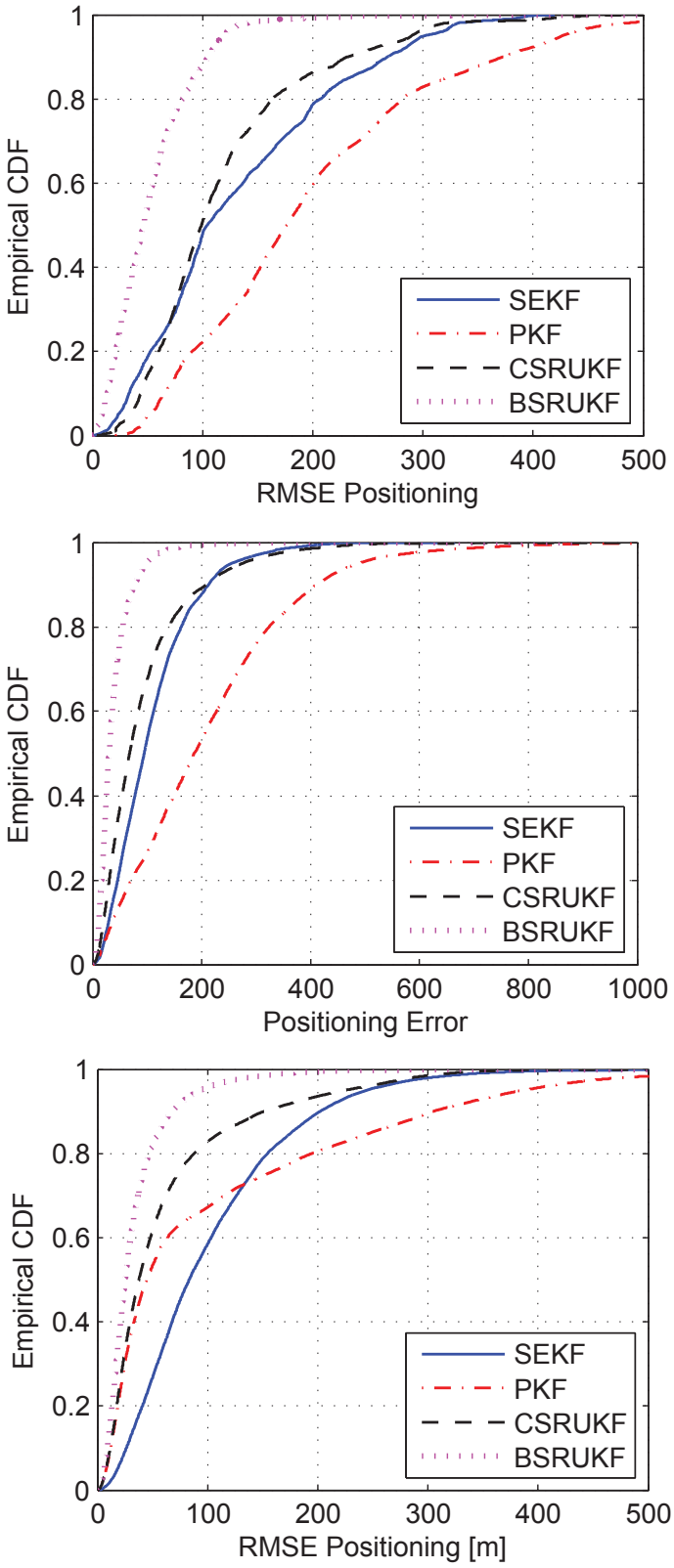

Fig. 3.3 From top to bottom: CDF for scenarios I, II, and III.

SEKF, while the RMSE of the PKF is relatively high. This shows that in order to obtain a decent localization performance, the projection of all the sigma points in CSRUKF is necessary as compared to projecting only the mean as done in PKF. The RMSE of all the techniques are lower bounded by the RMSE of 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 PKF and SEKF noticeably.

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 ranging applications with high SNR. The RMSE and CDF of the estimation error are illustrated in Fig. 3.4, and Fig. 3.5, respectively, for the scenarios I, II, and III.

As observed in Fig. 3.4 and Fig. 3.5, in all the scenarios, the proposed CSRUKF performs better than all the other methods, especially the BSRUKF. There are several reasons why 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 Kalman filter 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. 3.2 and Fig. 3.3, 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 BSRUKF was performing better than CSRUKF. Second, when σ_n is large, the feasible region $\mathcal{D}_k$ becomes larger compared to the case that σ_n is small. Therefore, it is more likely that most of the sigma points lie inside $\mathcal{D}_k$ and no projection is done, so 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.

Robustness to Errors in NLOS Identification

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

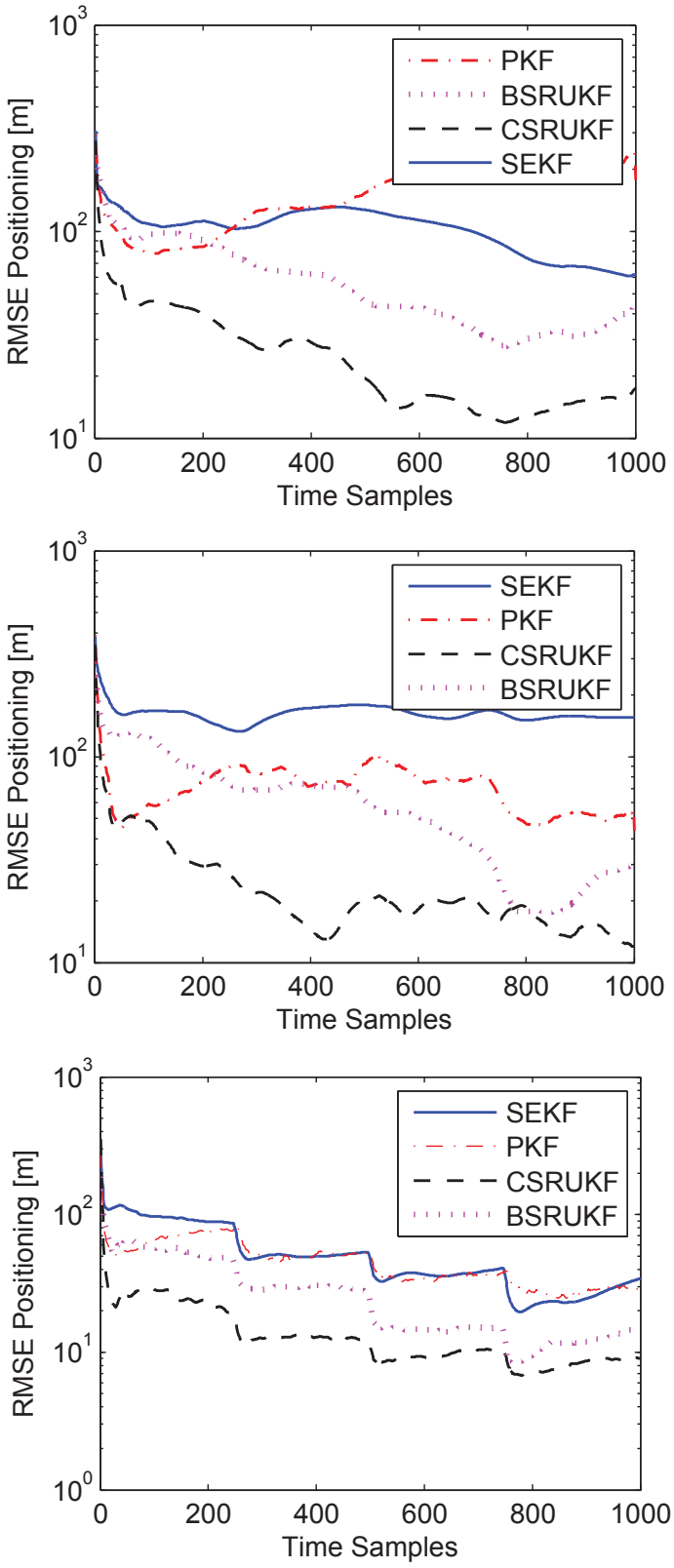

Fig. 3.4 From top to bottom: RMSE for scenarios I, II, and III.

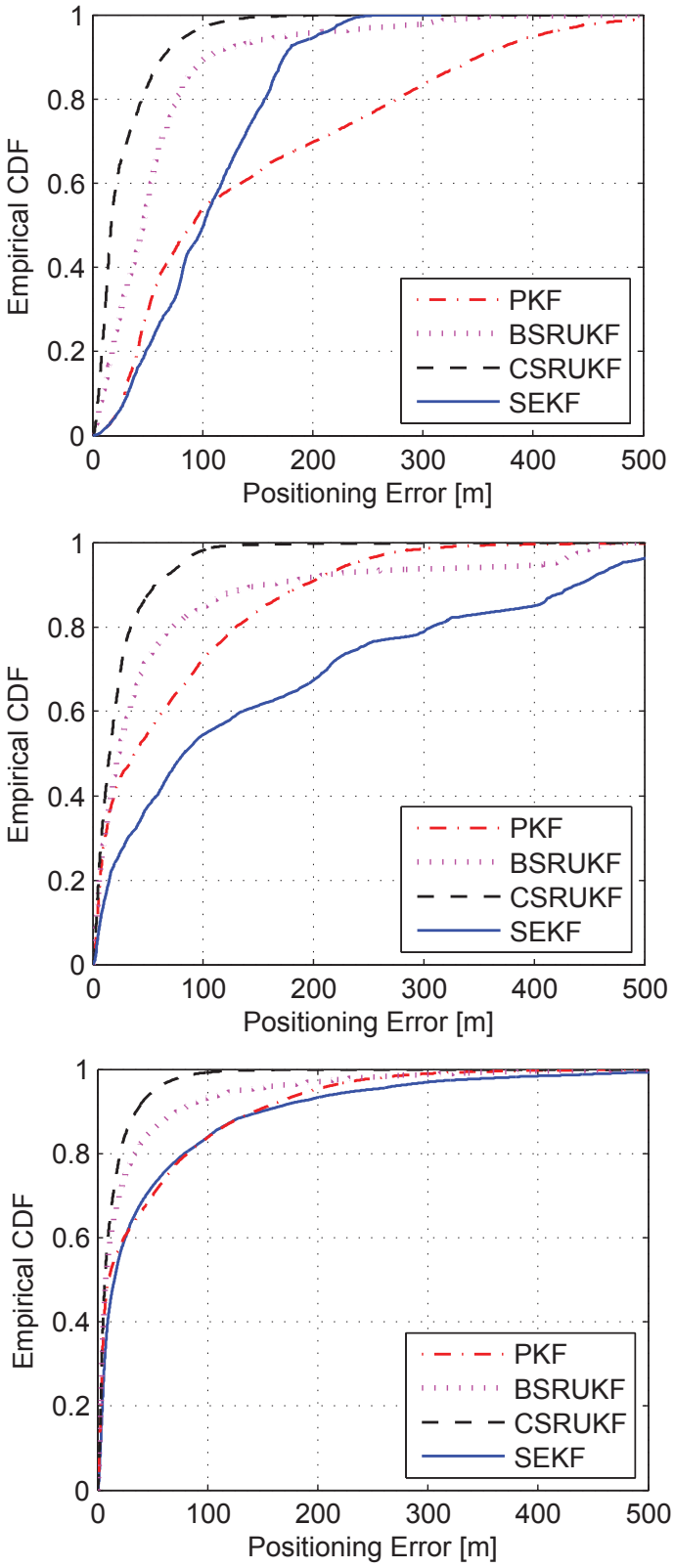

Fig. 3.5 From top to bottom: CDF for scenarios I, II, and III.

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, 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 ϵ can increase the chance that a LOS measurement also satisfies the constraint in (3.5), it is expected that FA does not severely degrade the performance of our proposed technique. The simulation results are shown in Fig. 3.6, where it is observed that CSRUKF is robust against FA error in NLOS identification and outperforms the SEKF. Note that the BSRUKF algorithm is evaluated under perfect NLOS identification, while its performance is still worse than our proposed technique.

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. We have avoided showing the simulation results for this scenario. As a remark, we note that for our algorithm to perform well most of the times, the threshold used for NLOS identification should change such that the probability of MD becomes very small.

Computation Time

The average computation time of each algorithm is calculated for each scenario and is shown in Table 3.4.1. Due to the use of optimization packages in Matlab, which use iterative methods, it may not be possible to express the computation time of the proposed CSRUKF in terms of the number of flops. Therefore, we compare the CPU time required for running each algorithm instead. The SEKF has a very small computation time, because it is essentially an ordinary Kalman filter. The computation time of PKF, where only the state vector needs to be projected onto the feasible region, is obviously lower than CSRUKF because only one QCQP problem might need to be solved at every time step. The most computationally demanding part of 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 CSRUKF. Still, the highest computation time of CSRUKF is lower than the total elapsed time for the entire trajectory (200 seconds), meaning that with the computer used here, the algorithm

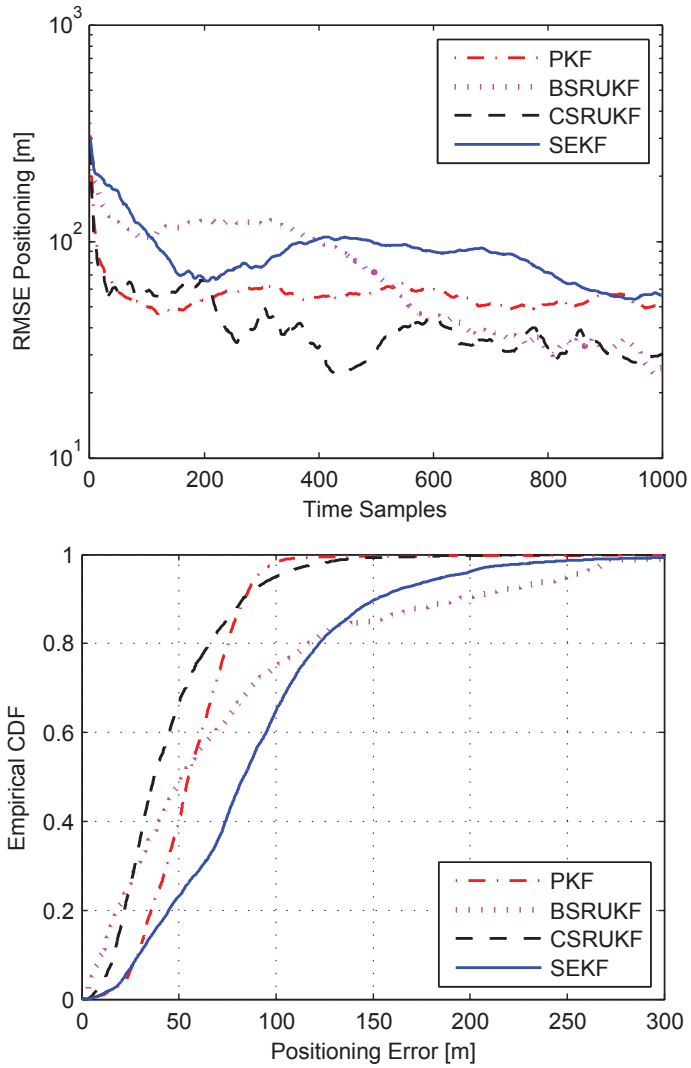

Fig. 3.6 Comparison with FA in identification of an NLOS anchor, top: RMSE, bottom CDF.

can be applied for online tracking. Note that the computational cost of many other popular methods such as KDE or particle filters is also much higher than ordinary EKF or SRUKF and therefore, our algorithm remains competitive in terms of computation time.

Table 3.1 Average running time (in seconds) of each algorithm in each scenario evaluated for the entire trajectory (200s)

	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

3.4.2 Cooperative Case

To evaluate the performance of the proposed technique, we consider a 2D area with $M = 4$ anchors and $N = 5$ mobile sensors. The sensors are initially placed uniformly on the plane, and move independently according to the considered motion model in (3.46) with $\delta t = 0.2$ for 100 time steps. The anchors are located at positions $\boldsymbol{a}^6 = [0, 0]$, $\boldsymbol{a}^7 = [0, 10]$, $\boldsymbol{a}^8 = [10, 0]$, and $\boldsymbol{a}^9 = [10, 10]$. We assume that if the true distance between the nodes is less than 10m then they are regarded as neighbours and they obtain pairwise range measurements. Although the communication range decreases if the link between two nodes is NLOS, we assume that the communication range is the same for all the links due to simplicity. The true range between neighbouring nodes is disturbed with a Gaussian noise with zero-mean and standard deviation $\sigma_n = 0.1$m in order to model the range measurement. Ranging with centimetres accuracy is in accordance with IEEE 802.15.4a in indoor environments with LOS connection [14]. For the NLOS links, a positive exponential random variable with mean and standard deviation of 5m is also added to the obtained ranges. The tuning parameters in CSRUKF are set as $\epsilon = 3$ and $\eta_\alpha = 0.8$. The convex QCQP problems are solved using Sedumi solver [109] and Yalmip optimization package [112].

To test the algorithm, we consider three different NLOS scenarios where the probability of a link being NLOS, denoted as $P_{\mathcal{N}}$ varies from a low to a high value, as follows:

- Small ratio of NLOS to LOS links: $P_{\mathcal{N}} = 0.05$
- Moderate ratio of NLOS to LOS links: $P_{\mathcal{N}} = 0.5$

- Large ratio of NLOS to LOS links: $P_{\mathcal{N}} = 0.95$

For comparison purposes, we consider an unconstrained SRUKF with rejection of NLOS links, denoted by "SRUKF Outlier Rejection". If there are enough LOS measurements available for each node, then outlier rejection is the right strategy, but in the absence of enough LOS nodes the performance of this method might be severely degraded. As a performance metric, the cumulative distribution function (CDF) of the network positioning error, defined as

$$\mathrm{CDF}(e_k) = \Pr\Big\{\frac{1}{N}\sum_{i=1}^{N} \|\boldsymbol{s}_{k|k}(2i-1{:}2i) - \boldsymbol{s}_k(2i-1{:}2i)\| \leq e_k\Big\}$$

is evaluated empirically for different values of $P_{\mathcal{N}}$.

As observed in Fig. 3.7, for low ratio of NLOS to LOS links, the SRUKF with outlier rejection performs almost the same as our proposed CSRUKF. This confirms that using a few NLOS links as constraints might not improve the localization performance. However, for $P_{\mathcal{N}} = 0.5$, the performance of Kalman filtering with outlier rejection is degraded noticeably (it sometimes even diverges) while that of the proposed CSRUKF is less than 1m with 90% chance. For large ratios of NLOS to LOS links, the outlier rejection technique diverges because it essentially uses the prediction step of the SRUKF (the measurement vector $\boldsymbol{z}_k$ is empty most of the times), while the proposed technique has a decent performance.

3.5 Conclusion and Future Work

A constrained square-root unscented Kalman filter (CSRUKF) with projection technique was proposed in this chapter for the aim of TOA-based localization of MNs in NLOS scenarios. The CSRUKF algorithm was first developed for tracking a single MN but was then extended to cooperative scenario where multiple MNs, which exchange information with each other, are tracked centrally by the filter. In these filters, 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 which violated the constraints were projected on the feasible region by solving a convex quadratically constrained quadratic program (QCQP). As compared to other constrained UKF techniques, we considered a square root filter and avoided computing the inverse

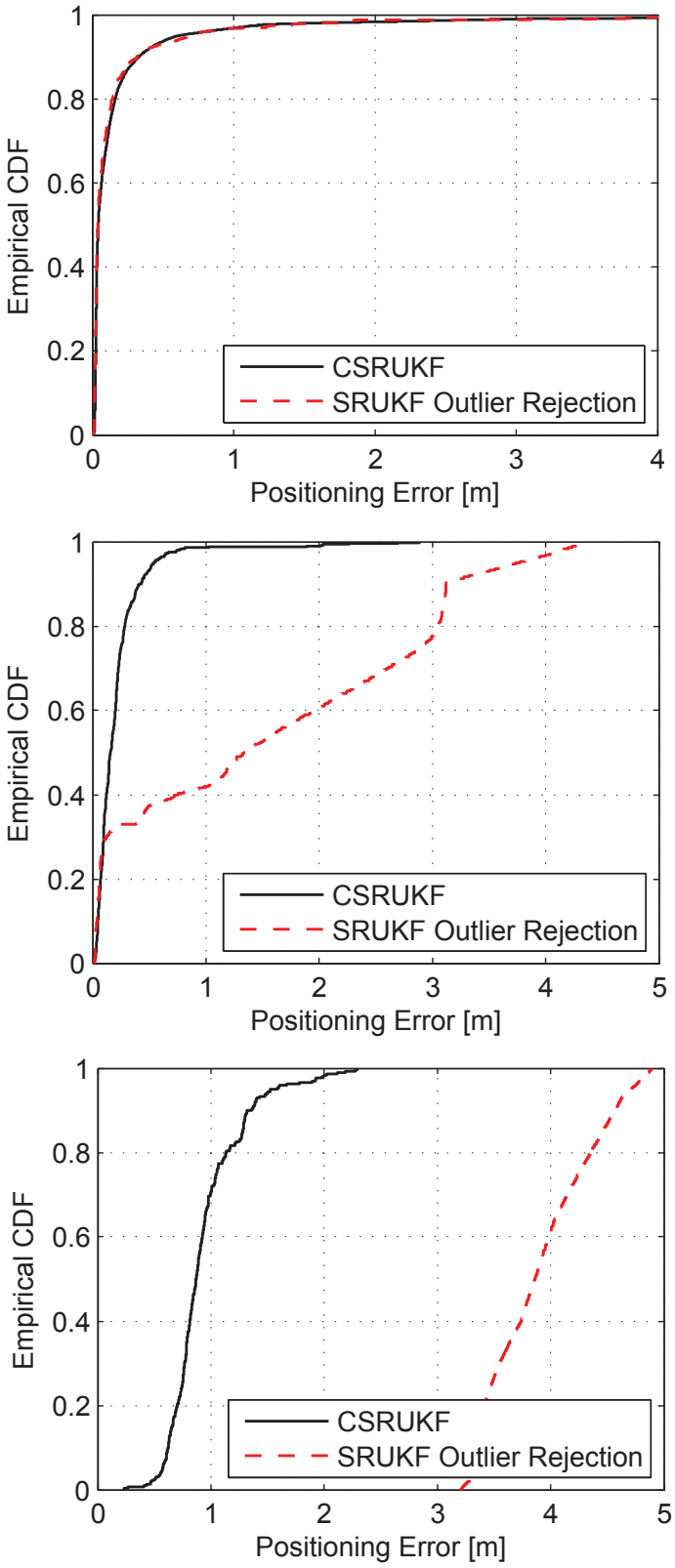

Fig. 3.7 CDFs of the network positioning error in different scenarios: (a) $P_{\mathcal{N}} = 0.05$; (b) $P_{\mathcal{N}} = 0.5$; (c) $P_{\mathcal{N}} = 0.95$.

of the state covariance matrix both in the Kalman filter 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 accurate TOA-based localization systems with high SNR. The computational cost of the proposed filter was relatively high especially for cooperative MN scenario due to the large number of optimization problems that needed to be solved for projecting the sigma points. However, the computation issue seems to be less of a problem for a cellular network with limited number of MNs and with the help of parallel processing on the servers, that are capable of doing large computations. 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.

Chapter 4

Robust Distributed Cooperative Localization in WSN

In this chapter, we propose a two-stage algorithm based on Huber M-estimation for distributed cooperative WSN localization in the presence of unidentified NLOS links. In Section 4.1, an introduction to the topic and state of the art is given. In Section 4.2 the system model is described and the problem is formulated. The proposed algorithm is described in Section 4.3. The numerical comparison with other methods is given in Section 4.4. Finally, Section 4.5 concludes this chapter.

4.1 Introduction

The WSN localization based on different measurements can be carried out for the entire network in a centralized or distributed fashion. Among the popular centralized algorithms are semi-definite programming (SDP) [23] and second-order cone programming (SOCP) [24] convex relaxations. Distributed algorithms have also been proposed, including distributed SOCP [25], the iterative parallel projection method (IPPM) [27] and other localization approaches that alternate between convex and non convex optimization problems [26] [89].

However, these approaches only consider the case where the pairwise range measurements are made under line of sight (LOS) condition. In practice, LOS measurements are limited and many links will face a non-line of sight (NLOS) condition. Due to the NLOS, the TOA measurements become positively biased [44] and consequently, the aforementioned techniques perform unsatisfactorily if the NLOS effects are not mitigated properly.

Many methods have been proposed for identification of NLOS links for non-cooperative networks (see in [44] and the references therein). Furthermore, many NLOS identification methods have been proposed especially for UWB systems [33], [100]. After detecting the NLOS measurements, the NLOS effect can be mitigated using different optimization techniques. A summary of the non-cooperative TOA-based NLOS localization methods is given in [44]. For cooperative localization, extension of the centralized SDP relaxation and the distributed IPPM for NLOS scenarios are considered in [64] and [1], respectively.

NLOS identification is however challenging for a large WSN with several pairwise measurements. Therefore, in many WSN applications, it is impractical to assume that all the NLOS links are identified accurately. In [116], a SDP relaxation is considered for non-cooperative localization in NLOS without prior detection of NLOS links. Although the SDP relaxation is a reliable centralized technique, it can not scale with the size of the network. Extension of SDP relaxation to a distributed manner needs to be further studied. In [2], a distributed cooperative POCS is employed to estimate the location of sensors, and is shown to be robust against NLOS errors. However, if only a portion of the measurements are affected by NLOS errors, the performance of POCS is far from being optimal. Another technique for robust estimation against outliers without prior outlier detection is to use Huber loss function, which is a trade-off between l_1 and l_2 norm minimizations [117]. In contrast to POCS, localization based on Huber cost function can achieve a good result only if it is well initialized and if a moderate or small portion of the measurements are contaminated by large errors, otherwise it may not necessarily give a good estimate.

In this chapter, we propose a two-stage algorithm based on Huber M-estimation for distributed cooperative localization in the presence of unidentified NLOS links. In the first stage, similar convex relaxation considered in [26] is applied on the Huber cost function, hence accurate sensor locations are iteratively estimated. Since the performance may not necessarily be good in low ratio of NLOS to LOS links, in the second stage, the original Huber cost function is minimized iteratively with a suitable choice of parameter. For iterative optimization in both stages, we use a simple gradient descent technique since it can be easily implemented in a distributed manner. Through simulations, we first show that the proposed convex relaxation gives a reliable estimate in different NLOS scenarios. Furthermore, we show that the position estimates are generally improved in the second stage as we minimize the original Huber cost function. The robustness of our algorithm to outliers is also evaluated by using real sensor measurement set obtained by measurement

campaign in [69].

4.2 System Model and Problem Formulation

4.2.1 System Model

We consider a sensor network, consisting of N sensor nodes with unknown locations denoted by $\boldsymbol{x}_i \in \mathbb{R}^2$, for $i = 1, \ldots, N$, and M anchors with known locations $\boldsymbol{x}_i \in \mathbb{R}^2$, for $i = N+1, \ldots, N+M$. We define $\mathcal{S}$ as the set of all indices (i, j) of all the neighbouring nodes that can communicate with each other and we assume $i < j$ to avoid repetition. We also define $\mathcal{S}_i$ as the set of indices of all the neighbouring nodes of $\boldsymbol{x}_i$. We assume that range measurement is obtained between each two pairs of neighbouring nodes with $(i, j) \in \mathcal{S}$. For accurate ranging, either the nodes have to be accurately synchronized or the two-way ranging (TWR) protocol can be exploited to remove the effect of clock parameters in the TOA measurements [7]. The range measurement model is represented as

$$r_{ij} = \begin{cases} d_{ij} + n_{ij}, & (i,j) \in \mathcal{L} \\ d_{ij} + b_{ij} + n_{ij}, & (i,j) \in \mathcal{N} \end{cases} \tag{4.1}$$

where $d_{ij} = \|\boldsymbol{x}_i - \boldsymbol{x}_j\|$, the sets are defined as

$$\mathcal{L} = \{(i,j) \in \mathcal{S} : \text{LOS link between } i\text{-th and } j\text{-th node}\}$$
$$\mathcal{N} = \{(i,j) \in \mathcal{S} : \text{NLOS link between } i\text{-th and } j\text{-th node}\}$$

hence it is followed that $\mathcal{S} = \mathcal{L} \cup \mathcal{N}$. The measurement noise n_{ij} is a zero mean independent and identically distributed Gaussian random variable with known variance σ_n^2, and b_{ij} is the NLOS bias between i-th and j-th nodes, which has been modelled differently depending on the environment and wireless channel, e.g., exponential [33] or uniform [100] distributions are generally used. However, in this work, we assume that we do not know *a priori* if a link is NLOS or not. Furthermore, we do not have any knowledge about the statistics of the NLOS biases, e.g., the mean and variance of b_{ij}.

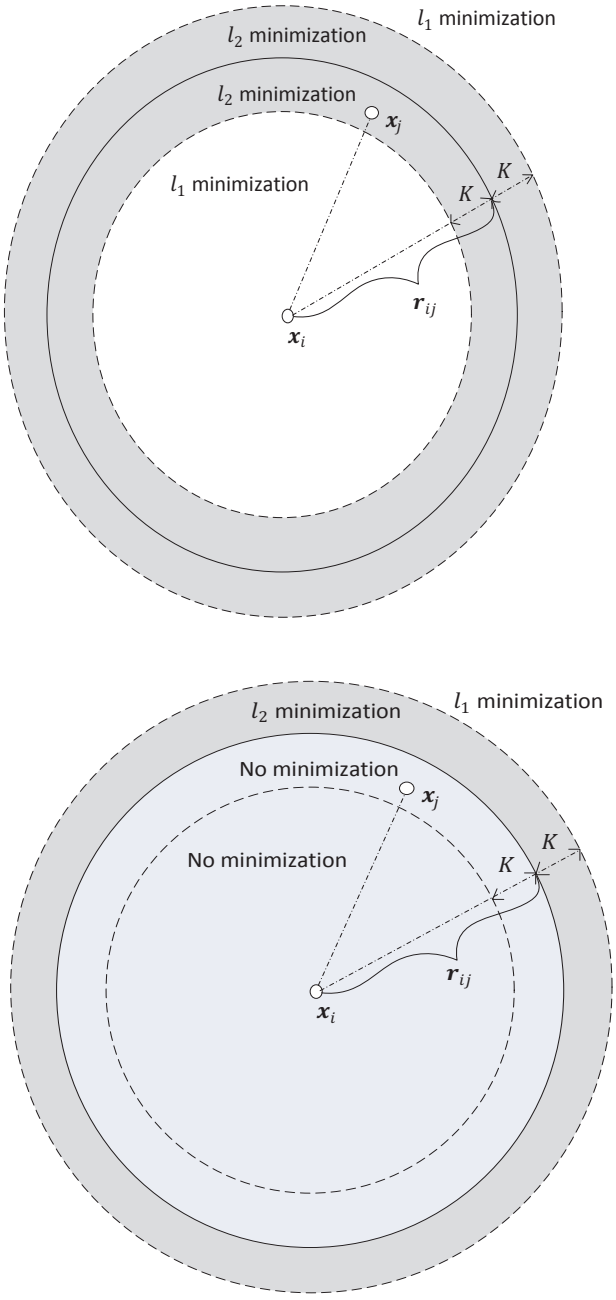

Fig. 4.1 Illustration of two nodes and their pairwise range measurement. The regions where l_1 and l_2 norm minimization are implemented. Top: Original Huber cost function. Bottom: Proposed convex cost function.

4.2.2 Problem Formulation

The aim is to find the locations of the N sensor nodes, represented by the matrix $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N] \in \mathbb{R}^{2\times N}$. To this end, we wish to minimize the average root mean square error (RMSE) in estimation of positions defined as

$$\Omega = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \|\hat{\boldsymbol{x}}_i - \boldsymbol{x}_i\|^2} \tag{4.2}$$

where $\hat{\boldsymbol{x}}_i$ is our estimate of $\boldsymbol{x}_i$. If there are no NLOS biases, due to the zero-mean Gaussian noises, the maximum likelihood estimation (MLE) is equivalent to the l_2 norm minimization, so the cost function to be minimized is

$$f(\boldsymbol{X}) = \sum_{(i,j)\in\mathcal{S}} \Big(\|\boldsymbol{x}_j - \boldsymbol{x}_i\| - r_{ij}\Big)^2 \tag{4.3}$$

which is a non-convex nonlinear least squares (NLS) problem with respect to $\boldsymbol{X}$ as discussed in [23, 26]. Since the NLOS biases exist in some measurements but can not be identified, using l_2 norm minimization might not yield robust estimates. In the presence of outliers, Huber cost function provides a suitable replacement for l_2 norm minimization, by interpolating between l_2 and l_1 norm minimizations. Therefore, instead of (4.3) one wishes to minimize

$$g(\boldsymbol{X}) = \sum_{(i,j)\in\mathcal{S}} \rho\Big(r_{ij} - \|\boldsymbol{x}_i - \boldsymbol{x}_j\|\Big) \tag{4.4}$$

where $\rho(\cdot)$ is the continuous and differentiable Huber function with parameter $u_{ij} = r_{ij} - \|\boldsymbol{x}_i - \boldsymbol{x}_j\|$, defined as

$$\rho(u_{ij}) = \begin{cases} u_{ij}^2, & |u_{ij}| < K \\ 2K|u_{ij}| - K^2, & |u_{ij}| \geq K \end{cases} \tag{4.5}$$

where K is a fixed parameter which is chosen to be proportional to σ_n, e.g., $K = \alpha\sigma_n$ and $1.5 \leq \alpha \leq 2$.

Although the Huber cost function is convex with respect to its argument, similar to the case of NLS, it is easy to show that it is not always convex with respect to the position coordinates $\boldsymbol{X}$. Therefore, if not initialized well, the estimates obtained by Huber estimation technique might not converge to stationary points close to the global minimum of the

Huber loss function. Furthermore, the Huber M-estimation can perform well if only a small portion of the measurements are outliers, otherwise it may not achieve a good estimation result. Therefore, in the following we propose a two-stage algorithm that is robust in any NLOS scenario.

4.3 Robust Distributed Algorithm

In this section, we first propose a convex relaxation of the Huber cost function. After converging to some stationary points, we then try to minimize the Huber cost function using iterative techniques.

4.3.1 Stage I: Convex Relaxation

A convex relaxation of the nonlinear least square problem in (4.3) has been proposed in [26]. This relaxation modifies the original cost function $f(\boldsymbol{X})$ as

$$\tilde{f}(\boldsymbol{X}) = \sum_{(i,j)\in\mathcal{S}} \Big((\|\boldsymbol{x}_j - \boldsymbol{x}_i\| - r_{ij})_+ \Big)^2 \tag{4.6}$$

where

$$(\|\boldsymbol{x}_j - \boldsymbol{x}_i\| - r_{ij})_+ = \begin{cases} 0, & \|\boldsymbol{x}_j - \boldsymbol{x}_i\| \leq r_{ij} \\ \|\boldsymbol{x}_j - \boldsymbol{x}_i\| - r_{ij}, & \|\boldsymbol{x}_j - \boldsymbol{x}_i\| > r_{ij} \end{cases}$$

Further explanation about the convexity of this cost function is given in [89]. The concept of this relaxation is similar to POCS proposed first in [90] and considered for cooperative sensor network localization in [2]. However, we instead propose to do a convex relaxation of Huber cost function as

$$\tilde{g}(\boldsymbol{X}) = \sum_{(i,j)\in\mathcal{S}} \tilde{\rho}\Big(r_{ij} - \|\boldsymbol{x}_i - \boldsymbol{x}_j\| \Big) \tag{4.7}$$

where $\tilde{\rho}(\cdot)$ is the convex relaxation of Huber function with respect to $\boldsymbol{X}$ defined as

$$\tilde{\rho}(u_{ij}) = \begin{cases} 2K_1 u_{ij} - K_1^2, & \|\boldsymbol{x}_i - \boldsymbol{x}_j\| \geq r_{ij} + K_1 \\ u_{ij}^2, & r_{ij} < \|\boldsymbol{x}_i - \boldsymbol{x}_j\| < r_{ij} + K_1 \\ 0, & \|\boldsymbol{x}_i - \boldsymbol{x}_j\| \leq r_{ij} \end{cases} \tag{4.8}$$

where $K_1 = \alpha_1 \sigma_n$ is the parameter of the Huber loss function. The illustration of the original and proposed Huber cost functions are expressed in Fig. 4.1 for the area between two nodes. The simulation result shows that, in many cases, this convex relaxation is more robust against large negative errors as well and gives a lower MSE for the network compared to the one in (4.6) or cooperative POCS [2].

The iterative gradient descent method for updating the position estimates can be stated at each node as

$$\boldsymbol{x}_i^{(l+1)} = \boldsymbol{x}_i^{(l)} - \mu_1 \sum_{(i,j)\in\mathcal{S}} \frac{\partial \tilde{\rho}(u_{ij})}{\partial \boldsymbol{x}_i^{(l)}}, \quad i = 1, \ldots, N \tag{4.9}$$

where μ_1 is a suitable step size and for every $j \in \mathcal{S}_i$

$$\frac{\partial \tilde{\rho}(u_{ij})}{\partial \boldsymbol{x}_i^{(l)}} = \begin{cases} 2K_1 \frac{\boldsymbol{x}_i^{(l)} - \boldsymbol{x}_j^{(l)}}{\|\boldsymbol{x}_i^{(l)} - \boldsymbol{x}_j^{(l)}\|}, & \|\boldsymbol{x}_i^{(l)} - \boldsymbol{x}_j^{(l)}\| \geq r_{ij} + K_1 \\ u_{ij}^{(l)} \frac{\boldsymbol{x}_i^{(l)} - \boldsymbol{x}_j^{(l)}}{\|\boldsymbol{x}_i^{(l)} - \boldsymbol{x}_j^{(l)}\|}, & r_{ij} < \|\boldsymbol{x}_i^{(l)} - \boldsymbol{x}_j^{(l)}\| < r_{ij} + K_1 \\ 0, & \|\boldsymbol{x}_i^{(l)} - \boldsymbol{x}_j^{(l)}\| \leq r_{ij} \end{cases}$$

After every sensor estimates its position, in the next time instant, it sends this estimate to its neighbours. Therefore, every sensor uses the current estimate about its own position, the known position of its neighbouring anchors, and the updated position of its neighbouring sensors to find a new estimate of its position. After a number of iterations, if the difference between the estimated positions of each sensor at two consecutive iterations is less than a predefined threshold, then the algorithm has reached near the global minimum of the relaxed Huber cost function in (4.7). The position estimates obtained at this stage may be in the proximity of the true sensor positions. However, as these position estimates are not the solution of the original problem in (4.4), we need to do further refinement of these estimates.

4.3.2 Stage II: Position Refinement

At this stage, we try to minimize the original Huber cost function in (4.4). The iterative gradient decent steps at each sensor node $\boldsymbol{x}_i$ is

$$\boldsymbol{x}_i^{(l+1)} = \boldsymbol{x}_i^{(l)} - \mu_2 \sum_{j \in \mathcal{S}_i} \frac{\partial \rho(u_{ij})}{\partial \boldsymbol{x}_i^{(l)}}, \quad i = 1, \ldots, N \tag{4.10}$$

where μ_2 is a suitable step size and for every $j \in \mathcal{S}_i$

$$\frac{\partial \rho(u_{ij})}{\partial \boldsymbol{x}_i^{(l)}} = \begin{cases} \frac{\boldsymbol{x}_i^{(l)} - \boldsymbol{x}_j^{(l)}}{\|\boldsymbol{x}_i^{(l)} - \boldsymbol{x}_j^{(l)}\|} u_{ij}^{(l)}, & |\|\boldsymbol{x}_i^{(l)} - \boldsymbol{x}_j^{(l)}\| - r_{ij}| < K_2 \\ 2K_2 \frac{\boldsymbol{x}_i^{(l)} - \boldsymbol{x}_j^{(l)}}{\|\boldsymbol{x}_i^{(l)} - \boldsymbol{x}_j^{(l)}\|}, & |\|\boldsymbol{x}_i^{(l)} - \boldsymbol{x}_j^{(l)}\| - r_{ij}| \geq K_2 \end{cases} \tag{4.11}$$

and $K_2 = \alpha_2 \sigma_n$ is the parameter of the Huber function, which is different from K_1. The algorithm continues iteratively for a limited number of iterations or until the difference between the estimates of each sensor's position at two consecutive iterations is less than a threshold.

Selecting a suitable K_2 is very important at this stage and can give a trade off between robustness and accuracy. If the ratio of NLOS to LOS links is high, then selecting K_2 as done usually for Huber M-estimation, i.e., $1.5\sigma_n \leq K_2 \leq 2\sigma_n$, might even result in deterioration of the position estimates. Thus, in this scenario, it is preferred to keep α_2 very small, so the second phase does not change the position estimates obtained in the first stage. On the other hand, if the ratio of the NLOS measurements is low compared to LOS ones, then the second stage can improve the positioning performance noticeably by selecting $1.5 \leq \alpha_2 \leq 2$. The performance is still improved if a small value of α_2 is chosen. Therefore, if in the network, we have an *a priori* estimate of the ratio of the NLOS to LOS measurements or the probability of a link being NLOS, then we can select K_2 according to the discussion above. However, if such information is not available, then we select α_2 to be small, e.g., $\alpha_2 = 0.1$, to achieve robust estimation result in all scenarios.

4.4 Test and Validation

4.4.1 Simulation Results

In this part, the performance of the proposed method is evaluated through simulations. We consider a network of $M = 4$ anchors and $N = 50$ sensors located on a 2D space. The sensors are randomly distributed on the plane while the anchors are at fixed locations $\boldsymbol{x}_{N+1} = [0, 0]^T$, $\boldsymbol{x}_{N+2} = [10, 0]^T$, $\boldsymbol{x}_{N+3} = [10, 10]^T$, and $\boldsymbol{x}_{N+4} = [0, 10]^T$, where the units are in meters. The range measurements were generated according to the model in (4.1) with $\sigma_n = 0.5$m and the NLOS bias is modelled as an exponential random variable with parameter $\gamma = 10$m. The Monte Carlo (MC) simulations are done under 500 runs.

We first consider the proposed convex relaxation and run this algorithm for 50 iterations with $\mu_1 = 0.04$ and $K = 2\sigma_n$. We compare the proposed technique with the relaxation of the NLS in [26] with similar parameters and the same number of iterations, denoted by Relaxation NLS. We also apply the mentioned iterative technique with the same parameters and iteration number on the original Huber cost function and denote it by Huber. Also we consider the cooperative POCS with parameter $\lambda^l = 0$, thus it becomes almost similar to the IPPM in [27], except that the projection is only implemented when $\|\boldsymbol{x}_i - \boldsymbol{x}_j\| \geq r_{ij}$. The initial sensor positions for all algorithms are selected to have a Gaussian distribution with mean equal to the true sensor positions and standard deviation of 10 meters. We define $P_{\mathcal{N}}$ as the probability of a link being NLOS. We now consider three scenarios where the probability that a link is in NLOS is chosen to be $P_{\mathcal{N}} = 0.9$, $P_{\mathcal{N}} = 0.5$, and $P_{\mathcal{N}} = 0.1$. In Fig. 4.2, the CDF of positioning error for different algorithms under various NLOS contamination level is shown. As observed in Fig. 4.2, the relaxation of Huber cost function is slightly better than the relaxation of NLS, and it has almost the same performance as POCS. The original Huber cost function does not achieve a good result due to the lack of convexity and poor initialization.

To do further position refinement, we also simulate the second phase of our algorithm for 50 iterations with $\mu_2 = 0.01$ and $K_2 = 0.1\sigma_n$. For the initialization, we use the position estimates obtained at the first stage by our proposed algorithm using convex relaxation of Huber cost function. To have a lower bound on the performance of our algorithm, we implement the IPPM proposed in [1] with the knowledge of perfect NLOS identification and denote it by IPPM NLOS. Since the IPPM algorithm may not necessarily converge to a good solution because of lack of convexity, we use the position estimates obtained by

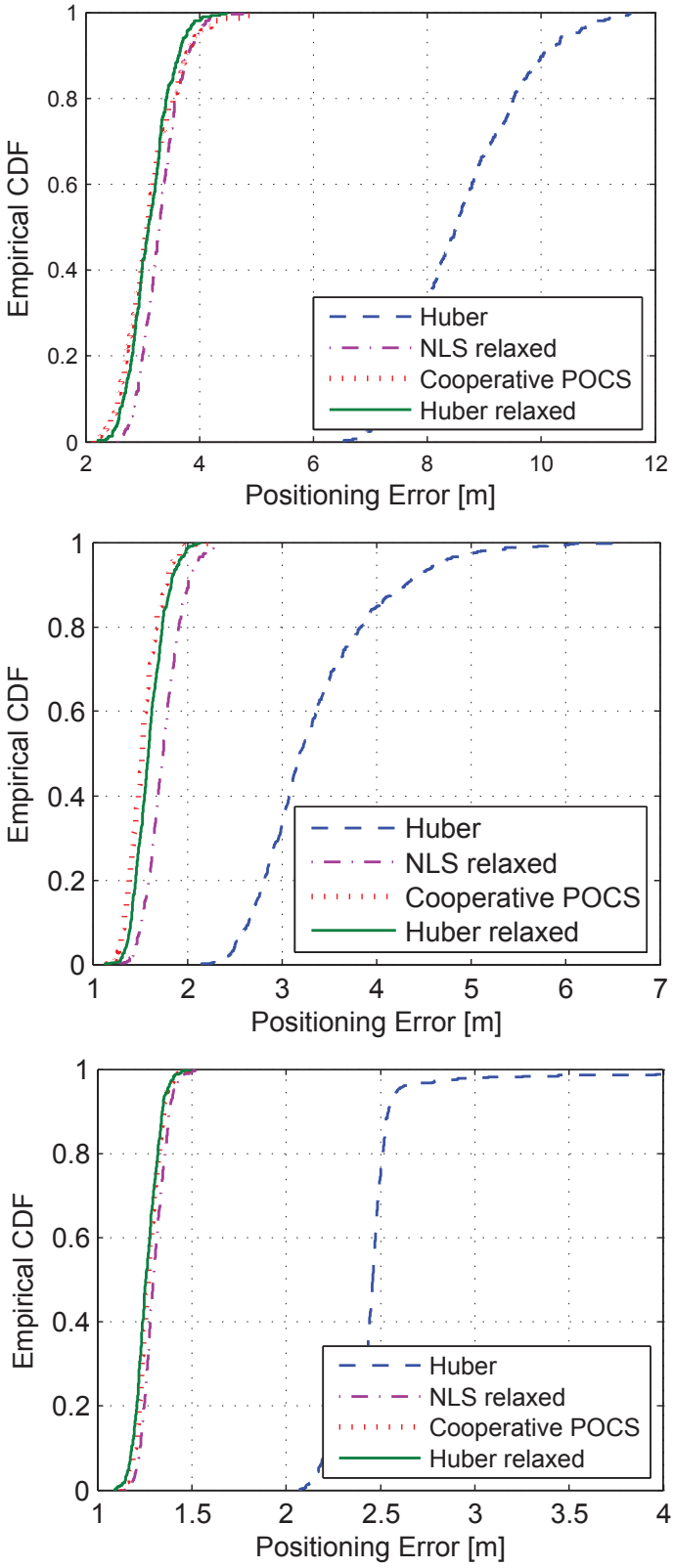

Fig. 4.2 CDF of different algorithms from top to bottom: $P_{\mathcal{N}} = 0.9$; $P_{\mathcal{N}} = 0.5$; $P_{\mathcal{N}} = 0.1$.

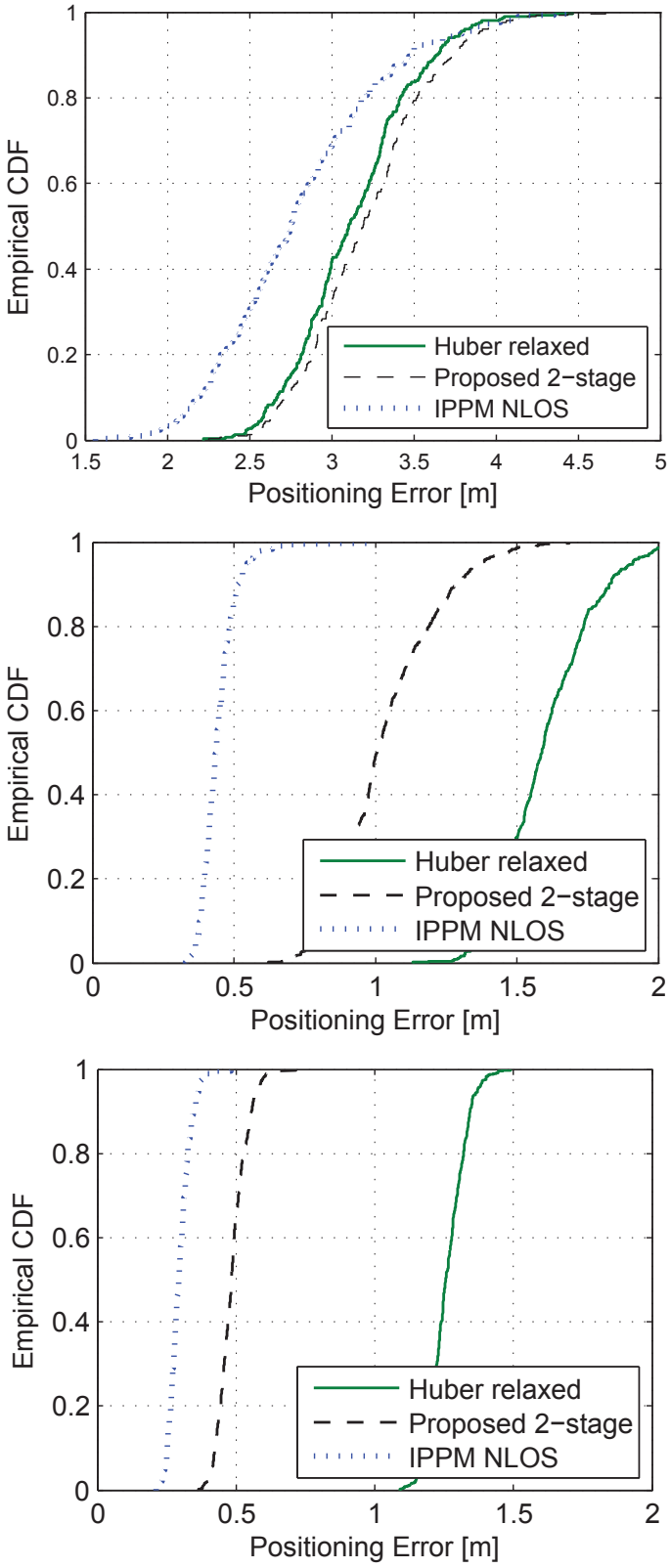

Fig. 4.3 CDF of the two stages of proposed algorithm and the IPPM in [1] from top to bottom: $P_{\mathcal{N}} = 0.9$; $P_{\mathcal{N}} = 0.5$; $P_{\mathcal{N}} = 0.1$.

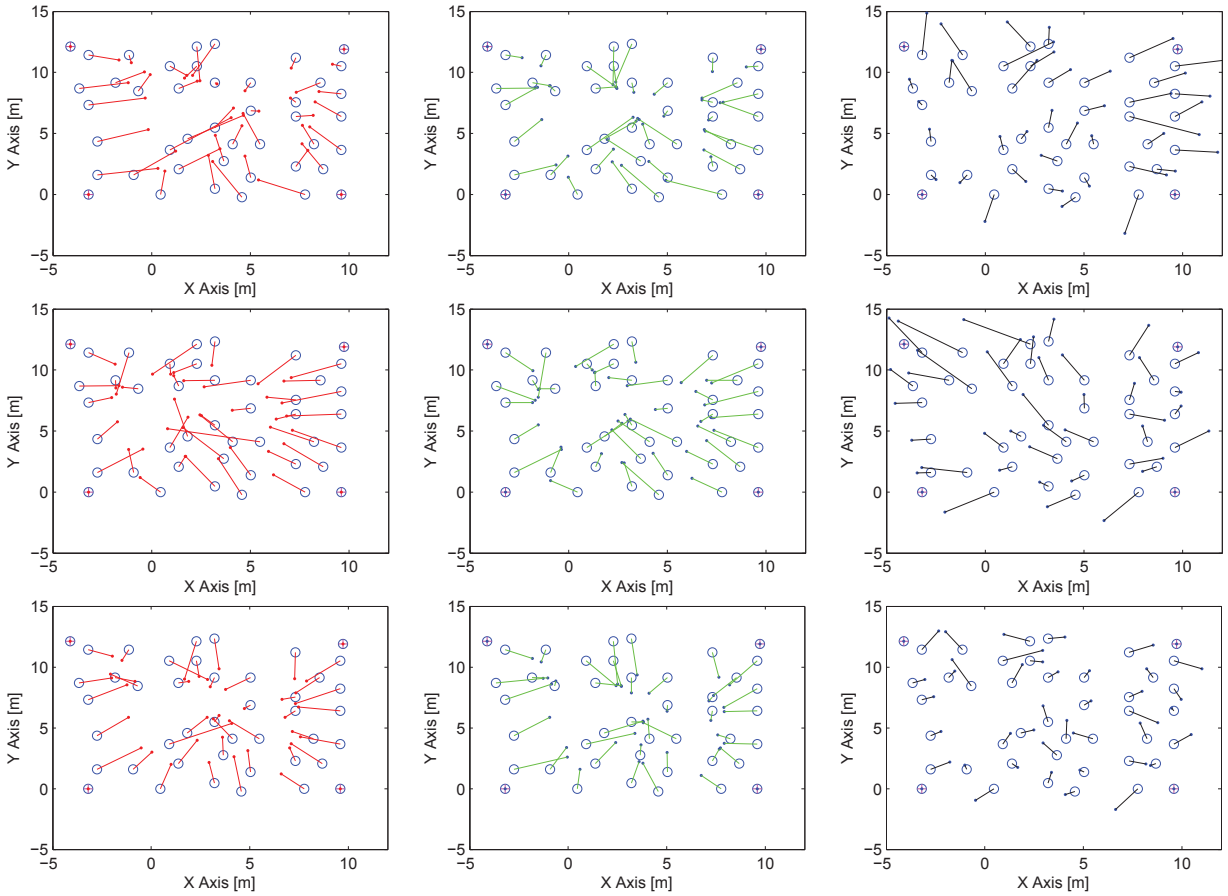

Fig. 4.4 Top: localization performance for large $P_{\mathcal{N}}$. left, cooperative POCS [2]; middle, first stage; right, second stage. Middle: localization performance for moderate $P_{\mathcal{N}}$: left, cooperative POCS [2]; middle, first stage; right, second stage. Bottom: localization performance for small $P_{\mathcal{N}}$: right, cooperative POCS [2]; middle, first stage; right, second stage.

cooperative POCS as initial points. The CDF of the error for our two-stage algorithm is illustrated in Fig. 4.3 along with the IPPM with prior NLOS identification. The CDF of the error of the proposed convex relaxation shown in Fig. 4.2 is also plotted in Fig. 4.3. The results show that when the ratio of the NLOS is high compared to LOS ones, the second stage of the algorithm may not improve the position estimates. However, when the ratio of the NLOS starts to decrease, the second stage can improve the estimates obtained in the first stage distinguishably. The performance of the proposed 2-stage algorithm is close to IPPM NLOS, which is based on perfect NLOS identification.

The performance of different methods are compared as a function of iteration number in Fig. 4.5 by averaging over all the sensors. For the first 50 iterations the first stage of our algorithm (Proposed 1-st stage) is shown along with the cooperative POCS. For the next 50 iterations, the second stage of our algorithm (Proposed 2-nd stage) is illustrated together with IPPM NLOS, where the latter is initialized using the estimates obtained by Cooperative POCS. As mentioned earlier, for large ratios of NLOS to LOS links, the relaxed problems yield decent solutions as they are suited for scenarios with positive measurement errors. The second stage of our algorithm may not necessarily improve the localization performance for large $\mathcal{P}_N$ and the improvement might be minor. In moderate ratios of NLOS to LOS links, however, the second stage can definitely improve the localization performance as it is based on Huber estimation technique. The 2-nd stage becomes even more useful when $\mathcal{P}_N$ is small and only a portion of links are outliers. Although here we used fixed number of iterations, it might be better to check the convergence rate of all the nodes and stop when the estimates can not be changed further. In this way, the efficiency of the proposed algorithm can be improved.

4.4.2 Experimental Results

In this part, we consider localization of sensors using real data obtained by the measurement campaign reported in [69]. The environment was an indoor office and there were 44 node locations where the transmitter and receiver were used at each location and pairwise range measurements were obtained. We consider four nodes in the corner as anchor nodes with perfect location information and the other 40 nodes as the sensors with unknown locations. Due to the scatterers and NLOS in the office, almost all of the measurements are affected by large positive errors as mentioned in [69]. It is mentioned that the average amount

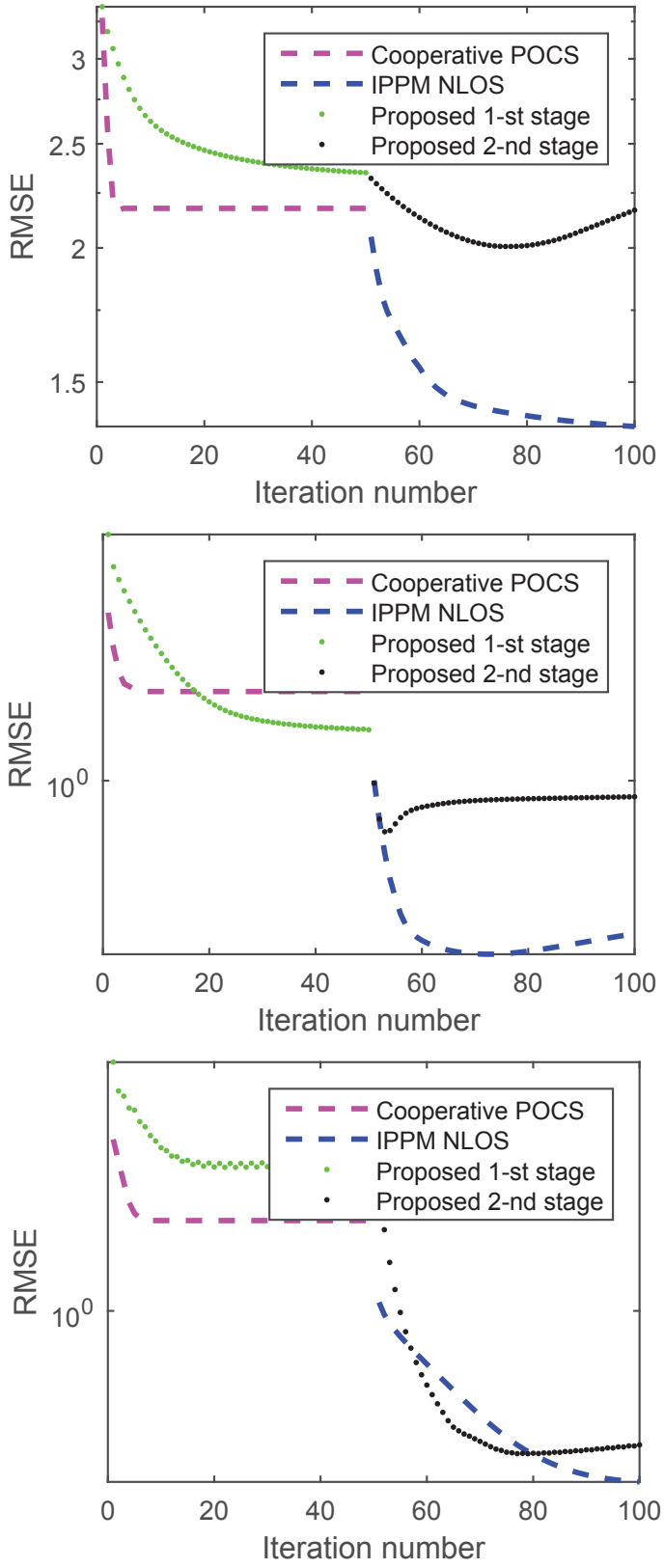

Fig. 4.5 RMSE versus the number of iterations. From top to bottom: $P_{\mathcal{N}} = 0.9$; $P_{\mathcal{N}} = 0.5$; $P_{\mathcal{N}} = 0.1$.

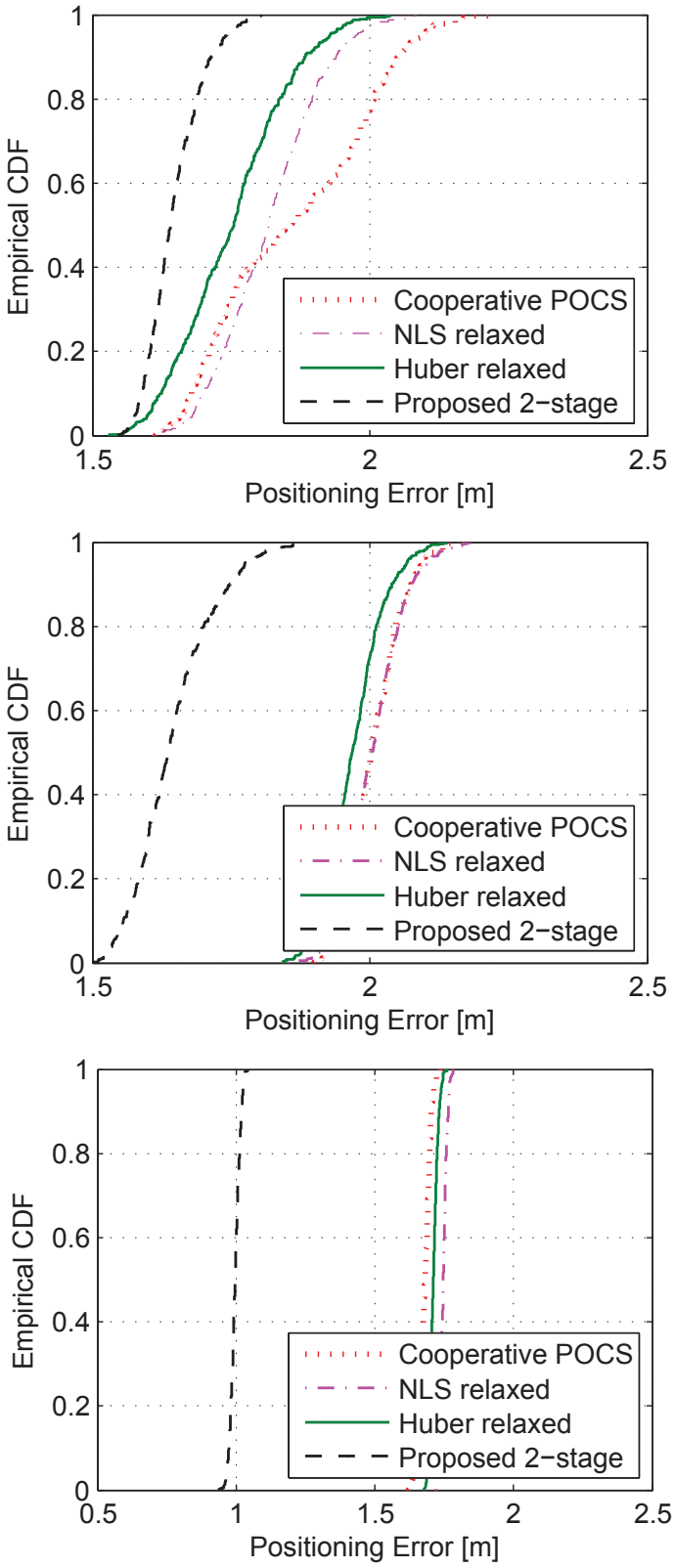

Fig. 4.6 RMSE of different techniques versus the number of iterations: (a) Large $P_{\mathcal{N}}$; (b) Moderate $P_{\mathcal{N}}$; (c) Small $P_{\mathcal{N}}$.

of error is also calculated for these measurements, therefore, by subtracting that quantity from the measurements, a less unbiased set of measurement is obtained. To evaluate the performance of our algorithm in different conditions we consider these scenarios:

- The raw measurements are considered, hence many of the measurements have positive errors, i.e., $P_{\mathcal{N}}$ is large.
- The positive bias is subtracted from half of the measurements randomly, hence $P_{\mathcal{N}}$ is moderate.
- The average bias is subtracted from all the raw measurements, thus $P_{\mathcal{N}}$ is small.

Using the unbiased measurements, the standard deviation of measurement noise is estimated roughly to be $\sigma_n = 1$m. By applying the iterative gradient descent technique on the proposed convex Huber cost function with $\mu_1 = 0.04$ and $K_1 = 2\sigma_n$, an estimate of the positions of sensors are obtained iteratively for 50 iterations. The position estimates are also refined in the second stage with $K_2 = 0.1\sigma_n$ and $\mu_2 = 0.01$ for 50 iterations. The final estimates at the end of each stage of our algorithm and the estimates obtained by cooperative POCS are shown along with the true sensor positions in Fig. 4.4. The CDF of the positioning error in different NLOS scenarios are also illustrated in Fig. 4.6 by running 500 MC trials.

The results show that in general the relaxed Huber function achieves a better result compared to the other approaches. Moreover, the second stage of the algorithm noticeably improves the position estimates obtained in the first stage, especially when $P_{\mathcal{N}}$ is small.

4.5 Conclusion

A robust distributed cooperative localization technique has been proposed in this chapter. We first applied a convex relaxation on the Huber cost function and preliminary position estimates were obtained iteratively. In the second stage of our algorithm, by iteratively minimizing the Huber loss function, it was shown that further position refinement could be generally obtained. For iterative optimization in each stage, a gradient descent method was used. By testing on real data set, the superiority of our algorithm was verified. We conclude that our two-stage algorithm performs robustly against outliers; in particular it significantly outperforms other distributed techniques when the ratio of NLOS to LOS

measurements is small. One issue with the distributed algorithms compared to centralized ones for localization is that a lot of communication needs to be done among the nodes until the positions can be estimated. This might take time and the communication load might be very high for the WSN. This is one of the disadvantages of distributed techniques compared to centralized ones. Therefore, faster iterative techniques based on Newton or Gauss-Newton methods are preferred to gradient descent technique. Implementing a distributed Gauss-Newton method has recently gained attention and seems to be an interesting idea worth consideration. Therefore, further improvements of the distributed localization algorithms might be possible in the near future.

Chapter 5

Distributed Outer-Approximation of Feasible Sets in WSNs under NLOS

In this chapter, a distributed technique based on SPAWN is applied to WSNs with positively biased range measurements to outer-approximate the convex hulls by ellipses. Our main contribution is to develop a novel method for outer-approximation of the intersection of ellipses using an ellipse in 2-dimensional space. This chapter is organized as follows: An introduction to the topic is given in Section 5.1. The system model and background are presented in Section 5.2. The distributed algorithm is presented in Section 5.3.1 and the proposed ellipse outer-approximation method is given in Section 5.3.2. The simulation results are given in Section 5.4. Finally, Section 5.5 concludes this chapter.

5.1 Introduction

Localization of sensor nodes in a wireless sensor network (WSN) is of great interest in many public safety and commercial applications [3]. In particular, cooperative localization has received special attention since it can improve localization accuracy and coverage [21]. In contrast to non-cooperative WSN, in which only measurements between the sensors being localized and anchors with known positions are performed, cooperative WSNs also use sensor-to-sensor measurements.

Localization in indoor places and dense urban areas is more challenging due to multipath propagation and non-line of sight (NLOS) conditions, which result in positively biased range measurements [44, 118, 119]. In these situations, the unknown location of each sensor is

restricted to the intersection of multiple balls (or discs in 2-dimensional (2-D) space), with centres corresponding to the locations of neighbouring nodes, e.g., anchors and sensors, and with radii equal to biased range measurements. The intersection of these balls is a convex feasible set, which can serve as a rough approximation of the uncertainty in the sensor's position. In many applications it is therefore desired to quantify the size of each feasible set. However, since the feasible set is complex and can not be generally described by a few parameters, outer-approximating it by a simple shape, e.g., a ball or an ellipsoid, is needed.

In cooperative WSNs, finding outer-approximations of these feasible sets is not straightforward as the centres of the balls corresponding to the locations of the neighbouring sensors are unknown. To address this issue, a distributed iterative algorithm was proposed in [2], where a ball is used for an outer-approximation of a feasible set. The algorithm has been improved in [65] by using ellipsoids instead of balls, on the basis that an ellipsoid can generally capture a complex convex set more tightly due to its additional degrees of freedom. In this algorithm, during the first iteration, each sensor finds a tight ellipsoidal outer-approximation of the intersection of the balls corresponding to its neighbouring anchors. Then at each sensor, the ellipsoids of neighbouring sensors are expanded along the semi-axes with the pairwise measured ranges between the former and the latter. Each sensor then needs to find an ellipsoid to outer-approximate the intersection of multiple balls and expanded ellipsoids, with centres corresponding to its neighbouring anchors and sensors, respectively. The tighter the obtained ellipsoid, the more accurate will be the outer-approximation of the feasible set containing each sensor.

Finding the tightest ellipsoidal outer-approximation of the intersection of multiple ellipsoids (and/or balls) is NP-hard [66], and to the best of our knowledge, there is no algorithm to find the optimal solution. However, there are several sub-optimal solutions, including methods considering two ellipsoids [120–123], as well as standard convex optimization methods for a larger number of ellipsoids [99, p.414], [66, p.44]. In [65], the method from [99] has been employed. As the methods from [66,99] are sub-optimal, and localization problem can usually be considered in a 2-D space (unknown latitude and longitude), there is a special interest in developing geometrical methods in a 2-D space that can find tighter ellipses.

In this letter, we propose a novel method for tight outer-approximation of the intersection of a number of ellipses in 2-D space. In this method, we first efficiently determine a tight polygon containing the intersection of ellipses, and then solve a convex optimization problem to obtain the tightest ellipse covering the vertices of the polygon. We employ the

proposed method in the distributed algorithms considered in [65] for outer-approximations of 2-D feasible sets in cooperative WSNs and show that it offers significant improvements in tightness with similar computational cost, compared to the case that the method from [99] is employed.

5.2 System Model and Background

5.2.1 System Model

We consider a 2-D WSN with N sensor nodes at unknown locations denoted by $\boldsymbol{x}_i \in \mathbb{R}^2$ for $i = \{1, \ldots, N\}$, and M anchors with known locations $\boldsymbol{a}_i \in \mathbb{R}^2$, for $i = \{N+1, \ldots, N+M\}$. Two nodes are regarded as neighbours if they are within communication range, i.e., they are within the given distance $R_{\max}$ of each other. For each sensor node j we define two sets $\mathcal{A}_j$ and $\mathcal{S}_j$ which include the indices of all the neighbouring anchors and sensors, respectively. The range measurements of j-th sensor are modelled as

$$r_{ij} = \|\boldsymbol{a}_i - \boldsymbol{x}_j\| + b_{ij} + n_{ij}, \quad i \in \mathcal{A}_j \tag{5.1}$$

$$r_{ij} = \|\boldsymbol{x}_i - \boldsymbol{x}_j\| + b_{ij} + n_{ij}, \quad i \in \mathcal{S}_j \tag{5.2}$$

where we assume that n_{ij} are independent identically distributed measurement errors having zero-mean, and $b_{ij} > 0$ represent the biases due to the NLOS, while for LOS measurements $b_{ij} = 0$. The noise is often assumed to have a Gaussian distribution with zero-mean and variance σ_n^2, while the bias term b_{ij} has been modelled as an exponential [33], or a uniformly distributed random variable [100]. We assume that (5.1) and (5.2) correspond to the NLOS measurements only, which can be identified from LOS ones using NLOS identification techniques, as done in [31, 33, 100].[1] Furthermore, to make our algorithm more robust, no knowledge is assumed about the distribution of n_{ij} and b_{ij}. In many applications, the bias dominates over the measurement noise [44], i.e., $b_{ij} + n_{ij} \geq 0$. Hence it follows that

[1] The LOS measurements can later be used in conventional localization algorithms by taking benefit of the bounds obtained in this work.

each sensor $\boldsymbol{x}_j$ is restricted to be inside the intersection area of multiple discs,[2] defined as

$$\mathcal{D}_{ij}^{\mathcal{A}} = \{\boldsymbol{x} \in \mathbb{R}^2 : \|\boldsymbol{x} - \boldsymbol{a}_i\| \leq r_{ij}\}, \quad i \in \mathcal{A}_j \tag{5.3}$$

$$\mathcal{D}_{ij}^{\mathcal{S}} = \{\boldsymbol{x} \in \mathbb{R}^2 : \|\boldsymbol{x} - \boldsymbol{x}_i\| \leq r_{ij}\}, \quad i \in \mathcal{S}_j \tag{5.4}$$

Therefore, $\boldsymbol{x}_j \in \mathcal{D}_j$ where

$$\mathcal{D}_j = \Big(\bigcap_{i \in \mathcal{A}_j} \mathcal{D}_{ij}^{\mathcal{A}} \Big) \bigcap \Big(\bigcap_{i \in \mathcal{S}_j} \mathcal{D}_{ij}^{\mathcal{S}} \Big). \tag{5.5}$$

Our objective is to determine an outer-approximation of the convex feasible set $\mathcal{D}_j$ for every sensor $\boldsymbol{x}_j$ through a distributed approach.[3] Note that each $\mathcal{D}_{ij}^{\mathcal{A}}$ is available to sensor j, while each $\mathcal{D}_{ij}^{\mathcal{S}}$ is not *a-priori* available since $\boldsymbol{x}_i$ is unknown. Therefore, the solution is not straightforward.

5.2.2 Notes on the Definition of Ellipsoids

An ellipsoid ξ_i in ν-dimensional space $\mathbb{R}^\nu$ can be defined in many different ways [99], including:

(i) The image of the unit ball under an affine transformation:

$$\xi_i = \Big\{ \boldsymbol{x} = \boldsymbol{P}_i \boldsymbol{y} + \boldsymbol{x}_{c,i} : \|\boldsymbol{y}\| \leq 1, \boldsymbol{y} \in \mathbb{R}^\nu \Big\}, \tag{5.6}$$

where $\boldsymbol{x}_{c,i}$ is the centre of the ellipsoid, and without loss of generality[4] $\boldsymbol{P}_i \in \mathbb{S}_{++}^\nu$, where $\mathbb{S}_{++}^\nu$ denotes the set of all symmetric positive definite matrices of size $\nu \times \nu$.

(ii) A quadratic form:

$$\xi_i = \Big\{ \boldsymbol{x} \in \mathbb{R}^\nu : \ \|\boldsymbol{B}_i \boldsymbol{x} + \boldsymbol{d}_i\| \leq 1 \Big\}, \tag{5.7}$$

where $\boldsymbol{B}_i \in \mathbb{S}_{++}^\nu$ and $\boldsymbol{d}_i \in \mathbb{R}^\nu$ is a translation vector. When $\boldsymbol{B}_i = \boldsymbol{P}_i^{-1}$ and $\boldsymbol{d}_i = \boldsymbol{P}_i^{-1} \boldsymbol{x}_{c,i}$, the two ellipsoids in (5.6) and (5.7) are identical.

[2] If $b_{ij} + n_{ij} \geq 0$ can not be guaranteed (e.g., due to large σ_n), a constant can be added to each r_{ij} in the right hand side of (5.3) and (5.4) to ensure that the position of each sensor is restricted to the discs with neighbouring nodes as centres.

[3] We assume that for every sensor j, there is at least one neighbouring node with pairwise measurement r_{ij} such that $\mathcal{D}_j$ is not empty.

[4] Note that we can always represent an ellipsoid by a symmetric positive definite matrix $\boldsymbol{P}_i$ [124, p.208].

5.3 Distributed Outer-approximation of Feasible Sets for Positioning

5.3.1 Distributed Outer-approximation for Positioning

The sum-product algorithm over a wireless network (SPAWN) is a distributed positioning technique, whereby each sensor iteratively determines and refines the statistical knowledge regarding its position [21]. This knowledge is described through a so-called *belief*, which is proportional to a probability density function (PDF). At each iteration, each sensor receives beliefs from neighbouring nodes. A received belief is combined with the ranging likelihood function to determine a so-called *message*. A message is a function describing the distribution of the receiving sensor's position, based solely on the belief of the transmitting node and the pairwise measured range between the two nodes. Multiplying the messages from the neighbouring nodes with the prior belief of sensor's position leads to the updated belief. In this section we will apply SPAWN for the model described in Section II and show how it is related to the distributed outer-approximating algorithm proposed in [65]. The initial belief of each node is either a Dirac delta function (for anchors) or a uniform PDF over a large selected area in the 2-D plane (for sensors).

Message Computation

Since in this work we assume $b_{ij} + n_{ij} \geq 0$, we can approximate the ranging likelihood by a uniform PDF on a disc with centre at the transmitting node and radius equal to the measured range r_{ij}. When the transmitting node is an anchor with known position $\boldsymbol{a}_i$, its belief about its position is a Dirac delta function. Based on the definition of the message in [21], the message from the i-th anchor to the j-th sensor will be a uniform PDF on a disc given by

$$\big\{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{a}_i\| \leq r_{ij}, \quad i \in \mathcal{A}_j\big\}. \tag{5.8}$$

In contrast, when the transmitting node is another sensor, its position $\boldsymbol{x}_i$ is uncertain and this affects the form of the message. Assume that the belief of i-th sensor is approximated to be a uniform PDF on an ellipse as

$$\big\{\boldsymbol{x} : \|\boldsymbol{B}_i\boldsymbol{x} + \boldsymbol{d}_i\| \leq 1\big\} \tag{5.9}$$

where $\boldsymbol{B}_i \in \mathbb{S}^2_{++}$. Therefore, it follows that the message of the i-th sensor to the j-th sensor[5] is uniform on an ellipse obtained by expansion of each semi-axis of the ellipse in (5.9) by r_{ij}. The semi-axes of the i-th ellipse in (5.9) are the eigenvalues of $\boldsymbol{P}_i = \boldsymbol{B}_i^{-1}$. Let $\boldsymbol{P}_i = \boldsymbol{V}_i \boldsymbol{\Gamma}_i \boldsymbol{V}_i^T$ be the eigendecomposition of $\boldsymbol{P}_i$ where $\boldsymbol{\Gamma}_i = \mathrm{diag}(\lambda_{1,i}, \lambda_{2,i})$. In order to expand the ellipse by r_{ij}, we replace $\boldsymbol{P}_i$ by $\tilde{\boldsymbol{P}}_{ij} = \boldsymbol{V}_i \tilde{\boldsymbol{\Gamma}}_{ij} \boldsymbol{V}_i^T$ where $\tilde{\boldsymbol{\Gamma}}_{ij} = \mathrm{diag}(\lambda_{1,i}, \lambda_{2,i}) + r_{ij}\boldsymbol{I}_2$. Then the expanded ellipse is

$$\big\{\boldsymbol{x} : \|\tilde{\boldsymbol{B}}_{ij}\boldsymbol{x} + \tilde{\boldsymbol{d}}_i\| \leq 1, \quad i \in \mathcal{S}_j\big\}, \tag{5.10}$$

where $\tilde{\boldsymbol{B}}_{ij} = \tilde{\boldsymbol{P}}_{ij}^{-1}$ and $\tilde{\boldsymbol{d}}_i = \tilde{\boldsymbol{P}}_{ij}^{-1}\boldsymbol{P}_i\boldsymbol{d}_i$.

Belief Update

To update the belief of sensor j, the messages corresponding to its neighbouring nodes and its prior belief (which are uniform functions herein) are multiplied. Therefore, the updated belief will also be uniform on the intersection region of (i) multiple discs as in (5.8), (ii) multiple ellipses as in (5.10) and (iii) a large 2-D area (due to the prior belief). The entire procedure continues until convergence or a predefined number of iterations. This intersection region, denoted by $\mathcal{E}_j$ is a complex convex set[6], and since describing it explicitly using a simple geometric shape is in general impossible, sensor j will need to outer-approximate it, e.g., by an ellipse.

In [65], first the largest volume ellipsoid contained in the intersection of M_e ellipsoids is determined by solving a convex optimization problem as mentioned in [99, p.414] as

$$\begin{aligned}
&\max_{\boldsymbol{P}_0, \boldsymbol{x}_{c_0}, \tau_1, \ldots, \tau_{M_e}} \quad \log\det \boldsymbol{P}_0 \\
&\text{subject to } \boldsymbol{P}_0 \succ 0, \\
&\tau_i \geq 0, \quad \begin{bmatrix} -\tau_i - c_i + \boldsymbol{b}_i^T \boldsymbol{A}_i^{-1}\boldsymbol{b}_i & \boldsymbol{0} & (\boldsymbol{x}_{c_0} + \boldsymbol{A}_i^{-1}\boldsymbol{b}_i)^T \\ \boldsymbol{0} & \tau_i \boldsymbol{I}_D & \boldsymbol{P}_0 \\ \boldsymbol{x}_{c_0} + \boldsymbol{A}_i^{-1}\boldsymbol{b}_i & \boldsymbol{P}_0 & \boldsymbol{A}_i^{-1} \end{bmatrix} \succeq 0, \ i = 1, \ldots, M_e
\end{aligned} \tag{5.11}$$

which is a convex SDP optimization problem and can be solved efficiently. Then by expanding the obtained ellipsoid with the dimension of the space ν, an ellipsoid that covers

[5] In SPAWN, unlike traditional belief propagation algorithm, sensor j computes the messages of all its neighbour nodes [21].

[6] Note that $\mathcal{E}_j$ should not be mistaken with $\mathcal{D}_j$, as the former is available to sensor j and contains the latter.

the intersection region can be found.

When $\nu = 2$, i.e., in a 2-D space we will show that in general a tighter approximation can be obtained using a geometrical approach.

5.3.2 Tight Outer-approximation of the Intersection of Ellipses

For each sensor j, we show how it is possible to efficiently find a polygon, represented by $\tilde{m}$ vertices $\boldsymbol{w}^{(l)}$ for $l = 1, \ldots, \tilde{m}$, which covers $\mathcal{E}_j$. The smallest area ellipse that contains these vertices (and hence contains $\mathcal{E}_j$) is found by solving the following convex optimization problem [99, 125]:

$$\begin{aligned} \min_{\boldsymbol{B}_j, \boldsymbol{d}_j} \quad & \log\det(\boldsymbol{B}_j^{-1}) \\ \text{subject to} \quad & \|\boldsymbol{B}_j \boldsymbol{w}^{(l)} + \boldsymbol{d}_j\| \leq 1, \quad l = 1, \ldots, \tilde{m}, \end{aligned} \tag{5.12}$$

where $\det(\boldsymbol{B}_j^{-1})$ is proportional to the area of the ellipse. Since each inequality in (5.12) can be written as a linear matrix inequality, this optimization problem can be formulated as a standard semi-definite programming (SDP) problem. For the ellipse to tightly bound $\mathcal{E}_j$, the polygon which bounds $\mathcal{E}_j$ has to be tight. Hence, the problem reverts to the determination of a polygon that covers $\mathcal{E}_j$ tightly. We propose below a method with three steps to achieve this:

Step 1 (generating discrete points)

We first generate a number of discrete points on the boundary of $\mathcal{E}_j$. One way to do so is to generate a fixed number of points on the boundary of each intersecting ellipse and then reject those that do not lie on $\mathcal{E}_j$. Harnessing the fact that an ellipse is an image of the unit disc under an affine transformation, we first generate m points $\boldsymbol{y}^{(l)}$ for $l = \{1, \ldots, m\}$, uniformly on a unit circle and then map these points onto the desired ellipse ξ_i as defined in (5.6), through the transformation $\boldsymbol{z}_i^{(l)} = \boldsymbol{P}_i \boldsymbol{y}^{(l)} + \boldsymbol{x}_{c,i}$. After rejecting the points not on $\mathcal{E}_j$, we denote the remaining points by $\tilde{\boldsymbol{z}}^{(l)}$ for $l = \{1, \ldots, \tilde{m}\}$ and the associated ellipse index for each point by $i^{(l)}$.

Step 2 (generating half planes)

Utilizing the form (5.7), the tangent lines to the i-th ellipse at the points $\tilde{\boldsymbol{z}}^{(l)}$ can be obtained, and hence the half planes are formed

$$(\boldsymbol{B}_{i^{(l)}}\tilde{\boldsymbol{z}}^{(l)} + \boldsymbol{d}_{i^{(l)}})^T(\boldsymbol{B}_{i^{(l)}}\boldsymbol{x} + \boldsymbol{d}_{i^{(l)}}) \leq 1, \quad l = 1\ldots, \tilde{m}. \tag{5.13}$$

Step 3 (determining the vertices of a polygon)

For a sufficient number of points $\tilde{\boldsymbol{z}}^{(l)}$, the intersection of these half planes forms a closed polygon covering $\mathcal{E}_j$. One way to find this polygon is to obtain the intersection point of every pair of tangent lines (by solving $\tilde{m}(\tilde{m}-1)/2$ linear systems of equations) and to verify if this point is inside the intersection region of all the remaining half-planes, i.e., if it satisfies all the remaining $\tilde{m}-2$ affine inequalities. The complexity of such a procedure scales as $\mathcal{O}(\tilde{m}^3)$. Hence, this procedure is intractable when m is large. Herein, we make use of the fact that $\nu = 2$ to develop a more efficient approach.

- *Step 3a:* Given the points $\tilde{\boldsymbol{z}}^{(l)}$ for $l = \{1, \ldots, \tilde{m}\}$, we first compute the average $\boldsymbol{z}_{\text{mean}} = \frac{1}{\tilde{m}}\sum_{l=1}^{\tilde{m}} \tilde{\boldsymbol{z}}^{(l)} \in \mathcal{E}_j$. The vectors $\boldsymbol{v}^{(l)} = \tilde{\boldsymbol{z}}^{(l)} - \boldsymbol{z}_{\text{mean}}$ connecting $\boldsymbol{z}_{\text{mean}}$ to the points $\tilde{\boldsymbol{z}}^{(l)}$ are sorted according to the angles $\alpha^{(l)} \in [0, 2\pi)$ with respect to the horizontal axis. This sorting imposes an order to the points $\tilde{\boldsymbol{z}}^{(l)}$.

- *Step 3b:* Given two sequential points $\tilde{\boldsymbol{z}}^{(l)}$ in the ordering, we determine the intersection point of the corresponding tangent lines. The obtained intersection points, $\boldsymbol{w}^{(l)}$ for $l = \{1, \ldots, \tilde{m}\}$ form the vertices of the polygon and are used as an input to (5.12).

In terms of complexity, in *Step 3*, the proposed technique requires solving $\tilde{m}$ linear systems of equations to find the polygon, hence the computational cost is reduced from $\mathcal{O}(\tilde{m}^3)$ to $\mathcal{O}(\tilde{m})$.

A closed polygon covering $\mathcal{E}_j$ obtained through the proposed method is illustrated in Fig. 5.1. However, we note that some degenerate cases can occur in *Step 3b*: (i) the intersection of the tangent lines of neighbouring points does not exist when the lines are parallel; (ii) it exists but does not satisfy other inequalities in (5.13). These two cases are illustrated in Fig. 5.2. In these cases, a closed convex polygon that covers the intersection

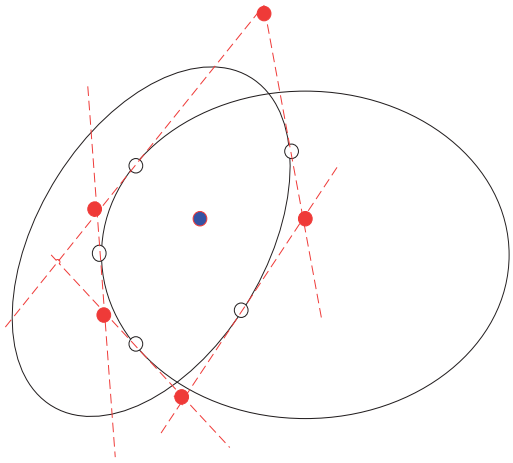

Fig. 5.1 The diagram of the intersecting ellipses and the half-planes forming a closed convex polygon. The white, blue, and red points correspond to $\tilde{\boldsymbol{z}}^{(l)}$, $\boldsymbol{z}_{\text{mean}}$, and $\boldsymbol{w}^{(l)}$, respectively.

of ellipses cannot be obtained. This problem can be easily avoided by selecting the number of discrete points m to be large enough.

The complete distributed bounding algorithm, which continues until convergence or when a predefined number of iterations K is reached, is summarized in Algorithm 1.

5.4 Numerical Performance Evaluation

We consider 4 anchors located on the four corners of a 10 m $\times$ 10 m 2-D area. We consider three scenarios where there are 10, 20, and 50 sensors, distributed uniformly on this 2-D area. The communication range is set to $R_{\max} = 12$ m and the measurement between each pair of neighbouring nodes is obtained by adding to the true range an exponentially distributed positive error with mean equal to 1 m. For solving the optimization problems we use Sedumi [109] and the CVX toolbox [126] in Matlab. In the proposed technique, we set $m = 128$, and the performance is evaluated in terms of the average area of the ellipses in each iteration, quantified by $\det(\boldsymbol{B}_j^{-1})$. As a benchmark, we use the method from [65].

In Fig. 5.3, we show the average area of the covering ellipses versus the iteration number for different number of sensors N. The results show that the distributed algorithm converges rapidly for both outer-approximation methods, although our proposed method converges to outer-approximating ellipses with almost half the area.

In Table I, we compare the computation time of each algorithm for the three scenarios

Algorithm 4 Distributed Outer-approximating Algorithm

for $k = 1$ until convergence (or predefined K) **do**
  for $j = 1, \ldots, N$ in parallel **do**
    if $k = 1$ **then**
      for all $i \in \mathcal{A}_j$ **do**
        Generate m points $\boldsymbol{z}_i^{(l)}$ on the discs in (5.8).
      end for
      Reject $\boldsymbol{z}_i^{(l)}$ outside $\mathcal{E}_j$, i.e., the intersection of (5.8).
    else
      for each $i \in \mathcal{S}_j$ **do**
        Exchange the updated $\boldsymbol{B}_i$ and $\boldsymbol{d}_i$.
        Expand the i-th ellipse in (5.9) to obtain (5.10).
        Generate m points $\boldsymbol{z}_i^{(l)}$ on the ellipses in (5.10).
      end for
      Reject the points outside $\mathcal{E}_j$, i.e., the intersection of (5.8) and (5.10).
    end if
    Find the half planes tangent to $\mathcal{E}_j$ at $\tilde{\boldsymbol{z}}^{(l)}$, i.e., (5.13).
    Calculate $\boldsymbol{z}_{\text{mean}}$, $\boldsymbol{v}^{(l)}$, and $\alpha^{(l)}$ for $l = \{1, \ldots, \tilde{m}\}$.
    Sort the vectors $\boldsymbol{v}^{(l)}$ according to the angles $\alpha^{(l)}$.
    Intersect the tangent lines of every two neighbouring points to obtain the polygon vertices $\boldsymbol{w}^{(l)}$.
    Update $\boldsymbol{B}_j$ and $\boldsymbol{d}_j$ by solving (5.12).
  end for
end for

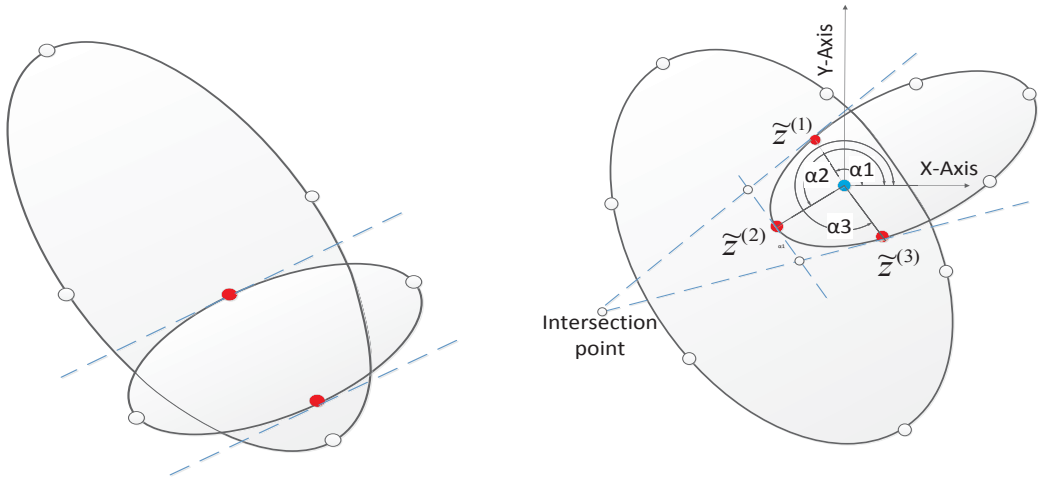

Fig. 5.2 Degenerate cases happening in finding a tight polygon. Left: parallel tangent lines can not form a closed region. Right: The obtained polygon does not contain the intersection of the ellipses.

after convergence, i.e., the CPU time required such that the difference between average areas in two consecutive iterations is less than 0.01m^2. Since the results are obtained by processing the information centrally on a CPU, we divide the computation time by the number of sensors N to have a better insight of the computation time in a distributed WSN. The results show that the proposed method has similar computation time compared to the one in [65]. Therefore, in terms of the trade-off between accuracy of outer-approximation and computational cost, the proposed method is clearly preferred.

Table 5.1 Comparison of computation times per sensor.

Methods	$N = 10$	$N = 20$	$N = 50$
Technique from [65]	2.8 s	4.3 s	11.3 s
Proposed	2.8 s	4.1 s	11.7 s

5.5 Conclusion and Future Work

In this chapter, we developed a tight outer-approximation of the intersection of ellipses in 2-D space. The proposed method finds applications in control, parameter estimation, and localization. It was used herein as part of a distributed bounding algorithm in a

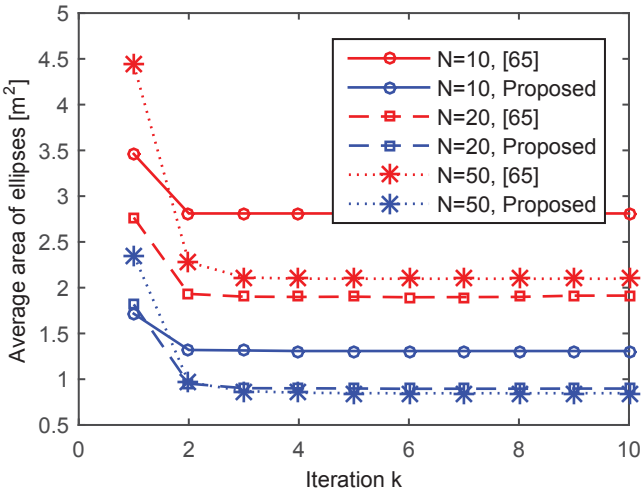

Fig. 5.3 Comparison of the average area of bounding ellipses as a function of the iteration index k for different N.

cooperative WSN with positive range measurement errors for outer-approximation of the convex sets containing the positions of the sensors. Through simulations, it was shown that the proposed method results in a tighter approximation of the convex sets compared to existing techniques. Improving the robustness of the proposed method against degeneracy problems and making it more computationally efficient will be considered in future work.

Chapter 6

Summary and Conclusion

In this chapter, we summarize the main contributions of this thesis and discuss some possible research paths for future work.

6.1 Summary of the Thesis

Network-based RF localization has gained considerable attention in the past decades due to the tremendous number of applications where location information is required, while the conventional localization systems such as the global positioning system (GPS) can not be employed due to its limitations. Positioning in harsh propagation environments such as dense urban areas, indoor places and underground areas, where the GPS satellites are not visible to the receiver and the GPS signal is attenuated, are among such applications. Network-based RF localization, where instead of satellites, a set of fixed reference nodes transmit and/or receive signal to/from a wireless device (target) can overcome the limitation of the GPS. Since the number of reference points is limited and might not be sufficient for unambiguous localization, it becomes necessary and beneficial for the targets to make pairwise measurements and exchange information with their neighbour targets; consequently cooperative localization has gained attention.

Nevertheless, localization in harsh propagation environments is challenging and can lead to large errors due to the multipaths and non-line of sight (NLOS) propagation. The problem of multipaths can be overcome using high resolution ultra wide-band (UWB) timing pulses so that accurate time of arrival (TOA) measurements can be obtained. However, the NLOS problem, in which the range measurements become positively biased still remain the

main challenge. In this thesis, the main goal was to develop efficient and reliable methods for network-based localization in harsh propagation environments. In particular the problem of NLOS which is one of the main sources of error in indoor places and urban canyons was considered.

In Chapter 3, a CSRUKF for tracking a mobile node in NLOS situation was proposed where first a non-cooperative scenario was considered and then the method was extended to a cooperative localization scenario. In this filter, 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 which 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 Kalman filter 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 significantly better when the measurement noise variance was small, 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.

In Chapter 4, a 2-stage distributed cooperative localization method which is robust against NLOS error was described. In the first stage, we applied a convex relaxation on the Huber cost function and preliminary position estimates were obtained iteratively. In the second stage of our algorithm, by iteratively minimizing the Huber loss function, it was shown that further refinement of the position estimates could be generally obtained. For iterative optimization in each stage, a gradient descent method was used. By testing our algorithm with real data set, its superiority over other competing algorithms was verified. The results showed that our two-stage algorithm performs robustly against outliers; in particular it significantly outperforms other distributed techniques when the ratio of NLOS to LOS measurements is low.

In Chapter 5, we focused on distributed localization techniques and outer-approximation of feasible sets in NLOS scenarios. We considered the problem of outer-approximation of convex sets, which can be formed as a result of positively biased range measurements in a WSN. We applied a distributed localization algorithm, which is based on SPAWN, to this problem, and developed a novel technique for outer approximation of the intersection

of ellipses in 2-D, which could give a tighter ellipsoidal outer-approximation compared to other benchmark approaches. The proposed outer-approximation method, when used in the SPAWN algorithm, could improve the performance of the distributed outer-approximation of convex sets in a WSN and could yield more accurate location estimates; it is thus of practical interest for 2-D localization.

In conclusion, in addition to being novel in theory, the proposed techniques can be of great importance in practice due to their performances and moderate computational costs.

6.2 Future Work

While the localization in a cellular network and a WSN has been considered for long and the research area seems saturated, there are still several challenges for improving the localization performance in practice. Some of these problems are briefly summarized below:

1. The effect of clock parameters on the localization performance under LOS condition has been considered in several works. However, mitigating the effect of clock errors on localization performance under NLOS conditions especially for cooperative WSNs is a potential topic that has been rarely considered except in a few recent works.

2. As we have seen in many occasions in this thesis, radio localization under NLOS conditions often leads to an NLS optimization problem with geometrical constraints. Developing faster iterative distributed algorithms based on Gauss-Newton or Newton methods for fast convergence to a local minimum of the cost function is still an open problem and may be considered in the future.

3. It is required by FCC that by 2020, all wireless devices be localized with an accuracy of less than 1 meter at all times in the next generation of cellular network 5G. Therefore, developing robust and novel techniques suitable for the specification of the 5G networks is of great importance. In addition to being accurate and robust, algorithms that are compatible with 5G standard and do not violate the security of the wireless devices, need to be developed. While developing novel algorithms for improving the localization accuracy has several applications, location information can be employed in different layers of wireless communication systems in 5G networks to reduce the delay and improve the performance of communication. Although in many different

algorithms, location information is assumed to be perfect, the effect of location estimation error on the communication systems can be tested and more novel techniques can be proposed to be robust against location errors. This is a promising research topic especially in 5G networks and is therefore worth consideration [127].

References

[1] T. Jia and R. Buehrer, "Collaborative position location with NLOS mitigation," in *IEEE 21st Int. Symp. on Personal, Indoor and Mobile Radio Communications Workshops*, pp. 267–271, Sep. 2010.

[2] M. Gholami, H. Wymeersch, E. Strom, and M. Rydstrom, "Wireless network positioning as a convex feasibility problem," *EURASIP J. on Wireless Communications and Networking*, no. 1, pp. 1–15, 2011.

[3] A. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location: challenges faced in developing techniques for accurate wireless location information," *IEEE Signal Process. Magazine*, vol. 22, pp. 24–40, Jul. 2005.

[4] Y. Gu, A. Lo, and I. Niemegeers, "A survey of indoor positioning systems for wireless personal networks," *IEEE Communications Surveys and Tutorials*, vol. 11, pp. 13–32, quarter 2009.

[5] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. on Antennas and Propagation*, vol. 34, pp. 276–280, Mar. 1986.

[6] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. on Acoustics, Speech and Signal Process.*, vol. 37, pp. 984–995, Jul. 1989.

[7] D. Dardari, R. D'Errico, C. Roblin, A. Sibille, and M. Win, "Ultrawide bandwidth RFID: The next generation?," *Proceedings of the IEEE*, vol. 98, pp. 1570–1582, Sep. 2010.

[8] K. Yu, I. Sharp, and G. Y.Jay, *Ground-Based Wireless Positioning*. John Wiley and Sons Ltd., 2009.

[9] S. Zhu and Z. Ding, "Joint synchronization and localization using TOAs: A linearization based WLS solution," *IEEE J. Selected Areas in Communications*, vol. 28, pp. 1017–1025, Sep. 2010.

[10] Y. Wang, G. Leus, and X. Ma, "Tracking a mobile node by asynchronous networks," in *IEEE Worshop on Signal Processing Advances in Wireless Communications*, pp. 161–165, Jun. 2011.

[11] E. Larsson, "Cramer-Rao bound analysis of distributed positioning in sensor networks," *IEEE Signal Process. Letters*, vol. 11, pp. 334–337, Mar. 2004.

[12] B. Denis, J.-B. Pierrot, and C. Abou-Rjeily, "Joint distributed synchronization and positioning in UWB ad hoc networks using TOA," *IEEE Trans. on Microwave Theory and Techniques*, vol. 54, pp. 1896–1911, Jun. 2006.

[13] S. Gezici, Z. Tian, G. Giannakis, H. Kobayashi, A. Molisch, H. Poor, and Z. Sahinoglu, "Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks," *IEEE Signal Process. Mag.*, vol. 22, pp. 70–84, Jul. 2005.

[14] A. Molisch, "Ultrawideband propagation channels-theory, measurement, and modeling," *IEEE Trans. on Vehicular Tech.*, vol. 54, pp. 1528–1545, Sep. 2005.

[15] L. Taponecco, A. D'Amico, and U. Mengali, "Joint TOA and AOA estimation for UWB localization applications," *IEEE Trans. on Wireless Commun.*, vol. 10, pp. 2207–2217, Jul. 2011.

[16] M. Navarro and M. Najar, "Frequency domain joint TOA and DOA estimation in IR-UWB," *IEEE Trans. on Wireless Commun.*, vol. 10, pp. 1–11, Oct. 2011.

[17] E. Lagunas, M. Najar, and M. Navarro, "Joint TOA and DOA estimation compliant with IEEE 802.15.4a standard," in *5th IEEE Int. Symp. on Wireless Pervasive Computing*, pp. 157–162, May 2010.

[18] E. Lagunas, M. Najar, and M. Navarro, "UWB joint TOA and DOA estimation," in *IEEE Int. Conf. on Ultra-Wideband*, pp. 839–843, Sep. 2009.

[19] F. Shang, I. Psaramiligkos, and B. Champagne, "A ML-based framework for joint TOA/AOA estimation of UWB pulses in dense multipath environments," *IEEE Trans. on Wireless Commun.*, vol. 13, pp. 5305 – 5318, Jul. 2014.

[20] N. Patwari, J. Ash, S. Kyperountas, A. Hero, R. Moses, and N. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *IEEE Signal Process. Magazine*, vol. 22, no. 4, pp. 54–69, 2005.

[21] H. Wymeersch, J. Lien, and M. Win, "Cooperative localization in wireless networks," *Proceedings of the IEEE*, vol. 97, pp. 427–450, Feb. 2009.

[22] R. Zekavat and M. Buehrer, *Handbook of Position Location: Theory, Practice and Advances*. Wiley-IEEE Press, 2011.

[23] P. Biswas and Y. Ye, "Semidefinite programming for ad hoc wireless sensor network localization," in *Proc. of the Third Int. Symp. on Information Processing in Sensor Networks*, (New York, NY, USA), pp. 46–54, ACM, 2004.

[24] P. Tseng, "Second-order cone programming relaxation of sensor network localization," *SIAM J. on Optimization*, vol. 18, pp. 156–185, Feb. 2007.

[25] G. Shirazi, M. Shenouda, and L. Lampe, "Second order cone programming for sensor network localization with anchor position uncertainty," in *8th Workshop on Positioning Navigation and Communication*, pp. 51–55, Mar. 2011.

[26] A. Abramo, F. Blanchini, L. Geretti, and C. Savorgnan, "A mixed convex/nonconvex distributed localization approach for the deployment of indoor positioning services," *IEEE Trans. on Mobile Computing*, vol. 7, pp. 1325–1337, Nov. 2008.

[27] T. Jia and R. Buehrer, "A set-theoretic approach to collaborative position location for wireless networks," *IEEE Trans. on Mobile Computing*, vol. 10, pp. 1264–1275, Sep. 2011.

[28] A. Ihler, J. Fisher, R. Moses, and A. Willsky, "Nonparametric belief propagation for self-localization of sensor networks," *IEEE J. on Selected Areas in Commun.*, vol. 23, pp. 809–819, Apr. 2005.

[29] Y. Qi, H. Kobayashi, and H. Suda, "Analysis of wireless geolocation in a non-line-of-sight environment," *IEEE Trans. on Wireless Commun.*, vol. 5, pp. 672–681, Mar. 2006.

[30] Y.-T. Chan, W.-Y. Tsui, H.-C. So, and P.-C. Ching, "Time-of-arrival based localization under NLOS conditions," *IEEE Trans. on Vehicular Tech.*, vol. 55, pp. 17–24, Jan. 2006.

[31] N. Alsindi, C. Duan, J. Zhang, and T. Tsuboi, "NLOS channel identification and mitigation in ultra wideband TOA-based wireless sensor networks," in *Proc. 6th Workshop on Positioning, Navigation and Communication*, pp. 59–66, Mar. 2009.

[32] I. Guvenc, C.-C. Chong, and F. Watanabe, "NLOS identification and mitigation for UWB localization systems," in *Proc. IEEE Wireless Communications and Networking Conference*, pp. 1571–1576, Mar. 2007.

[33] 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. A. Commun.*, vol. 28, pp. 1026–1035, Sep. 2010.

[34] F. Montorsi, F. Pancaldi, and G. Vitetta, "Statistical characterization and mitigation of NLOS errors in UWB localization systems," in *Proc. IEEE Int. Conf. on Ultra-Wideband*, pp. 86–90, Sep. 2011.

[35] L. Cong and W. Zhuang, "Non-line-of-sight error mitigation in TDOA mobile location," in *Proc. IEEE Global Telecommunications Conf.*, vol. 1, pp. 680–684, 2001.

[36] L. Cong and W. Zhuang, "Nonline-of-sight error mitigation in mobile location," in *Proc. Twenty-third Annual Joint Conf. of the IEEE Computer and Communications Societies*, vol. 1, Mar. 2004.

[37] J. Figueiras and S. Frattasi, *Mobile Positioning and Tracking from Conventional to Cooperative Techniques*. John Wiley and Sons Ltd., 2010.

[38] X. Wang, Z. Wang, and B. O'Dea, "A TOA-based location algorithm reducing the errors due to non-line-of-sight (NLOS) propagation," *IEEE Trans. on Vehicular Tech.*, vol. 52, pp. 112–116, Jan. 2003.

[39] S. Venkatesh and R. Buehrer, "A linear programming approach to NLOS error mitigation in sensor networks," in *Proc. Fifth Int. Conf. on Information Processing in Sensor Networks*, pp. 301–308, 2006.

[40] S. Venkatesh and R. Buehrer, "NLOS mitigation using linear programming in ultrawideband location-aware networks," *IEEE Trans. on Vehicular Tech.*, vol. 56, pp. 3182–3198, Sep. 2007.

[41] C.-L. Chen and K.-T. Feng, "An efficient geometry-constrained location estimation algorithm for NLOS environments," in *Proc. Int. Conf. on Wireless Networks, Communications and Mobile Computing*, vol. 1, pp. 244–249, Jun. 2005.

[42] Y. Kim, J. Lee, and G.-I. Jee, "The interior-point method for an optimal treatment of bias in trilateration location," *IEEE Trans. on Vehicular Tech.*, vol. 55, pp. 1291–1301, Jul. 2006.

[43] K. Yu and Y. Guo, "Improved positioning algorithms for nonline-of-sight environments," *IEEE Trans. on Vehicular Tech.*, vol. 57, pp. 2342–2353, Jul. 2008.

[44] I. Guvenc and C.-C. Chong, "A survey on TOA based wireless localization and NLOS mitigation techniques," *IEEE Communications Surveys Tutorial*, vol. 11, pp. 107–124, quarter 2009.

[45] H. Miao, K. Yu, and M. Juntti, "Positioning for NLOS propagation: Algorithm derivations and Cramer-Rao bounds," *IEEE Trans. on Vehicular Tech.*, vol. 56, pp. 2568–2580, Sep. 2007.

[46] H. Miao, M. Juntti, and K. Yu, "2-D unitary ESPRIT based joint AOA and AOD estimation for MIMO system," in *IEEE 17th Int. Symp. on Personal, Indoor and Mobile Radio Communications*, pp. 1–5, Sep. 2006.

[47] K. Papakonstantinou and D. Slock, "Hybrid TOA/AOD/doppler-shift localization algorithm for NLOS environments," in *IEEE 20th Int. Symp. on Personal, Indoor and Mobile Radio Communications*, pp. 1948–1952, Sep. 2009.

[48] K. Papakonstantinou and D. Slock, "ESPRIT-based estimation of location and motion dependent parameters," in *IEEE 69th Vehicular Technology Conf., Spring*, pp. 1–5, Apr. 2009.

[49] S. Al-Jazzar and M. Ghogho, "A joint TOA/AOA constrained minimization method for locating wireless devices in non-line-of-sight environment," in *IEEE 66th Vehicular Technology Conf., Fall*, pp. 496–500, Oct. 2007.

[50] N. Thomas, D. Cruickshank, and D. Laurenson, "Performance of a TDOA-AOA hybrid mobile location system," in *Proc. 2nd Int. Conf. on 3G Mobile Communication Technologies*, pp. 216–220, 2001.

[51] C.-D. Wann, Y.-J. Yeh, and C.-S. Hsueh, "Hybrid TDOA/AOA indoor positioning and tracking using extended Kalman filters," in *Proc. IEEE Vehicular Tech. Conf.-Spring*, vol. 3, pp. 1058–1062, May 2006.

[52] J. Hol, F. Dijkstra, H. Luinge, and T. Schon, "Tightly coupled UWB/IMU pose estimation," in *Proc. IEEE Int. Conf. on Ultra-Wideband*, pp. 688–692, Sep. 2009.

[53] 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. on Ultra-Wideband*, pp. 160–164, Sep. 2011.

[54] K. Yu and E. Dutkiewicz, "NLOS identification and mitigation for mobile tracking," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 49, no. 3, pp. 1438–1452, 2013.

[55] B. 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. on Vehicular Tech.*, vol. 58, pp. 1157–1169, Mar. 2009.

[56] 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. on Acoustics, Speech and Signal Process.*, pp. 3049–3052, 2009.

[57] U. Hammes, E. Wolsztynski, and A. Zoubir, "Robust tracking and geolocation for wireless networks in NLOS environments," *IEEE J. of Selected Topics in Signal Process.*, vol. 3, pp. 889–901, Oct. 2009.

[58] U. Hammes and A. Zoubir, "Robust MT tracking based on M-estimation and interacting multiple model algorithm," *IEEE Trans. on Signal Process.*, vol. 59, pp. 3398–3409, Jul. 2011.

[59] M. Najar, J. Huerta, J. Vidal, and J. Castro, "Mobile location with bias tracking in non-line-of-sight," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 956–9, May 2004.

[60] D. Jourdan, J. Deyst, J.J., M. Win, and N. Roy, "Monte Carlo localization in dense multipath environments using UWB ranging," in *Proc. IEEE Int. Conf. on Ultra-Wideband*, pp. 314–319, Sep. 2005.

[61] J. González, J. L. Blanco, C. Galindo, A. Ortiz-de Galisteo, J. A. Fernández-Madrigal, F. A. Moreno, and J. L. Martínez, "Mobile robot localization based on ultra-wideband ranging: A particle filter approach," *Robot. Auton. Syst.*, vol. 57, pp. 496–507, May 2009.

[62] S. Yousefi, X. Chang, and B. Champagne, "Improved extended Kalman filter for mobile localization with NLOS anchors," in *Proc. Int. Conf. on Wireless and Mobile Communications*, pp. 25–30, Jul. 2013.

[63] H. Chen, G. Wang, Z. Wang, H. So, and H. Poor, "Non-line-of-sight node localization based on semi-definite programming in wireless sensor networks," *IEEE Trans. on Wireless Commun.*, vol. 11, pp. 108–116, Jan. 2012.

[64] R. Vaghefi and R. Buehrer, "Cooperative sensor localization with NLOS mitigation using semidefinite programming," in *9th Workshop on Positioning Navigation and Communication*, pp. 13–18, Mar. 2012.

[65] M. Gholami, H. Wymeersch, S. Gezici, and E. Strom, "Distributed bounding of feasible sets in cooperative wireless network positioning," *IEEE Commun. Letters*, vol. 17, pp. 1596–1599, Aug. 2013.

[66] S. P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*, vol. 15. SIAM, 1994.

[67] R. Van der Merwe and E. Wan, "The square-root unscented Kalman filter for state and parameter-estimation," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, vol. 6, pp. 3461–3464, 2001.

[68] R. Kandepu, B. Foss, and L. Imsland, "Applying the unscented Kalman filter for nonlinear state estimation," *J. of Process Control*, vol. 18, no. 78, pp. 753–768, 2008.

[69] N. Patwari, A. Hero, M. Perkins, N. Correal, and R. O'Dea, "Relative location estimation in wireless sensor networks," *IEEE Trans. on Signal Process.*, vol. 51, no. 8, pp. 2137–2148, 2003.

[70] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.

[71] Q. Liu, X. Liu, J. Zhou, G. Zhou, G. Jin, Q. Sun, and M. Xi, "Adasynch: A general adaptive clock synchronization scheme based on kalman filter for WSNs," *Wireless Personal Communications*, vol. 63, pp. 217–239, 2012.

[72] Y. Chan and K. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. on Signal Process.*, vol. 42, pp. 1905–1915, Aug. 1994.

[73] Z. Yang, J. Pan, and L. Cai, "Adaptive clock skew estimation with interactive multimodel kalman filters for sensor networks," in *Proc. IEEE Int. Conf. on Communications*, pp. 1–5, May 2010.

[74] K. C. Ho, "Bias reduction for an explicit solution of source localization using TDOA," *IEEE Trans. on Signal Process.*, vol. 60, pp. 2101–2114, May 2012.

[75] Y. Zhou and L. Lamont, "Constrained linear least squares approach for TDOA localization: A global optimum solution," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pp. 2577–2580, Apr. 2008.

[76] K. W. Cheung, H. C. So, W.-K. Ma, and Y. T. Chan, "A constrained least squares approach to mobile positioning: algorithms and optimality," *EURASIP J. Appl. Signal Process.*, vol. 2006, pp. 150–150, Jan. 2006.

[77] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 37, pp. 984–995, Jul. 1989.

[78] R. Vaghefi, M. Gholami, and E. Strom, "RSS-based sensor localization with unknown transmit power," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pp. 2480–2483, May 2011.

[79] H. Lohrasbipeydeh, T. Gulliver, and H. Amindavar, "A minimax SDP method for energy based source localization with unknown transmit power," *IEEE Wireless Communications Letters*, vol. 3, pp. 433–436, Aug. 2014.

[80] M. Gholami, R. Vaghefi, and E. Strom, "RSS-based sensor localization in the presence of unknown channel parameters," *IEEE Trans. on Signal Process.*, vol. 61, pp. 3752–3759, Aug. 2013.

[81] B. Li, I. Quader, and A. G. Dempster, "On outdoor positioning with Wi-Fi," *J. of GPS*, pp. 18–26, 2008.

[82] I. Quader, B. Li, W.-P. Peng, and A. Dempster, "Use of fingerprinting in Wi-Fi based outdoor positioning," in *Int. Global Navigation Satellite System Society*, Dec. 2007.

[83] A. Kushki, K. Plataniotis, and A. Venetsanopoulos, *WLAN Positioning Systems: Principles and Applications in Location-Based Services*. Cambridge Press., 2012.

[84] K. W. Cheung, H. C. So, W.-K. Ma, and Y. T. Chan, "A constrained least squares approach to mobile positioning: algorithms and optimality," *EURASIP J. Appl. Signal Process.*, pp. 150–150, Jan. 2006.

[85] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, pp. 401–422, Mar. 2004.

[86] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Trans. on Signal Process.*, vol. 50, pp. 174–188, Feb. 2002.

[87] L. Doherty, K. pister, and L. El Ghaoui, "Convex position estimation in wireless sensor networks," in *Proc. IEEE 20th Annual Joint Conf. of the Computer and Communications Societies*, vol. 3, pp. 1655–1663, 2001.

[88] A. M.-C. So and Y. Ye, "Theory of semidefinite programming for sensor network localization," in *Proc. of the sixteenth annual ACM-SIAM Symp. on Discrete Algorithms*, (Philadelphia, PA, USA), pp. 405–414, 2005.

[89] S. Zhu and Z. Ding, "Distributed cooperative localization of wireless sensor networks with convex hull constraint," *IEEE Trans. on Wireless Commun.*, vol. 10, pp. 2150–2161, Jul. 2011.

[90] H. Trussell and M. Civanlar, "The landweber iteration and projection onto convex sets," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 33, pp. 1632–1634, Dec. 1985.

[91] A. Hero and D. Blatt, "Sensor network source localization via projection onto convex sets (POCS)," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, vol. 3, pp. iii/689–iii/692, Mar. 2005.

[92] S. Mazuelas, F. Lago, J. Blas, A. Bahillo, P. Fernandez, R. Lorenzo, and E. Abril, "Prior NLOS measurement correction for positioning in cellular wireless networks," *IEEE Trans. on Vehicular Tech.*, vol. 58, pp. 2585–2591, Jun. 2009.

[93] A.-J. van der Veen, M. Vanderveen, and A. Paulraj, "Joint angle and delay estimation using shift-invariance properties," *IEEE Signal Process. Letters*, vol. 4, no. 5, pp. 142–145, 1997.

[94] M. Najar and J. Vidal, "Kalman tracking for mobile location in NLOS situations," in *IEEE Proc. on Personal, Indoor and Mobile Radio Communications*, vol. 3, pp. 2203–2207 vol.3, Sep. 2003.

[95] Y. Xie, Y. Wang, P. Zhu, and X. You, "Grid-search-based hybrid TOA/AOA location techniques for NLOS environments," *IEEE Communications Letters*, vol. 13, pp. 254–256, Apr. 2009.

[96] D. Simon, "Kalman filtering with state constraints: a survey of linear and nonlinear algorithms," *IET Control Theory Applications*, vol. 4, no. 8, pp. 1303–1318, 2010.

[97] H. Liu, F. K. W. Chan, and H. C. So, "Non-line-of-sight mobile positioning using factor graphs," *IEEE Trans. on Vehicular Tech.*, vol. 58, pp. 5279–5283, Nov. 2009.

[98] S. Van de Velde, H. Wymeersch, and H. Steendam, "Comparison of message passing algorithms for cooperative localization under NLOS conditions," in *Proc. 9th Workshop on Positioning Navigation and Communication*, pp. 1–6, Mar. 2012.

[99] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[100] S. Venkatesh and R. Buehrer, "Non-line-of-sight identification in ultra-wideband systems based on received signal statistics," *IET Microwaves, Antennas Propagation*, vol. 1, pp. 1120–1130, Dec. 2007.

[101] R. M. Vaghefi and R. M. Buehrer, "Target tracking in NLOS environments using semidefinite programming," in *IEEE Military Communications Conf.*, pp. 169–174, Nov. 2013.

[102] D. Simon and D. L. Simon, "Constrained Kalman filtering via density function truncation for turbofan engine health estimation," *Int. J. of Systems Science*, vol. 41, no. 2, pp. 159–171, 2010.

[103] J. Lan and X. Li, "State estimation with nonlinear inequality constraints based on unscented transformation," in *Proc. 14th Int. Conf. on Information Fusion*, pp. 1–8, Jul. 2011.

[104] D. Simon, *Optimal State Estimation: Kalman, H-Infinity, and Nonlinear Approaches*. Wiley-Interscience, Aug. 2006.

[105] D. Zachariah, I. Skog, M. Jansson, and P. Handel, "Bayesian estimation with distance bounds," *IEEE Signal Process. Letters*, vol. 19, pp. 880–883, Dec. 2012.

[106] D. Simon and D. Simon, "Kalman filtering with inequality constraints for turbofan engine health estimation," *IEE Proc. on Control Theory and Applications*, vol. 153, pp. 371–378, May 2006.

[107] C. Paige, "Computer solution and perturbation analysis of generalized linear least squares problems," *J. Mathematics of Computation*, vol. 33, no. 145, pp. 171–183, 1979.

[108] D. Luenberger and Ye, *Linear and Nonlinear Programming*. Springer, third ed., 2008.

[109] J. F. Strum, "Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones," *Optimization methods and software*, vol. 11, no. 1-4, pp. 625–653, 1999.

[110] U. Hammes and A. Zoubir, "Robust mobile terminal tracking in NLOS environments based on data association," *IEEE Trans. on Signal Process.*, vol. 58, pp. 5872–5882, Nov. 2010.

[111] B. L. Le, K. Ahmed, and H. Tsuji, "Mobile location estimator with NLOS mitigation using Kalman filtering," in *Proc. IEEE Wireless Communications and Networking Conference*, vol. 3, pp. 1969–1973, Mar. 2003.

[112] J. Lofberg, "Yalmip : a toolbox for modeling and optimization in matlab," in *Proc. IEEE Int. Symp. on Computer Aided Control Systems Design*, pp. 284–289, 2004.

[113] K. Yu and E. Dutkiewicz, "Improved kalman filtering algorithms for mobile tracking in NLOS scenarios," in *Proc. IEEE Wireless Communications and Networking Conference*, pp. 2390–2394, Apr. 2012.

[114] L. Chen, S. Ali-Lytty, R. Pich, and L. Wu, "Mobile tracking in mixed line-of-sight/non-line-of-sight conditions: Algorithm and theoretical lower bound," *Wireless Personal Communications*, vol. 65, no. 4, pp. 753–771, 2012.

[115] C. Fritsche, A. Klein, and F. Gustafsson, "Bayesian Cramer-Rao bound for mobile terminal tracking in mixed LOS/NLOS environments," *IEEE Wireless Commun. Letters*, vol. PP, pp. 1–4, Jun. 2013.

[116] R. Vaghefi, J. Schloemann, and R. Buehrer, "NLOS mitigation in TOA-based localization using semidefinite programming," in *10th Workshop on Positioning Navigation and Commabunication*, pp. 1–6, Mar. 2013.

[117] P. J. Huber, "Robust estimation of a location parameter," *The Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 73–101, 1964.

[118] H. Wymeersch, S. Marano, W. Gifford, and M. Win, "A machine learning approach to ranging error mitigation for UWB localization," *IEEE Trans. on Commun.*, vol. 60, pp. 1719–1728, Jun. 2012.

[119] K. Pahlavan, F. O. Akgul, M. Heidari, A. Hatami, J. M. Elwell, and R. D. Tingley, "Indoor geolocation in the absence of direct path," *IEEE Wireless Communications*, vol. 13, pp. 50–58, Dec. 2006.

[120] F. L. Chernousko, "Ellipsoidal bounds for sets of attainability and uncertainty in control problems," *Optimal Control Applications and Methods*, vol. 3, no. 2, pp. 187–202, 1982.

[121] W. Kahan, "Circumscribing an ellipsoid about the intersection of two ellipsoids," *Can. Math. Bull*, vol. 11, no. 3, pp. 437–441, 1968.

[122] D. Maksarov and J. Norton, "State bounding with ellipsoidal set description of the uncertainty," *Int.l J. of Control*, vol. 65, no. 5, pp. 847–866, 1996.

[123] L. Ros, A. Sabater, and F. Thomas, "An ellipsoidal calculus based on propagation and fusion," *IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 32, pp. 430–442, Aug. 2002.

[124] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization*. Society for Industrial and Applied Mathematics, 2001.

[125] N. Z. Shor and O. Berezovski, "New algorithms for constructing optimal circumscribed and inscribed ellipsoids," *Optimization Methods and Software*, vol. 1, no. 4, pp. 283–299, 1992.

[126] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab software for disciplined convex programming, version 2.1," Sep. 2014.

[127] R. Di Taranto, S. Muppirisetty, R. Raulefs, D. Slock, T. Svensson, and H. Wymeersch, "Location-aware communications for 5G networks: How location information can improve scalability, latency, and robustness of 5G," *IEEE Signal Process. Magazine*, vol. 31, pp. 102–112, Nov. 2014.