TW-TOA based positioning in the presence of clock imperfections

TW-TOA Based Positioning in the Presence of Clock Imperfections

Mohammad Reza Gholami[a,*], Sinan Gezici[b], Erik G. Ström[c]

[a] *Campanja AB, Stockholm SE-111 57, Sweden*
[b] *Department of Electrical and Electronics Engineering, Bilkent University, Ankara 06800, Turkey.*
[c] *Division of Communication Systems, Department of Signals and Systems, Chalmers University of Technology.*

Abstract

This manuscript studies the positioning problem based on two-way time-of-arrival (TW-TOA) measurements in semi-asynchronous wireless sensor networks in which the clock of a target node is unsynchronized with the reference time. Since the optimal estimator for this problem involves difficult nonconvex optimization, two suboptimal estimators are proposed based on the squared-range least squares and the least absolute mean of residual errors. We formulated the former approach as an extended general trust region subproblem (EGTR) and propose a simple technique to solve it approximately. The latter approach is formulate as a *difference of convex functions programming* (DCP), which can be solved using a concave-convex procedure. Simulation results illustrate the high performance of the proposed techniques, especially for the DCP approach.

Keywords: Positioning, two-way time-of-arrival (TW-TOA), clock imperfection, convex optimization, trust region subproblems, concave-convex procedure.

1. Introduction

Location aware services are becoming vital requirements for many wireless systems. Due to some drawbacks of using GPS receivers at wireless nodes for some scenarios, self-position recovery has been proposed as an alternative approach and extensively investigated in the literature [1, 2, 3, 4, 5, 6]. Positioning based on range estimates between nodes is a popular technique in the literature. For synchronous networks, the time-of-arrival (TOA) technique provides a good estimate of the distance between two nodes for reasonable signal-to-noise ratios. A huge number of algorithms have been proposed in the literature to address the positioning problem based on range measurements, e.g., the maximum likelihood estimator [2], linear least-squares [7, 8, 9], squared-range least squares [10], projection onto convex sets [11, 12, 13], and convex relaxation techniques [14, 15, 16, 17, 18].

*Corresponding author
Email addresses: mohrg@kth.se (Mohammad Reza Gholami), gezici@ee.bilkent.edu.tr (Sinan Gezici), erik.strom@chalmers.se (Erik G. Ström)

In asynchronous networks, the range estimate based on the TOA is highly sensitive to clock imperfections. Therefore, the positioning accuracy can be considerably degraded in the presence of clock imperfections. In particular, for an affine model describing the clock behavior, the accuracy of the positioning techniques based on the TOA measurements is affected by non-ideal clock offset and clock skew. The clock of a target node can be synchronized with a reference time (clock) with a synchronization technique using, e.g., the MAC layer time stamp exchange, e.g., see, [19, 20, 21, 22, 23] and references therein. Motivated by pairwise synchronization techniques, the authors in [24] formulate a joint synchronization and positioning problem in the MAC layer. If the major part of the delay is the fixed delay due to propagation through the radio channel, the joint position and timing estimation technique works well. In [25], the positioning problem is studied in the presence of clock imperfection, which is only due to the clock offset. Considering the effects of an imperfect clock on distance estimates in the physical layer, the authors in [26] investigate the positioning problem using time-difference-of-arrival (TDOA) measurements in the presence of clock imperfections. The TDOA technique effectively removes the clock offset, but still suffers from the clock skew. Another popular approach for estimating the distance between sensor nodes is to use a so-called two-way time-of-arrival (TW-TOA) or time-of-flight based technique, which is an elegant approach in removing the effect of the clock offset on range measurements [9]. TW-TOA based positioning has an important advantage over the conventional TOA and TDOA based positioning techniques in terms of implementation complexity. This is due to the fact that TOA based positioning requires synchronization among all reference nodes and the target node, and TDOA based positioning requires synchronization among all reference (anchor) nodes. On the other hand, no synchronization is required for TW-TOA based positioning.

Range estimates obtained via TW-TOA are affected by the clock skew and a processing delay called the turn-around time [27]. A number of researchers have tackled the positioning problem or distance estimation based on TW-TOA in fully or partially asynchronous networks [28, 29, 30, 31]. The authors in [32] propose an approach to refine the position and clocks of the reference nodes during positioning and synchronization. To improve the range estimation via TW-TOA, an effective technique based on a new clock counter scheme is proposed in [33]. The authors in [29] study the positioning problem in the presence of clock imperfections for a TW-TOA based technique and propose a linear least squares based approach to solve the problem. The proposed approaches work well in some scenarios, e.g., when there is a sufficient number of reference nodes at known positions. In general, the previously proposed approaches require modifications to be effectively applied to the positioning problem in which the clock skew and turn around times are also unknown. In addition, for practical applications the proposed algorithms may not be robust against outliers and non-line-of-sight errors. In this study, we consider the positioning of a single target node based on TW-TOA

measurements in the presence of clock imperfections. In this approach, a target node transmits a signal to a reference node located at a known position and the reference node responds to the received signal after an unknown turn-around time delay. As it is common in the literature, we assume that the reference node measures the turn-around time by a loop back test and transmits the estimate to the target node [34, 33]. The target node then computes the round-trip delay based on an estimate of the turn-around time. Assuming an affine model for the clock of the oscillator, it is observed that the range estimation using the TW-TOA measurement is affected by an unknown clock skew of the target node. Modeling the measurement errors as Gaussian random variables, we obtain the optimal estimator to find the clock skew, and the location of the target node, and the turn-around times for the reference nodes. The optimal estimator poses a high dimensional optimization problem and needs more than one distance estimate for every link to provide good estimates of the unknown parameters. We, then, omit the effect of the turn-around times using a linear transformation and consequently obtain a near-optimal estimator to find the location and clock skew of the target node. Both the optimal and near-optimal estimators for the positioning problem considered in this study are nonconvex and difficult to solve. Using some approximations, we obtain two suboptimal estimators. In the first approach, we consider the squared-range least-squares approach and formulate the problem as an extended general trust region subproblem (EGTR)– a quadratic programming with two nonconvex constraints. In general, EGTR is difficult to solve; hence, we modify the proposed technique in [35] to approximately solve EGTR. In the second approach, we minimize the residual errors based on the ℓ_1 norm and arrive at a nonconvex problem in the form of the *difference of convex functions programming* (DCP). The estimator based on ℓ_1 norm minimization of the residual can be an effective approach when there are outliers or when the measurement errors deviate from the Gaussian distribution. For example in practical scenarios, the direct path may be blocked and the measured distance may be larger than the actual distance, resulting in positive bias and non-Gaussian errors. In the positioning literature the DCP approach was first applied to TDOA based positioning in [36]. We employ a similar concave-convex procedure as in [36] to solve the problem. Note that the latter approach is robust against outliers. Simulation results indicate the high performance of the proposed techniques, especially the DCP.

In summary, we extend our previous work [35] with the following main contributions:

- an approximate MLE (AMLE) to estimate the location and clock skew of the target node; (The MLE was first investigated in the previous work [35].)
- a suboptimal estimator based on extended general trust region subproblem (EGTR) for squared-range measurements; (The proposed approach is different from GTR in [35] since EGTR considers two

constraints as opposed to a single constraint in GTR proposed in [35].)

- a suboptimal estimator formulated as DCP that can be solved using a concave-convex procedure.

The remainder of the manuscript is organized as follows. Section 2 explains the signal model considered in this study. In Section 3 and 4 the localization algorithms are studied. Complexity analyses of different approaches are discussed in Section 5. Simulation results are presented in Section 6. Finally, Section 7 makes some concluding remarks.

Notation: The following notations are used in this manuscript. Lowercase Latin/Greek letters, e.g., a, b, β, denote scalar values and bold lowercase Latin/Greek letters represent vectors. Matrices are shown by bold uppercase Latin/Greek letters. $\mathbf{I}_M$ is the M by M identity matrix. The operator $\mathbb{E}\{\cdot\}$ is used to denote the expectation of a random variable (or vector). The ℓ_p norm of a vector is denoted by $\|\cdot\|_p$. The $\text{diag}(X_1, \ldots, X_N)$ is a diagonal matrix with diagonal elements $X_1, \ldots, X_N$. For two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, $\boldsymbol{A} \succeq \boldsymbol{B}$ means $\boldsymbol{A} - \boldsymbol{B}$ is positive semidefinite. $\bigtriangledown g(\boldsymbol{a})$ denotes the gradient of $g(\boldsymbol{x})$ at $\boldsymbol{x} = \boldsymbol{a}$. The set of all N-vector with positive components are denoted by $\mathbb{R}^N_+$. We use $\otimes$ to denote the Kronecker product.

2. System Model

Consider a two dimensional network [1] with N reference (anchor) nodes located at known positions $\boldsymbol{a}_i = [a_{i,1}\ a_{i,2}]^T \in \mathbb{R}^2,\ i = 1, ..., N$. Suppose that one target node is placed at unknown position $\boldsymbol{x} = [x_1\ x_2]^T \in \mathbb{R}^2$. We assume that the target node estimates the distance to a reference node by performing a TW-TOA measurement. We further assume that the clock value of an imperfect clock follows an affine relation with the true (global) time t [37, 19, 20, 22, 23]. That is, the clock value of reference node i is

$$h_i(t) \triangleq w_i t + \theta_i \tag{1}$$

where w_i is the skew and θ_i is the offset associated with the ith node clock. Note that a perfectly synchronized clock has $w_i = 1$ and $\theta_i = 0$. In practice, w_i is a number close to 1. For convenience, we denote the target clock as $h(t)$, where $h(t) = wt + \theta$.

A TW-TOA measurement between the target node and the ith reference node for the kth round (time) is carried out as follows: (a) the target sends a message to the reference node at (global) time $t^k_{i,1}$, (b) the message arrives at the reference node at time $t^k_{i,2}$, (c) the reference node sends a return message at time $t^k_{i,3}$, and (d) the return message arrives at the target node at time $t^k_{i,4}$. Clearly, $t^k_{i,2} - t^k_{i,1} = t^k_{i,4} - t^k_{i,3} = d_i/c$, where c is the propagation speed and $d_i \triangleq d(\boldsymbol{x}, \boldsymbol{a}_i) \triangleq \|\boldsymbol{x} - \boldsymbol{a}_i\|_2$ is the distance between the target and ith

[1] The generalization to a three dimensional network is straightforward, but is not explored in this study.

reference node. Moreover, $t^k_{i,3} = t^k_{i,2} + T_i$, where T_i is the turn-around time in the ith reference node, which is assumed to be fixed during the positioning process. The TWO-TOA measurement is computed in the target local clock as

$$z^k_i = \frac{1}{2}\big[h(t^k_{4,i}) - h(t^k_{1,i}) + n^k_i\big] = w\frac{d_i}{c} + w\frac{T_i}{2} + \frac{n^k_i}{2},\ k = 1, 2, \ldots, K \tag{2}$$

where n^k_i is the TW-TOA measurement error, which we model it as a zero-mean Gaussian with standard deviation σ_i, i.e., $n^k_i \sim \mathcal{N}(0, \sigma^2_i)$, and K as the number of the TW-TOA measurements during the positioning process.

The unknown parameter T_i either might be extremely small and can be neglected [20], (e.g., for a small network when there are no strict constraints on the MAC layer delay) or it needs to be estimated. One way to deal with the unknown parameter T_i is to jointly estimate it along with the location of the target node [38]. It can also be estimated by reference node i using a loop back test and is sent back to the target node [33]. In this study, we consider the latter approach. The estimate of T_i is

$$\tilde{T}^k_i = h_i(t^k_{3,i}) - h_i(t^k_{2,i}) + \epsilon^k_i = w_i T_i + \epsilon^k_i,\ k = 1, 2, \ldots, K \tag{3}$$

where we model the estimation error as $\epsilon \sim \mathcal{N}(0, \gamma^2_i)$.

In the sequel, we assume that the reference nodes are synchronized with a reference clock, e.g., via a GPS signal.[2] Therefore $w_i \approx 1$ and we can write $\tilde{T}^k_i \approx \hat{T}^k_i$, where

$$\hat{T}^k_i \triangleq T_i + \epsilon^k_i \tag{4}$$

or equivalently

$$T_i = \hat{T}^k_i - \epsilon^k_i. \tag{5}$$

Estimating the turn around time in reference nodes involves TOA measurements (in a loopback based test); hence, it is subject to TOA estimation errors [33].

We now replace T_i in (5) with that in (2) and obtain an approximate (transformed) model for measurements

$$z^k_i = w\frac{d_i}{c} + w\frac{\hat{T}^k_i}{2} + \frac{n^k_i}{2} - w\frac{\epsilon^k_i}{2}. \tag{6}$$

As mentioned, the approximation is good in the (reasonable) case when the reference nodes are equipped with accurate clocks.

[2] There error of synchronization using GPS signals is on the order of 10 nanoseconds or less [39, 40].

In the following sections, we use the input data $\left\{\{z_i^k, \hat{T}_i^k\}_{i=1}^N\right\}_{k=1}^K$ to obtain the optimal estimator based on models (2)–(4) or suboptimal estimators according to (6). The parameters w and T_i, $i = 1, 2, \ldots, N$, are considered as unknown nuisance parameters, while σ_i, γ_i, and $\boldsymbol{a}_i$ are assumed to be known for $i = 1, 2, \ldots, N$.

3. Maximum Likelihood Estimator

We define the measurement vector

$$\boldsymbol{m} \triangleq \begin{bmatrix} \boldsymbol{m}^1 \\ \boldsymbol{m}^2 \\ \vdots \\ \boldsymbol{m}^K \end{bmatrix}, \tag{7}$$

where

$$\boldsymbol{m}^k \triangleq \begin{bmatrix} z_1^k & z_2^k & \cdots & z_N^k & \hat{T}_1^k & \hat{T}_2^k & \cdots & \hat{T}_N^k \end{bmatrix}^T. \tag{8}$$

To obtain the MLE for joint estimation of the position and clock skew of the target node, the following optimization problem needs to be solved [41]:

$$\begin{bmatrix} \hat{\boldsymbol{x}}^T & \hat{w} & \hat{\boldsymbol{t}}_a^T \end{bmatrix} = \underset{w\in\mathbb{R}_+;\, \boldsymbol{t}_a\in\boldsymbol{R}_+^N;\, \boldsymbol{x}\in\mathbb{R}^2}{\arg\max} p(\boldsymbol{m}; w, \boldsymbol{t}_a, \boldsymbol{x}) \tag{9}$$

where $p(\boldsymbol{m}; w, \boldsymbol{t}_a, \boldsymbol{x})$ is the probability density function (pdf) of vector $\boldsymbol{m}$ indexed by the vector $[\hat{\boldsymbol{x}}^T \; w \; \boldsymbol{t}_a^T]^T$ and $\boldsymbol{t}_a = [T_1 \; T_2 \ldots T_N]^T$. Since the TOA measurement errors are assumed to be independent and identically distributed random variables, the pdf of $\boldsymbol{m}$ can be calculated from (2) and (4) as

$$p(\boldsymbol{m}; w, \boldsymbol{t}_a, \boldsymbol{x}) = \prod_{k=1}^K \prod_{i=1}^N \sqrt{\frac{2}{\pi\sigma_i^2}} \exp\left(-\frac{2(z_i^k - wT_i/2 - wd(\boldsymbol{x}, \boldsymbol{a}_i)/c)^2}{\sigma_i^2}\right) \sqrt{\frac{1}{2\pi\gamma_i^2}} \exp\left(-\frac{(\hat{T}_i^k - T_i)^2}{2\gamma_i^2}\right). \tag{10}$$

Then, the MLE is obtained as

$$\begin{aligned} \begin{bmatrix} \hat{\boldsymbol{x}}^T & \hat{w} & \hat{\boldsymbol{t}}_a \end{bmatrix}^T &= \underset{\boldsymbol{x}\in\mathbb{R}^2;\, w\in\mathbb{R}_+;\, \boldsymbol{t}_a\in\boldsymbol{R}_+^N}{\arg\max} p(\boldsymbol{m}; w, \boldsymbol{t}_a, \boldsymbol{x}) \\ &= \underset{\boldsymbol{x}\in\mathbb{R}^2;\, w\in\mathbb{R}_+;\, \boldsymbol{t}_a\in\boldsymbol{R}_+^N}{\arg\min} \sum_{k=1}^K \sum_{i=1}^N \frac{2}{\sigma_i^2}\left(z_i^k - w\frac{T_i}{2} - w\frac{d_i}{c}\right)^2 + \frac{(\hat{T}_i^k - T_i)^2}{2\gamma_i^2}. \end{aligned} \tag{11}$$

For the MLE formulated in (11) there are $N + 3$ unknowns. Therefore, for low numbers of messages K, the MLE problem can be ill-posed. To alleviate the difficulty for solving the optimal MLE in (11), we investigate another approximate MLE based on the model obtained in (6). In fact in the MLE we use both measurements and prior knowledge about T_i to jointly estimate the location, clock skew, and turn-around times, while, relying on the estimate of turn-around time, we can use the approximate model in (6) (a transformed model) to only estimate the location and clock skew.

We collect z_i^k from (6) in a vector $\boldsymbol{m}_a = [z_1^1 \ \dots \ z_N^1 \ \dots \ z_1^K \ \dots \ z_N^K]^T$. Next, we compute the pdf of $\boldsymbol{m}_a$ as

$$p(\boldsymbol{m}_a; w, \boldsymbol{x}) = \prod_{k=1}^{K}\prod_{i=1}^{N}\sqrt{\frac{2}{\pi(\sigma_i^2 + w^2\gamma_i^2)}}\exp\left(-\frac{2(z_i^k - wd_i/c - w\hat{T}_i^k/2)^2}{(\sigma_i^2 + w^2\gamma_i^2)}\right). \tag{12}$$

We then find an approximate MLE (AMLE)[3] as

$$\begin{aligned}\left[\hat{\boldsymbol{x}}^T \ \hat{w}\right]^T &= \underset{\boldsymbol{x}\in\mathbb{R}^2;\, w\in\mathbb{R}_+}{\arg\max}\ p(\boldsymbol{m}_a; w, \boldsymbol{x}) \\ &= \underset{\boldsymbol{x}\in\mathbb{R}^2;\, w\in\mathbb{R}_+}{\operatorname{argmin}} \sum_{k=1}^{K}\sum_{i=1}^{N}\frac{2}{(\sigma_i^2 + w^2\gamma_i^2)}\left(z_i^k - w\frac{\hat{T}_i^k}{2} - w\frac{d_i}{c}\right)^2 + \frac{1}{2}\ln(\sigma_i^2 + w^2\gamma_i^2). \end{aligned}\tag{13}$$

It is observed that the search domain in the AMLE in (13) is limited to the location $\boldsymbol{x}$ and the clock skew w, thus a lower dimensional search compared to that of the MLE in (11).

It is also noted that both the MLE and AMLE formulations in (11) and (13) pose difficult global optimization problems. To avoid the drawbacks in solving these problems, we propose two suboptimal estimators in the next section.

4. Proposed techniques

In this section, we propose two techniques based on squared-range least squares and ℓ_1 norm minimization of residuals. First, we divide both sides of (6) by w (we safely assume that $w \neq 0$) and express the model as

$$z_i^k\alpha - \frac{\hat{T}_i^k}{2} = \frac{d_i}{c} + \frac{n_i^k}{2}\alpha - \frac{\epsilon_i^k}{2},\ i = 1, 2, \dots, N,\ k = 1, 2, \dots, K \tag{14}$$

where $\alpha = 1/w$.

In the following, the model in (14) is employed in order to derive the proposed suboptimal estimators.

4.1. Squared-Range measurement Least Squares

In this section, we assume that the measurement noise $\alpha n_i^k/2 - \epsilon_i^k/2$ is small compared to d_i/c. Then, taking the square of both sides of (14) and dropping the small terms yield

$$(z_i^k\alpha)^2 + \frac{(\hat{T}_i^k)^2}{4} - z_i^k\hat{T}_i^k\alpha \simeq \frac{1}{c^2}(\boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{a}_i^T\boldsymbol{x} + \|\boldsymbol{a}_i\|_2^2) + \nu_i^k, \tag{15}$$

where $\nu_i^k = d_i(\alpha n_i^k - \epsilon_i^k)/c$. Now, we apply a weighted least squares criterion to the model in (15) and obtain the following minimization problem:

$$\underset{\boldsymbol{x}\in\mathbb{R}^2;\ \alpha\in\mathbb{R}_+}{\text{minimize}} \ \sum_{k=1}^{K}\sum_{i=1}^{N}\frac{1}{d_i^2(\alpha^2\sigma_i^2 + \gamma_i^2)}\left(\frac{1}{c^2}\boldsymbol{x}^T\boldsymbol{x} - \frac{2}{c^2}\boldsymbol{a}_i^T\boldsymbol{x} - (z_i^k)^2\alpha^2 + z_i^k\hat{T}_i^k\alpha + \frac{1}{c^2}\|\boldsymbol{a}_i\|_2^2 - \frac{(\hat{T}_i^k)^2}{4}\right)^2. \tag{16}$$

[3]We call the MLE in (13) as AMLE because it is based on the approximate model (6) instead of the original measurements in (7).

The problem in (16) can be expressed as a quadratic programming problem

$$\begin{aligned} &\underset{\boldsymbol{y}}{\text{minimize}} \quad \|\boldsymbol{W}^{1/2}(\boldsymbol{A}\boldsymbol{y}-\boldsymbol{b})\|_2^2 \\ &\text{subject to } \boldsymbol{y}^T\boldsymbol{D}_1\boldsymbol{y}+2\boldsymbol{f}_1^T\boldsymbol{y}=0 \\ &\qquad\qquad \boldsymbol{y}^T\boldsymbol{D}_2\boldsymbol{y}+2\boldsymbol{f}_2^T\boldsymbol{y}=0 \end{aligned} \tag{17}$$

where matrices $\boldsymbol{W}$, $\boldsymbol{A}$, $\boldsymbol{D}_1$, and $\boldsymbol{D}_2$ and vectors $\boldsymbol{b}$, $\boldsymbol{f}_1$, $\boldsymbol{f}_2$, and $\boldsymbol{y}$ are defined as

$$\begin{aligned} &\boldsymbol{W} = \boldsymbol{I}_K \otimes \operatorname{diag}\left(\frac{1}{d_1^2(\alpha^2\sigma_1^2+\gamma_1^2)},\ldots,\frac{1}{d_N^2(\alpha^2\sigma_N^2+\gamma_N^2)}\right), \\ &\boldsymbol{A} \triangleq \begin{bmatrix} \frac{1}{c^2} & -\frac{2}{c^2}\boldsymbol{a}_1^T & -(z_1^1)^2 & z_1^1\hat{T}_1^1 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{c^2} & -\frac{2}{c^2}\boldsymbol{a}_N^T & -(z_N^1)^2 & z_N^1\hat{T}_N^1 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{c^2} & -\frac{2}{c^2}\boldsymbol{a}_1^T & -(z_1^K)^2 & z_1^K\hat{T}_1^K \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{c^2} & -\frac{2}{c^2}\boldsymbol{a}_N^T & -(z_N^K)^2 & z_N^K\hat{T}_N^K \end{bmatrix}, \\ &\mathbf{b} \triangleq \begin{bmatrix} -\frac{1}{c^2}\|\boldsymbol{a}_1\|_2^2+\frac{(\hat{T}_1^1)^2}{4} \\ \vdots \\ -\frac{1}{c^2}\|\boldsymbol{a}_N\|_2^2+\frac{(\hat{T}_N^1)^2}{4} \\ \vdots \\ -\frac{1}{c^2}\|\boldsymbol{a}_1\|_2^2+\frac{(\hat{T}_1^K)^2}{4} \\ \vdots \\ -\frac{1}{c^2}\|\boldsymbol{a}_N\|_2^2+\frac{(\hat{T}_N^K)^2}{4} \end{bmatrix}, \\ &\boldsymbol{D}_1 \triangleq \operatorname{diag}(0,1,1,0,0), \\ &\boldsymbol{f}_1 \triangleq \begin{bmatrix} -\frac{1}{2} & 0 & 0 & 0 & 0 \end{bmatrix}^T \\ &\boldsymbol{D}_2 \triangleq \operatorname{diag}(0,0,0,0,1) \\ &\boldsymbol{f}_2 \triangleq \begin{bmatrix} 0 & 0\,0 & -\frac{1}{2} & 0 \end{bmatrix}^T \\ &\boldsymbol{y} \triangleq \begin{bmatrix} \|\boldsymbol{x}\|_2^2 & \boldsymbol{x}^T & \alpha^2 & \alpha \end{bmatrix}^T. \end{aligned} \tag{18}$$

Note that since the weighting matrix $\boldsymbol{W}$ depends on the unknown distance d_i and α, we first replace $\boldsymbol{W}$ with the identity matrix and find an estimate of the location and α as described above. Then, we reconstruct the distance considering the estimate $\hat{\boldsymbol{x}}$ as $\hat{d}_i = \|\hat{\boldsymbol{x}}-\boldsymbol{a}_i\|_2$ and form a new approximate weighting

matrix. This approach can be continued for a number of iterations; however, as we have observed through simulations, after two updates, the estimation accuracy improves only slightly via additional iterations. For i.i.d. measurement errors, $\sigma_i = \sigma$ and $\gamma_i = \gamma$, the weighing matrix will be simplified and only be dependent on distances.

The constraints in (17) equivalently express two constraints on elements of $\boldsymbol{y}$, i.e., $\boldsymbol{x}^T\boldsymbol{x} = \|\boldsymbol{x}\|_2^2$ and $\alpha\alpha = \alpha^2$. The problem in (17) minimizes a quadratic function over two quadratic constraints. This type of problems is called the extended trust region problem (EGTR) or two trust region problem and is generally difficult to solve [42, 43]. For special cases, the EGTR problem can be exactly solved [44]. In the previous work, by dropping the second constraint in (17), the problem was formulated as a trust region subproblem (GTR) [35] that can be solved under mild conditions [45]. It has also been known that the GTR has zero duality gap and the optimal solution can be extracted from the dual solution [46, 44, 45]. However, in this study, we consider both constraints and propose an algorithm to approximately solve the problem in (17) by modifying the GTR approach in [35]. The performance of the proposed approach requires further investigations in future work, but as investigated through the simulations, the algorithm provides good performance in various situations. The proposed approach relies on a fact about the structure of the problem and tries to adjust iterations toward a reasonable solution. To this end, we first omit the second constraint and consider a GTR similar to [35]. For GTR, a necessary and sufficient condition for $\boldsymbol{y}^*$ to be optimal in (17) is that there exists a $\mu \in \mathbb{R}$ such that [46]

$$\begin{aligned}
&(\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{A} + \mu\boldsymbol{D}_1)\boldsymbol{y}^* = (\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{b} - \mu\boldsymbol{f}_1),\\
&(\boldsymbol{y}^*)^T\boldsymbol{D}_1\boldsymbol{y}^* + 2\boldsymbol{f}_1^T\boldsymbol{y}^* = 0,\\
&(\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{A} + \mu\boldsymbol{D}_1) \succeq 0.
\end{aligned} \tag{19}$$

Under the conditions considered in (19), the solution to the problem of (17) is given by

$$\boldsymbol{y}(\mu) = (\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{A} + \mu\boldsymbol{D}_1)^{-1}(\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{b} - \mu\boldsymbol{f}_1). \tag{20}$$

In such a situation to find μ, we simply replace (20) into constraint $\boldsymbol{y}^T\boldsymbol{D}_1\boldsymbol{y} + 2\boldsymbol{f}_1^T\boldsymbol{y} = 0$, i.e.,

$$\phi(\mu) = \boldsymbol{y}^T(\mu)\boldsymbol{D}_1\boldsymbol{y}^T(\mu) + 2\boldsymbol{f}_1^T\boldsymbol{y}(\mu) = 0, \quad \mu \in \mathcal{I} \tag{21}$$

where the interval $\mathcal{I}$ consists of all μ such that $\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{A} + \mu\boldsymbol{D}_1 \succeq 0$. The interval $\mathcal{I}$ is given by [10]

$$\mathcal{I} = (-1/\mu_1, \infty), \tag{22}$$

with μ_1 representing the largest eigenvalue of $(\mathbf{A}^T\boldsymbol{W}\mathbf{A})^{-1/2}\boldsymbol{D}_1(\mathbf{A}^T\boldsymbol{W}\mathbf{A})^{-1/2}$ [45]. Next in order to force the solution to satisfy the second constraint, we check the last two components of $\boldsymbol{y}(\mu)$ to see if they

are close to each other. If not, we replace the values of $[\boldsymbol{y}(\mu)]_4$ and $[\boldsymbol{y}(\mu)]_5$ by their average, $[\boldsymbol{y}(\mu)]_4 = ([\boldsymbol{y}(\mu)]_4 + ([\boldsymbol{y}(\mu)]_5)^2)/2$ and $[\boldsymbol{y}(\mu)]_5 = \sqrt{[\boldsymbol{y}(\mu)]_4}$.

In summary, the suggested algorithm to (approximately) solve the problem in (17) is expressed as

- Use a bisection search to find a root of $\phi(\mu) = 0$, say μ^*. Note that $\phi(\mu)$ is a strictly decreasing function with respect to μ [45]. In every step of the bisection search, if $|[\boldsymbol{y}(\mu)]_4 - [\boldsymbol{y}(\mu)]_5| \geq \varsigma$ (ς is a predetermined small value), then replace $[\boldsymbol{y}(\mu)]_4$ by $([\boldsymbol{y}(\mu)]_4 + ([\boldsymbol{y}(\mu)]_5)^2)/2$ and $[\boldsymbol{y}(\mu)]_5$ by $\sqrt{[\boldsymbol{y}(\mu)]_4}$.
- Replace μ^* into (20) to obtain $\boldsymbol{y}^* = \boldsymbol{y}(\mu^*)$.
- Estimate the unknown parameters as $\hat{\boldsymbol{x}} = [\boldsymbol{y}^*]_{2:3}$ and $\hat{w} = 1/[\boldsymbol{y}^*]_4$, with $[\boldsymbol{v}]_{i:j}$ denoting the ith to the jth elements of vector $\boldsymbol{v}$.

The details of the algorithm are shown in Algorithm 1.

Algorithm 1 ***EGTR***

Initialization: $\lambda_1 = -1/\mu_1$ and $\lambda_2 = 1/\mu_1$ and set values of ς and s
for $k = 0$ until convergence or predefined number K **do**
  Compute $\boldsymbol{y}(\lambda_i),\ i = 1, 2$ from (20)
  If $|[\boldsymbol{y}(\lambda_i)]_4 - [\boldsymbol{y}(\lambda_i)]_5| \geq \varsigma$ (ς is a predetermined small value), then replace $[\boldsymbol{y}(\lambda_i)]_4$ by $([\boldsymbol{y}(\lambda_i)]_4 + ([\boldsymbol{y}(\lambda_i)]_5)^2)/2$ and $[\boldsymbol{y}(\lambda_i)]_5$ by $\sqrt{[\boldsymbol{y}(\lambda_i)]_4}$.
  Compute $\phi(\lambda_i),\ i = 1, 2$ from (21)
  if $\phi(\mu_1)\phi(\mu_1) > 0$ **then**
    $\lambda_1 = \lambda_2$ and $\lambda_2 = s\lambda_2$
  else
    $\lambda' = (\lambda_1 + \lambda_2)/2$
    compute $\phi(\lambda')$
    if $\phi(\lambda')>0$ **then**
      $\lambda_1 = \lambda'$
    else
      $\lambda_2 = \lambda'$
    end if
  end if
end for

The main difference between the GTR approach in [35] and the EGTR in this work is in the step four of Algorithm 1. Namely, this step is not present in our previous work. In the simulation section, we compare

the performance of the EGTR in this study and the GTR in [35] and observe that the proposed approach generally outperforms the previous algorithm.

Remark 1. *Yet another approach to approximately solve the problem in* (17) *is to break the problem into two GTR problems. That is, we consider two sperate GTRs with the same objective functions but different constraints, one with the first constraint and the other with the second constraint. We then solve GTRs in parallel, but for every iteration, we make sure both solutions are close to each other. That is, we force both solutions to agree on their components. We will not investigate this technique since it is more complex than the approach proposed above.*

Another estimator based on a linear least squares (LLS) approach obtained in Appendix 8.1 can be alternatively applied to the model in (15). Note that the algorithm derived in Appendix 8.1 is similar to the one proposed in [29], except the correction technique introduced in this study. As will be observed in the simulations section the proposed approach in this section, i.e., EGTR, has better performance than the LLS approach, especially for low number of reference nodes.

4.2. A concave-convex procedure (CCCP)

In this section, we take the ℓ_1 norm minimization of residual errors into account and propose a technique to solve the positioning problem. Namely, based on (14), we consider the following ℓ_1 norm minimization problem:

$$\underset{\boldsymbol{x}\in\mathbb{R}^2;\ \alpha\in\mathbb{R}_+}{\text{minimize}} \quad \|\boldsymbol{r}\|_1 \tag{23}$$

where $\boldsymbol{r} = [r_1^1 \dots r_N^1 \dots r_1^K \dots r_N^K]^T$ with $r_i^k = z_i^k\alpha - \hat{T}_i^k/2 - d_i/c$. Note that for high signal-to-noise ratios (low standard deviations of noise), the ℓ_2 and ℓ_1 minimization approaches have similar performance [47]. Moreover, the ℓ_1 norm minimization in (23) is robust against outliers [47]. Outliers in the positioning process may arise due to various reasons; e.g., blockage of the direct path can lead to outliers in some situations. The optimization problem in (23) can be written (in the epigraph form) as [47, 36, 48]

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^2;\alpha\in\mathbb{R}_+;\boldsymbol{t}\in\mathbb{R}_+^N}{\text{minimize}} \quad & \sum_{k=1}^{K}\sum_{i=1}^{N} t_i^k \\ \text{subject to} \quad & z_i^k\alpha - \hat{T}_i^k/2 - d_i/c \le t_i^k \\ & z_i^k\alpha - \hat{T}_i^k/2 - d_i/c \ge -t_i^k. \end{aligned} \tag{24}$$

The nonconvex problem in (24) is reminiscent of a well-known nonconvex problem called difference of convex functions programming (DCP) [49]. The general form of DCP is as follows:

$$\begin{aligned} &\underset{\boldsymbol{x}}{\text{minimize}} \;\; f_0(\boldsymbol{x}) - g_0(\boldsymbol{x}) \\ &\text{subject to } w_i(\boldsymbol{x}) - g_i(\boldsymbol{x}) \leq 0, \;\; i = 1, \ldots, M \end{aligned} \tag{25}$$

where $f_0(\boldsymbol{x})$, $w_i(\boldsymbol{x})$, and $g_i(\boldsymbol{x})$ are smooth convex functions for $i = 1, \ldots, M$. A method to solve (25) is to sequentially solve the problem. That is, we first approximate the concave function $(-g_i(\boldsymbol{x}))$ with a convex one by an affine approximation. Let us consider a point $\boldsymbol{x}^j$ in the domain of the problem in (25), linearize the concave function around $\boldsymbol{x}^j$ and write the optimization problem in (25) as

$$\begin{aligned} &\underset{\boldsymbol{x}}{\text{minimize}} \;\; f_0(\boldsymbol{x}) - g_0(\boldsymbol{x}^j) - \bigtriangledown g_0(\boldsymbol{x}^j)^T(\boldsymbol{x} - \boldsymbol{x}^j) \\ &\text{subject to} \;\; w_i(\boldsymbol{x}) - g_i(\boldsymbol{x}^j) - \bigtriangledown g_i(\boldsymbol{x}^j)^T(\boldsymbol{x} - \boldsymbol{x}^j) \leq 0. \end{aligned} \tag{26}$$

The convex problem in (26) can now be solved efficiently. Denote the solution of (26) as $\boldsymbol{x}^{j+1}$. Next we go for further improving the solution by convexifing (25) for the new point $\boldsymbol{x}^{j+1}$ similar to the procedure employed for $\boldsymbol{x}^j$. This sequential programming procedure, called concave-convex programming (CCCP), continues for a number of iterations. The convergence of the CCCP to a stationary point has been shown in the literature, e.g., [49, 50] and references therein.

Applying the CCCP technique to the problem in (24), we get the following optimization problem:

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^2;\alpha\in\mathbb{R}_+;\boldsymbol{t}\in\mathbb{R}_+^N}{\text{minimize}} \;\; & \sum_{k=1}^{K}\sum_{i=1}^{N} t_i^k \\ \text{subject to} \;\; & z_i^k\alpha - \boldsymbol{h}_{i,j}^T\boldsymbol{x} - b_{i,k}^j - t_i^k \leq 0 \\ & \frac{1}{c}\|\boldsymbol{x} - \boldsymbol{a}_i\|_2 - z_i^k\alpha + \frac{\hat{T}_i^k}{2} - t_i^k \leq 0 \end{aligned} \tag{27}$$

where

$$\begin{aligned} \boldsymbol{h}_{i,j} &= (\boldsymbol{x}^j - \boldsymbol{a}_i)/(cd(\boldsymbol{a}_i, \boldsymbol{x}^j)) \\ b_{i,k}^j &= \hat{T}_i^k/2 - \boldsymbol{h}_{i,j}^T\boldsymbol{x}^j + d(\boldsymbol{a}_i, \boldsymbol{x}^j)/c. \end{aligned} \tag{28}$$

The optimization problem in (27), which is called second order cone programming (SOCP), can be efficiently solved. We call the corresponding CCCP as CCCP-SOCP. Algorithm 2 shows a high level implementation of the algorithm.

Algorithm 2 ***CCCP-SOCP***

1: Initialization: choose initial value for $\boldsymbol{x}^0$

2: **for** $j = 0$ until convergence or predefined number J **do**

3: Compute $\boldsymbol{h}_{i,j}$ and $b_{i,k}^j$ from (28)

4: Solve problem (27) and denote the solution $\boldsymbol{x}^o$

5: set $\boldsymbol{x}^{j+1} = \boldsymbol{x}^o$

6: **end for**

5. Complexity analysis

In this section, we study the complexity of the proposed techniques in terms of floating point operations (*flops*) and also running time in Matlab. We compare the complexity of the MLE, LLS, EGTR, and CCCP-SOCP. To compute the complexity of the MLE, we assume that a good initial point is available, and an iterative algorithm such as the Gauss-Newton (GN) method converges to the global minimum after a number of iterations. Of course, finding a good initial point for the MLE is a challenging problem and this study also aims to tackle it. For the problem at hand the complexity of the MLE for every Newton step can be computed as $O(K^2N^3))$. For the AMLE, the complexity for every Newton step can be computed as $O((KN)^2)$. The corresponding LLS needs an order of $O(KN)$ to implement. For the EGTR, we need to use a bisection search to solve (21), which is the most complex part of the algorithm. Suppose the bisection search takes k_{sq} steps, then the total cost of the the proposed approach can be approximated as $O(k_{\text{sq}}KN)$. In the simulations, we have observed that the bisection search algorithm usually takes 20 to 30 iterations to find the optimal value of γ. Note that we need to run the LLS and EGTR twice. Thus the corresponding complexities are increased by a factor of two. Finally, the complexity of the CCCP-SOCP can be computed as follows. Consider a general form of the SOCP problem as

$$\begin{aligned} &\underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} \ \boldsymbol{c}^T\boldsymbol{x} \\ &\text{subject to } \|\boldsymbol{A}_i\boldsymbol{x}+\boldsymbol{b}_i\|_2 \le \boldsymbol{c}_i^T\boldsymbol{x}+d_i,\ i=1,\ldots,m, \\ &\qquad\qquad \|\boldsymbol{x}\|_2 \le R \end{aligned} \tag{29}$$

where $\boldsymbol{A}_i \in \mathbb{R}^{k_i\times n}, \boldsymbol{b}_i \in \mathbb{R}^{k_i}$, and $d_i \in \mathbb{R}$. Note that the constraint on the norm of $\boldsymbol{x}$ ensures the strong convexity of the centering problem in the barrier approach [47]. The worst-case complexity of the problem in (29) can be computed as $O((1+m)^{1/2}n(n^2+m+\sum_{i=1}^m k_i^2)\log(1/\epsilon))$ [51], where ϵ is an accuracy tolerance in solving the problem.

The complexity of the CCCP-SOCP for every estimate $\boldsymbol{x}^j$ can now be approximated as

$$O((KN)^{3.5}\log(1/\epsilon)).$$

As mentioned before we need to solve the problem in k_{cccp} steps, hence the total cost is $O(k_{\text{cccp}}(KN)^{3.5}\log(1/\epsilon))$. As we observe, a small number of updatings, usually three, $k_{\text{cccp}} = 3$, is enough to obtain the solution. Table 1 summarizes the complexity of the different approaches.

We have also measured the average running time of different algorithms for a network consisting of 6 reference nodes as considered in Section 6. In the simulations, we set $K = 2$ and $\sigma_i = \gamma_i = 10$. The algorithms have been implemented in Matlab on a MacBook Pro (Processor 2.3 GHz Intel Core i7, Memory 8 GB 1600 MHz DDR3). The MLEs are implemented by Matlab function *fminsearch* initialized with the true values of the target position, the clock skew, and turn-around times. It is noted that the function *fminsearch* is based on the Nelder-Mead simplex algorithms, which does not compute gradients or Hessians to find; hence, its complexity might not be the same as the complexity of Newton-type algorithms. The CCCP-SOCP is implemented by the *CVX* toolbox [52]. We use three updatings to get an estimate.

We run the algorithms for 500 realizations of the network and compute the average running time in ms. The results are shown in Table 2. Considering the complexity analysis and average running time in Tables 1 and 2, respectively, we can conclude that the proposed approach has reasonable complexity and running time. Although CCCP-SOCP takes a longer amount of time than MLE, it does not need a good initial point. While for the MLE with an arbitrary initial point, the algorithm may converge to a local minimum resulting in a large position error. As we will see in the next section, the CCCP-SOCP outperforms both the LLS and EGTR approaches in terms of the root-mean-squared-error.

6. Numerical results

In this section, we evaluate the performance of the proposed approaches through computer simulations. We consider a 1600 m by 1600 m area and a number of reference nodes that are located at fixed positions $\boldsymbol{a}_1 = [800\ 800]$, $\boldsymbol{a}_2 = [800\ -800]$, $\boldsymbol{a}_3 = [-800\ 800]$, $\boldsymbol{a}_4 = [-800\ -800]$, $\boldsymbol{a}_5 = [800\ 0]$, $\boldsymbol{a}_6 = [0\ 800]$, $\boldsymbol{a}_7 = [-800\ 0]$, and $\boldsymbol{a}_8 = [0\ -800]$. In the simulations, we pick the first N reference nodes, i.e., $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N$. One target node is randomly distributed inside the area. The turn-around time is set to 0.001 ms. The clock skew is assumed to be unknown and is set to 100 PPM, i.e., $w = 1.0001$. Such a value for clock skew is common for a practical oscillator. For example for a center carrier frequency $f_c = 100$ MHz and a frequency offset $\Delta f = 10$ kHz, the actual frequency is given by $f_c \pm \Delta f$. Therefore, the period of the oscillating signal

(versus the nominal period $T_0 = 1/f_c$) is given by

$$T = \frac{1}{f_c \mp \Delta f} \approx \frac{1}{f_c}(1 \pm \frac{\Delta f}{f_c}) = T_o(1 \pm 0.0001) \tag{30}$$

which shows how the clock of the oscillator is scaled with respect to the nominal clock T_o.

To compare different approaches, we use the root-mean-squared-errors (RMSEs) defined as

$$\text{RMSE} \triangleq \sqrt{\mathbb{E}\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|_2^2}. \tag{31}$$

We compare the proposed techniques (CCCP-SOCP, which is implemented using CVX, and EGTR) with the MLE and AMLE in (11) and (13), respectively, (which are implemented by Matlab function *fminsearch* initialized with the true values of the target location, turn-around times, and clock skew), the GTR in [35], the LLS derived in Appendix 8.1, and the Cramér-Rao lower bound (CRLB) as derived in Appendix 8.2. In the simulations, we assume that $\sigma_i = \gamma_i = \sigma,\ i = 1, \ldots, N$. We randomly initialize the CCCP-SOCP inside the network and we also set $k_{\text{cccp}} = 3$. We use 2000 Monte-Carlo simulations to generate the results. To simulate the range measurements and estimates of turn around times, we use models (4) and (2), respectively. To implement the bisection search, we consider an interval defined by I_{lower} and I_{upper} and investigate if the zero crossing of $\phi(\mu)$ in (21) occurs in the interval. To check if the solution lies in the interval, we simply check the sign of $\phi(\mu)$ at I_{lower} and I_{upper}. No change in sign means that the solution lies outside of the current interval. For initialization, we set $I_{\text{lower}} = -1/\mu_1$ and $I_{\text{upper}} = 1/\mu_1$. If the solution of $\phi(\mu) = 0$ is not found in the interval, we change the interval as $I'_{\text{lower}} = I_{\text{upper}}$ and $I'_{\text{upper}} = 10I_{\text{upper}}$ ($s = 10$ in Algorithm 1). If the solution lies in an interval, we bisect the interval and investigate which subinterval contains the solution. We also set $\varsigma = 0.05$.

Fig. 1 shows the RMSEs of location estimates for different approaches for various numbers of reference nodes. In the simulations, we set $K = 2$. It is observed that the proposed approach, CCCP-SOCP, achieves good performance very close to the optimal estimator MLE and the CRLB, especially for low number of reference nodes and high signal-to-noise ratios. From the figure, it is noted that the EGTR proposed in this study generally outperforms the previous GTR and LLS, especially for low numbers of reference nodes. As the number of reference nodes increases, the LLS, GTR, and EGTR show similar performance. We have observed a similar behavior for other network deployments. In general, EGTR provides more robust and accurate estimates than GTR, especially for small number of measurements (small K or N). It is also observed that for large numbers of reference nodes, the least squares based approach shows better performance compared to the CCCP-SOCP approach for the low standard deviation of noise. The reason is that for low measurement errors, the squared term $(v_i^k)^2$ is negligible and thus the approximation in (15) is

more likely to be valid. In addition for a larger number of reference nodes, matrix $\boldsymbol{A}$ will be well-conditioned and thus numerical roundoff errors will decrease.

Next, we study the effects of NLOS measurements on the performance of estimators. We assume that a range measurement can be affected by NLOS errors with probability 0.2. For every NLOS measurement, we add a uniform noise to the measurements as follows:

$$z_i^k = w\left(\frac{d_i}{c} + \frac{T_i}{2}\right) + q_i u_i^k + \frac{n_i^k}{2} \tag{32}$$

where we assume that $u_i^k \sim \mathcal{U}[0, 5/c]$ and $q_i \in \{0, 1\}$ are iid Bernoulli random variables with $\Pr\{q_i = 1\} = 0.2$. The uniform distribution is commonly used to model NLOS error, e.g., [53, 54, 12].

Fig. 2 depicts the performance of different approaches in NLOS conditions for $K = 2$. It is observed that the CCCP-SOCP achieves high performance compared to the other approaches, especially for low standard deviations of noise, and it is robust against outliers as expected. For small σ, the dominant perturbation is outlier disturbance and consequently the MLE derived in this study is not optimal, explaining why the MLE is worse than the CCCP-SOCP approach. For large standard deviations of noise, which indicates the Gaussian measurement noise is dominant, the CCCP-SOCP seems to outperform the MLE. This can be explained by the fact that the MLE is only guaranteed to be asymptotically optimal, i.e., for low noise standard deviation or large number of measurements. Note that we have employed the MLE computed in (11) and (13) to study their robustness against NLOS conditions. It may be possible to derive an MLE to deal with NLOS measurements if the distribution of outliers is known. From the figure, it is observed that the proposed EGTR outperforms GTR and LLS, especially for low numbers of reference nodes. In general, for a fixed network, the performance of algorithms is affected by two perturbations: measurement noise and NLOS errors. As the measurement error becomes smaller, the performance is mainly affected by NLOS errors. Since NLOS statistics are fixed in the simulation, we expect a kind of flat behavior for RMSE.

We now study the convergence of CCCP-SOCP through simulations. Fig. 3 depicts the convergence speed of the proposed approach for 50 random initializations. In the simulations, we set $K = 2$. For every estimate given by CCCP-SOCP, we compute the residual $\|\boldsymbol{r}\|_1$, where $\boldsymbol{r}$ is given by (23). It is observed that the CCCP-SOCP approach converges very fast, approximately in three sequential updatings.

Finally we briefly compare the performance of EGTR with GTR-based algorithm without considering the clock skew. The signal model of (2) can be expressed as

$$z_i^k = \frac{d_i}{c} + \frac{T_i}{2} + \frac{n_i^k}{2} + \underbrace{\rho(\frac{d_i}{c} + \frac{T_i}{2})}_{\triangleq b_i}, \quad k = 1, 2, \ldots, K \tag{33}$$

where ρ is the clock skew deviation from one, i.e., $w = 1 + \rho$. It is observed that imperfect clock skew

causes a bias b_i in the measurement compared to the ideal scenario, i.,e., for $\rho = 0$. For very small b_i, estimating the clock skew parameter along with the location estimate may result in large location errors. However, when the bias is considerable compared to the measurement error $\frac{n_i^k}{2}$, the joint estimation leads to improve accuracy. Considering the model in (33), we may need to take into account different issues to design an algorithm. For example, since the mean of the perturbation is nonzero, we may need to estimate and subtract the mean from the measurement for designing an algorithm such as least squares. Here, we employ a localization algorithm proposed for the ideal scenario when the model actually comes form (33). In particular, the proposed GTR algorithm for ideal clock is compared with EGTR for different values of ρ.

Fig. 4 shows the RMSE of EGTR and GTR without clock skew consideration. It is observed that the joint estimation of the location and clock skew for high SNR improves the accuracy of localization. For low SNRs, the accuracy mainly depends on the variance of measurement noise. It is also observed that as the clock skew deviation from one increases, the performance of the traditional approach without considering the clock parameter degrades drastically, especially for high SNRs. The performance of EGTR remains almost the same as ρ changes. This figure also shows an improved performance for EGTR compared to GTR. In fact, it can be observed that EGTR is more robust than GTR.

7. Conclusions

In this manuscript, TW-TOA based positioning has been studied in a semi-asynchronous network in which the clock of the target node is not synchronized with a perfect clock. Since the optimal ML estimator is highly nonconvex and difficult to solve, two efficient suboptimal estimators have been obtained for the problem under some approximations and conditions. The first method is based on the squared-range least squares that is formulated as an extended general trust region subproblem (EGTR). A simple approach has been proposed to solve EGTR. The second approach is derived by replacing the ℓ_2 norm minimization of residuals by an ℓ_1 norm minimization, which in turn can be formulated as difference of convex programming (DCP). A concave-convex procedure has been employed to solve the resulting DCP. Simulation results show the high performance of the proposed techniques, especially the DCP approach. It has also been observed through simulations that the DCP approach is robust against NLOS errors. The future work considers the extension of the approaches studied in this manuscript to cooperative scenarios. Simulation results show a promising performance for EGTR, in terms of robustness and accuracy. Although we have not encountered any convergence problems of the proposed EGTR approach, an extensive study of the convergence properties of EGTR is left as a topic for future work.

8. Appendices

8.1. Linear Least Squares (LLS)

In this section we obtain an LLS estimator similar to [8, 28]. We consider the following linear model (originated from (15)):

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{\nu}, \tag{34}$$

where $\boldsymbol{\nu} = [\nu_1^1 \ \dots \ \nu_N^1 \dots \nu_1^K \ \dots \ \nu_N^K]^T$, $\boldsymbol{A}$, $\boldsymbol{b}$, and $\boldsymbol{y}$ are given in (18). We assume that $\boldsymbol{A}$ has full column rank. A necessary condition for this is that $KN \geq 5$.

The unconstrained weighted least squares solution to (34) is given by [41]

$$\hat{\boldsymbol{y}} = (\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{b}. \tag{35}$$

where $\boldsymbol{W}$ is as in (18). The covariance matrix of $\hat{\boldsymbol{y}}$ can be computed as

$$\boldsymbol{C}_{\hat{\boldsymbol{y}}} = (\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{A})^{-1}. \tag{36}$$

Note that for a large network, matrix $\boldsymbol{A}$ can be ill-conditioned [28]. Then, we can use a regularization technique to resolve the drawback in the least squares solution [47, 28].

We can further improve the location estimate by applying a correction technique similar to [8, 28]. We consider the following relations:

$$\begin{aligned}
[\boldsymbol{y}]_1 &= \|x\|_2^2 + \xi_1, \\
[\boldsymbol{y}]_4 &= \alpha^2 + \xi_4, \\
[\boldsymbol{y}]_2 &= x_1 + \xi_2, \\
[\boldsymbol{y}]_3 &= x_2 + \xi_3, \\
[\boldsymbol{y}]_5 &= \alpha + \xi_5,
\end{aligned} \tag{37}$$

where $\boldsymbol{\xi} = [\xi_1 \ \dots \ \xi_5]^T$ is the estimation error. Assuming small estimation errors, we take the squares of both sides of last three equations in (37) and obtain the following expressions:

$$\begin{aligned}
[\boldsymbol{y}]_2^2 &\simeq x_1^2 + 2x_1\xi_2, \\
[\boldsymbol{y}]_3^2 &\simeq x_2^2 + 2x_2\xi_3, \\
[\boldsymbol{y}]_5^2 &\simeq \alpha^2 + 2\alpha\xi_5.
\end{aligned} \tag{38}$$

Based on (37) and (38), we obtain a linear model as

$$\boldsymbol{h} = \boldsymbol{B}\boldsymbol{\theta} + \boldsymbol{P}\boldsymbol{\xi}, \tag{39}$$

where

$$\boldsymbol{B} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 2x_1 & 0 & 0 & 0 \\ 0 & 0 & 2x_2 & 0 & 0 \\ 0 & 0 & 0 & 2\alpha & 0 \end{bmatrix}$$

$$\boldsymbol{h} = \begin{bmatrix} [\boldsymbol{y}]_1 + [\boldsymbol{y}]_4 \\ [\boldsymbol{y}]_2^2 \\ [\boldsymbol{y}]_3^2 \\ [\boldsymbol{y}]_5^2 \end{bmatrix}, \quad \boldsymbol{\theta} = [x_1^2 \; x_2^2 \; \alpha^2]^T. \tag{40}$$

The least squares solution to (39) is given by

$$\hat{\boldsymbol{\theta}} = (\boldsymbol{B}^T \boldsymbol{C}_{\hat{\boldsymbol{\theta}}}^{-1} \boldsymbol{B})^{-1} \boldsymbol{B}^T \boldsymbol{C}_{\hat{\boldsymbol{\theta}}}^{-1} \boldsymbol{h}, \tag{41}$$

where the covariance matrix $\boldsymbol{C}_{\boldsymbol{\theta}}$ can be computed as

$$\boldsymbol{C}_{\hat{\boldsymbol{\theta}}} = \boldsymbol{P} \boldsymbol{C}_{\hat{\boldsymbol{y}}} \boldsymbol{P}^T. \tag{42}$$

To compute the matrix $\boldsymbol{P}$, we use the estimate of $\hat{\boldsymbol{x}}$ obtained in (35) instead of unknown vector $\boldsymbol{x}$.

Finally the location estimate can be obtained as

$$\tilde{x}_i = \text{sgn}([\boldsymbol{y}]_{i+1}) \sqrt{|[\hat{\boldsymbol{\theta}}]_i|}, \quad i = 1, 2, \tag{43}$$

where sgn denotes the signum function defined as

$$\text{sgn}(x) = \begin{cases} 1 & \text{if } x \geq 0; \\ -1 & \text{if } x < 0. \end{cases} \tag{44}$$

The covariance matrix of the estimate in (43) can be obtained similar to [28].

8.2. Cramér-Rao Lower Bound (CRLB)

Considering the measurement vector in (7) with mean $\boldsymbol{\mu}_K = \boldsymbol{1}_K \otimes \boldsymbol{\mu}$ and covariance matrix $\boldsymbol{C}_K = \boldsymbol{I}_K \otimes \boldsymbol{C}$ where

$$\begin{aligned} \boldsymbol{\mu} &= \left[f\left(\frac{d_1}{c} + \frac{T_1}{2}\right) \dots f\left(\frac{d_N}{c} + \frac{T_N}{2}\right) \; T_1 \dots T_N \right]^T, \\ \boldsymbol{C} &= \text{diag}\left(\frac{\sigma_1^2}{4}, \dots, \frac{\sigma_N^2}{4}, \gamma_1^2, \dots, \gamma_N^2\right), \end{aligned} \tag{45}$$

the elements of the Fisher information matrix can be computed as [41, Ch. 3]

$$J_{nm} = [\boldsymbol{J}]_{nm} = \left[\frac{\partial \boldsymbol{\mu}_K}{\partial \psi_n}\right]^T \boldsymbol{C}_K^{-1} \left[\frac{\partial \boldsymbol{\mu}_K}{\partial \psi_m}\right], \quad n, m = 1, 2, \dots, N+3, \tag{46}$$

where

$$\psi_n = \begin{cases} x_n, & \text{if } n = 1, 2 \\ w, & \text{if } n = 3 \\ T_n, & \text{if } n > 3. \end{cases} \tag{47}$$

From (45), $\partial \boldsymbol{\mu}_K / \partial \psi_n$ can be obtained as follows:

$$\left[\frac{\partial \boldsymbol{\mu}_K}{\partial \psi_n}\right] = \mathbf{1}_K \otimes \left[\frac{\partial \mu_1}{\partial \psi_n} \ \cdots \ \frac{\partial \mu_N}{\partial \psi_n}\right]^T, \quad n = 1, 2, \ldots, N+3, \tag{48}$$

where

$$\frac{\partial \mu_i}{\partial \psi_n} = \begin{cases} w \frac{x_n - a_{1,n}}{c\, d(\boldsymbol{a}_i, \boldsymbol{x})}, & \text{if } n = 1, 2, \ i \leq N \\ \frac{d_i}{c} + \frac{T_i}{2}, & \text{if } n = 3, \ i \leq N \\ 0, & \text{if } n = 1, 2, \text{or } 3, \ i > N \\ \frac{w}{2}, & \text{if } n > 3, \ i \leq N \\ 1, & \text{if } n > 3, \ i > N. \end{cases} \tag{49}$$

After some calculations, the entries of the Fisher information matrix can be computed as follows:

$$
\begin{aligned}
J_{11} &= 4Kw^2 \sum_{i=1}^{N} \left(\frac{x_1 - a_{i,1}}{\sigma_i c d(\boldsymbol{a}_i, \boldsymbol{x})} \right)^2 , \\
J_{22} &= 4Kw^2 \sum_{i=1}^{N} \left(\frac{x_2 - a_{i,2}}{\sigma_i c d(\boldsymbol{a}_i, \boldsymbol{x})} \right)^2 , \\
J_{33} &= 4K \sum_{i=1}^{N} \left(\frac{d_i/c + T_i/2}{\sigma_i} \right)^2 , \\
J_{jj} &= K \left(\frac{2w^2}{\sigma_j^2} + \frac{1}{\gamma_j^2} \right), \ j > 3 \\
J_{12} &= J_{21} = 4Kw^2 \sum_{i=1}^{N} \left(\frac{x_1 - a_{i,1}}{\sigma_i c\, d(\boldsymbol{a}_i, \boldsymbol{x})} \right) \left(\frac{x_2 - a_{i,2}}{\sigma_i c\, d(\boldsymbol{a}_i, \boldsymbol{x})} \right), \\
J_{13} &= 4Kw \sum_{i=1}^{N} \left(\frac{x_1 - a_{i,1}}{\sigma_i c d(\boldsymbol{a}_i, \boldsymbol{x})} \right) \left(\frac{d_i/c + T_i/2}{\sigma_i} \right), \\
J_{23} &= 4Kw \sum_{i=1}^{N} \left(\frac{x_2 - a_{i,2}}{\sigma_i c d(\boldsymbol{a}_i, \boldsymbol{x})} \right) \left(\frac{d_i/c + T_i/2}{\sigma_i} \right), \\
J_{j1} &= J_{1j} = K \left(w \frac{x_1 - a_{i,1}}{\sigma_i^2 c d(\boldsymbol{a}_i, \boldsymbol{x})} \right), \\
J_{j2} &= J_{2j} = K \left(w \frac{x_2 - a_{i,2}}{\sigma_i^2 c(\boldsymbol{a}_i, \boldsymbol{x})} \right), \\
J_{j3} &= J_{3j} = 4Kw \sum_{i=1}^{N} \left(\frac{x_1 - a_{i,1}}{\sigma_i c\, d(\boldsymbol{a}_i, \boldsymbol{x})} \right) \left(\frac{d_i/c + T_i/2}{\sigma_i} \right), \\
J_{ij} &= J_{ji} = 0, \quad i \neq j, \ i, \ j > 3
\end{aligned}
\tag{50}
$$

The CRLB, which is a lower bound on the variance of any unbiased estimator, is given as

$$
\mathrm{Var}(\hat{\psi}_i) \geq [\boldsymbol{J}^{-1}]_{i,i} \, . \tag{51}
$$

9. Acknowledgment

This work was supported in part by the Swedish Research Council (contract no. 2007-6363) and in part by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMmunications #(contract no. 318306). S. Gezici's research was also supported in part by the Distinguished Young Scientist Award of Turkish Academy of Sciences (TUBA-GEBIP 2013). Part of this work was presented in IEEE ICASSP 2013 [35].

References

[1] D. Dardari, M. Luise, E. Falletti, Satellite and Terrestrial Radio Positioning Techniques: A Signal Processing Perspective, Academic Press, 2011.

[2] G. Mao, B. Fidan, Localization Algorithms and Strategies for Wireless Sensor Networks, Information Science reference, Hershey. New York, 2009.

[3] N. Patwari, J. Ash, S. Kyperountas, A. O. Hero, N. C. Correal, Locating the nodes: Cooperative localization in wireless sensor network, IEEE Signal Process. Mag. 22 (4) (2005) 54–69.

[4] S. Gezici, Z. Tian, G. B. Giannakis, H. Kobayashi, A. F. Molisch, H. V. Poor, Z. Sahinoglu, Localization via ultra-wideband radios: A look at positioning aspects for future sensor networks, IEEE Signal Process. Mag. 22 (4) (2005) 70–84.

[5] A. H. Sayed, A. Tarighat, N. Khajehnouri, Network-based wireless location: challenges faced in developing techniques for accurate wireless location information, IEEE Signal Process. Mag. 22 (4) (2005) 24–40.

[6] M. R. Gholami, Positioning algorithms for wireless sensor networks, Licentiate Thesis, Chalmers University of Technology (Mar. 2011).
URL `http://publications.lib.chalmers.se/records/fulltext/138669.pdf`

[7] X. Wang, Z. Wang, B. O'Dea, A TOA-based location algorithm reducing the errors due to non-line-of-sight (NLOS) propagation, IEEE Trans. Veh. Technol. 52 (1) (2003) 112–116.

[8] M. Sun, K. C. Ho, Successive and asymptotically efficient localization of sensor nodes in closed-form, IEEE Trans. Signal Process. 57 (11) (2009) 4522–4537.

[9] M. R. Gholami, S. Gezici, E. G. Ström, M. Rydström, Hybrid TW-TOA/TDOA positioning algorithms for cooperative wireless networks, in: Proc. IEEE International Conference on Communications (ICC), 2011.

[10] A. Beck, P. Stoica, J. Li, Exact and approximate solutions of source localization problems, IEEE Trans. Signal Process. 56 (5) (2008) 1770–1778.

[11] A. O. Hero, D. Blatt, Sensor network source localization via projection onto convex sets (POCS), in: Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 3, Philadelphia, USA, 2005, pp. 689–692.

[12] M. R. Gholami, H. Wymeersch, E. G. Ström, M. Rydström, Wireless network positioning as a convex feasibility problem, EURASIP Journal on Wireless Communications and Networking 2011.

[13] M. R. Gholami, S. Gezici, E. G. Ström, M. Rydström, A distributed positioning algorithm for cooperative active and passive sensors, in: Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 2010.

[14] P. Tseng, Second-order cone programming relaxation of sensor network localization, SIAM J. Optim. 18 (1) (2007) 156–185.

[15] P. Biswas, T.-C. Lian, T.-C. Wang, Y. Ye, Semidefinite programming based algorithms for sensor network localization, ACM Trans. Sens. Netw. 2 (2) (2006) 188–220.

[16] J. Nie, Sum of squares method for sensor network localization, Comput. Optim. Appl. 43 (2009) 151–179.

[17] E. Xu, Z. Ding, S. Dasgupta, Reduced complexity semidefinite relaxation algorithms for source localization based on time difference of arrival, IEEE Transactions on Mobile Computing 10 (9) (2011) 1276–1282.

[18] E. Xu, Z. Ding, S. Dasgupta, Source localization in wireless sensor networks from signal time-of-arrival measurements, IEEE Transactions on Signal Processing 59 (6) (2011) 2887–2897.

[19] E. Serpedin, Q. M. Chaudhari, Synchronization in wireless sensor networks: Parameter estimation, performance benchmarks and protocols, New York, NY, USA: Cambridge University Press, 2009.

[20] Q. M. Chaudhari, E. Serpedin, K. Qaraqe, Some improved and generalized estimation schemes for clock synchronization of listening nodes in wireless sensor networks, IEEE Trans. Commun. 58 (1) (2010) 63–67.

[21] S. Wanlu, E. G. Ström, F. Brännström, Long-term clock synchronization in wireless sensor networks with arbitrary delay distributions, in: IEEE Global Telecommunications Conference, 2012.

[22] Q. Chaudhari, E. Serpedin, J. Kim, Energy-efficient estimation of clock offset for inactive nodes in wireless sensor network, IEEE Trans. Inf. Theory 56 (1) (2010) 582–596.

[23] Y.-C. Wu, Q. Chaudhari, E. Serpedin, Clock synchronization of wireless sensor networks, IEEE Signal Process. Mag. 28 (1) (2011) 124–138.

[24] J. Zheng, Y. Wu, Joint time synchronization and localization of an unknown node in wireless sensor networks, IEEE Trans. Signal Process. 58 (3) (2010) 1309–1320.

[25] X. Cheng, A. Thaeler, G. Xue, D. Chen, TPS: A time-based positioning scheme for outdoor wireless sensor networks, in: IEEE INFOCOM 2004, Vol. 4, 2004, pp. 2685–2696.

[26] M. R. Gholami, S. Gezici, E. G. Ström, TDOA based positioning in the presence of unknown clock skew, IEEE Trans. Commun. 61 (6) (2013) 2522–2534.

[27] M. R. Gholami, S. Gezici, E. G. Ström, M. Rydström, Positioning algorithms for cooperative networks in the presence of an unknown turn-around time, in: SPAWC, 2011, pp. 166–170.

[28] M. R. Gholami, S. Gezici, E. G. Ström, Improved position estimation using hybrid TW-TOA and TDOA in cooperative networks, IEEE Trans. Signal Process. 60 (7) (2012) 3770–3785.

[29] Y. Wang, X. Ma, G. Leus, Robust time-based localization for asynchronous networks, IEEE Trans. Signal Process. 59 (9) (2011) 4397–4410.

[30] M. R. Gholami, S. Dwivedi, M. Jansson, P. Händel, Ranging without time stamps exchanging, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015, pp. 3981–3985.

[31] A. De Angelis, G. De Angelis, P. Carbone, Using Gaussian-uniform mixture models for robust time-interval measurement, IEEE Transactions on Instrumentation and Measurement 64 (12) (2015) 3545–3554.

[32] L. Rui, S. Chen, K. C. Ho, Anchor nodes refinement in joint localization and synchronization of a sensor node, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015, pp. 2834–2838.

[33] Z. Sahinoglu, Improving range accuracy of IEEE 802.15.4a radios in the presence of clock frequency offsets, IEEE Commun. Lett. 15 (2) (2011) 244–246.

[34] S. Gezici, Z. Sahinoglu, Enhanced position estimation via node cooperation, in: Proc. IEEE International Conference on Communications (ICC), 2010.

[35] M. R. Gholami, S. Gezici, E. G. Ström, Range based sensor node localization in the presence of unknown clock skews, in: Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, 2013, pp. 4046–4050.

[36] M. R. Gholami, S. Gezici, E. G. Ström, A concave-convex procedure for TDOA based positioning, IEEE Communications Letters 17 (2013) 765–768.

[37] D. Dardari, A. Conti, U. Ferner, A. Giorgetti, M. Win, Ranging with ultrawide bandwidth signals in multipath environments, Proceedings of the IEEE 97 (2009) 404–426.

[38] M. R. Gholami, H. Wymeersch, E. G. Ström, M. Rydström, Robust distributed positioning algorithms for cooperative networks, in: SPAWC, 2011, pp. 156–160.

[39] R. Kim, T. Nagayama, H. Jo, B. Spencer Jr, Preliminary study of low-cost GPS receivers for time synchronization of wireless sensors, in: SPIE Smart Structures and Materials+ Nondestructive Evaluation and Health Monitoring, 2012, pp. 83451A–83451A.

[40] D. C. Ganskopp, D. D. Johnson, GPS error in studies addressing animal movements and activities, Rangeland Ecology & Management 60 (4) (2007) 350–358.

[41] S. M. Kay, Fundamentals of Statistical Signal Processing: Estimation theory, Englewood Cliffs, NJ: Prentice-Hall, 1993.

[42] K. Anstreicher, H. Wolkowicz, On lagrangian relaxation of quadratic matrix constraints, SIAM Journal on Matrix Analysis and Applications 22 (1) (2000) 41–55.

[43] Y. Ye, S. Zhang, New results on quadratic minimization, SIAM Journal on Optimization 14 (1) (2003) 245–267.

[44] A. Beck, Y. C. Eldar, Strong duality in nonconvex quadratic optimization with two quadratic constraints, SIAM Journal on Optimization 17 (3) (2006) 844–860.

[45] J. J. More, Generalization of the trust region problem, Optimization Methods and Software 2 (1993) 189–209.

[46] D. Sorensen, Newton's method with a model trust region modification, SIAM J. Numer. Anal. 19 (2) (1982) 409–426.

[47] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.

[48] M. R. Gholami, Wireless sensor network positioning techniques, Ph.D. thesis, Chalmers (Nov. 2013). URL `http://publications.lib.chalmers.se/records/fulltext/185092/185092.pdf`

[49] R. Horst, N. V. Thoai, DC programming: Overview, Journal of Optimization Theory and Applications 103 (1999) 1–43.

[50] T. D. Quoc, M. Diehl, Sequential convex programming methods for solving nonlinear optimization problems with DC constraints, arXiv preprint arXiv:1107.5841.

[51] A. Ben-Tal, A. Nemirovski, Lectures on modern convex optimization (2012). URL `http://www2.isye.gatech.edu/~nemirovs/Lect_ModConvOpt.pdf`

Table 1: Complexity of different approaches.

Method	Complexity
MLE (GN initialized with good initial points)	$O(k_{\text{mle}}K^2N^3)$
AMLE (GN initialized with good initial points)	$O(k_{\text{amle}}(KN)^2)$
CCCP-SOCP	$O(k_{\text{cccp}}(KN)^{3.5}\log 1/\epsilon)$
LLS	$O(KN)$
EGTR	$O(k_{\text{sq}}KN)$

Table 2: Running time of Different Algorithms.

Method	Time (ms)
AMLE	32
MLE	317
LLS	0.74
EGTR	6.2
CCCP-SOCP	976

[52] M. Grant, S. Boyd, CVX: Matlab software for disciplined convex programming, version 1.21 (Feb. 2011).
URL http://cvxr.com/cvx

[53] M. Rydström, Algorithms and models for positioning and scheduling in wireless sensor networks, Ph.D. thesis, Chalmers University of Technology (2008).

[54] A. Urruela, Signal processing techniques for wireless locationing, Ph.D. thesis, Technical University of Catalonia (2006).
URL http://spcom.upc.edu/documents/T_2006_urruela.pdf

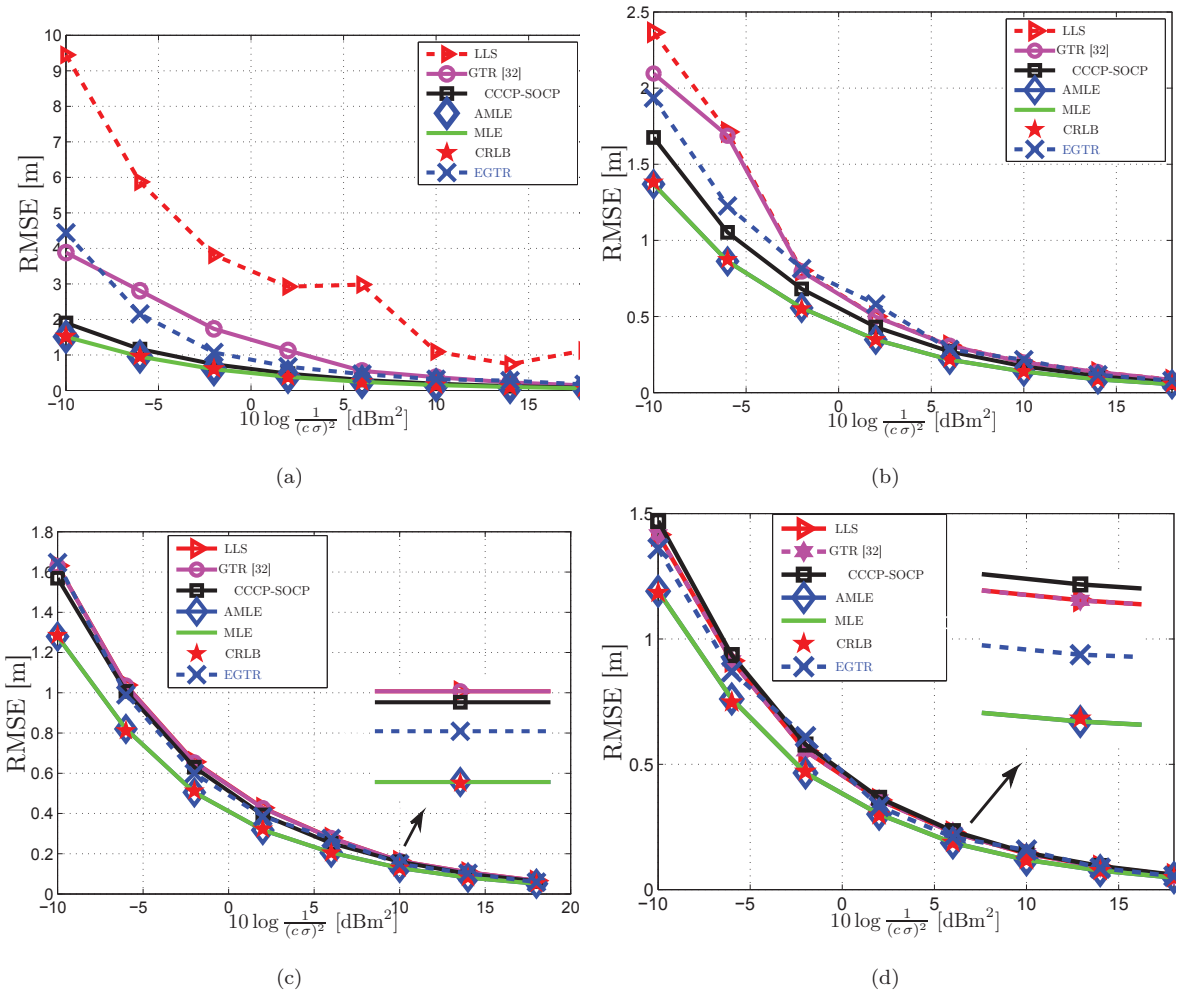

Figure 1: The RMSE of different approaches for $K = 2$ for (a) five reference nodes, (b) six reference nodes, (c) seven reference nodes, and (d) eight reference nodes.

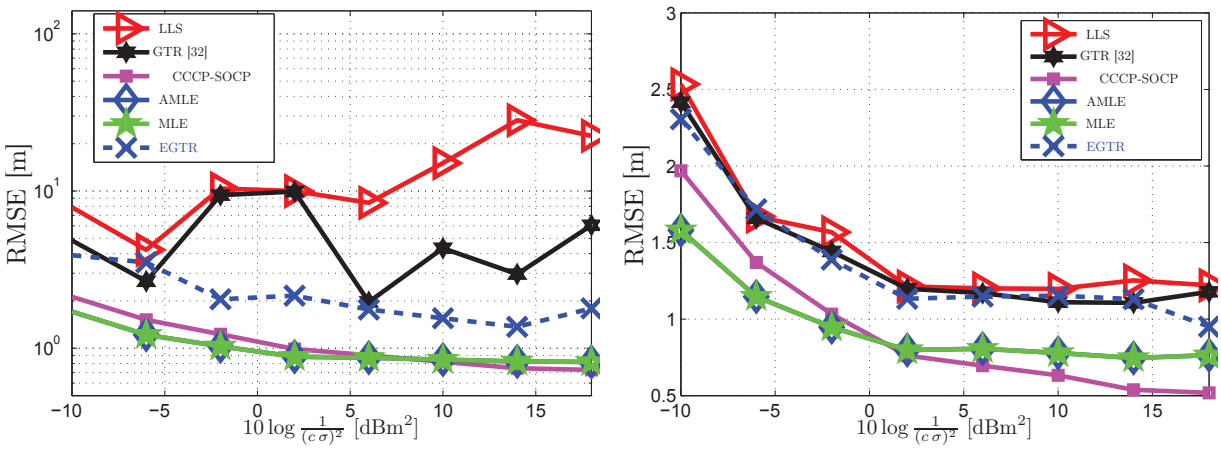

Figure 2: The RMSE of difference approaches for NLOS conditions ($K = 2$) for (a) five reference nodes and (b) six reference nodes.

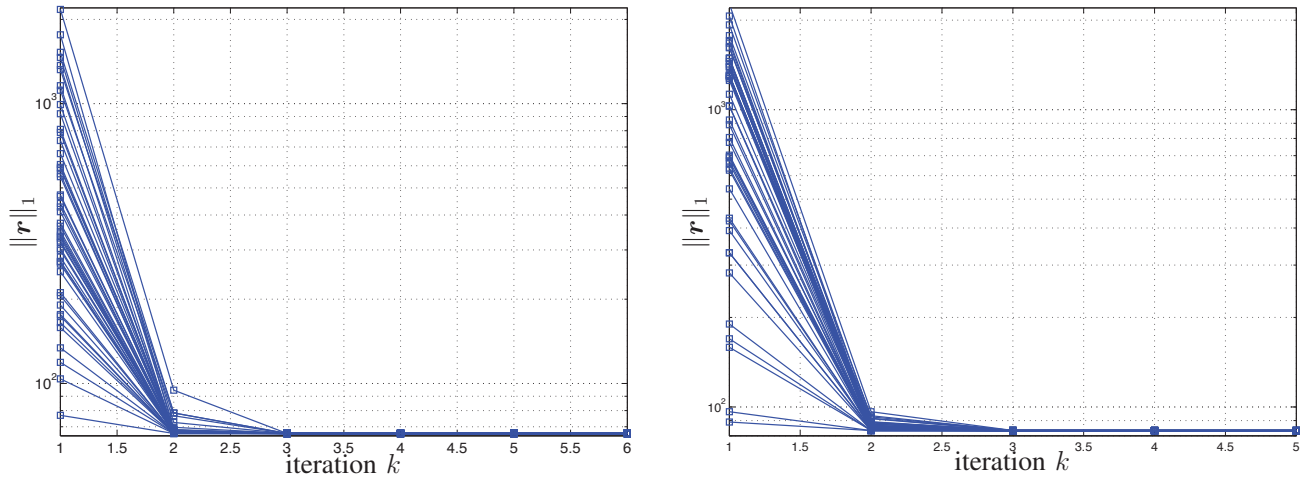

Figure 3: Convergence of proposed approaches for 50 random initializations for $c\sigma = 10$ and $K = 2$ for (a) 6 reference nodes, (b) 8 reference nodes.

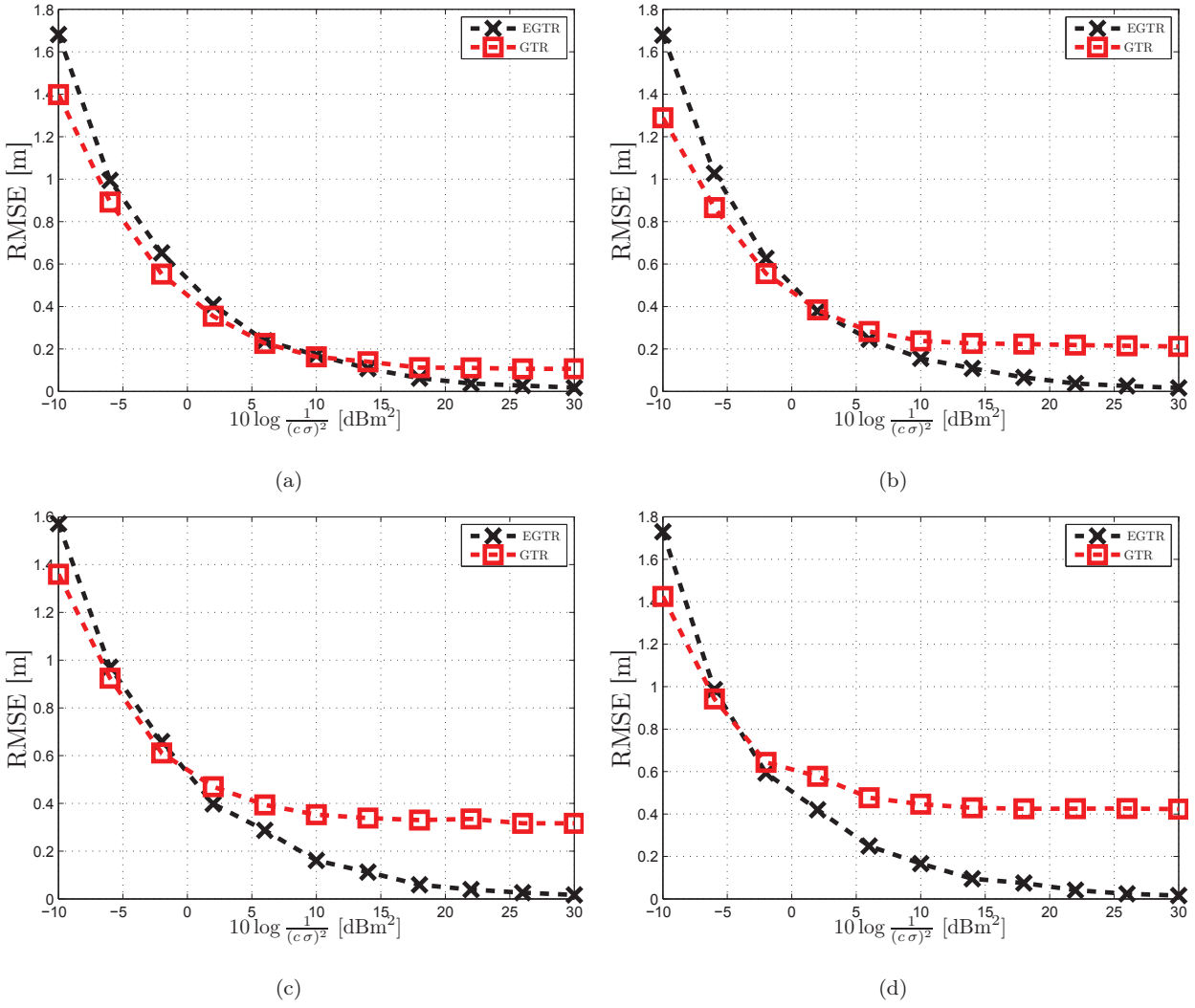

Figure 4: The RMSE of EGTR and GTR omitting clock skew for $K = 2$ for different values of ρ (a) $\rho = 0.0001$, (b) $\rho = 0.0002$, (c) $\rho = 0.0003$, and (d) $\rho = 0.0004$.