WILEY | Hindawi

Research Article

Robust Bearing-Only Localization Using Total Least Absolute Residuals Optimization

Ji-An Luo (ID),[1,2] **Chang-Cheng Xue,**[2] **and Dong-Liang Peng** (ID)[2]

[1]*Key Lab for IOT and Information Fusion Technology of Zhejiang, Hangzhou Dianzi University, Hangzhou 310018, China*
[2]*Institute of Information and Control, School of Automation, Hangzhou Dianzi University, Hangzhou 310018, China*

Correspondence should be addressed to Dong-Liang Peng; dlpeng@hdu.edu.cn

Robust techniques critically improve bearing-only target localization when the relevant measurements are being corrupted by impulsive noise. Resistance to isolated gross errors refers to the conventional least absolute residual (LAR) method, and its estimate can be determined by linear programming when pseudolinear equations are set. The LAR approach, however, cannot reduce the bias attributed to the correlation between system matrices and noise vectors. In the present study, perturbations are introduced into the elements of the system matrix and the data vector simultaneously, and the total optimization problem is formulated based on least absolute deviations. Subsequently, an equivalent form of total least absolute residuals (TLAR) is obtained, and an algorithm is developed to calculate the robust estimate by dual ascent algorithms. Moreover, the performance of the proposed method is verified through the numerical simulations by using two types of localization geometries, i.e., random and linear. As revealed from the results, the TLAR algorithm is capable of exhibiting significantly higher localization accuracy as compared with the LAR method.

1. Introduction

Bearing-only target localization (BOTL) by exploiting spatially distributed sensors can have extensive applications in vehicle [1] or gunshot [2] positioning, animal habit monitoring [3], network localization [4], multiagent systems [5, 6], recently massive arrays [7], etc. BOTL algorithms have received broad studies under the measurement noise of Gaussian [8–11]. For numerous applications, however, especially for military applications, the bearing sensors exhibit vulnerability to external interference, enemy attack, or node failure. The bearing measurements may be subject to impulsive noises [12–16], thereby dramatically reducing the localization performance. How to maintain high positioning performance under unreliable measurements should be considered in this paper.

On the whole, the existing robust BOTL methods to process outlier data can fall to the outlier detection [17] and the M-estimate [18]. The outlier detection method aims to first detect suspected outlier data, separate them from the

original data set, and then exploit the remaining data to complete the localization task. Picard and Weiss and Picard [17] proposed a sparse representation method to detect the outlier data of time-of-arrival, time-difference-of-arrival, and direction-of-arrival, as solved by linear programming. Xiong et al. [19] developed a robust expectation-maximization algorithm for distance outlier detection. Though the outlier detection method is intuitive and effective, it does not apply to large data sets or complex application scenarios.

The other important aspect refers to M-estimate, capable of estimating robust positions without preprocessing data. M-estimate primarily aims to comply with some other criteria, instead of the least-squares criterion, which is more robust to impulsive noise, as an attempt to improve estimators to be less sensitive to model errors. The robust criteria consist of Huber [20], Bi-square [18], the negative log-likelihood of the Cauchy distribution [21], Wilcoxon [22], Chebyshev norm [23], L_p-norms ($1 \le p < 2$) [24], etc. Based on energy information for sensor networks, Liu et al. [18] presented a distributed robust localization algorithm.

Such algorithm employs Bi-square function as the cost function of M-estimate. Panigrahi et al. [22] proposed a distributed incremental least mean square algorithm based on Wilcoxon norm for parameter estimation of sensor networks. Wu et al. [25] proposed a robust structure total least-squares algorithm for passive localization. The algorithm adopts the optimized Danish weight function to reduce the effect of outlier data on the localization performance. Furthermore, the L_p-norms $(1 \leq p < 2)$ are useful for robust estimation since less weight is given to isolated deviations. Under the application of the L_1-norm, such M-estimate is termed as the least absolute residuals (LAR) [26] method. For outlier suppression, the L_1-norm appears to be markedly superior among the L_p-norms $(1 \leq p < \infty)$ [27].

From a statistical perspective, the least-squares optimization equals the maximum likelihood (ML) estimation when the measurements are corrupted with independent and identically distributed Gaussian noise. However, the noise distribution is altered with the presence of outliers. The Laplace distribution refers to a probability distribution that accommodates large residuals. The ML estimator of this distribution leads to the LAR solution [28]. Though the ML method is optimal for statistics, its cost function is nonlinear and nonconvex with respect to the target location parameters. The iterative numerical search is inevitable for ML, and the absolute norm should be considered in the respective iteration. The iterative algorithms tend to diverge when poorly initialized and computationally expensive. To remedy the defects of the ML method, [29] proposed a pseudolinear estimator (PLE) by lumping the nonlinearities into the noise term. However, the PLE is subjected to severe bias [30], and the bias remains with the increase in the number of sensors due to the correlations between system matrices and measurement noises. Various algorithms have been proposed to reduce the bias, including instrumental variable (IV) [31, 32] and total least-squares (TLS) [33]. The IV method [31] is capable of reducing biases by setting an instrumental matrix that is asymptotically uncorrelated with the noise vector. Inconsistent with the IV method, the TLS algorithm [33] attempts to reduce biases by minimizing the errors in the system matrix and the measurement vector. However, the IV and TLS estimators fail to improve the bias performance if the measuring angles have gross errors.

In brief, robust pseudolinear algorithms for BOTL aim to reduce biases attributed to both large residuals and the correlation between system matrix and noise vector. The least absolute residuals (LAR) minimization can be adopted to reduce the bias attributed to outlier data. However, LAR faces a major problem of the correlation bias that remains with the increase in the number of sensors. In the present study, these two types of bias are reduced by conducting total least absolute residuals (TLAR) optimization [34]. We first formulate the problem of BOTL subject to outlier data.

Moreover, the pseudolinear measurement model for BOTL is reviewed, and the bias of PLE is analyzed. Subsequently, the TLAR algorithm is developed for BOTL with significantly reduced bias and root mean square error as compared with the LAR method.

The main contributions of the proposed method can be summarized as follows:

(i) Development of a new bias reduced estimator based on TLAR for BOTL under bearing gross errors

(ii) Development of an algorithm for TLAR optimization using dual ascent algorithms

(iii) Demonstration of the performance improvement achieved by the TLAR estimator with respect to the LAR method

The rest of this paper is organized as follows. In Section 2, the measurement model is described. In Section 3, the pseudolinear equations from the BOTL problem are reviewed, the two types of bias for the PLE method are analyzed, and a LAR solution is presented when bearing measurements are subjected to large deviations. In Section 4, the TLAR approach is presented, and an algorithm is developed for TLAR based on dual ascent algorithms in Section 5. In Section 6, numerical examples illustrating the performance of PLE, TLS, LAR, and TLAR are presented. Lastly, conclusions are drawn in Section 7.

2. Model Description

A group of sensors are formed by K_n normal sensors and K_a abnormal sensors with K_n significantly larger than K_a. The total number of sensors reaches $K = K_n + K_a$. Each node is capable of measuring a bearing in the sensor between the positive horizontal direction and the straight line from the target to the node. It is noteworthy that the normal sensors conduct the effective measurements, while the abnormal sensors collect the wrong observations as impacted by object occlusion, interference or network attack, etc. It is assumed that whether the bearing measurements have outliers cannot be distinguished. Thus, robust localization methods should be designed to prevent the location performance degradation attributed to outlier data. The problem of robust BOTL is to estimate an unknown target position in $\mathbb{R}^2$ as accurately as possible by all K bearing measurements.

The localization geometry is illustrated in Figure 1, where $\mathbf{p} = [p_x, p_y]^T$ denotes the target position vector, $\mathbf{s}_k = [s_{x,k}, s_{y,k}]^T$ represents the sensor location vector for the kth measurement, θ_k is the true bearing at sensor k, and the angle θ_k is positive to the counterclockwise direction, $k = 1, 2, \ldots, K$. The relationship between the bearing angle, target position, and sensor location is expressed as the following nonlinear equation:
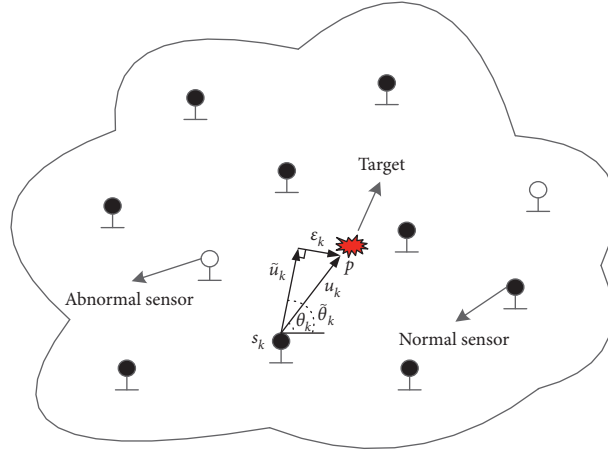
FIGURE 1: Illustration of BOTL geometry. The circle in black color is normal sensor and the circle in white color is abnormal sensor.

$$\theta_k = \tan^{-1}\frac{p_y - s_{y,k}}{p_x - s_{x,k}}, \tag{1}$$

where $\tan^{-1}$ denotes the four-quadrant arctangent and $\theta_k \in (0, 2\pi]$.

In fact, the observed bearings have errors, and the kth measurement can be described as

$$\widetilde{\theta}_k = \theta_k + n_k, \tag{2}$$

where $n_k = \delta \cdot b_k + e_k$ denotes the measurement error; e_k represents the independent and identically distributed (i.i.d.) zero mean Gaussian noise with variance σ^2; b_k represents a bias term; and δ is a binary random variable defined as follows:

$$\begin{cases} \delta = 1, & \text{if } \widetilde{\theta}_k \text{ is outlier,} \\ \delta = 0, & \text{otherwise.} \end{cases} \tag{3}$$

Since the prior knowledge of δ is unknown, which measurement is reliable and which is not cannot be distinguished in advance. Accordingly, it is necessary to develop a robust localization algorithm when both reliable and unreliable measurements are used.

3. The Least Absolute Residuals Method

To develop a robust localization method for BOTL, we first review the PLE method by converting nonlinear bearing measurements to pseudolinear equations. We then derive the least absolute residual (LAR) algorithm for BOTL. Lastly, two types of bias for PLE are analyzed.

3.1. Pseudolinear Equation. The measurement equation in (2) is nonlinear with respect to the unknown target location which makes BOTL a nontrivial task. A natural option as we will illustrate in the following would be to model a pseudolinear equation by lumping the nonlinearities into the

noise term. For this end, an orthogonal vector sum is first established between the measured angle vector and the true angle vector from Figure 1 given by

$$\mathbf{u}_k = \mathbf{p} - \mathbf{s}_k = \widetilde{\mathbf{u}}_k + \boldsymbol{\varepsilon}_k, \tag{4}$$

where $\mathbf{u}_k$ denotes the true angle vector between $\mathbf{s}_k$ and $\mathbf{p}$; $\widetilde{\mathbf{u}}_k$ represents the measured angle vector starting from $\mathbf{s}_k$ and generates the noisy bearing $\widetilde{\theta}_k$ according to the horizontal direction; and $\boldsymbol{\varepsilon}_k$ indicates the error vector. By defining $\boldsymbol{\alpha}_k = [\cos\widetilde{\theta}_k, \sin\widetilde{\theta}_k]^T$ and $\boldsymbol{\beta}_k = [\sin\widetilde{\theta}_k, -\cos\widetilde{\theta}_k]^T$ as two orthogonal unit trigonometric vectors, $\widetilde{\mathbf{u}}_k$ and $\boldsymbol{\varepsilon}_k$ are written in terms of $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$:

$$\begin{aligned} \widetilde{\mathbf{u}}_k &= \|\mathbf{u}_k\|_2 \cos n_k \boldsymbol{\alpha}_k, \\ \boldsymbol{\varepsilon}_k &= \|\mathbf{u}_k\|_2 \sin n_k \boldsymbol{\beta}_k. \end{aligned} \tag{5}$$

Using the fact that $\widetilde{\mathbf{u}}_k^T \boldsymbol{\beta}_k = 0$, multiplying (4) with $\boldsymbol{\beta}_k^T$ yields

$$\xi_k = \boldsymbol{\beta}_k^T \mathbf{p} - \boldsymbol{\beta}_k^T \mathbf{s}_k, \tag{6}$$

where $\xi_k = \|\mathbf{u}_k\|_2 \sin n_k$ is a nonlinear transformed measurement error. Collecting the pseudolinear equation errors as a vector $\boldsymbol{\xi} = [\xi_1, \xi_2, \ldots, \xi_K]^T$, we obtain

$$\boldsymbol{\xi} = \mathbf{A}\mathbf{p} - \mathbf{h}, \tag{7}$$

where $\mathbf{A} = [\beta_1, \beta_2, \ldots, \beta_K]^T$ and $\mathbf{h} = [\beta_1^T \mathbf{s}_1, \beta_2^T \mathbf{s}_2, \ldots, \beta_K^T \mathbf{s}_K]^T$ are the measurement matrix and vector, respectively. The PLE requires that $\mathbf{p}$ be estimated by minimizing $\boldsymbol{\xi}$ with respect to $\mathbf{p}$ in the least-squares sense. The position estimate can be obtained by solving

$$\begin{aligned} &\min_{\mathbf{p}} \|\boldsymbol{\xi}\|_2, \\ &s.t \quad \boldsymbol{\xi} = \mathbf{A}\mathbf{p} - \mathbf{h}. \end{aligned} \tag{8}$$

The solution to (8) is given by

$$\widehat{p} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{h}. \tag{9}$$

It is termed as pseudolinear estimator (PLE).

3.2. Bias Analysis.

The bias of $\widehat{p}$ obtained from (9) is defined by

$$\boldsymbol{\eta} = E\{\widehat{p}\} - \mathbf{p} = -E\left\{\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\boldsymbol{\xi}\right\}. \tag{10}$$

The bias of $\widehat{p}$ includes two parts. The first part is attributed to the correlation between $\mathbf{A}$ and ξ. The second part is formed by large residuals. Let $\mathbf{A}_n$ and $\mathbf{h}_n$ be the pseudomeasurement matrix and vector obtained from normal sensors. Without considering the abnormal measurements, the PLE solution becomes

$$\widehat{\mathbf{p}_n} = \left(\mathbf{A}_n^T\mathbf{A}_n\right)^{-1}\mathbf{A}_n^T\mathbf{h}_n. \tag{11}$$

If the number of normal sensors is significantly bigger than that of abnormal sensors, the first type of bias can be approximated as

$$\boldsymbol{\eta}_1 = E\{\widehat{\mathbf{p}_n}\} - \mathbf{p} = -E\left\{\left(\mathbf{A}_n^T\mathbf{A}_n\right)^{-1}\mathbf{A}_n^T\boldsymbol{\xi}_n\right\}, \tag{12}$$

where ξ_n denotes the pseudomeasurement noise with its kth entry given by $\xi_{n,k} = \|\mathbf{u}_k\|_2 \sin e_k$. Based on the Slutsky theorem [33], the first type of bias can be asymptotically computed by

$$\boldsymbol{\eta}_1 \approx -E\left\{\frac{\mathbf{A}_n^T\mathbf{A}_n}{K_n}\right\}^{-1} E\left\{\frac{\mathbf{A}_n^T\boldsymbol{\xi}_n}{K_n}\right\}. \tag{13}$$

As K_n goes to infinity, (13) becomes an equality. For finite K_n, η_1 obtained by (13) is a good approximation to (12) [35]. After η_1 is calculated, the second type of bias is $\eta_2 = \eta - \eta_1$. An example of PLE bias is depicted in Figure 2(b), Section 6. The second type of bias appears to dominate if the standard deviation of e_k is small enough. Thus, the performance of PLE will degrade dramatically under impulsive noise since L_2-norm optimization can be severely affected by pseudolinear errors with large residuals. To ensure such items have less influence, we could instead minimize a cost function that gives less weight to large deviations.

3.3. Least Absolute Residual.

A common choice to alleviate the effect of gross errors is the absolute value metric denoted by $\chi(\xi) = \|\xi\|_1$, where $\|\cdot\|_1$ represents L_1-norm. As such, the LAR optimization can be achieved by

$$\min_{\mathbf{p}} \chi(\boldsymbol{\xi}), \\ \text{s.t.} \quad \boldsymbol{\xi} = \mathbf{Ap} - \mathbf{h}. \tag{14}$$

The derivative of function $\chi(\xi)$ is bounded for all ξ by the value ± 1, demonstrating that the cost function $\chi(\xi)$ is robust for all deviations. For the PLE criterion, the derivative is not bounded, and it increases linearly with ξ.

In the literature, numerous algorithms have been developed to solve the minimization problem of $\chi(\xi)$ (e.g., iteratively reweighted least-squares (IRWLS) [36], expectation-maximization (EM) procedure [37], and linear programming [38]). A weight matrix is defined as

$$\mathbf{W}^{(i)} = \mathrm{diag}\left(\frac{1}{\left|\xi_1^{(i)}\right|}, \frac{1}{\left|\xi_2^{(i)}\right|}, \cdots, \frac{1}{\left|\xi_K^{(i)}\right|}\right). \tag{15}$$

The WLS algorithm is explicitly given by

$$\widehat{\mathbf{p}^{(i+1)}} = \left(\mathbf{A}^T\mathbf{W}^{(i)}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{W}^{(i)}\mathbf{h}. \tag{16}$$

The problem of (16) is that the weights become extraordinarily large for $\xi_k \approx 0$ or numerically indeterminate for $\xi_k = 0$. Benefiting from convex optimization, LAR problem can be relaxed to identify the minimum bound of the absolute value:

$$\min_{\mathbf{p},\mathbf{v}} \quad \mathbf{1}^T\mathbf{v}, \\ \text{s.t.} \quad -\mathbf{v} \le \mathbf{Ap} - \mathbf{h} \le \mathbf{v}, \tag{17}$$

where $\mathbf{v}$ denotes the upper bound of ξ; 1 is a column vector of ones. To be specific, two nonnegative vectors are denoted, i.e., $\mathbf{r} = 0.5(\mathbf{v} + \mathbf{Ap} - \mathbf{h})$ and $\mathbf{t} = 0.5(\mathbf{v} - \mathbf{Ap} + \mathbf{h})$. The equivalent form of (17) is

$$\min_{\mathbf{p},\mathbf{v}} \mathbf{1}^T\mathbf{r} + \mathbf{1}^T\mathbf{t}, \\ \text{s.t.} \quad \mathbf{Ap} - \mathbf{h} = \mathbf{r} - \mathbf{t}, \quad \mathbf{r} \ge 0, \mathbf{t} \ge 0. \tag{18}$$

Note that (23) refers to a standard linear programming problem, which can be solved by using the existing CVX software [39]. When the pseudolinear errors follow i.i.d. Laplacian distribution with zero mean, (14) is equivalent to the ML estimator. However, the LAR method implicitly assumes that only $\mathbf{h}$ is subjected to errors. This is not the case since the system matrix $\mathbf{A}$ is corrupted with measurement noises as well. The correlation between $\mathbf{A}$ and ξ causes the LAR estimator to be inconsistent. Though the LAR estimator has bias, it gives a reasonable estimate and provides an initial guess for other robust estimators.

4. Total LAR Optimization

The LAR algorithm expressed in (10) implicitly indicates that only $\mathbf{h}$ has errors and the gross errors are reduced by using the weights all restricted in $\mathbf{h}$. In fact, matrix $\mathbf{A}$ is also subject to measurement errors. When both $\mathbf{A}$ and $\mathbf{h}$ are disturbed with noise, the LAR solution of (9) will inevitably cause large bias as $\mathbf{A}$ and ξ are statistically dependent. To increase the accuracy of the LAR estimator, the idea of total least absolute
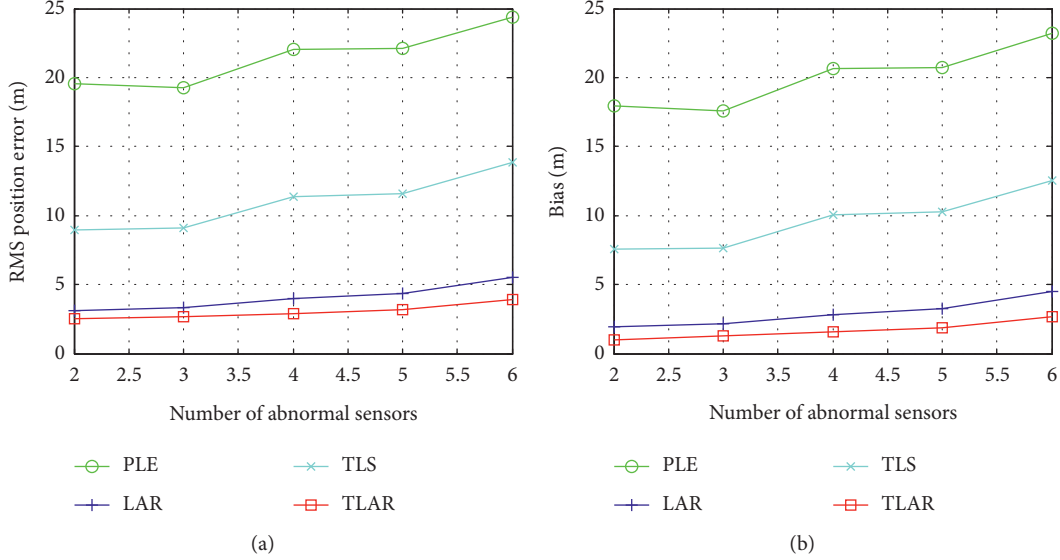
(a)

(b)

FIGURE 2: Performance comparison of PLE, LAR, TLS, and TLAR estimates with various outliers. (a) RMSE results of various estimators. (b) Bias results of various estimators.

residual (TLAR) can be exploited to reduce the errors in both **A** and **h**.

The concept of TLAR is that the disturbance vector $\Delta \mathbf{h}$ is adopted to correct the data vector **h**, while the disturbance matrix $\Delta \mathbf{A}$ is employed to disturb the data matrix **A**. In other words, we use the following equations in the TLAR problem:

$$(\mathbf{A} + \Delta \mathbf{A})\mathbf{p} = \mathbf{h} + \Delta \mathbf{h},$$
$$(\mathbf{M} + \boldsymbol{\Sigma})\mathbf{x} = 0, \tag{19}$$

where $\mathbf{M} = [\mathbf{A}, \mathbf{h}]$ denotes a $K \times 3$ augmented matrix; $\boldsymbol{\Sigma} = [\Delta \mathbf{A}, \Delta \mathbf{h}]$ represents a $K \times 3$ perturbation matrix; $\Delta \mathbf{A}$ is a perturbation matrix of **A**; $\Delta \mathbf{h}$ is a perturbation vector of **h**; $\mathbf{x} = \tau \cdot [p_x, p_y, -1]^T$ is a 3×1 vector; and τ is a scaling factor. Notably, both **A** and **h** are corrupted with noise. Next, the error statistics for **A** and **h** are examined. The error matrix $\boldsymbol{\Sigma}$ is defined as

$$\boldsymbol{\Sigma} = \begin{bmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ \vdots & \vdots & \vdots \\ \rho_{K1} & \rho_{K2} & \rho_{K3} \end{bmatrix}, \tag{20}$$

where the entries of $\boldsymbol{\Sigma}$ are given by the difference between **M** and the noiseless augmented matrix

$$\rho_{k1} = \sin \theta_k - \sin \widetilde{\theta}_k = \sin \theta_k - \mu_k \cos e_k + v_k \sin e_k,$$
$$\rho_{k2} = -\cos \theta_k + \cos \widetilde{\theta}_k = -\cos \theta_k - v_k \cos e_k - \mu_k \sin e_k,$$
$$\rho_{k3} = [\sin \theta_k, -\cos \theta_k]\mathbf{s}_k - \boldsymbol{\beta}_k^T \mathbf{s}_k = [e_{k1}, e_{k2}]\mathbf{s}_k, \tag{21}$$

where $\mu_k = \sin \theta_k \cos (a \cdot b_k) + \cos \theta_k \sin (a \cdot b_k)$, $v_k = \sin \theta_k \sin (a \cdot b_k) - \cos \theta_k \cos (a \cdot b_k)$. If e_k is sufficiently small, it yields $\sin e_k \approx e_k$, $\cos e_k \approx 1$. Thus, ρ_{k1} and ρ_{k2} become

$$\rho_{k1} \approx \sin \theta_k - \mu_k + v_k e_k,$$
$$\rho_{k2} \approx -\cos \theta_k - v_k - \mu_k e_k. \tag{22}$$

The mean and second-order moments of the error items are written as

$$E\{\rho_{k1}\} \approx \sin \theta_k - \mu_k,$$
$$E\{\rho_{k2}\} \approx -\cos \theta_k - v_k, \tag{23}$$
$$E\{\rho_{k3}\} \approx [\sin \theta_k - \mu_k, -\cos \theta_k - v_k]\mathbf{s}_k.$$

$$E\{\rho_{k1}^2\} \approx (\sin \theta_k - \mu_k)^2 + \sigma_k^2 v_k^2,$$
$$E\{\rho_{k2}^2\} \approx (\cos \theta_k + v_k)^2 + \sigma_k^2 \mu_k^2, \tag{24}$$

$$E\{\rho_{k3}^2\} \approx ([\sin \theta_k - \mu_k, -\cos \theta_k - v_k]\mathbf{s}_k)^2 + \sigma_k^2 ([v_k, -\mu_k]\mathbf{s}_k)^2. \tag{25}$$

Note that from (23), the mean value of errors in **M** is nonzero as impacted by the outlier data. In addition, (24) and (25) are nonzero even if b_k equals zero. It is therefore suggested that the error terms in the identical row of $\boldsymbol{\Sigma}$ are correlated with each other. To more effectively reduce the bias of LAR, TLAR metric is adopted to minimize the disturbance matrix and vector simultaneously, and the TLAR problem for BOTL is formulated as

$$\min_{\Sigma,\mathbf{x}} \|\Sigma\|_1, \tag{26}$$
$$\text{s.t.} \quad (\mathbf{M} + \Sigma)\mathbf{x} = 0.$$

If the L_1-norm in (26) is replaced by the Frobenius matrix norm, (26) becomes the well-known TLS problem. Both TLAR and TLS fall to the domain of "total approximation problem," known as the total least pth problem for $p \geq 1$. To solve the TLAR problem effectively, an equivalent form of (26) is explored. The following result holds.

Theorem 1. *Suppose that* $(\mathbf{M} + \Sigma)\mathbf{x} = 0$, *where* $\mathbf{M} = [\mathbf{A}, \mathbf{h}]$ *denotes the augment matrix with* $\mathbf{A}$ *and* $\mathbf{h}$ *defined in (7),* $\Sigma = [\Delta\mathbf{A}, \Delta\mathbf{h}]$ *is the error matrix with* $\Delta\mathbf{A}$ *and* $\Delta\mathbf{h}$ *defined in (19), and* $\mathbf{x}$ *is a* 3×1 *vector. Subsequently, the estimation of* $\mathbf{p}$ *from the minimization of* $\|\Sigma\|_1$ *can be achieved by*

$$\min_{\mathbf{x}} \|\mathbf{M}\mathbf{x}\|_1, \tag{27}$$
$$\text{s.t.} \quad \|\mathbf{x}\|_\infty = 1.$$

Proof. From $(\mathbf{M} + \Sigma)\mathbf{x} = 0$, it yields $\|\mathbf{M}\mathbf{x}\|_1 = \|\Sigma\mathbf{x}\|_1$. Let us denote $\Sigma = [\rho_1, \rho_2, \rho_3]$ and $\mathbf{x} = [x_1, x_2, x_3]^T$. Then, $\|\Sigma\mathbf{x}\|_1$ can be computed as

$$\|\Sigma\mathbf{x}\|_1 = \sum_{i=1}^{3} |\rho_i x_i| = \sum_{i=1}^{3} (|\rho_i| \cdot |x_i|). \tag{28}$$

Note that $|x_i| \leq \|\mathbf{x}\|_\infty$, for $i = 1, 2, 3$. Accordingly, it yields $\|\sum\mathbf{x}\|_1 \leq \sum_{i=1}^{3} |\rho_i| \cdot \|\mathbf{x}\|_\infty = \|\Sigma\|_1 \|\mathbf{x}\|_\infty$. If the infinite norm of $\mathbf{x}$ satisfies $\|\mathbf{x}\|_\infty = 1$, the above inequality becomes $\|\sum\mathbf{x}\|_1 \leq \|\Sigma\|_1$. Lastly, we get the conclusion that the minimization of $\|\Sigma\|_1$ subject to $(\mathbf{M} + \Sigma)\mathbf{x} = 0$ can be obtained from (27).

Theorem 1 presents a practicable solution for BOTL since $\widehat{p} = -\mathbf{x}(1:2)/x_3$. However, (27) is not a convex problem since the feasible region does not fall to $\|\mathbf{x}\|_\infty = 1$. Thus, the global minimum is not ensured. Indeed, its local minima can be computed by adopting the Lagrange multiplier formulation of (27).

5. Algorithm Development

In the present section, an algorithm is derived to solve (27). Since the optimization of (27) is not convex, a stationary point satisfying the first-order necessary conditions is calculated. First, (27) is transformed into an unconstrained minimization problem by leveraging the Lagrange multiplier method. The Lagrange objective function is defined as

$$L(\mathbf{x}, \lambda) = \|\mathbf{M}\mathbf{x}\|_1 + \lambda(\|\mathbf{x}\|_\infty - 1). \tag{29}$$

And the dual problem of (27) is given by

$$\max_{\lambda} G(\lambda), \tag{30}$$
$$\text{s.t.} \quad G(\lambda) = \inf_{\mathbf{x}} L(\mathbf{x}, \lambda), \quad \lambda \geq 0,$$

where λ denotes the Lagrange multiplier. Set λ^* as the optimal value of (30). By substituting λ^* into (29), the minimization of $L(\mathbf{x}, \lambda^*)$ generates the primal optimal point $\mathbf{x}^*$. Such dual problem can be solved by a dual ascent (DS) algorithm [40] as expressed below.

(1) λ_j is assumed as the optimal solution of the dual problem (30) at the jth step.

(2) The primal optimal point $\mathbf{x}_{j+1}$ can be determined from λ_j by minimizing $L(\mathbf{x}, \lambda_j)$:

$$\mathbf{x}_{j+1} = \arg\min_{\mathbf{x}} L(\mathbf{x}, \lambda_j). \tag{31}$$

(3) The dual variable is updated by

$$\lambda_{j+1} = \lambda_j + \gamma_j \left(\|\mathbf{x}_{j+1}\|_\infty - 1 \right), \tag{32}$$

where $\gamma_j > 0$ is a step size. The λ-update is realized by using gradient ascent. With proper choice of γ_j, $G(\lambda)$ increases for each step, i.e., $G(\lambda_{j+1}) > G(\lambda_j)$.

(4) Next, we investigate the problem of minimizing $L(\mathbf{x}, \lambda_j)$. Based on Taylor's series expansion of $L(\mathbf{x} + \alpha\mathbf{q}, \lambda_j)$ around $\mathbf{x}$ up to the first order, it yields

$$L(\mathbf{x}, \lambda_j) - L(\mathbf{x} + \alpha\mathbf{q}, \lambda_j) \approx -\alpha \nabla L(\mathbf{x}, \lambda_j)^T \mathbf{q}. \tag{33}$$

(5) Note that $\nabla L(\mathbf{x}, \lambda_j)^T \mathbf{q} = \|\nabla L(\mathbf{x}, \lambda_j)\|_2 \|\mathbf{q}\|_2 \cos\beta$, where β is the angle between $\nabla L(\mathbf{x}, \lambda_j)$ and $\mathbf{q}$. Thus, the steepest decent direction is $\mathbf{q} = -\nabla L(\mathbf{x}, \lambda_j)$. Let $\partial\|\mathbf{M}\mathbf{x}\|_1$ denote the subdifferential of $\|\mathbf{M}\mathbf{x}\|_1$, where

$$\partial\|\mathbf{M}\mathbf{x}\|_1 = \left\{ \mathbf{g} \big| \|\mathbf{g}\|_\infty \leq 1, \mathbf{g}^T \mathbf{M}\mathbf{x} = 1 \right\}. \tag{34}$$

(6) $\mathbf{g}$ is a $K \times 1$ subgradient vector and its kth element is given by

$$g_k = \begin{cases} 1, & \text{if } (\mathbf{M}\mathbf{x})_k > 0, \\ -1, & \text{if } (\mathbf{M}\mathbf{x})_k < 0, \\ -[1,1] & \text{if } (\mathbf{M}\mathbf{x})_k = 0. \end{cases} \tag{35}$$

(7) The subdifferential of $\|\mathbf{x}\|_\infty$ at $\|\mathbf{x}\|_\infty$, denoted by $\partial\|\mathbf{x}\|_\infty$, is defined by

$$\partial \|\mathbf{x}\|_\infty = \begin{cases} \text{conv}\{\text{sgn}(x_i)\mathbf{o}_i \colon |x_i| = \|\mathbf{x}\|_\infty\}, & \mathbf{x} \neq 0, \\ \{\mathbf{y} \colon \|\mathbf{y}\|_1 \leq 1\}, & \mathbf{x} = 0. \end{cases} \quad (36)$$

where conv denotes the convex hull and $\mathbf{o}_i$ is a vector whose ith element is 1 and all other elements are 0. Based on (34) and (36), $\nabla L(\mathbf{x}, \lambda_j)$ is written as

$$\nabla L(\mathbf{x}, \lambda_j) = \partial \|\mathbf{M}\mathbf{x}\|_1 + \lambda_j \partial \|\mathbf{x}\|_\infty \quad (37)$$

6. Experimental Results

In the present section, numerical examples are studied to compare the localization performance of the proposed TLAR algorithm with PLE, TLS, and LAR under abnormal sensor and nonabnormal sensor. In addition, two typical localization geometries are applied in the simulations. The first one is randomly distributed sensors, and the other one is linearly distributed sensors. In the respective scenario, four cases are considered. To be specific, (1) the number of abnormal sensors varies and the total number of sensors is fixed, (2) both the numbers of normal and abnormal sensors vary, (3) the numbers of normal and abnormal sensors are fixed, and (4) all sensors are normal. The bearing data achieved from abnormal sensors exhibit uniform distribution $U[-\pi, \pi]$. For normal sensors, the bearing measurement errors are assumed as i.i.d. zero mean Gaussian with standard deviation σ. Simulation comparisons are drawn based on $Mc = 1000$ Monte-Carlo simulation runs. This study employs bias and root mean square error (RMSE) for localization performance comparison, which are written as

$$\text{BIAS} = \sqrt{\left(p_x - \frac{1}{Mc}\sum_{m=1}^{Mc}\widehat{p_x}(m)\right)^2 + \left(p_y - \frac{1}{Mc}\sum_{m=1}^{Mc}\widehat{p_y}(m)\right)^2},$$

$$\text{RMSE} = \sqrt{\frac{1}{Mc}\sum_{m=1}^{Mc}\left[(p_x - \widehat{p_x}(m))^2 + (p_y - \widehat{p_y}(m))^2\right]},$$

$$(38)$$

where $\widehat{p_x}(m)$ and $\widehat{p_y}(m)$ denote the estimates of target location parameters for the mth Monte-Carlo run of the simulation.

6.1. Randomly Distributed Sensors.
In this section, all sensors are randomly placed in a 100×100 m^2 region centered at (50, 50) m (Figure 2(a)). The unknown target is placed at (100, 100) m. In the first example, the number of abnormal sensors is elevated from 2 to 6, and that of normal sensors decreases from 18 to 14. The total number of sensors is fixed at 20. σ is set to $3\pi/180$ (3°). Figure 2(a) presents the RMSE results with the increase in the number of abnormal sensors. The LAR and TLAR methods are capable of reducing the outliers, as indicated in the two plots of Figure 2(a), the blue line with "+" and the red line with "square." The RMSE of LAR is 0.618 m above that of TLAR when the number of outliers is set to 2. Such value increases to 1.551 m when the number of outliers is kept at 6. The estimation bias is presented in Figure 2(b). As the number of outliers is elevated, the value of bias turns more significant.

In the next example, the number of outliers is fixed at 3. The sensors have a total number of twenty. σ ranges from $\pi/180$ to $5\pi/180$ (1° to 5°). The number of abnormal sensors (red circles) is set to three, and the rest are normal sensors (blue circles), as illustrated in Figure 3(a). The simulated biases of PLE are plotted in Figure 3(b). Under small σ, the first type of bias attributed to the correlation between $\mathbf{A}$ and ξ keeps at low level, and the second type of bias formed by large residuals dominates the theoretical bias of PLE. With the increase in the measurement noise variance, the effect of the second type of bias turns out to be less significant. Figures 4(a) and 4(b) illustrate the RMSE and bias curves of various methods in the presence of abnormal sensors. The PLE method and the TLS estimator fail to give accurate target location estimates since they are not robust to outlier data. The blue line with "+" in Figure 4(a) represents the RMSE value determined by using the LAR method, and the red line with "square" represents the RMSE curve for the TLAR algorithm. They are well separated and the gap between these two lines increases as σ becomes large, suggesting the reduction of the first type of bias. Furthermore, this phenomenon is verified in Figure 4(b), where the bias of TLAR is significantly lower than that of the LAR method.

Specific to the third example, the number of sensors is elevated from ten to thirty, five at a time, in which one is abnormal, and the other four are normal. σ is set to $3\pi/180$ (3°). Figures 5(a) and 5(b) draw a comparison of the RMSE and bias of various methods. With the increase in the number of sensors, the RMSE and bias are reduced gradually since the numbers of abnormal and normal sensors increase simultaneously and are proportional. It is again proved that the TLAR method exhibits the optimal RMSE and bias performance.

Subsequently, the RMSE and bias performance are determined according to different bearing noise standard deviations when all sensors are normal sensors, as presented in Figures 6(a) and 6(b). In this scenario, the TLS method exhibits the optimal RMSE performance. The RMSE curve of TLAR is slightly higher than that of TLS since the bearing
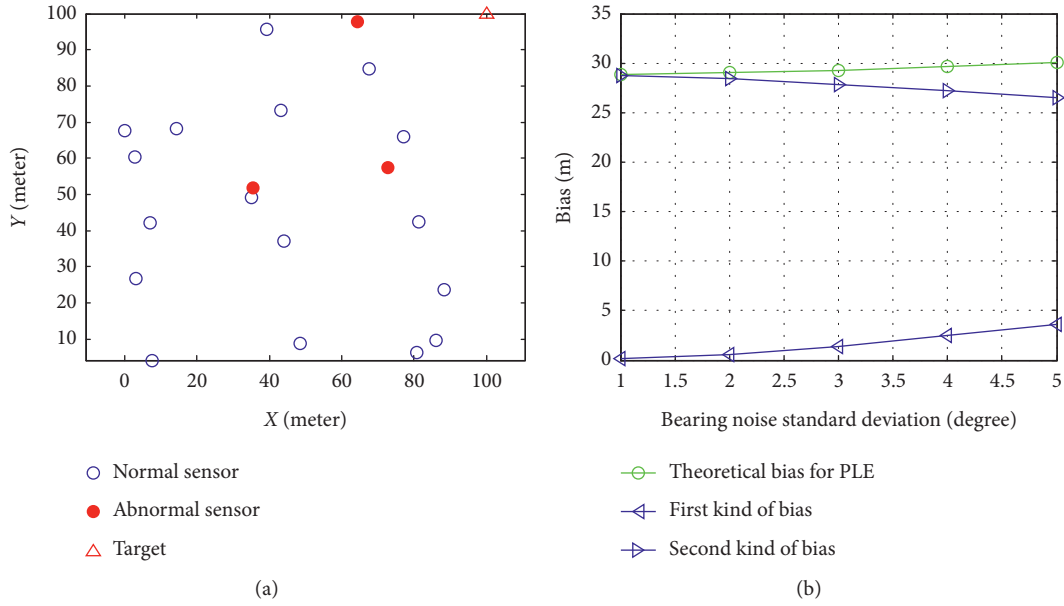
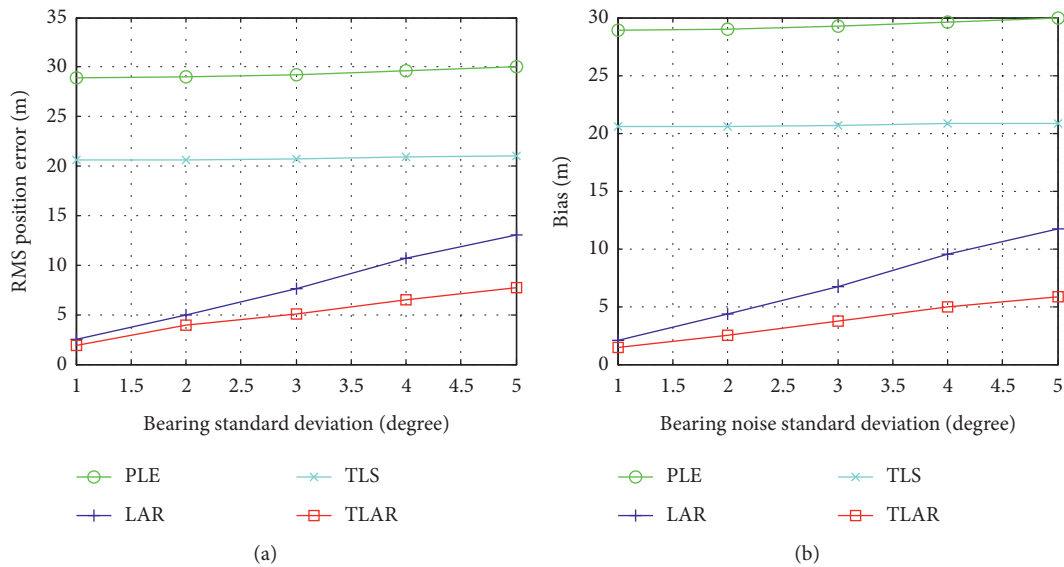FIGURE 3: (a) Locations of sensors and target. (b) Theoretical bias of PLE.

FIGURE 4: Performance comparison of PLE, LAR, TLS, and TLAR estimates with fixed outliers. (a) RMSE results of various estimators. (b) Bias results of various estimators.

measurement errors overall comply with the Gaussian distribution. In such scenario, the L_2-norm criterion is optimal. Unlike the RMSE performance, the bias of TLAR is comparable with that of TLS and much lower than that of LAR or PLE. When all sensors are normal, the LAR and TLAR methods are capable of achieving effective results,

compared with PLE and TLS, thereby demonstrating the robustness of the proposed algorithm.

6.2. Linearly Distributed Sensors. In the present section, all sensors are linearly distributed with the equal spacing from
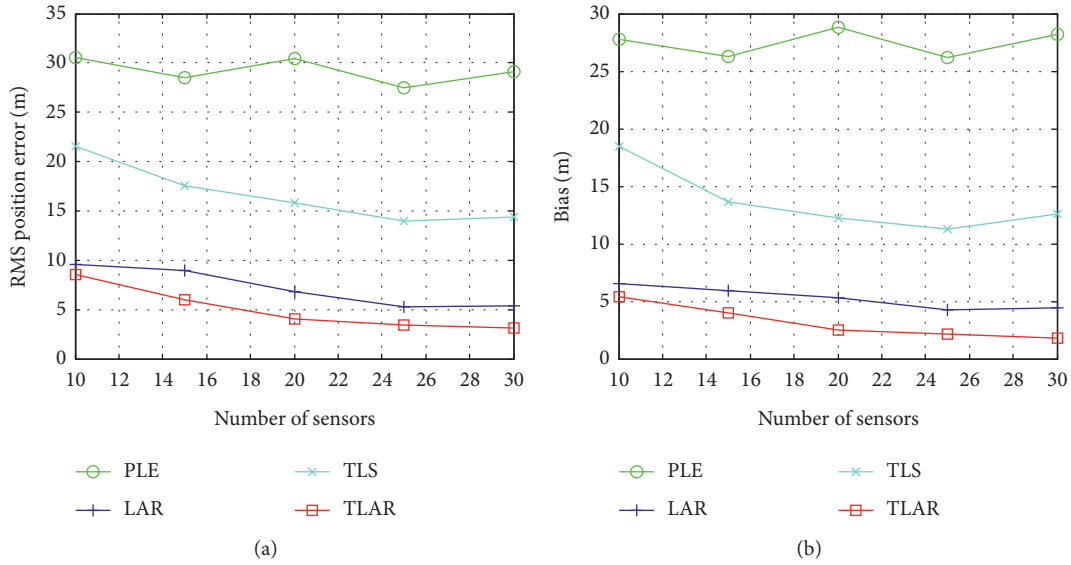
FIGURE 5: Performance comparison of PLE, LAR, TLS, and TLAR estimates with different number of sensors. (a) RMSE results of various estimators. (b) Bias results of various estimators.
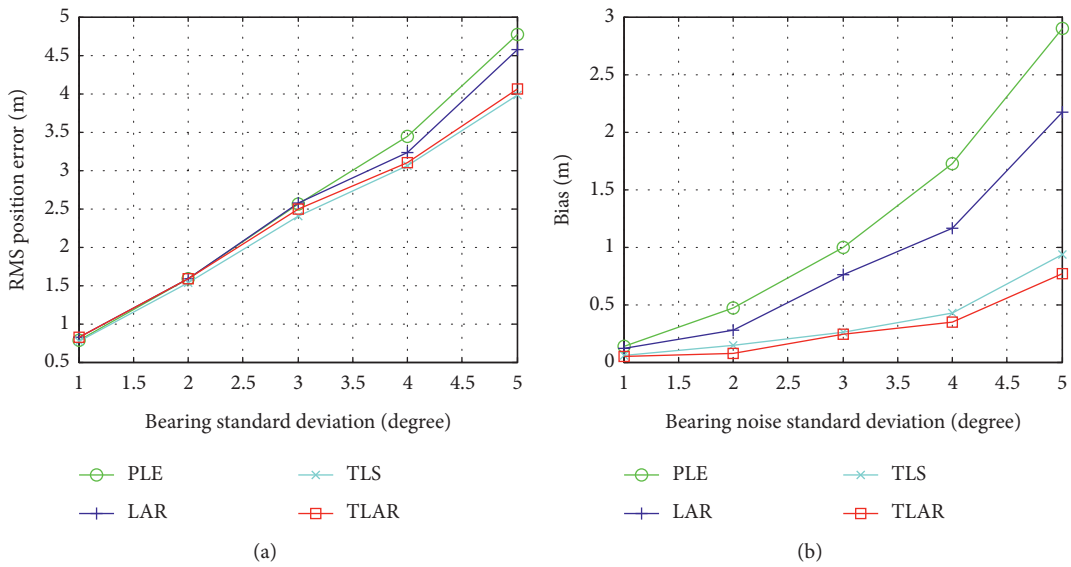

FIGURE 6: Performance comparison of PLE, LAR, TLS, and TLAR estimates without outliers. (a) RMSE results of various estimators. (b) Bias results of various estimators.

(5, 26) m to (155, −34) m (Figure 7(a)). The target is placed at (200, 40) m. In the first example, the number of abnormal sensors ranges from 2 to 6, and that of sensors is set to 31. σ is set at $3\pi/180$ (3°). Figures 8(a) and 8(b) illustrate the RMSE and bias performance when the number of

abnormal sensors is altered. It is therefore demonstrated that the TLAR algorithm exhibits better RMSE and bias performance than the LAR method. As the number of outliers is elevated, the performance is more significantly improved.
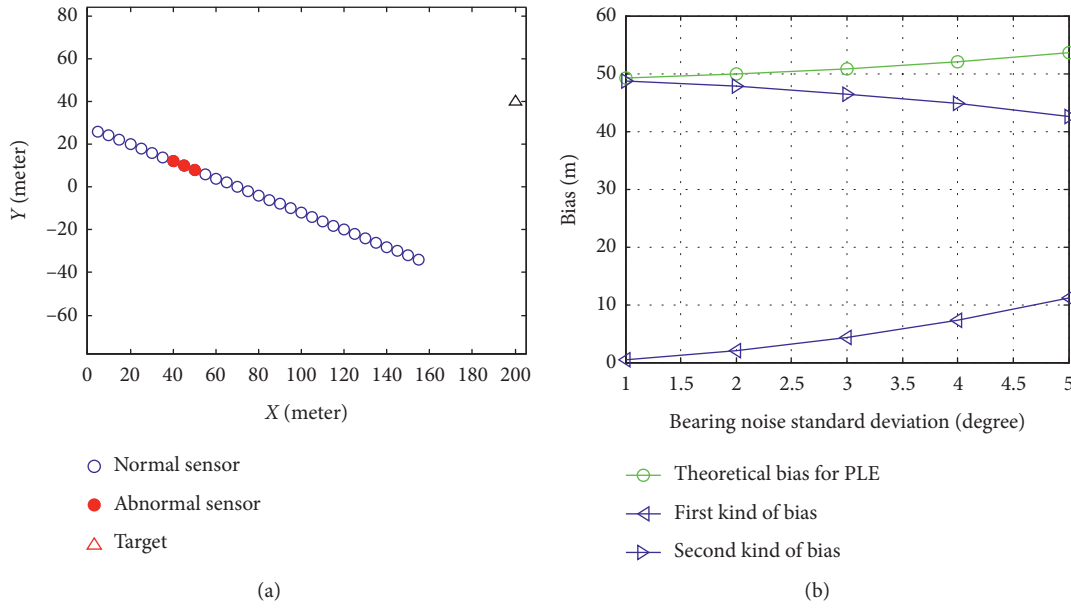
(a)

(b)

Figure 7: (a) Locations of the target, normal sensors, and abnormal sensors. (b) Theoretical bias of PLE.
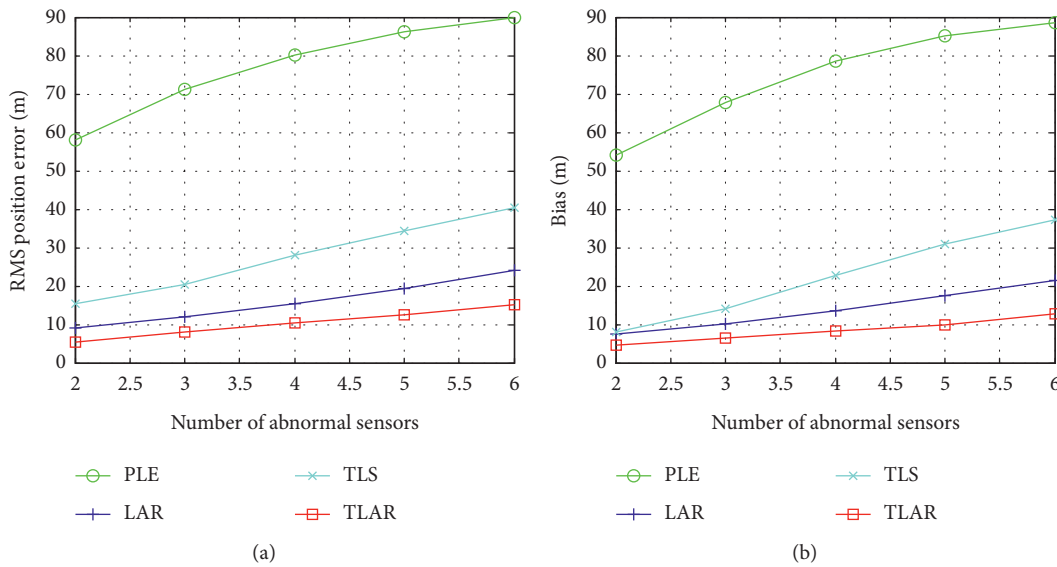

(a)

(b)

Figure 8: Performance comparison of PLE, LAR, TLS, and TLAR estimates with various outliers. (a) RMSE results of various estimators. (b) Bias results of various estimators.

In the next example, all sensors are working properly except for the 8th to 10th sensors. Consistent with Section 6.1, the theoretical bias of PLE is plotted in Figure 7(b). The results comply with those presented in Figure 3(b).

The RMSE and bias results for such example are given in Figures 9(a) and 9(b), respectively. The RMSE value of TLAR is lower than that of the LAR method for the case of linearly distributed sensors, even though only a 0.502 m

reduction for $\sigma = 1°$ and a 4.56 m reduction for $\sigma = 5°$. With the increase in the bearing measurement noise variance, the effect of large residuals turns out to be less significant. The amount of bias reduction is 5.04 m for $\sigma = 5°$ if TLAR is used.

Specific to the third example, the RMSE and bias results are shown in Figures 10(a) and 10(b) when the number of sensors ranges from 16 to 31, five at a time, one of which is
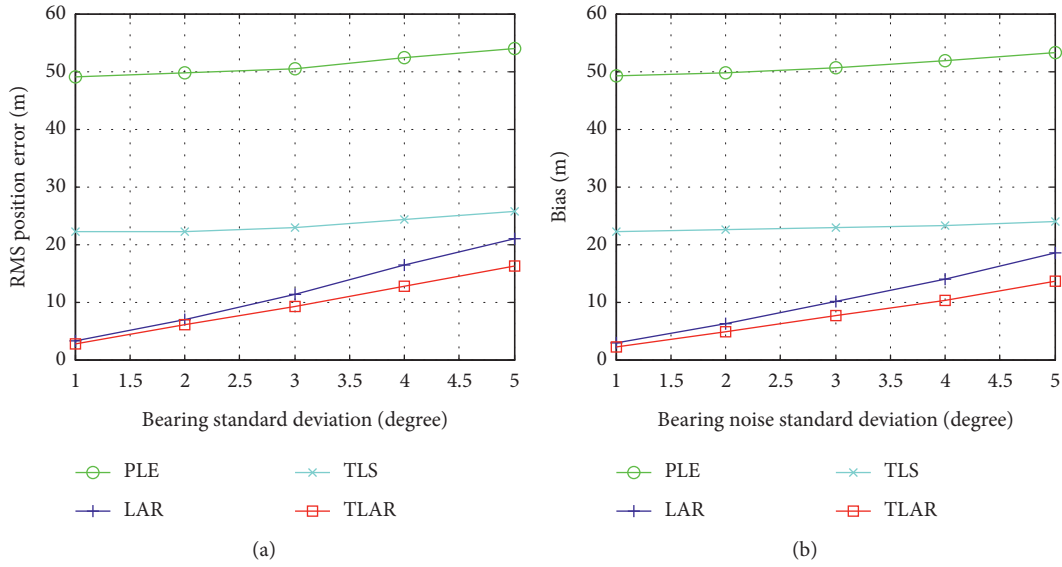
FIGURE 9: Performance comparison of PLE, LAR, TLS, and TLAR estimates with fixed outliers. (a) RMSE results of various estimators. (b) Bias results of various estimators.
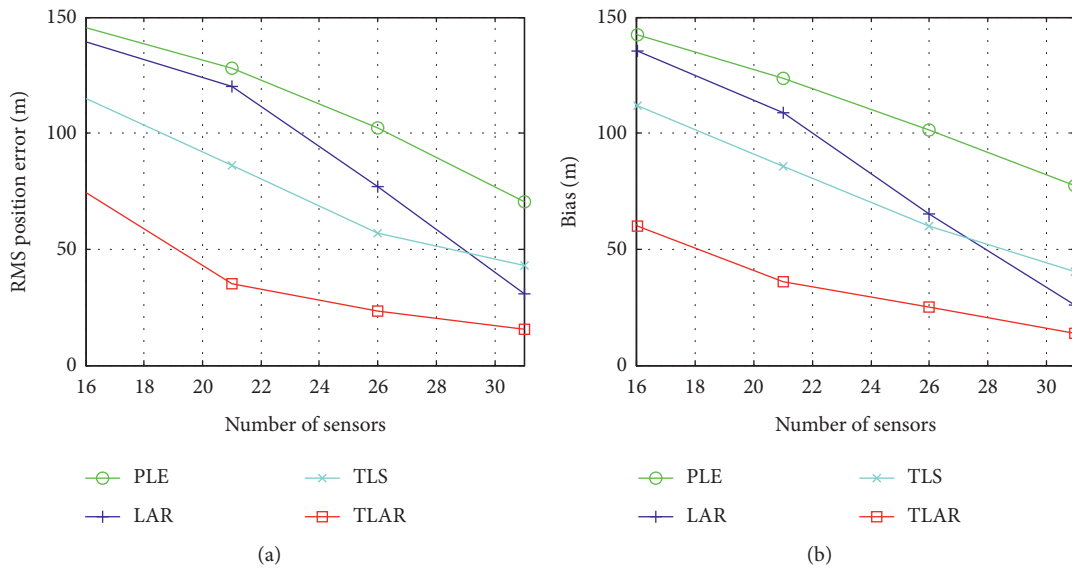

FIGURE 10: Performance comparison of PLE, LAR, TLS, and TLAR estimates with different number of sensors. (a) RMSE results of various estimators. (b) Bias results of various estimators.

abnormal, and the other four are normal. σ is set to $3\pi/180$ ($3°$). Figures 10(a) and 10(b) verify that the solution of TLAR has better RMSE and bias performance than that of PLE, TLS, and LAR. Based on more than 21 sensors, the RMSE and bias of LAR suddenly decrease. This is not unexpected since the sensor observation angles are more significantly discriminated.

In addition, the RMSE and bias curves of PLE, LAR, TLS, and TLAR without abnormal sensors are plotted. Figure 11(a) gives the RMSE results as the bearing standard deviation increases, and Figure 11(b) compares the bias of various solutions. As clearly demonstrated by the RMSE and bias results, TLAR outperforms LAR. The amounts of reduction in terms of RMSE and bias are 1.85 m and 2.845 m, respectively.
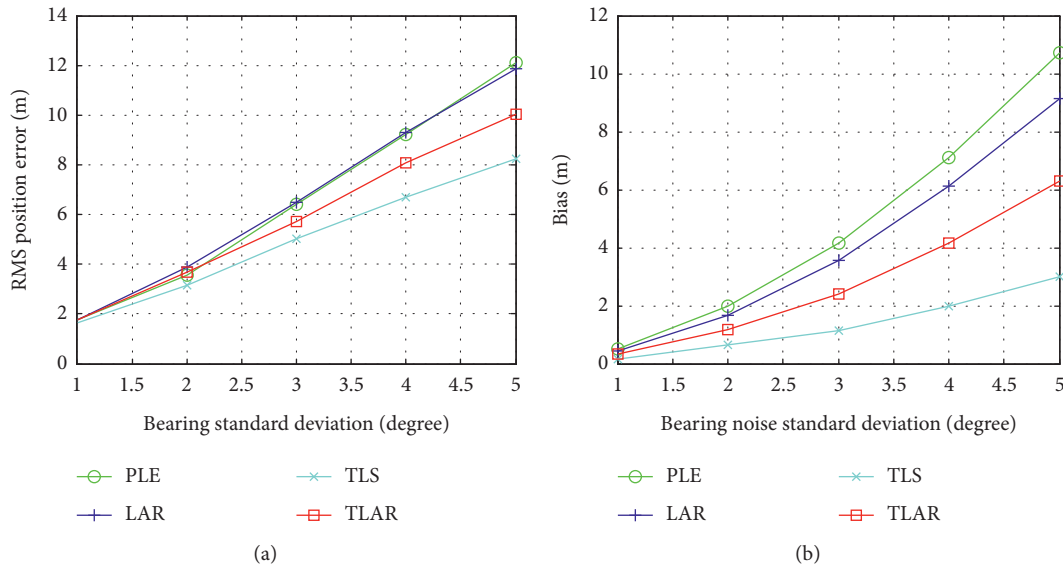
Figure 11: Performance comparison of PLE, LAR, TLS, and TLAR estimates without outliers. (a) RMSE results of various estimators. (b) Bias results of various estimators.

7. Conclusions

The present study presents a TLAR algorithm to solve the BOTL problem in the presence of outlier data. Though the conventional LAR is robust to significant deviations, it has bias formed by the correlation between the system matrix and the noise vector and such bias remains with the increase in the number of sensors. To increase the accuracy of the LAR estimator, the TLAR is proposed by adding minimal perturbations to both system matrix and data vector in least absolute residuals sense, so the perturbed matrix is consistent. As revealed from the experimental results, the TLAR algorithm outperforms the LAR method. When the bearing noise power becomes more significant, the performance of RMSE and bias reduction is more obviously improved.

Data Availability

The simulation data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] L. M. Kaplan and Q. Le, "On exploiting propagation delays for passive target localization using bearings-only measurements," *Journal of the Franklin Institute*, vol. 342, no. 2, pp. 193–211, 2005.

[2] P. Volgyesi, G. Balogh, A. Nadas, C. B. Nash, and A. Ledeczi, "Shooter localization and weapon classification with soldier-wearable networked sensors," in *Proceedings of International Conference on Mobile Systems, Applications and Services*, pp. 113–126, San Juan, Puerto Rico, June 2007.

[3] A. M. Ali, S. Asgari, T. C. Collier et al., "An empirical study of collaborative acoustic source localization," *Journal of Signal Processing Systems*, vol. 57, no. 3, pp. 415–436, 2009.

[4] I. Shames, A. N. Bishop, and B. D. O. Anderson, "Analysis of noisy bearing-only network localization," *IEEE Transactions on Automatic Control*, vol. 58, no. 1, pp. 247–252, 2013.

[5] R. Sharma, R. W. Beard, C. N. Taylor, and S. Quebe, "Graph-based observability analysis of bearing-only cooperative localization," *IEEE Transactions on Robotics*, vol. 28, no. 2, pp. 522–529, 2012.

[6] A. N. Bishop, B. Fidan, B. D. O. Anderson, K. Doğançay, and P. N. Pathirana, "Optimality analysis of sensor-target localization geometries," *Automatica*, vol. 46, no. 3, pp. 479–492, 2010.

[7] N. Garcia, H. Wymeersch, E. G. Larsson, A. M. Haimovich, and M. Coulon, "Direct localization for massive MIMO," *IEEE Transactions on Signal Processing*, vol. 65, no. 10, pp. 2475–2487, 2017.

[8] M. Gavish and A. J. Weiss, "Performance analysis of bearing-only target location algorithms," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 3, pp. 817–828, 1992.

[9] A. N. Bishop, B. D. O. Anderson, B. Fidan, P. N. Pathirana, and G. Guoqiang Mao, "Bearing-only localization using geometrically constrained optimization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 1, pp. 308–320, 2009.

[10] Z. Wang, J.-A. Luo, and X.-P. Zhang, "A novel location-penalized maximum likelihood estimator for bearing-only target localization," *IEEE Transactions on Signal Processing*, vol. 60, no. 12, pp. 6166–6181, 2012.

[11] J.-A. Luo, X.-H. Shao, D.-L. Peng, and X.-P. Zhang, "A novel subspace approach for bearing-only target localization," *IEEE Sensors Journal*, vol. 19, no. 18, pp. 8174–8182, 2019.

[12] Q. Liu, C. Yang, Y. Gu, and H. C. So, "Robust sparse recovery via weakly convex optimization in impulsive noise," *Signal Processing*, vol. 152, pp. 84–89, 2018.

[13] A. Swami and B. Sadler, "TDE DOA and related parameter estimation problems in impulsive noise," in *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics*, pp. 273–277, Banff, Alberta, Canada, July 1997.

[14] R. J. Kozick and B. M. Sadler, "Maximum-likelihood array processing in non-gaussian noise with gaussian mixtures," *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3520–3535, 2000.

[15] Q. Liu, Y. Gu, and H. C. So, "DOA estimation in impulsive noise via low-rank matrix approximation and weakly convex optimization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 6, pp. 3603–3616, 2019.

[16] Q. Liu, J. Xu, Z. Ding, and H. C. So, "Target localization with jammer removal using frequency diverse array," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11685–11696, 2020.

[17] J. S. Picard and A. J. Weiss, "Bounds on the number of identifiable outliers in source localization by linear programming," *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2884–2895, 2010.

[18] Y. Liu, Y. H. Hu, and Q. Pan, "Distributed, robust acoustic source localization in a wireless sensor network," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4350–4359, 2012.

[19] Y. F. Xiong, N. Wu, H. Wang, and J. M. Kuang, "Joint localization and cooperative detection in location-aware wireless networks in the presence of ranging outliers," in *Proceedings of 2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications*, pp. 1–5, Edinburgh, UK, June 2016.

[20] B. R. A. Maronna, D. R. Martin, and V. J. Yohai, *Robust Statistics: Theory and Methods*, Wiley, Hoboken, NJ, USA, 2006.

[21] Y. Jiang and M. R. Azimi-Sadjadi, "A robust source localization algorithm applied to acoustic sensor network," in *Proceedings of 2007 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1233–1236, Honolulu, HI, USA, April 2007.

[22] T. Panigrahi, G. Panda, B. Mulgrew, and B. Majhi, "Robust incremental LMS over wireless sensor network in impulsive noise," in *Proceedings of 2010 International Conference on Computational Intelligence and Communication Networks*, pp. 205–209, Bhopal, India, November 2010.

[23] M. Hladík and M. Černý, "Total least squares and chebyshev norm," *Procedia Computer Science*, vol. 51, no. 1, pp. 1791–1800, 2015.

[24] Y. Chen, H. C. So, and E. E. Kuruoglu, "Variance analysis of unbiased least lp-norm estimator in non-Gaussian noise," *Signal Processing*, vol. 122, pp. 190–203, 2016.

[25] H. Wu, S. Chen, Y. Zhang, H. Zhang, and J. Ni, "Robust structured total least squares algorithm for passive location," *Journal of Systems Engineering and Electronics*, vol. 26, no. 5, pp. 946–953, 2015.

[26] A. Agresti, *Foundation of Linear and Generalized Linear Models*, Wiley, " Hoboken, NJ, USA, 2015.

[27] J. R. Rice and J. S. White, "Norms for smoothing and estimation," *SIAM Review*, vol. 6, no. 5, pp. 243–256, 1964.

[28] Q. Ke and T. Kanade, "Robust L1 norm factorization in the presence of outliers and missing data by alternative convex programming," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 739–746, San Diego, CA, USA, June 2005.

[29] A. G. Lindgren and K. F. Gong, "Position and velocity estimation via bearing observations," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-14, no. 4, pp. 564–577, 1978.

[30] K. Dogancay, "Bias compensation for the bearings-only pseudolinear target track estimator," *IEEE Transactions on Signal Processing*, vol. 54, no. 1, pp. 59–68, 2006.

[31] Y. J. Zhang and G. Z. Xu, "Bearings-only target motion analysis via instrumental variable estimation," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 5523–5533, 2010.

[32] K. Doğançay, "Passive emitter localization using weighted instrumental variables," *Signal Processing*, vol. 84, no. 3, pp. 487–497, 2004.

[33] G. Gu, "A novel power-bearing approach and asymptotically optimum estimator for target motion analysis," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 912–922, 2011.

[34] M. R. Osborne and G. A. Watson, "An analysis of the total approximation problem in separable norms, and an algorithm for the total l_1 problem," *SIAM Journal on Scientific and Statistical Computing*, vol. 6, no. 2, pp. 410–424, 1985.

[35] K. Doğançay, "On the bias of linear least squares algorithms for passive target localization," *Signal Processing*, vol. 84, no. 3, pp. 475–486, 2004.

[36] C. Marx, "On resistant Lp-norm estimation by means of iteratively reweighted least squares," *Journal of Applied Geodesy*, vol. 7, no. 1, pp. 1–10, 2013.

[37] J. Shi, K. Chen, and W. Song, "Robust errors-in-variables linear regression via Laplace distribution," *Statistics & Probability Letters*, vol. 84, pp. 113–120, 2014.

[38] P. Bloomfield and W. L. Steiger, *Least Absolute Deviations, Theory, Applications and Algorithms*, Birkhäuser, Basel, Switzerland, 1983.

[39] M. Grant and S. Boyd, "CVX user's guide for CVX version 1.21," 2010.

[40] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2010.