Article

A Total Lp-Norm Optimization for Bearing-Only Source Localization in Impulsive Noise with $S\alpha S$ Distribution

Ji-An Luo ¹ , Chang-Cheng Xue ¹, Ying-Jiao Rong ^{2,*} and Shen-Tu Han ¹

¹ School of Automation, Hangzhou Dianzi University, Hangzhou 310018, China; luojian@hdu.edu.cn (J.-A.L.); xcc@hdu.edu.cn (C.-C.X.); hansentu@hdu.edu.cn (S.-T.H.)

² Science and Technology on Near-Surface Detection Laboratory, Wuxi 214035, China

* Correspondence: ryingjiao@hdu.edu.cn

Abstract: This paper considers the problem of robust bearing-only source localization in impulsive noise with symmetric α -stable distribution based on the Lp-norm minimization criterion. The existing Iteratively Reweighted Pseudolinear Least-Squares (IRPLS) method can be used to solve the least LP-norm optimization problem. However, the IRPLS algorithm cannot reduce the bias attributed to the correlation between system matrices and noise vectors. To reduce this kind of bias, a Total Lp-norm Optimization (TLPO) method is proposed by minimizing the errors in all elements of system matrix and data vector based on the minimum dispersion criterion. Subsequently, an equivalent form of TLPO is obtained, and two algorithms are developed to solve the TLPO problem by using Iterative Generalized Eigenvalue Decomposition (IGED) and Generalized Lagrange Multiplier (GLM), respectively. Numerical examples demonstrate the performance advantage of the IGED and GLM algorithms over the IRPLS algorithm.

Citation: Luo, J.-A.; Xue, C.-C.; Rong, Y.-J.; Han, S.-T. A Total Lp-Norm Optimization for Bearing-Only Source Localization in Impulsive Noise with $S\alpha S$ Distribution. *Sensors* **2021**, *21*, 6471. <https://doi.org/10.3390/s21196471>

Academic Editors: Mahendra Mallick and Ratnasingham Tharmarasa

Received: 30 August 2021

Accepted: 23 September 2021

Published: 28 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: bearing-only; source localization; robust estimation; least Lp-norm; total Lp-norm optimization

1. Introduction

Bearing-Only Source Localization (BOSL) using spatially distributed sensors can be widely applied in network localization [1], vehicle localization [2], gunshot localization [3], animal behavior monitoring [4] and rigid body localization [5], to name but a few. The problem of BOSL is to estimate the source location from a set of noise-corrupted bearing measurements where its main challenge originates in the highly nonlinear nature of the angle observations with the true source position. Under the assumption of Gaussian measurement noise, many estimation approaches have been proposed to handle the non-linearity, including the grid search method, the pseudolinear estimators [6,7], the iterative maximum likelihood methods [8] and the subspace approaches [9]. However, in the complex field environment, the sensor is vulnerable to external interference, enemy attack or node failure. The bearing measurements may suffer impulse noise [10–13] and those outlier data can degrade the localization performance dramatically. Therefore, it is necessary to develop new estimators that are robust to impulsive noise.

In fact, non-Gaussian models corresponding to impulsive noise have been extensively studied in the literature [13–15]. These studies have demonstrated that the Symmetric α -Stable ($S\alpha S$) distribution is more suitable to model impulsive noise than the Gaussian distribution. The family of stable distribution is a generalization of the Gaussian distribution under the stable law, including a large range of distributions with mutable values of impulsiveness (α), skewness (β) and dispersion (γ). In particular, two special cases of α -Stable distribution can be obtained by letting impulsiveness parameter α take values of 1 and 2. One is Cauchy distribution ($\alpha = 1$), and the other is Gaussian distribution ($\alpha = 2$).

In the present study, we focus on the problem of robust BOSL with impulsive noise modeled as $S\alpha S$ distribution. In this situation, the methods derived using L2-norm opti-

mization exhibit unreliable estimates because they are not robust to outliers. Therefore, robust statistics need to be used to improve the localization performance. Under the assumption of $S\alpha S$ measurement noise, a Maximum Likelihood Estimator (MLE) was proposed in [16] to produce optimal estimates. Unfortunately, no closed-form expressions exist for the general likelihood function in the cases of $1 < \alpha < 2$ and therefore no explicit solution for MLE is possible. Although the $S\alpha S$ distribution does not have finite variance, it has finite Fractional Lower Order Moments (FLOM) [13], which can be calculated by its dispersion γ and its characteristic exponent α . Moreover, minimizing the FLOM of estimation errors is equivalent to minimizing the dispersion. It is well known as the minimum dispersion criterion, which minimizes the L_p -norm of the estimation residuals. Unlike L_2 -norm minimization, the least L_p -norm estimator ($1 < p < 2$) does not have a closed-form solution and consequently needs to be solved in an iterative manner [17].

The least L_p -norm estimator belongs to the M-estimator. The main idea of M-estimate is to replace the sum of squares of least squares residuals with cost functions more robust to impulsive noise, so as to reduce the sensitivity of estimators with respect to model errors. These functions include Huber [18], Bi-square [19], the negative log-likelihood of the Cauchy distribution [20], Wilcoxon [21], L_1 -norm [22,23], L_p -norm [24] and L_∞ -norm [25], etc. Ref. [19] presented a distributed robust localization algorithm based on energy information for sensor networks. The algorithm uses Bi-square function as the cost function of M-estimate. A distributed incremental least mean square algorithm based on Wilcoxon norm was proposed in [21] for parameter estimation of sensor networks. Ref. [26] proposed a robust structure total least-squares algorithm for BOSL by using the improved Danish weight function to suppress the impact of outlier data on the localization performance. In addition to the M-estimator, there are other algorithms that can handle outlier data, such as outlier detection [27], clustering [28] and game theoretic techniques [29]. The outlier detection method [27] is to detect suspected outlier data first, and separate it from the original data set. The clustering based techniques [28] can be used to classify normal and abnormal sensors. In game theory [29], the defense strategies could be employed to detect the outlier data and adaptive threshold selection can be also introduced.

Recently, an Iteratively Reweighted Pseudolinear Least-Squares (IRPLS) method was proposed in [14] to reduce biases attributed to the impulsive noise. However, IRPLS still suffers from a major bias problem caused by the correlation between system matrix and the noise vector. This bias can be reduced by exploiting the use of an Instrumental-Variable (IV) matrix [14] that is approximately uncorrelated with the noise vector. Inconsistent with the IV method, we present a robust total least-squares method using L_p -norm minimization that can reduce biases by minimizing the errors in the system matrix and the data vector simultaneously. We first formulate the problem of BOSL subjected to impulsive noise modeled as $S\alpha S$ distribution and review the pseudo-linear measurement model for BOSL. Next, we present the Total L_p -norm Optimization (TLPO) method for BOSL to minimize the errors in the system matrix and the data vector under the minimum dispersion criterion. Two algorithms, named Iterative Generalized Eigenvalue Decomposition (IGED) and Generalized Lagrange Multiplier (GLM), are designed to solve the TLPO problem. The performance advantage of the proposed algorithms is demonstrated by numerical simulations in terms of both bias and Root-Mean-Square-Error (RMSE) performance. The main contributions of the proposed method can be summarized as follows:

- Development of a new bias reduced estimator based on TLPO for BOSL when the measurement noise is modeled as $S\alpha S$ distribution;
- Development of two algorithms for TLPO optimization using the IGED approach and the GLM method, respectively.

The rest of this paper is organized below. Section 2 briefly summarizes the $S\alpha S$ distribution, presents the measurement model for BOSL and discusses the nonlinear least L_p -norm for BOSL. In Section 3, the pseudolinear estimator and the iteratively reweighted pseudolinear least-squares algorithm are reviewed. In Section 4, a new TLPO method

is presented, and two iterative algorithms are developed to solve the TLPO problem. Section 5 presents the Cramér-Rao Lower Bound and Section 6 illustrates the bias and RMSE performance of PLE, TLS, LAR and TLAR using various numerical examples. Lastly, conclusions are drawn in Section 7.

2. Lp-Norm Optimization for Robust BOSL

2.1. Symmetric Alpha-Stable Distribution

The impulsive noise is more likely to exhibit outliers than normal noise. Studies [13] have shown that $S\alpha S$ distribution can model the impulsive noise better than Gaussian distribution due to the reason that the $S\alpha S$ densities have heavier tails than the Gaussian density. The characteristic function of $S\alpha S$ distribution is described as:

$$\chi(\tau) = \exp\{i\delta\tau - \gamma|\tau|^\alpha\} \quad (1)$$

where α ($0 < \alpha \leq 2$) denotes the characteristic exponent, γ indicates the dispersion parameter and δ stands for the location parameter. The $S\alpha S$ distribution is completely determined by these three parameters. The value of α indicates the heaviness of the tails of the density function. A small positive value of α implies high impulsiveness, while a value of α close to 2 shows a type of Gaussian-like shape. The dispersion γ performs like the variance and δ controls the location of the density function.

A $S\alpha S$ distribution is called standard if $\delta = 0$, $\gamma = 1$. Let x be a random variable that follows the standard $S\alpha S$ distribution. By taking the inverse Fourier transform of its characteristic function, the density function of x is of the form:

$$f_\alpha(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(-ix\tau - \gamma|\tau|^\alpha) d\tau \quad (2)$$

It is known that a $S\alpha S$ distribution with characteristic exponent α only has finite moments for orders less than α , which are called the Fractional Lower Order Moments (FLOM). In particular, the FLOM of a standard $S\alpha S$ random variable x is given by

$$\mathbb{E}\{|x|^p\} = C(p, \alpha) \gamma^{p/\alpha}, \quad 0 < p < \alpha \quad (3)$$

where $\mathbb{E}\{\cdot\}$ denotes expectation operator,

$$C(p, \alpha) = \frac{2^{p+1} \Gamma\left(\frac{p+1}{2}\right) \Gamma\left(-\frac{p}{\alpha}\right)}{\alpha \sqrt{\pi} \Gamma\left(-\frac{p}{2}\right)} \quad (4)$$

is a constant that depends only on q and α and $\Gamma(a) = \int_0^{+\infty} y^{(a-1)} \exp(-y) dy$ is the Gamma function. It is worth mentioning that the linear space of a $S\alpha S$ process is a Banach space for $\alpha \in [1, 2)$, and it is only a metric space for $\alpha \in (0, 1)$ [30]. Therefore, the tools of Hilbert space are not applicable when one solves a linear estimation problem with the $S\alpha S$ process. In the present study, we only focus on the case of $\alpha \in (1, 2)$.

2.2. Measurement Model

The problem of robust BOSL is depicted in Figure 1, where $\mathbf{t} = [t_x, t_y]^T$ denotes the unknown target location vector, $\mathbf{r}_m = [r_{x,m}, r_{y,m}]^T$ represents the m th sensor location vector and θ_m is the true bearing at sensor m . The nonlinear relationship between θ_m , $\mathbf{t}$ and $\mathbf{r}_m$ is given by [31]

$$\theta_m = \tan^{-1}(t_y - r_{y,m}, t_x - r_{x,m}), \quad \theta_m \in [0, 2\pi) \quad (5)$$

where $\tan^{-1}$ denotes the two-argument inverse tangent function and $m = 1, 2, \dots, M$.

Figure 1. Illustration of the BOSL system.

The bearing measurement taken at sensor m is given by

$$\tilde{\theta}_m = \theta_m + n_m \quad (6)$$

where n_m follows independent zero mean $S\alpha S$ distribution with γ_m representing the dispersion parameter and α accounting for the characteristic exponent. The dispersion γ_m can vary with m . If the noise dispersion γ_m is known *a priori*, then the noise term n_m can be normalized. Let $e_m = \gamma_m^{-1/\alpha} n_m$ denote the normalized noise that has unit dispersion. The normalized measurement equation can be written as

$$\tilde{\phi}_m = \phi_m + e_m \quad (7)$$

where $\tilde{\phi}_m = \gamma_m^{-1/\alpha} \tilde{\theta}_m$ is the normalized bearing measurement and $\phi_m = \gamma_m^{-1/\alpha} \theta_m$ is the normalized true bearing. Stacking (7) in the vector form yields

$$\tilde{\boldsymbol{\phi}} = \boldsymbol{\phi} + \boldsymbol{e} \quad (8)$$

where $\tilde{\boldsymbol{\phi}} = [\tilde{\phi}_1, \tilde{\phi}_2, \dots, \tilde{\phi}_M]^T$ denotes the normalized measurement vector, $\boldsymbol{\phi} = [\phi_1, \phi_2, \dots, \phi_M]^T$ represents the true bearing vector and $\boldsymbol{e} = [e_1, e_2, \dots, e_M]^T$ indicates the measurement noise vector.

2.3. Minimum Dispersion Criterion

The main difficulty for parameter estimation with the α -stable process is that all non-Gaussian α -stable distributions have infinite variance. As such, the traditional nonlinear least-squares techniques [32], which rely on the second order moments, are not suitable for solving BOSL problems with impulsive noise. Fortunately, we can use the minimum dispersion criterion instead of minimizing the variance. To do this, we first define the norm of the $S\alpha S$ random variable x as

$$\|x\|_\alpha = \gamma^{\frac{1}{\alpha}}, \quad \alpha \in [1, 2) \quad (9)$$

Thus, a suitable measure of the dispersion could be $\gamma = \|x\|_\alpha^\alpha$. Combining (3) and (9), one can easily obtain

$$\left(\frac{\mathbb{E}\{|x|^p\}}{C(p, \alpha)}\right)^{\frac{1}{p}} \equiv \|x\|_\alpha, \quad p \in (0, \alpha), \quad \alpha \in [1, 2) \quad (10)$$

Therefore, we may estimate the source location by solving the following nonlinear Lp-norm optimization problem:

$$\hat{\mathbf{t}} = \arg \min_{\mathbf{t}} J(\mathbf{t}), \quad J(\mathbf{t}) = \|\tilde{\boldsymbol{\phi}} - \boldsymbol{\phi}\|_p^p = \sum_{m=1}^M |\tilde{\phi}_m - \phi_m(\mathbf{t})|^p \quad (11)$$

where $p \in (1, \alpha)$. The definition of the Lp-norm of the vector $\boldsymbol{\zeta}$ is

$$\|\boldsymbol{\zeta}\|_p = \left(\sum_k |\zeta_k|^p\right)^{1/p} \quad (12)$$

In the present study, we do not consider the case of $0 < p \leq 1$, since the corresponding Lp-norm cost function is not differentiable. For non-normalized bearing measurements, the Lp-norm objective function in (11) becomes

$$J(\mathbf{t}) = \sum_{m=1}^M |\gamma_m^{-1/\alpha}(\tilde{\theta}_m - \theta_m(\mathbf{t}))|^p \quad (13)$$

The optimization problem listed in (11) can be solved numerically for a given 2D space of interest. To be more specific, we can perform a grid search over that 2D space. The global solution of maximum likelihood (ML) estimator is guaranteed for the given set of data as long as the spacing between grids is small enough. However, if the range of the parameter of interest is not confined to a relatively small interval or the dimension of unknown parameter vector is high, the grid search approach is computationally infeasible. Instead, one may resort to iterative optimization methods, such as gradient decent, and Gauss–Newton [33], etc. In particular, the Gauss–Newton algorithm is

$$\hat{\mathbf{t}}^{(i+1)} = \hat{\mathbf{t}}^{(i)} - \left(\nabla J(\mathbf{t})^T \nabla J(\mathbf{t})\right)^{-1} \nabla J(\mathbf{t})^T (\tilde{\boldsymbol{\phi}} - \boldsymbol{\phi}(\hat{\mathbf{t}}^{(i)})) \quad (14)$$

where $\nabla J(\mathbf{t})$ denotes the Jacobian matrix,

$$\nabla J(\mathbf{t}) = p \sum_{m=1}^M \frac{\partial \phi_m(\mathbf{t})}{\partial \mathbf{t}} |e_m|^{p-1} \text{sign}(e_m) \quad (15)$$

If $e_m > 0$, $\text{sign}(e_m) = 1$ and -1 otherwise. The statistical properties of the nonlinear Lp-norm minimizer have been studied in the literature [24]. The theoretical covariance $\mathbf{B}$ is related to the value of p , and it is given by [24]:

$$\mathbf{B} \approx \frac{C(2p-2, \alpha)}{(p-1)^2 C^2(p-2, \alpha)} (\nabla \boldsymbol{\phi}^T \nabla \boldsymbol{\phi})^{-1} \quad (16)$$

where $\nabla \boldsymbol{\phi}$ denotes the Jacobian matrix of $\boldsymbol{\phi}$ with respect to $\mathbf{t}$. By minimizing the scalar term of $\mathbf{B}$, we obtain the optimal choice of p

$$p^o = \arg \min_p \frac{C(2p-2, \alpha)}{(p-1)^2 C^2(p-2, \alpha)} \quad (17)$$

where p^o denotes the optimal value of p .

3. Pseudolinear Lp-Norm Minimization

3.1. Pseudolinear Estimator

The robust BOSL problem is nontrivial because the measurement Equation in (7) is nonlinearly related to the unknown source location. An attractive solution is to set up a pseudolinear equation by lumping the nonlinearities into the noise term. As illustrated in Figure 1, an orthogonal vector sum between the measured angle vector and the true angle vector can be geometrically described by

$$\mathbf{u}_m = \tilde{\mathbf{u}}_m + \boldsymbol{\varepsilon}_m = \mathbf{t} - \mathbf{r}_m \quad (18)$$

where $\mathbf{u}_m$ is the true angle vector between $\mathbf{r}_m$ and $\mathbf{t}$, $\tilde{\mathbf{u}}_m$ is the measured angle vector starting from $\mathbf{r}_m$ and produces the noisy bearing $\tilde{\theta}_m$ with respect to the horizontal direction and $\boldsymbol{\varepsilon}_m$ is the error vector. Let $\boldsymbol{\mu}_m = [\cos \tilde{\theta}_m, \sin \tilde{\theta}_m]^T$ and $\mathbf{v}_m = [\sin \tilde{\theta}_m, -\cos \tilde{\theta}_m]^T$ denote two orthogonal unit trigonometric vectors. Then, $\tilde{\mathbf{u}}_m$ and $\boldsymbol{\varepsilon}_m$ are given by

$$\tilde{\mathbf{u}}_m = \|\mathbf{u}_m\| \cos n_m \cdot \boldsymbol{\mu}_m \quad (19)$$

$$\boldsymbol{\varepsilon}_m = \|\mathbf{u}_m\| \sin n_m \cdot \mathbf{v}_m \quad (20)$$

where $\|\cdot\|$ denotes Euclidean norm. Note from (19) that $\tilde{\mathbf{u}}_m^T \mathbf{v}_m = 0$. Substituting (19) and (20) into (18) and multiplying (18) with $\mathbf{v}_m^T$ yields

$$\zeta_m = \mathbf{v}_m^T \mathbf{t} - \mathbf{v}_m^T \mathbf{r}_m \quad (21)$$

where $\zeta_m = \|\mathbf{u}_m\| \sin n_m$ is a nonlinear transformed measurement error. Collecting the pseudolinear equation errors as a vector $\boldsymbol{\zeta} = [\zeta_1, \zeta_2, \dots, \zeta_K]^T$, we obtain

$$\boldsymbol{\zeta} = \mathbf{A}\mathbf{t} - \mathbf{b} \quad (22)$$

where $\mathbf{A} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_M]^T$, $\mathbf{b} = [\mathbf{v}_1^T \mathbf{r}_1, \dots, \mathbf{v}_M^T \mathbf{r}_M]^T$ are the measurement matrix and vector, respectively.

The PLE requires that $\mathbf{t}$ be estimated by minimizing $\|\mathbf{A}\mathbf{t} - \mathbf{b}\|_2^2$ with respect to $\mathbf{t}$ in an L2-norm optimization sense, and the solution is

$$\hat{\mathbf{t}}_{\text{PLE}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} \quad (23)$$

The above PLE solution has a large bias because the abnormal nodes can produce significantly large bearing measurement errors. As a typical robust estimation algorithm, Lp-norm minimization has been widely used for parameter estimation using measurements corrupted by impulse noise. In this paper, we aim to develop robust estimators using Lp-norm optimization to reduce the bias as much as possible.

3.2. Iteratively Reweighted Pseudolinear Least-Squares Algorithm

This subsection reviews an Iteratively Reweighted Pseudolinear Least-Squares (IRPLS) algorithm, which has been proposed in [14]. The IRPLS algorithm is derived from the following Lp-norm optimization problem:

$$\min_{\mathbf{t}} \|\mathbf{G}(\mathbf{A}\mathbf{t} - \mathbf{b})\|_p^p = \min_{\mathbf{t}} \sum_{m=1}^M |g_m((\mathbf{A}\mathbf{t})_m - b_m)|^p \quad (24)$$

where $\mathbf{G} = \text{diag}(g_1, g_2, \dots, g_M)$, $g_m = (\gamma_m^{1/\alpha} \|\mathbf{t} - \mathbf{r}_m\|)^{-1}$, $(\mathbf{A}\mathbf{t})_m$ and b_m are the m -th element of $\mathbf{A}\mathbf{t}$ and $\mathbf{b}$, respectively. When the measurement noise is small, we have $\sin n_m \approx n_m$. Then, $|g_m((\mathbf{A}\mathbf{t})_m - b_m)|^p$ can be expressed approximately as

$$|g_m((\mathbf{A}\mathbf{t})_m - b_m)|^p \approx \gamma_m^{-p/\alpha} \|\mathbf{t} - \mathbf{r}_m\|^{-2} |\tilde{\theta}_m - \theta_m|^{(p-2)} |(\mathbf{A}\mathbf{p})_m - b_m|^2 \quad (25)$$

The above interpretation suggests an iterative reweighted pseudolinear least-squares with the weight matrix for the i -th iteration given by

$$\mathbf{W}^{(i)} = \text{diag} \left(\gamma_1^{-\frac{p}{2\alpha}} \|\hat{\mathbf{t}}^{(i-1)} - \mathbf{r}_1\|^{-1} |\tilde{\theta}_1 - \hat{\theta}_1^{(i-1)}|^{\frac{p-2}{2}}, \dots, \gamma_M^{-\frac{p}{2\alpha}} \|\hat{\mathbf{t}}^{(i-1)} - \mathbf{r}_M\|^{-1} |\tilde{\theta}_M - \hat{\theta}_M^{(i-1)}|^{\frac{p-2}{2}} \right) \quad (26)$$

where $\hat{\mathbf{t}}^{(i-1)}$ and $\hat{\theta}_m^{(i-1)}$ denote the estimated source location and the m th bearing obtained from the previous iteration $i - 1$. Using the weight matrix, the weight error for the k -th iteration is given by

$$(\mathbf{W}^{(i)} \boldsymbol{\zeta}^{(i)})^T \mathbf{W}^{(i)} \boldsymbol{\zeta}^{(i)} = \sum_{m=1}^M |g_m((\mathbf{A}\mathbf{t})_m - b_m)|^p \quad (27)$$

The homotopy method can be applied in the iterations by starting with a value for p equals to 2 (the weight matrix is unit matrix) and decreasing it each iteration until it reaches the designated value. Thus,

$$p^{(i)} = \max(p, \kappa p^{(i-1)}) \quad (28)$$

where $\kappa < 1$ is the homotopy parameter. The source location can be estimated by performing least-squares estimation:

$$\hat{\mathbf{t}}^{(i)} = \left((\mathbf{W}^{(i)} \mathbf{A})^T \mathbf{W}^{(i)} \mathbf{A} \right)^{-1} (\mathbf{W}^{(i)} \mathbf{A})^T \mathbf{W}^{(i)} \mathbf{b} \quad (29)$$

The IRPLS algorithm is stopped when $\|\hat{\mathbf{t}}^{(i-1)} - \hat{\mathbf{t}}^{(i)}\| \leq \epsilon$, where ϵ is a tolerance parameter.

Finally, the whole process of the IRPLS algorithm is summarized in Algorithm 1.

Algorithm 1 The IRPLS algorithm.

- 1: **Initialization:** Set $\hat{\mathbf{t}}^{(0)} = \hat{\mathbf{t}}_{\text{PLE}}$.
 - 2: **for** $i = 0; i++$ **do**
 - 3: Compute the weight matrix using (31).
 - 4: Determine the value of p for each iteration using (28).
 - 5: Perform weighted least squares estimation using (29).
 - 6: if $\|\hat{\mathbf{t}}^{(i-1)} - \hat{\mathbf{t}}^{(i)}\| \leq \epsilon$, we obtain the final solution, and stop the iteration. Otherwise $i = i + 1$, go to step 3.
 - 7: **end for**
-

The Lp-norm criterion ensures that the cost function in (24) gives less weight to large deviations, and therefore reduces the bias formed by pseudolinear errors with large residuals. However, the IRPLS algorithm implicitly assumes that only $\mathbf{b}$ 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 $\mathbf{b}$ causes the IRPLS estimator to be inconsistent. In this subsection, we analyze such bias attributed to the correlation between $\mathbf{A}$ and $\mathbf{b}$.

Let $\hat{\mathbf{t}}^*$ denote the final solution of the IRPLS algorithm. This solution satisfies

$$\hat{\mathbf{t}}^* = \left(\mathbf{A}^T \mathcal{W}(\hat{\mathbf{t}}^*) \mathbf{A} \right)^{-1} \mathbf{A}^T \mathcal{W}(\hat{\mathbf{t}}^*) \mathbf{b} \quad (30)$$

where

$$\mathcal{W}(\hat{\mathbf{t}}^*) = \text{diag} \left(\gamma_1^{-\frac{p}{\alpha}} \|\hat{\mathbf{t}}^* - \mathbf{r}_1\|^{-2} |\tilde{\theta}_1 - \hat{\theta}_1^*|^{(p-2)}, \dots, \gamma_M^{-\frac{p}{\alpha}} \|\hat{\mathbf{t}}^* - \mathbf{r}_M\|^{-2} |\tilde{\theta}_M - \hat{\theta}_M^*|^{(p-2)} \right) \quad (31)$$

and $\hat{\theta}_m^*$ denotes the m th bearing obtained from (5) by using $\hat{\mathbf{t}}^*$.

According to (22) and (30), we can get

$$\Delta \mathbf{t} = \hat{\mathbf{t}}^* - \mathbf{t} = -\left(\mathbf{A}^T \mathbf{W}(\hat{\mathbf{t}}^*) \mathbf{A}\right)^{-1} \mathbf{A}^T \mathbf{W}(\hat{\mathbf{t}}^*) \boldsymbol{\xi} \quad (32)$$

The analytical bias of IRPLS is given by [14]

$$\mathbb{E}\{\Delta \mathbf{t}\} = -\frac{1}{p-1} \mathbb{E}\left\{\frac{\mathbf{A}^T \mathbf{W}(\mathbf{t}) \mathbf{A}}{M}\right\}^{-1} \mathbb{E}\left\{\frac{\mathbf{A}^T \mathbf{W}(\mathbf{t}) \boldsymbol{\xi}}{M}\right\} \quad (33)$$

where the weighting matrix $\mathbf{W}(\mathbf{t})$ is computed from (31) using true source location parameter $\mathbf{t}$.

4. Total Lp-Norm Optimization

In this section, we present a method of TLPO for BOSL. The IRPLS algorithm only considers the deviation of the system vector $\mathbf{b}$. But in fact, the system matrix $\mathbf{A}$ also has a measurement residual. The IRPLS algorithm will inevitably cause a large bias, because the system matrix $\mathbf{A}$ and the system vector $\mathbf{b}$ are statistically correlated.

4.1. Method Description

In order to improve the performance of the IRPLS algorithm and reduce the bias caused by the correlation between $\mathbf{A}$ and $\mathbf{b}$, we develop the TLPO method in this subsection. Unlike the IRPLS method, the TLPO method uses the correction matrix $\Delta \mathbf{A}$ and the correction vector $\Delta \mathbf{b}$ to compensate the system matrix $\mathbf{A}$ and system vector $\mathbf{b}$, respectively. The normalized equation can be written as

$$\mathbf{G}(\mathbf{A} + \Delta \mathbf{A})\mathbf{t} = \mathbf{G}(\mathbf{b} + \Delta \mathbf{b}) \quad (34)$$

where $\Delta \mathbf{A}$ is the perturbation matrix of $\mathbf{A}$ and $\Delta \mathbf{b}$ is the perturbation vector of $\mathbf{b}$. Let $\mathbf{K} = \mathbf{G}[\mathbf{A}, \mathbf{b}]$ and $\mathbf{v} = [\mathbf{t}, -1]^T$ be the augmented matrix and vector, and Equation (34) can be rewritten as

$$(\mathbf{K} + \boldsymbol{\Omega})\mathbf{v} = 0 \quad (35)$$

where $\boldsymbol{\Omega} = \mathbf{G}[\Delta \mathbf{A}, \Delta \mathbf{b}]$ is the error matrix.

In order to reduce the bias of IRPLS, we use the TLPO approach to minimize the perturbation matrix $\Delta \mathbf{A}$ and perturbation vector $\Delta \mathbf{b}$ simultaneously, and the TLPO problem for BOTL can be formulated as

$$\min_{\mathbf{E}, \mathbf{t}} \|\boldsymbol{\Omega}\|_p, \quad \text{s. t. } (\mathbf{K} + \boldsymbol{\Omega})\mathbf{v} = 0 \quad (36)$$

for $1 < p < 2$. Note that if $p = 1$, (36) becomes the total least absolute residual method, and if $p = 2$, it is the well-known TLS method. To solve the TLPO problem conveniently, it is necessary to develop an equivalent form of (36). The following proposition holds.

Proposition 1. *Given the TLPO problem defined in (36), the estimation of $\mathbf{t}$ from the minimization of $\|\boldsymbol{\Omega}\|_p$ subject to $(\mathbf{K} + \boldsymbol{\Omega})\mathbf{v} = 0$ is equivalent to*

$$\min_{\mathbf{v}} \|\mathbf{K}\mathbf{v}\|_p, \quad \text{s. t. } \|\mathbf{v}\|_q - 1 = 0 \quad (37)$$

where $\mathbf{v} = [\mathbf{t}, -1]^T$, $\frac{1}{p} + \frac{1}{q} = 1$ and $1 < p < \alpha$.

Proof. From the constraint $(\mathbf{K} + \mathbf{\Omega})\mathbf{v} = 0$, we have $\|\mathbf{K}\mathbf{v}\|_p = \|\mathbf{\Omega}\mathbf{v}\|_p$. Let ω_m^T denote the m th row of $\mathbf{\Omega}$. $\|\mathbf{\Omega}\mathbf{v}\|_p^p$ becomes

$$\begin{aligned}\|\mathbf{\Omega}\mathbf{v}\|_p^p &= \sum_{m=1}^M |\omega_m^T \mathbf{v}|^p \\ &\leq \sum_{m=1}^M \|\omega_m\|_p^p \|\mathbf{v}\|_q^p\end{aligned}\quad (38)$$

for p and q satisfy the equation $\frac{1}{p} + \frac{1}{q} = 1$ and $1 < p < \alpha$. To derive (38), we used the properties of Hölder's inequality [34]. If the L_q-norm of $\mathbf{v}$ satisfies $\|\mathbf{v}\|_q = 1$, then the inequality (38) becomes $\|\mathbf{\Omega}\mathbf{v}\|_p \leq \|\mathbf{\Omega}\|_p$. Therefore, we can conclude that the minimization of $\|\mathbf{\Omega}\|_p$ subject to $(\mathbf{K} + \mathbf{\Omega})\mathbf{v} = 0$ is equivalent to (37). $\square$

Proposition 1 provides a facilitated way to solve the TLPO problem. In the following two subsections, we concentrate on deriving the solution on this particular problem and developing the IGED and GLM algorithms.

4.2. The IGED Algorithm

The optimization problem in Proposition 1 is equivalent to

$$\min_{\mathbf{v}} \|\mathbf{K}\mathbf{v}\|_p^p, \quad \text{s. t. } \|\mathbf{v}\|_q^q - 1 = 0 \quad (39)$$

To solve (39), we first transform it into an unconstrained optimization problem by using the Lagrange multiplier method. The appropriate Lagrangian function of (39) can be written as

$$\begin{aligned}\mathcal{L}(\mathbf{v}, \lambda) &= \|\mathbf{K}\mathbf{v}\|_p^p + \lambda(1 - \|\mathbf{v}\|_q^q) \\ &= \mathbf{v}^T \mathbf{K}^T \mathbf{D} \mathbf{K} \mathbf{v} + \lambda(1 - \mathbf{v}^T \mathbf{C} \mathbf{v})\end{aligned}\quad (40)$$

where $\mathbf{D} = \text{diag}(|(\mathbf{K}\mathbf{v})_m|^{p-2})$, $(\mathbf{K}\mathbf{v})_m$ represents the m th element of vector $\mathbf{K}\mathbf{v}$ for $m = 1, 2, \dots, M$, $\mathbf{C} = \text{diag}(|v_1^{q-2}|, |v_2^{q-2}|, |v_3^{q-2}|)$ and λ is the Lagrangian multiplier. Taking the partial derivative of $\mathcal{L}(\mathbf{v}, \lambda)$ with respect to $\mathbf{v}$ and setting it to zero yields

$$\mathbf{K}^T \mathbf{D} \mathbf{K} \mathbf{v} = \lambda \mathbf{C} \mathbf{v} \quad (41)$$

After multiplying both sides of (41) with $\mathbf{v}^T$, it can be found that only $\lambda = \mathbf{v}^T \mathbf{K}^T \mathbf{D} \mathbf{K} \mathbf{v}$ is the cost to be minimized. So, the optimal solution $\mathbf{v}^*$ is the eigenvector of the smallest generalized eigenvalue of $(\mathbf{K}^T \mathbf{D} \mathbf{K}, \mathbf{C})$.

Figure 2 gives a flowchart to depict the whole process of the IGED algorithm. The algorithm starts from the initialization procedure, where $\mathbf{v}_0$ is set as $[\hat{\mathbf{t}}_{\text{PLE}}^T, -1]^T$ and $\hat{\mathbf{t}}_{\text{PLE}}$ is obtained from (23). Given p and q , matrices $\mathbf{C}$ and $\mathbf{D}$ can be computed, respectively. After performing generalized eigenvalue decomposition from the pair $(\mathbf{K}^T \mathbf{D} \mathbf{K}, \mathbf{C})$, the solution of $\mathbf{v}$ in (41) is the generalized eigenvector of $(\mathbf{K}^T \mathbf{D} \mathbf{K}, \mathbf{C})$ that gives the minimum generalized eigenvalue. Since IGED is an iterative algorithm, it requires the convergence check. Let λ_i and λ_{i+1} be the minimum generalized eigenvalues at the i th and $i + 1$ th iterations, respectively. If there exists ε such that

$$\left| \frac{\lambda_{i+1} - \lambda_i}{\lambda_{i+1}} \right| \leq \varepsilon \quad (42)$$

then we would get the final solution $\mathbf{v}^*$ and the estimate for source location is

$$\hat{\mathbf{t}} = \frac{\mathbf{v}^*(1:2)}{\mathbf{v}^*(3)}. \quad (43)$$

Figure 2. The flowchart of the IGED algorithm.

In summary, the IGED Algorithm 2 can be implemented by the following procedures.

Algorithm 2 The IGED algorithm.

- 1: **Initialization:** $v_0 = [\hat{t}_{\text{PLE}}^T - 1]^T$.
 - 2: **for** $i = 0; i++$ **do**
 - 3: Compute C and D .
 - 4: Perform generalized eigenvalue decomposition of the pair $(K^T DK, C)$.
 - 5: Select the smallest generalized eigenvalue and its corresponding generalized eigenvector.
 - 6: Set λ_i and λ_{i+1} as the minimum generalized eigenvalues for iterations i and $i + 1$.
If $|(\lambda_{i+1} - \lambda_i) / \lambda_{i+1}| > \varepsilon$, go to step 3. Otherwise, obtain the final solution v^* and stop.
 - 7: Get the source location estimate $\hat{t}$ using (43).
 - 8: **end for**
-

4.3. The GLM Algorithm

The GLM method [35] is widely used for constrained optimization. Its basic principle is to add a penalty term to the Lagrangian function to form an augmented Lagrangian function, which can impose a larger penalty on infeasible points. Thus, the constrained optimization problem (37) would be transformed into a new unconstrained optimization problem that can be solved efficiently.

To solve (37) with the GLM method, we formulate an augmented Lagrangian function as follows:

$$\mathcal{L}(v, \lambda, s) = g(v) + \frac{s}{2} \|h(v)\|_2^2 - \lambda h(v) \quad (44)$$

where $g(v) = \|Kv\|_p$ and $h(v) = 1 - \|v\|_q$. Compared with the Lagrangian function, the GLM cost function has an additional term $s/2 \|h(v)\|_2^2$. This item is a punishment for violation of constraint $h(v) = 0$. The punishment parameter s determines the degree of punishment, and it is generally sufficiently large. λ is a positive scalar, and the term $\lambda h(v)$ is to ensure that the optimal solution is a strict local minimum point of $F(v) = g(v) + \frac{s}{2} \|h(v)\|_2^2$ under the condition of obeying the constraint $h(v) = 0$.

Figure 3 shows the flow diagram of the GLM algorithm. For the initialization step, we set $v_0 = [\hat{t}_{\text{PLE}}^T - 1]^T$. Given initial multiplier λ , penalty factor s , admissible error ε , scaling parameter $a > 1$ and parameter $\rho \in (0, 1)$, the problem (44) can be solved by using the existing unconstrained optimization methods, such as Newton, Interior point method, simplex method, etc. We use fminsearch function in MATLAB to estimate v from (44). Then, we check whether the termination criteria are satisfied. If $\|h(v_i)\| \leq \delta$, the iteration terminates and v_i is the near optimal solution of (37). Otherwise, go to the next step which is to control the convergence speed. If there exists $\rho \in (0, 1)$ such that

$$\frac{\|h(v_i)\|}{\|h(v_{i-1})\|} \geq \rho \quad (45)$$

we can increase the penalty factor s by setting $s = a \cdot s$ ($a > 1$). If not, then we leave the value of s unchanged. The penalty factor is used for multiplier update $\lambda = \lambda + s\|h(v_i)\|$. The updated λ and s are used to solve the unconstrained optimization problem (44) again.

Figure 3. The flow diagram of the GLM algorithm.

Finally, the implementation process of the GLM algorithm is summarized in Algorithm 3.

Algorithm 3 The GLM algorithm.

- 1: **Initialization:** $v_0 = [\hat{f}_{PLE}^T, -1]^T$, $\delta, \lambda, s, a > 1, \rho \in (0, 1)$.
 - 2: **for** $i = 0; i++$ **do**
 - 3: Using v_i as the initial point, find the minimum value of $\mathcal{L}(v, \lambda, s)$.
 - 4: If $\|h(v_i)\| \leq \delta$ holds, stop the iterations and get the final solution v^* , and go to step 7. Otherwise, go to step 5.
 - 5: If $\|h(v_i)\| / \|h(v_{i-1})\| \geq \rho$, update the penalty coefficient $s = a \cdot s$, and go to step 6. Otherwise, go to step 6 directly.
 - 6: Update $\lambda = \lambda - \rho\|h(v_i)\|$ and go to step 3.
 - 7: Find the target position $\hat{f} = v^*(1 : 2) / v^*(3)$.
 - 8: **end for**
-

4.4. Computational Complexity Analysis

In this subsection, we analyze the computational complexities of the proposed algorithms and compare with those of PLE and IRPLS. We only consider the asymptotic computational cost. For each algorithm, we separate the complexity calculation to small steps, e.g., multiplying two matrices of size $n \times m$ and $m \times p$ costs $\mathcal{O}(nmp)$. Let L_1 , L_2 and L_3 denote the number of iterations with regarding to IRPLS, IGED and GLM, respectively. The iterative implementation of IRPLS, IGED and GLM, however, could be very sensitive to initialization. In this paper, we use the solution of PLE for initialization.

The computational costs are listed in Table 1. Table 1 summarizes the computational complexities of the PLE and IRPLS, IGED and GLM algorithms. PLE requires the least amounts of computation. IRPLS and IGED exhibit the similar computational costs. For the GLM algorithm, we employ the Broyden–Fletcher–Goldfarb–Shanno (BFGS) [33] algorithm to solve the unconstrained problem presented in (44). BFGS is an iterative method for solving unconstrained nonlinear optimization problems by providing an approximation to the Hessian matrix. We assume that the number of iterations for BFGS is N_i for the i th iteration of GLM. For each iteration, we assume the order of computational complexity of BFGS is $O_i(M)$. The GLM algorithm needs to estimate v which is computationally demanding. The estimation procedure needs to repeat L_3 times since it is iterative.

Table 1. Computational complexity.

Method	Operation	Cost
PLE	$A^T A, A^T b$	$\mathcal{O}(4M) + \mathcal{O}(2M)$
	$(A^T A)^{-1}$	$\mathcal{O}(2^3)$
	$(A^T A)^{-1} A^T b$	$\mathcal{O}(4M)$
IRPLS	Compute $W^{(i)} A$ L_1 times	$\mathcal{O}(2L_1 M)$
	Compute $W^{(i)} b$ L_1 times	$\mathcal{O}(L_1 M)$
	Compute (29) L_1 times	$\mathcal{O}(L_1(10M + 8))$
IGED	Compute DK L_2 times	$\mathcal{O}(3L_2 M)$
	Perform GED L_2 times	$\mathcal{O}(3^3 L_2)$
GLM	L_3 evaluations of BFGS	$\mathcal{O}(\sum_{i=1}^{L_3} N_i \mathcal{O}_i(M))$

5. Performance Bound

In this section, we present the Cramér-Rao Lower Bound (CRLB) for robust BOSL when the measurement noise is modeled as $S\alpha S$ distribution. We then discuss how the CRLB is calculated. The CRLB for any unbiased estimator of t is derived by the inverse of the Fisher Information Matrix (FIM) [36], which is computed as

$$\text{CRLB}(t) = J^{-1} \quad (46)$$

where J is the FIM given by the 2×2 matrix

$$J = \mathbb{E} \left\{ \begin{bmatrix} \left(\frac{\partial f(\tilde{\phi}|t)}{\partial t_x} \right)^2 & \frac{\partial f(\tilde{\phi}|t)}{\partial t_x} \frac{\partial f(\tilde{\phi}|t)}{\partial t_y} \\ \frac{\partial f(\tilde{\phi}|t)}{\partial t_x} \frac{\partial f(\tilde{\phi}|t)}{\partial t_y} & \left(\frac{\partial f(\tilde{\phi}|t)}{\partial t_y} \right)^2 \end{bmatrix} \right\} \quad (47)$$

where $f(\tilde{\phi}|t)$ denotes the Probability Density Function (PDF) of $\tilde{\phi}$. Since $e_1, e_2, \dots, e_M$ have the same standard dispersion, the PDF of $\tilde{\phi}_m$ is

$$f_\alpha(e) = f_\alpha(\tilde{\phi}_m - \phi_m) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(-i(\tilde{\phi}_m - \phi_m)\tau - \gamma|\tau|^\alpha) d\tau \quad (48)$$

According to [37], the FIM J for the parameters t_x and t_y is given by

$$J = \kappa_\alpha \sum_{m=1}^M \begin{bmatrix} \left(\frac{\partial \phi_m}{\partial t_x} \right)^2 & \frac{\partial \phi_m}{\partial t_x} \frac{\partial \phi_m}{\partial t_y} \\ \frac{\partial \phi_m}{\partial t_x} \frac{\partial \phi_m}{\partial t_y} & \left(\frac{\partial \phi_m}{\partial t_y} \right)^2 \end{bmatrix} \quad (49)$$

where κ_α is the Fisher information for the location of $f_\alpha(e)$,

$$\kappa_\alpha = \int_{-\infty}^{+\infty} \frac{(f'_\alpha(e))^2}{f_\alpha(e)} de \quad (50)$$

and the expressions of $\frac{\partial \phi_m}{\partial t_x}$ and $\frac{\partial \phi_m}{\partial t_y}$ are given by

$$\frac{\partial \phi_m}{\partial t_x} = -\frac{t_y - r_{y,m}}{\|t - r_m\|}, \quad \frac{\partial \phi_m}{\partial t_y} = \frac{t_x - r_{x,m}}{\|t - r_m\|} \quad (51)$$

Note that the value of κ_α depends on the parameter α . If $\alpha = 1$, $f_\alpha(e)$ follows a Cauchy PDF and κ_α equals to $\frac{3}{5}$ for $\gamma = 1$. In particular, when $\alpha = 2$, $f_\alpha(e)$ is Gaussian with variance 2γ and $\kappa_\alpha = \gamma$. Therefore, the CRLB of t in the case of $\alpha = 2$ is consistent with the CRLB result given in [8].

However, when $1 < \alpha < 2$, $f_\alpha(e)$ has no closed-form expressions. Hence, computing $f_\alpha(e)$ involves numerically evaluating the integrals in (48). We use the MATLAB function *STBLPDF* to compute the pdf of the $S\alpha S$ distribution. The derivative of $f_\alpha(e)$ is also calculated by using numerical differentiation methods.

6. Simulations

In this section, we present three simulation examples to illustrate the performance of the proposed IGED and GLM algorithms and compare them with the PLE and IRPLS algorithm, as well as the CRLB. The localization performance is characterized with the bias norm and the root-mean-square-error (RMSE). The RMSE for source localization is defined as

$$\text{RMSE} = \left(\frac{1}{N} \sum_{n=1}^N \|\hat{\mathbf{t}}_n - \mathbf{t}\|^2 \right)^{1/2} \quad (52)$$

where N represents the total number of Monte Carlo runs, and $\hat{\mathbf{t}}_n$ denotes the source location estimate at the n th Monte Carlo run. The expression of the bias norm is given by

$$\text{BIAS} = \left\| \frac{1}{N} \sum_{n=1}^N (\hat{\mathbf{t}}_n - \mathbf{t}) \right\| \quad (53)$$

The RMSE given in (52) is bounded by the square root of the trace of the Cramér-Rao Lower Bound matrix

$$\text{RMSE} \geq \sqrt{\text{tr}(\text{CRLB}(\mathbf{t}))} \quad (54)$$

where $\text{tr}(\cdot)$ denotes the trace of a matrix.

In the simulations, we examine the localization accuracy versus three kinds of parameters: (1) noise dispersion; (2) number of sensors; and (3) noise impulsiveness. All the sensors are uniformly placed in a $100 \times 100 \text{ m}^2$ plane centered at (50, 50) m. The source is located at (100, 100) m. The IRPLS algorithm and the proposed IGED and GLM algorithms are iterative and they are all initialized by the PLE derived from (23) for ensuring a fair comparison. The termination parameters ϵ , ε and δ are set as $\epsilon = 10^{-5}$, $\varepsilon = 10^{-10}$ and $\delta = 10^{-5}$, respectively. For IRPLS, IGED and GLM, the maximum iteration is fixed at 200.

6.1. Various Levels of Noise Dispersion

In this example, we illustrate the RMSE and bias performance of PLE, IRPLS, IGED and GLM versus noise dispersion. For convenience, we use the generalized signal-to-noise ratio (GSNR) to characterize different noise levels. The GSNR is inversely proportional to the noise dispersion γ and is calculated as $1/\gamma$ after normalized. The range of $\gamma^{1/\alpha}$ is from $2\pi/180$ radian to $6\pi/180$ radian. The characteristic exponent α is set to 1.5 and the corresponding optimum value of p is 1.225.

Figure 4 plots the RMSE and bias curves of the PLE, IRPLS, IGED and GLM algorithms together with the CRLB. For RMSE curves shown in Figure 4a, the green line with Asterisk is the RMSE value for the PLE algorithm, the blue line with Point is for IRPLS, the black line with Cross is for IGED, the red line with Square is for GLM and the blue dash line is the CRLB. The PLE method can not give accurate target location estimates in the presence of impulsive noise. The large bias norm formed by the outlier data in turn leads to a poor performance for the PLE. The least-norm estimators, including IRPLS, IGED and GLM, yield much better localization performance. However, the IRPLS method still has a large estimation bias because it cannot overcome the bias attributed to the correlation between the system matrix A and the data vector $\mathbf{b}$. On the other hand, the IGED and GLM algorithms utilizes the total Lp-norm optimization technique that can minimize the errors in A and $\mathbf{b}$ simultaneously. Therefore, the IGED and GLM algorithms are capable of outperforming the IRPLS in much lower RMSEs and biases.

Figure 4. RMSE and bias performance comparison of PLE, IRPLS, IGED and GLM estimates with various GSNRs. (a) RMSE results. (b) Bias norm results.

Beyond the RMSE and bias performance, we need to explore the number of iterations and the computation time of IRPLS, IGED and GLM. The results are depicted in Figure 5. The algorithms run on a laptop with CPU i5-7200U @ 2.5 GHz and RAM 8 GB. The version of software is MATLAB 2017a. Collectively, the average number of iterations decreases when the GSNR increases. The GLM appears to have the least number of iterations. Its average number of iterations reduces from 1.05 to 1.01 when the GSNR ranges from 29.51 to 153.3. However, its computation time for single iteration is high (i.e., 0.1343 s), because the GLM algorithm has the steps of unconstrained optimization. On the contrary, the IGED algorithm has the least computing time since its computational complexity mainly lies in generalized eigenvalue decomposition. The IRPLS algorithm has high computing time due to the large number of iterations (over 30 iterations).

Figure 5. Number of iterations and computing time versus GSNR for the IRPLS, IGED and GLM algorithms. (a) Number of iterations. (b) Computing time.

6.2. Different Number of Sensors

In this subsection, we compare the RMSE and bias performance of the PLE, IRPLS, IGED and GLM algorithms over the number of sensors M ranging from 10 to 30, five at a time, when $\gamma^{1/\alpha}$ is kept at $4\pi/180$ radian. The other parameters remain the same. The simulation results are shown in Figures 6 and 7.

Figure 6. RMSE and bias performance comparison of PLE, IRPLS, IGED and GLM estimates with different number of sensors. (a) RMSE results. (b) Bias norm results.

Figure 7. Number of iterations and computing time versus number of sensors for the IRPLS, IGED and GLM algorithms. (a) Number of iterations. (b) Computing time.

From Figure 6, we observe that the GLM algorithm has better RMSE and bias performance than that of PLE, IRPLS and IGED. When $M \geq 20$, the RMSE of GLM is close to the CRLB. The IGED algorithm slightly deviates from the CRLB, because the algorithm has not fully converged. Moreover, we also observe from Figure 7 that the average iterations is kept at 2.5 and is almost the same as that of the GLM algorithm. Meanwhile, the bias of IGED is higher than that of GLM. Setting more iterations would help the IGED algorithm achieve better performance. As expected, the PLE exhibits unreliable results since it is not robust to the impulsive noise. The bias of IRPLE does not vanish as the number of sensors increases. This phenomenon is also validated in (33). In terms of computational complexity, the total computing time for GLM is about 0.3378 s when the number of sensors is fixed at 20. The IGED has the least amounts of computation and it only requires 0.01032 s at $N = 20$.

6.3. Various Values of Noise Impulsiveness

To further verify the effectiveness of the proposed algorithms, we examine the performance of the algorithms for different levels of noise impulsiveness. In the following figures drawing the simulation results, the value of the noise impulsiveness level α deviates from

1.9 to 1.1 and the corresponding optimum values of p are set as 1.546, 1.430, 1.348, 1.282, 1.225, 1.174, 1.127, 1.083, 1.041 from (17). In this example, $\gamma^{1/\alpha}$ is fixed at $4\pi/180$ radian. The other parameters remain unchanged.

Figure 8 depicts the RMSE and bias performance with respect to various α . As shown in Figure 8, the RMSE and the bias of PLE is much higher than that of IRPLS, IGED and GLM, which is caused by the impulsive noise. As α decreases, the influence of the noise impulsiveness becomes more significant. The level of noise impulsiveness also affects the bias and RMSE performance of IRPLS. The GLM algorithm has relatively small bias and its RMSE is very close to the CRLB when $\alpha \geq 1.4$. The IGED has the comparable RMSE performance with the GLM. However, the bias of IGED slightly deviates from that of GLM as α decreases. From Figure 9, we observe that the number of iterations of IRPLS, GLM and IGED decrease as α increases, i.e., the number of average iterations reduces from 38.8 to 29.5 for IRPLS, from 4.4 to 1 for GLM and from 2.9 to 2.2 for IGED. The computation time also decreases when α increases. Again, the computation time of GLM is much higher than that of IGED.

Figure 8. RMSE and bias performance comparison of PLE, IRPLS, IGED and GLM estimates with various noise impulsiveness. (a) RMSE results. (b) Bias norm results.

Figure 9. Number of iterations and computing time versus noise impulsiveness for the IRPLS, IGED and GLM algorithms. (a) Number of iterations. (b) Computing time.

6.4. Different Number of Iterations

In order to further verify the influence of the number of iterations on the proposed algorithm, we tested the performance of the algorithm for the increasing number of iterations. The simulation results are shown in Figure 10, where the number of iterations is set 50 for IRPLS and IGED, and 6 for GLM. The RMSE and bias are recorded every 10 iterations for IRPLS and IGED. The RMSE and bias performance of GLM is plotted for each iteration. In this example, $\gamma^{1/\alpha}$ is fixed at $4\pi/180$ radian. The number of sensors is kept at 20. The characteristic exponent α is set to 1.5 and the corresponding optimum value of p is 1.225.

Figure 10 describes the RMSE and bias performance of the IRPLS, IGED and GLM algorithms versus different iterations. As shown in Figure 10a, the RMSE result of IRPLS remains at 2.1 m after 30 iterations. We observe that the bias of the IRPLS algorithm decreases as the number of iterations increasing, as shown in Figure 10b. However, the bias does not vanish. It still has 0.41 m after 50 iterations. The RMSE value of IGED keeps at 1.8 m after 30 iterations and the bias of IGED reduces to 0.12 m using about 40 iterations. Differing from the IRPLS and IGED algorithms, the RMSE result of GLM reaches at 1.8 m using only four iterations. After 5 iterations, the bias of GLM reduces to 0.12 m. From Figure 10, we can observe that the RMSE and bias performance of IGED is almost the same as that of GLM if IGED runs a sufficient number of iterations.

Figure 10. RMSE and bias performance versus the number of iterations for IRPLS, IGED and GLM. (a) RMSE results. (b) Bias norm results.

6.5. Scalability Evaluation

We analyze the scalability of the IRPLS, IGED and GLM algorithms in this subsection for increasing number of sensors, from 20 to 100, and decreasing levels of noise impulsiveness, from $\alpha = 1.1$ to $\alpha = 1.9$. To examine the scalability of the iterative algorithms, the IRPLS, IGED and GLM algorithms share the same stopping condition. Let $\hat{t}_i$ denote estimation result for the i th iteration. The criterion for stopping these iterative algorithms is given by $\|\hat{t}_i - \hat{t}_{i-1}\| < 10^{-5}$.

Figure 11 compares the number of iterations, computing time for each iteration, RMSE and bias performance of the IRPLS, IGED and GLM algorithms versus different numbers of sensors. The other parameters remain the same as Section 6.4. It can be observed from Figure 11 that the number of iterations of these algorithms does not reduce as the number of sensors increases. Among them, the number of iterations for IRPLS and IGED maintains at 35, while that of the GLM algorithm keeps at 10. The computing time for each iteration, however, slightly increases as the number of sensors increasing. The RMSE and bias performance decreases as the number of sensors increases due to the fact that more normal sensors can be utilized. This result is also demonstrated in Figure 6. We also

observe that the RMSE and bias curves of IGED and GLM are consistently lower than those of IRPLS, which shows the performance advantage of the proposed algorithms.

Figure 12 shows the number of iterations, computing time for each iteration, RMSE and bias performance of the IRPLS, IGED and GLM algorithms versus different levels of noise impulsiveness. The number of sensors is set at 20. Other parameters remain unchanged. From Figure 12a, we observe that the number of iterations for IRPLS scales from 88 to 11, 43 to 7 for IGED. However, the iteration value does not change for GLM and the number keeps at 7. That is because GLM performs optimization at each iteration to reduce the impact of impulsive noise. Moreover, the RMSE and bias performance of these algorithms decreases as increasing the value of α . As expected, IGED and GLM exhibit almost the same performance. A small value of α results in a poorer IRPLE performance, as illustrated in Figure 12b. When α is greater than 1.6, the RMSE and bias curves of IRPLS reach those of IGED and GLM due to the fact that the level of noise impulsiveness has less impact on the bias performance.

Figure 11. Scalability analysis for various number of sensors. (a) Number of iterations and computing time. (b) RMSE and bias performance.

Figure 12. Scalability analysis for different levels of noise impulsiveness. (a) Number of iterations and computing time. (b) RMSE and bias performance.

7. Conclusions

In this paper, a total Lp-norm optimization method is presented to solve the BOSL problem and two algorithms, named IGED and GLM, are proposed to fulfill the optimization of the total least Lp-norm. By minimizing the errors in the system matrix and the

data vector simultaneously for the least Lp-norm optimization, the proposed algorithms can overcome the bias arising from the correlation between the system matrix and the pseudolinear noise vector.

Simulation results show that the proposed IGED and GLM algorithms have much better RMSE and bias performance than the IRPLS algorithm. The number of iterations for GLM remains at a low level. With only a few iterations, the RMSEs of GLM can approach the CRLB. However, GLM expends more computation time than IGED under the same number of iterations, since GLM requires unconstrained optimization in each iteration.

Author Contributions: J.-A.L. studied the total Lp-norm optimization problem for bearing-only source localization and wrote the draft of the manuscript. C.-C.X. helped solve the total Lp-norm optimization problem and did the simulations, S.-T.H. and Y.-J.R. were involved in the discussion of the problem and revised the manuscript. All authors read and approved the final version of the manuscript.

Funding: This work was supported by the National Science Foundation for Young Scientists of China (No. 61703129) and the Science and Technology on Near-Surface Detection Laboratory Foundation, China (No. 6142414200203).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare that there is no conflict of interest regarding the publication of this article.

References

- Li, X.; Luo, X.; Zhao, S. Globally Convergent Distributed Network Localization Using Locally Measured Bearings. *IEEE Trans. Control. Netw. Syst.* **2020**, *7*, 245–253. [\[CrossRef\]](#)
- Kaplan, L.M.; Le, Q. On exploiting propagation delays for passive target localization using bearings-only measurements. *J. Frankl. Inst.* **2005**, *342*, 193–211. [\[CrossRef\]](#)
- Volgyesi, P.; Balogh, G.; Nadas, A.; Nash, C.B.; Ledeczi, A. Shooter localization and weapon classification with soldier-wearable networked sensors. In Proceedings of the International Conference on Mobile Systems, Applications and Services, San Juan, Puerto Rico, 11–13 June 2007; pp. 113–126.
- Ali, A.M.; Yao, K.; Collier, T.C.; Taylor, C.E.; Blumstein, D.T.; Girod, L. An empirical study of collaborative acoustic source localization. *J. Signal Process. Syst.* **2009**, *57*, 415–436. [\[CrossRef\]](#)
- Wang, Y.; Wang, G.; Chen, S.; Ho, K.C.; Huang, L. An Investigation and Solution of Angle Based Rigid Body Localization. *IEEE Trans. Signal Process.* **2020**, *68*, 5457–5472. [\[CrossRef\]](#)
- Doğançay, K. Bias compensation for the bearings-only pseudolinear target track estimator. *IEEE Trans. Signal Process.* **2006**, *54*, 59–68. [\[CrossRef\]](#)
- Wang, Y.; Ho, K.C. An Asymptotically Efficient Estimator in Closed-Form for 3-D AOA Localization Using a Sensor Network. *IEEE Trans. Wirel. Commun.* **2015**, *14*, 6524–6535. [\[CrossRef\]](#)
- Gavish, M.; Weiss, A.J. Performance analysis of bearing-only target location algorithms. *IEEE Trans. Aerosp. Electron. Syst.* **1992**, *28*, 817–828. [\[CrossRef\]](#)
- Luo, J.A.; Shao, X.H.; Peng, D.L.; Zhang, X.P. A Novel Subspace Approach for Bearing-Only Target Localization. *IEEE Sens. J.* **2019**, *19*, 8174–8182. [\[CrossRef\]](#)
- Swami, A.; Sadler, B.M. TDE, DOA and related parameter estimation problems in impulsive noise. In Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics, Banff, AB, Canada; 21–23 July 1997; pp. 273–277.
- Kozick R.J.; Sadler, B.M. Maximum-likelihood array processing in non-Gaussian noise with Gaussian mixtures. *IEEE Trans. Signal Process.* **2000**, *48*, 3520–3535.
- Oh, H.; Seo, D.; Nam, H. Design of a Test for Detecting the Presence of Impulsive Noise. *Sensors* **2020**, *20*, 7135. [\[CrossRef\]](#)
- Shao, M.; Nikias, C.L. Signal processing with fractional lower order moments: Stable processes and their applications. *Proc. IEEE* **1993**, *81*, 986–1010. [\[CrossRef\]](#)
- Nguyen, N.H.; Doğançay, K.; Kuruoğlu, E.E. An Iteratively Reweighted Instrumental-Variable Estimator for Robust 3-D AOA Localization in Impulsive Noise. *IEEE Trans. Signal Process.* **2019**, *67*, 4795–4808. [\[CrossRef\]](#)
- Zhong, X.; Premkumar, A.B.; Madhukumar, A.S. Particle Filtering for Acoustic Source Tracking in Impulsive Noise with Alpha-Stable Process. *IEEE Sens. J.* **2013**, *13*, 589–600. [\[CrossRef\]](#)
- Tsakalides, P.; Nikias, C.L. Maximum likelihood localization of sources in noise modeled as a stable process. *IEEE Trans. Signal Process.* **1995**, *43*, 2700–2713. [\[CrossRef\]](#)

17. Luo, J.A.; Fang, F.; Shi, Y.F.; Shen-Tu, H.; Guo, Y.F. L1-Norm and Lp-Norm Optimization for Bearing-Only Positioning in Presence of Unreliable Measurements. In Proceedings of the Chinese Control And Decision Conference (CCDC), Hefei, China, 22–24 August 2020; pp. 1201–1205.
18. Maronna, B.R.A.; Martin, D.R.; Yohai, V.J. *Robust Statistics: Theory and Methods*; Wiley: New York, NY, USA, 2006.
19. Liu, Y.; Hu, Y.H.; Pan, Q. Distributed, robust acoustic source localization in a wireless sensor network. *IEEE Trans. Signal Process.* **2012**, *60*, 4350–4359. [[CrossRef](#)]
20. Jiang, Y.; Azimi-Sadjadi, M.R. A robust source localization algorithm applied to acoustic sensor network. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Honolulu, HI, USA, 15–20 April 2007; pp. 1233–1236.
21. Panigrahi, T.; Panda, G.; Mulgrew, B.; Majhi, B. Robust incremental LMS over wireless sensor network in impulsive noise. In Proceedings of the International Conference on Computational Intelligence and Communication Networks, Bhopal, India, 26–28 November 2010; pp. 205–209.
22. Luo, J.A.; Xue, C.C.; Peng, D.L. Robust Bearing-Only Localization Using Total Least Absolute Residuals Optimization. *Complexity* **2020**, *2020*, 3456923. [[CrossRef](#)]
23. Satar, B.; Soysal, G.; Jiang, X.; Efe, M.; Kirubarajan, T. Robust Weighted $l_{1,2}$ Norm Filtering in Passive Radar Systems. *Sensors* **2020**, *20*, 3270. [[CrossRef](#)]
24. Chen, Y.; So, H.C.; Kuruoglu, E.E. Variance analysis of unbiased least lp-norm estimator in non-Gaussian noise. *Signal Process.* **2015**, *122*, 190–203. [[CrossRef](#)]
25. Jiang, X.; Chen, J.; So, H.C.; Liu, X. Large-Scale Robust Beamforming via ℓ_∞ -Minimization. *IEEE Trans. Signal Process.* **2018**, *66*, 3824–3837. [[CrossRef](#)]
26. Wu, H.; Chen, S.X.; Zhang, Y.H.; Zhang, H.Y.; Ni, J. Robust structured total least squares algorithm for passive location. *J. Syst. Eng. Electron.* **2015**, *26*, 946–953. [[CrossRef](#)]
27. Picard, J.S.; Weiss, A.J. Bounds on the number of identifiable outliers in source localization by linear programming. *IEEE Trans. Signal Process.* **2010**, *58*, 2884–2895. [[CrossRef](#)]
28. Fragkos, G.; Apostolopoulos, P.A.; Tsiropoulou, E.E. ESCAPE: Evacuation strategy through clustering and autonomous operation in public safety systems. *Future Internet* **2019**, *11*, 20. [[CrossRef](#)]
29. Zhang, Y.; Guizani, M. *Game Theory for Wireless Communications and Networking*; CRC Press: Boca Raton, FL, USA, 2011.
30. Cambanis, S.; Miller, G. Linear Problems in Linear Problems in pth Order and Stable Processes. *SIAM J. Appl. Math.* **1981**, *41*, 43–69. [[CrossRef](#)]
31. Mallik, M. A Note on Bearing Measurement Model; ResearchGate. 2018. Available online: https://www.researchgate.net/publication/325214760_A_Note_on_Bearing_Measurement_Model (accessed on 16 September 2021).
32. Bishop, A.N.; Anderson, B.D.O.; Fidan, B.; Pathirana, P.N.; Mao, G. Bearing-Only Localization using Geometrically Constrained Optimization. *IEEE Trans. Aerosp. Electron. Syst.* **2009**, *45*, 308–320. [[CrossRef](#)]
33. Roger, F. *Practical Methods of Optimization*; John Wiley & Sons: New York, NY, USA, 1987.
34. Grinshpan, A.Z. Weighted inequalities and negative binomials. *Adv. Appl. Math.* **2010**, *45*, 564–606. [[CrossRef](#)]
35. Hestenes, M.R. Multiplier and gradient methods. *Optim. Theory Appl.* **1969**, *4*, 303–320. [[CrossRef](#)]
36. Kay, S.M. *Fundamentals of Statistical Signal Processing: Estimation Theory*; Prentice Hall PTR: Englewood Cliffs, NJ, USA, 1993.
37. Sadler, B.M.; Kozick, R.J.; Moore, T. Performance analysis for direction finding in non-Gaussian noise. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Phoenix, AZ, USA, 15–19 March 1999; pp. 2857–2860.