IEEE Signal Processing Society | IEEE Transactions on Signal Processing

An Iteratively Reweighted Instrumental-Variable Estimator for Robust 3D AOA Localization in Impulsive Noise

Journal:	*Transactions on Signal Processing*
Manuscript ID	T-SP-24561-2019
Manuscript Type:	Regular Paper
Date Submitted by the Author:	09-Jan-2019
Complete List of Authors:	Nguyen, Ngoc Hung; University of South Australia, School of Engineering Dogancay, Kutluyil; University of South Australia, School of Engineering Kuruoglu, Ercan; Istituto di Scienza e Tecnologie dell'Informazione, CNR
EDICS:	92. RAS-LCLZ Source localization < RAS RADAR AND SONAR SIGNAL PROCESSING, 104. SAM-DOAE Direction of arrival estimation and source localization < SAM SENSOR ARRAY AND MULTICHANNEL PROCESSING

An Iteratively Reweighted Instrumental-Variable Estimator for Robust 3D AOA Localization in Impulsive Noise

Ngoc Hung Nguyen, *Member, IEEE*, Kutluyıl Doğançay, *Senior Member, IEEE*, and Ercan Engin Kuruoğlu, *Senior Member, IEEE*

***Abstract*—This paper considers the problem of robust three-dimensional (3D) angle-of-arrival (AOA) source localization in the presence of impulsive α-stable noise based on the l_p-norm minimization criterion. The iteratively reweighted least-squares algorithm (IRLS) is a well-known technique for solving l_p-norm minimization with the desirable global convergence property. Adopting the IRLS for 3D AOA localization requires nonlinear-to-pseudolinear transformation of azimuth and elevation angle measurement equations, thus resulting in a new variant of the IRLS, called the iteratively reweighted pseudolinear least-squares estimator (IRPLE). Unfortunately, there exists correlation between the measurement matrix and noise vector in the pseudolinear measurement equations, which consequently makes the IRPLE biased. To counter the bias problem of the IRPLE, a new iteratively reweighted instrumental-variable estimator (IRIVE) is proposed based on the exploitation of instrumental variables. The IRIVE is analytically shown to achieve the theoretical covariance of the general least l_p-norm estimation. Extensive simulation studies are presented to demonstrate the performance advantages of the IRIVE over the IRPLE as well as other existing least-squares and least l_p-norm estimators. The IRIVE is observed to produce nearly unbiased estimates with mean squared error performance very close to the Cramér-Rao lower bound.**

***Index Terms*—Angle of arrival, target localization, robust estimation, least l_p-norm, iteratively reweighted least-squares, pseudolinear estimation, instrumental variables.**

I. Introduction

SOURCE localization by angle-of-arrival (AOA) has received continuous attention for several decades owing to its applications in many diverse areas such as radar, sonar, navigation, wireless communications and indoor acoustic localization, to name but a few. In a general three-dimensional (3D) setting, the objective of AOA localization is to estimate the unknown position of a source using noisy azimuth and elevation angle measurements collected by several spatially distributed sensors.

AOA localization is essentially a nonlinear estimation problem where the main challenge arises from the highly nonlinear relationship of the azimuth and elevation angle measurements with the true source position. Various estimation techniques have been proposed in the literature to tackle the nonlinearity ranging from the direct grid search method, the iterative maximum likelihood estimator to the closed-form pseudolinear estimators (see e.g., [1]–[5] and the references therein). These methods were developed under the assumption of Gaussian noise based on second-order statistics. The Gaussian distribution has traditionally been the most widely accepted and used noise model in the signal processing and communications literature, mainly for the purpose of developing analytical solutions.

In the last two decades, tremendous attention has been devoted to non-Gaussian noise models due to the fact that various natural phenomena are more appropriately represented by distributions with a more impulsive (heavy-tailed) nature [6]–[23]. The α-stable distribution, a generalization of the Gaussian distribution, has been known to provide a better model for impulsive noise than the Gaussian distribution [6]–[10]. The α-stable distribution family is comprised of a wide range of distributions with varying degrees of impulsiveness and skewness, having the Gaussian, Cauchy and Lévy distributions as special cases. The applications of α-stable distribution can be found in different fields ranging from finance, econometrics, astronomy to signal processing and communications. Specifically, for the AOA localization problem under consideration, the α-stable distribution has been used to model the ambient noise in various applications including (i) underwater acoustic noise in sonar applications [19], [20], (ii) clutter returns in radar applications [18], [21], man-made and reverberation noise in indoor acoustic localization [15]–[17], and burst noise in indoor localization via wireless communications [18], [23].

This paper focuses on the problem of robust AOA localization in 3D space in the presence of impulsive α-stable noise. Under such a scenario, the aforementioned least-squares methods like those presented in [1]–[5] exhibit unreliable estimation performance since they are very sensitive to outliers [22]–[25]. This necessitates the need of developing new AOA localization algorithms to tackle impulsive noise. The l_p-norm minimization approach (with $1 \leq p < 2$) is an attractive alternative to the least-squares (i.e., l_2-norm minimization) approach and has been widely used to achieve robust estimation in impulsive noise [22]–[25]. The l_p-norm error minimization criterion does not admit a closed-form solution and must be solved in an iterative manner. The iteratively reweighted least-squares algorithm (IRLS) is one of the most widely-used solvers for l_p-norm minimization [19], [22], [25]. The main attraction of the IRLS is that it enjoys a desirable global convergence property under weak conditions [26]. In principle, the

N. H. Nguyen (e-mail: ngoc.nguyen@mymail.unisa.edu.au) and K. Doğançay are with School of Engineering, University of South Australia, Mawson Lakes, SA 5095, Australia. E. E. Kuruoğlu is with Institute of Information Science and Technology, "A. Faedo," (ISTI), Pisa, National Council of Research (CNR), Italy.

IRLS starts from the conventional least-squares solution and at each iteration solves a new weighted least-squares problem using the weights computed from the residuals in the previous iteration. Unfortunately, the IRLS is not directly applicable to the AOA localization problem under consideration due to the nonlinear relationship between the angle measurements and the unknown source position.

In this paper, we develop a new variant of the IRLS based on the pseudolinear form of the azimuth and elevation angle measurement equations. We call the new algorithm the iteratively reweighted pseudolinear least-squares estimator (IRPLE). However, it turns out from our analysis that the IRPLE suffers from a severe bias problem due to the correlation between the measurement matrix and the pseudolinear noise vector in a similar way to the pseudolinear estimator [2]. We then propose a new iteratively reweighted instrumental-variable estimator (IRIVE) to overcome the bias problem of the IRPLE by exploiting the use of an instrumental-variable (IV) matrix [27], [28] that is approximately uncorrelated with the pseudolinear noise vector. Pseudolinear estimation and IV estimation are well-known techniques under the least-squares framework for source localization in Gaussian noise (see e.g., [2]–[4], [29]–[34]). However, to the best of our knowledge, these techniques have not yet been applied to l_p-norm minimization for robust source localization in impulsive noise. It should be emphasized that applying the pseudolinear and IV estimation approaches to the least l_p-norm estimation problem presents a number of new challenges that are not encountered in the conventional least-squares estimation. First, the distinct nature of the α-stable distribution, e.g., lacking second-order moments and having infinite variance, makes the bias and covariance analysis of the estimators under study much more challenging than the estimators presented in [2]–[4] that rely on the Gaussian noise assumption. Moreover, in contrast to the pseudolinear and IV based estimators [2]–[4] which are closed-form under the Gaussian least-squares framework, the incorporation of pseudolinear and IV techniques into the IRLS produces new estimators that are iterative in nature. This creates difficulties in terms of analyzing the IRPLE bias, and makes the development and analysis of the proposed IRIVE not only challenging but also distinct from the existing least-squares IV estimators.

In this work, we also present a theoretical covariance analysis for the IRIVE and show that the IRIVE can achieve the analytical covariance of the general least l_p-norm estimation. The performance advantages of the IRIVE over the IRPLE, other existing least-squares pseudolinear and IV based estimators, and the least l_p-norm solver based on the Nelder-Mead method are numerically demonstrated via extensive simulation examples in terms of both bias and root-mean-square-error performance. The simulation results demonstrate the poor performance of the least-squares estimators in impulsive noise, confirm the severe bias problem of the IRPLE, and verify the effectiveness of the IRIVE in tackling the IRPLE bias. In addition to producing almost no bias, the IRIVE is observed to closely attain the derived analytical covariance and the theoretical Cramér-Rao lower bound. We also observe that the IRIVE converges well and finds accurate source position estimates even for large noise levels at which the Nelder-Mead based least l_p-norm solver exhibits divergence.

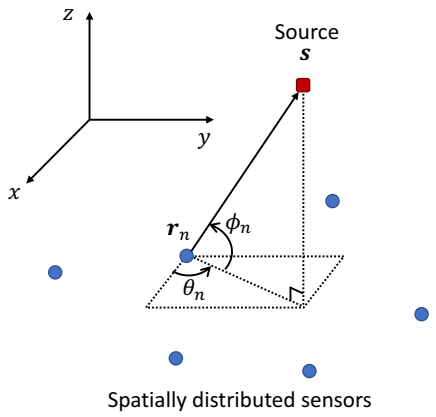


Fig. 1. 3D AOA target localization geometry.

The paper is organized as follows. Section II introduces the problem of robust 3D AOA localization in impulsive α-stable noise. Section III develops the IRPLE, and its bias analysis is presented in Section IV. In Section V, the IRIVE is derived. Section VI provides the theoretical covariance analysis for the IRIVE. Comparative simulation studies are presented in Section VII and concluding remarks are drawn in Section VIII.

II. Robust 3D AOA Localization in Impulsive Noise

A. Measurement Model

Fig. 1 depicts the problem of robust AOA source localization in 3D-space using azimuth and elevation angle measurements collected from N spatially distributed sensors. Let $\boldsymbol{s} = [s_x, s_y, s_z]^T$ denote the unknown source position and $\boldsymbol{r}_n = [r_{x,n}, r_{y,n}, r_{z,n}]^T$ the position of sensor n ($n \in \{1, \ldots, N\}$). Here, the superscript ^{T} stands for transpose operation.

The azimuth and elevation angle measurements taken at sensor n are given by

$$\tilde{\theta}_n = \theta_n(\boldsymbol{s}) + e_{\theta,n}, \quad \theta_n = \tan^{-1}\frac{s_y - r_{y,n}}{s_x - r_{x,n}} \tag{1a}$$

$$\tilde{\phi}_n = \phi_n(\boldsymbol{s}) + e_{\phi,n}, \quad \phi_n = \sin^{-1}\frac{s_z - r_{z,n}}{\|\boldsymbol{s} - \boldsymbol{r}_n\|} \tag{1b}$$

where $\tan^{-1}$ stands for the 4-quadrant arctangent and $\|\cdot\|$ the Euclidean norm. Here, $e_{\theta,n} \sim S\alpha S(\gamma_{\theta,n})$ and $e_{\phi,n} \sim S\alpha S(\gamma_{\phi,n})$ are independent zero-mean symmetric α-stable ($S\alpha S$) distributed random variables accounting for the impulsive measurement noise at the sensors. The noise dispersions $\gamma_{\theta,n}$ and $\gamma_{\phi,n}$ can vary with n and are assumed to be known *a priori*. In practice, the parameters of α-stable impulsive noise, including the dispersion γ and the impulsiveness parameter α, can be effectively estimated in real-time with high accuracy based on fractional lower order moments, logarithmic moments, or extreme value statistics (see e.g., [14], [35] and the references therein).

The characteristic function, i.e., the Fourier transform of the probability density function, of zero-mean $S\alpha S$ distribution is given by [6]–[10]

$$\varsigma(v) = \exp(-\gamma|v|^\alpha) \tag{2}$$

where $\alpha \in (0, 2]$ is the characteristic exponent parameter controlling the impulsiveness of the distribution and $\gamma > 0$ is the dispersion parameter measuring the deviation around the mean (analogous to the variance of the Gaussian distribution). Note that $\alpha = 2$ and 1 are special cases corresponding to the zero-mean Gaussian and Cauchy distributions, respectively.

$S\alpha S$ distribution only has finite moments for orders less than α, which are commonly known as fractional lower order moments. The fractional lower order moment of a zero-mean $S\alpha S$ distributed variable v is given by [14]

$$\mathbb{E}\{|v|^q\} = C_\alpha(q, \alpha)\gamma^{q/\alpha}, \quad q \in (-1, 1) \cup (1, \alpha) \tag{3}$$

where

$$C_\alpha(q, \alpha) = \frac{\Gamma\left(\frac{q+1}{2}\right)\Gamma(-\frac{q}{\alpha})}{\alpha\sqrt{\pi}\,\Gamma\left(-\frac{q}{2}\right)}\, 2^{(q+1)}. \tag{4}$$

Here, $\cup$ stands for the union operator and $\Gamma(v) = \int_0^\infty \exp(-t)t^{(v-1)}dt$ is the gamma function. In addition, we also have [25]

$$\mathbb{E}\{|v|^{(q-1)}\,\mathrm{sign}(v)\} = 0. \tag{5}$$

B. Localization Objective Function

As motivated in Section I, the objective of the AOA localization problem under consideration is to robustly estimate the unknown target position $\boldsymbol{p}$ from N impulsive noise-corrupted azimuth/elevation angle measurements based on l_p-norm minimization. To formulate the l_p-norm objective function, the azimuth and elevation angle measurements in (1) are normalized as

$$\tilde{\psi}_{\theta,n} = \psi_{\theta,n}(\boldsymbol{s}) + \epsilon_{\theta,n} \tag{6a}$$
$$\tilde{\psi}_{\phi,n} = \psi_{\phi,n}(\boldsymbol{s}) + \epsilon_{\phi,n} \tag{6b}$$

with

$$\begin{aligned}&\tilde{\psi}_{\theta,n} = \gamma_{\theta,n}^{-1/\alpha}\tilde{\theta}_n,\ \psi_{\theta,n}(\boldsymbol{s}) = \gamma_{\theta,n}^{-1/\alpha}\theta_n(\boldsymbol{s}),\ \epsilon_{\theta,n} = \gamma_{\theta,n}^{-1/\alpha}e_{\theta,n}\\ &\tilde{\psi}_{\phi,n} = \gamma_{\phi,n}^{-1/\alpha}\tilde{\phi}_n,\ \psi_{\phi,n}(\boldsymbol{s}) = \gamma_{\phi,n}^{-1/\alpha}\phi_n(\boldsymbol{s}),\ \epsilon_{\phi,n} = \gamma_{\phi,n}^{-1/\alpha}e_{\phi,n}\end{aligned} \tag{7}$$

so that the normalized noise $\epsilon_{\theta,n}$ and $\epsilon_{\phi,n}$ are i.i.d. with unit noise dispersion. Rewriting (6) in vector form gives

$$\tilde{\boldsymbol{\psi}} = \boldsymbol{\psi}(\boldsymbol{s}) + \boldsymbol{\epsilon} \tag{8}$$

where

$$\tilde{\boldsymbol{\psi}} = [\tilde{\boldsymbol{\psi}}_\theta^T, \tilde{\boldsymbol{\psi}}_\phi^T]^T,\ \boldsymbol{\psi} = [\boldsymbol{\psi}_\theta^T, \boldsymbol{\psi}_\phi^T]^T,\ \boldsymbol{\epsilon} = [\boldsymbol{\epsilon}_\theta^T, \boldsymbol{\epsilon}_\phi^T]^T \tag{9}$$

with

$$\tilde{\boldsymbol{\psi}}_\theta = [\tilde{\psi}_{\theta,1}, \ldots, \tilde{\psi}_{\theta,N}]^T,\ \tilde{\boldsymbol{\psi}}_\phi = [\tilde{\psi}_{\phi,1}, \ldots, \tilde{\psi}_{\phi,N}]^T \tag{10a}$$
$$\boldsymbol{\psi}_\theta = [\psi_{\theta,1}, \ldots, \psi_{\theta,N}]^T,\ \boldsymbol{\psi}_\phi = [\psi_{\phi,1}, \ldots, \psi_{\phi,N}]^T \tag{10b}$$
$$\boldsymbol{\epsilon}_\theta = [\epsilon_{\theta,1}, \ldots, \epsilon_{\theta,N}]^T,\ \boldsymbol{\epsilon}_\phi = [\epsilon_{\phi,1}, \ldots, \epsilon_{\phi,N}]^T. \tag{10c}$$

The AOA localization problem is now stated as solving the following l_p-norm minimization:

$$\hat{\boldsymbol{s}} = \arg\min_{\boldsymbol{s}} J(\boldsymbol{s}) \text{ with } J(\boldsymbol{s}) = \|\tilde{\boldsymbol{\psi}} - \boldsymbol{\psi}(\boldsymbol{s})\|_p^p. \tag{11}$$

for $p \in [1, 2)$. Here, the l_p-norm of a vector $\boldsymbol{v}$ is defined as $\|\boldsymbol{v}\|_p = (\sum |v_i|^p)^{1/p}$ where v_i is the ith element of $\boldsymbol{v}$.

III. Iteratively Reweighted Pseudolinear Least-Squares Estimator

In order to apply the IRLS technique, the nonlinear azimuth and elevation angle measurement equations (1) are rewritten in pseudolinear forms as

$$\boldsymbol{A}_{\theta,n}\boldsymbol{s} = b_{\theta,n} + \eta_{\theta,n} \tag{12}$$

where

$$\boldsymbol{A}_{\theta,n} = [\sin\tilde{\theta}_n, -\cos\tilde{\theta}_n, 0] \tag{13a}$$
$$b_{\theta,n} = [\sin\tilde{\theta}_n, -\cos\tilde{\theta}_n, 0]\,\boldsymbol{r}_n \tag{13b}$$
$$\begin{aligned}\eta_{\theta,n} &= \|\boldsymbol{d}_n\|\cos\phi_n \sin e_{\theta,n}\\ &\approx (\|\boldsymbol{d}_n\|\cos\phi_n)e_{\theta,n}\end{aligned} \tag{13c}$$

and

$$\boldsymbol{A}_{\phi,n}\boldsymbol{s} = b_{\phi,n} + \eta_{\phi,n} \tag{14}$$

with

$$\boldsymbol{A}_{\phi,n} = [\sin\tilde{\phi}_n\cos\tilde{\theta}_n, \sin\tilde{\phi}_n\sin\tilde{\theta}_n, -\cos\tilde{\phi}_n] \tag{15a}$$
$$b_{\phi,n} = [\sin\tilde{\phi}_n\cos\tilde{\theta}_n, \sin\tilde{\phi}_n\sin\tilde{\theta}_n, -\cos\tilde{\phi}_n]\,\boldsymbol{r}_n \tag{15b}$$
$$\begin{aligned}\eta_{\phi,n} &= \|\boldsymbol{d}_n\|(\sin e_{\phi,n} - 2\sin\tilde{\phi}_n\cos\phi_n\sin^2(e_{\theta,n}/2))\\ &\approx \|\boldsymbol{d}_n\|e_{\phi,n}.\end{aligned} \tag{15c}$$

Here, $\boldsymbol{d}_n = \boldsymbol{s} - \boldsymbol{r}_n$ and the approximations of pseudolinear noise in (13c) and (15c) are obtained for sufficiently small noise. These pseudolinear azimuth/elevation angle equations were derived based on the orthogonal vector method (see e.g., [36] for details).

Stacking (12) for $n = 1, \ldots, N$ after multiplying both sides of (12) with $u_{\theta,n} = \left(\gamma_{\theta,n}^{1/\alpha}\|\boldsymbol{d}_n\|\cos\phi_n\right)^{-1}$ yields

$$\boldsymbol{U}_\theta\boldsymbol{A}_\theta\boldsymbol{s} \approx \boldsymbol{U}_\theta\boldsymbol{b}_\theta + \boldsymbol{\epsilon}_\theta \tag{16}$$

where $\boldsymbol{U}_\theta = \mathrm{diag}(u_{\theta,1}, \ldots, u_{\theta,N})$, $\boldsymbol{A}_\theta = [\boldsymbol{A}_{\theta,1}^T, \ldots, \boldsymbol{A}_{\theta,N}^T]^T$, $\boldsymbol{b}_\theta = [b_{\theta,1}, \ldots, b_{\theta,N}]^T$ and $\boldsymbol{\epsilon}_\theta = [\epsilon_{\theta,1}, \ldots, \epsilon_{\theta,N}]^T$. Similarly for (14) with $u_{\phi,n} = \left(\gamma_{\phi,n}^{1/\alpha}\|\boldsymbol{d}_n\|\right)^{-1}$, we have

$$\boldsymbol{U}_\phi\boldsymbol{A}_\phi\boldsymbol{s} \approx \boldsymbol{U}_\phi\boldsymbol{b}_\phi + \boldsymbol{\epsilon}_\phi \tag{17}$$

where $\boldsymbol{U}_\phi = \mathrm{diag}(u_{\phi,1}, \ldots, u_{\phi,N})$, $\boldsymbol{A}_\phi = [\boldsymbol{A}_{\phi,1}^T, \ldots, \boldsymbol{A}_{\phi,N}^T]^T$, $\boldsymbol{b}_\phi = [b_{\phi,1}, \ldots, b_{\phi,N}]^T$ and $\boldsymbol{\epsilon}_\phi = [\epsilon_{\phi,1}, \ldots, \epsilon_{\phi,N}]^T$. By stacking (16) and (17), we obtain

$$\boldsymbol{U}\boldsymbol{A}\boldsymbol{s} \approx \boldsymbol{U}\boldsymbol{b} + \boldsymbol{\epsilon} \tag{18}$$

where $\boldsymbol{U} = \mathrm{diag}(\boldsymbol{U}_\theta, \boldsymbol{U}_\phi)$, $\boldsymbol{A} = [\boldsymbol{A}_\theta^T, \boldsymbol{A}_\phi^T]^T$, $\boldsymbol{b} = [\boldsymbol{b}_\theta^T, \boldsymbol{b}_\phi^T]^T$ and $\boldsymbol{\epsilon} = [\boldsymbol{\epsilon}_\theta^T, \boldsymbol{\epsilon}_\phi^T]^T$.

From (8) and (18), the l_p-norm objective function in (11) is approximated as

$$J(\boldsymbol{s}) \approx \|\boldsymbol{U}\boldsymbol{A}\boldsymbol{s} - \boldsymbol{U}\boldsymbol{b}\|_p^p \tag{19}$$

which we refer to as the *pseudolinear l_p-norm objective function*. Applying the IRLS technique to this pseudolinear l_p-norm objective function leads to a new IRPLE method:

- Initialization: $\hat{\boldsymbol{s}}_0 = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$
- At iteration k:
  1) Compute the weighting matrix $\boldsymbol{W}$ using $\hat{\boldsymbol{s}}_{k-1}$ from the previous iteration $k-1$:

$$\boldsymbol{W}(\hat{\boldsymbol{s}}_{k-1}) = \mathrm{diag}(\boldsymbol{W}_\theta, \boldsymbol{W}_\phi) \tag{20}$$

where

$$\boldsymbol{W}_\theta = \mathrm{diag}(w_{\theta,1}, \ldots, w_{\theta,N}) \tag{21a}$$

$$w_{\theta,n} = \gamma_{\theta,n}^{-p/\alpha}(\|\boldsymbol{d}_n(\hat{\boldsymbol{s}}_{k-1})\|\cos\phi_n(\hat{\boldsymbol{s}}_{k-1}))^{-2} \times |\tilde{\theta}_n - \theta_n(\hat{\boldsymbol{s}}_{k-1})|^{p-2} \tag{21b}$$

$$\boldsymbol{W}_\phi = \mathrm{diag}(w_{\phi,1}, \ldots, w_{\phi,N}) \tag{21c}$$

$$w_{\phi,n} = \gamma_{\phi,n}^{-p/\alpha}\|\boldsymbol{d}_n(\hat{\boldsymbol{s}}_{k-1})\|^{-2}|\tilde{\phi}_n - \phi_n(\hat{\boldsymbol{s}}_{k-1})|^{p-2} \tag{21d}$$

  2) Perform least-squares estimation:

$$\hat{\boldsymbol{s}}_k = (\boldsymbol{A}^T\boldsymbol{W}(\hat{\boldsymbol{s}}_{k-1})\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{W}(\hat{\boldsymbol{s}}_{k-1})\boldsymbol{b}. \tag{22}$$

Note that $\boldsymbol{U}$ does not appear in (22) because it has already been incorporated into the weighting matrix $\boldsymbol{W}$ in (21). The IRPLE is halted when the relative change in the l_p-norm of position estimates between two consecutive iterations falls below than a threshold or when a preset number of iterations is reached.

IV. Bias Problem of the IRPLE

The IRPLE developed in the previous section suffers from a bias problem due to the injection of noise into the pseudolinear measurement matrix $\boldsymbol{A}$ as a result of the nonlinear-to-pseudolinear transformation of azimuth/elevation angle measurement equations. In what follows, we present a detailed analysis of the IRPLE bias.

Theorem 1. *The IRPLE is biased. The bias of the IRPLE is given by*

$$\boldsymbol{\delta}_{\mathrm{IRPLE}} \approx -\frac{1}{p-1}\mathbb{E}\left\{\frac{\boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{A}}{N}\right\}^{-1}\mathbb{E}\left\{\frac{\boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{\eta}}{N}\right\}. \tag{23}$$

for sufficiently large N. Here, $\boldsymbol{W}^o = \boldsymbol{W}(\boldsymbol{s})$.

Proof. On the convergence of the IRPLE, we have

$$\hat{\boldsymbol{s}}^*_{\mathrm{IRPLE}} = \left(\boldsymbol{A}^T\boldsymbol{W}(\hat{\boldsymbol{s}}^*_{\mathrm{IRPLE}})\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{W}(\hat{\boldsymbol{s}}^*_{\mathrm{IRPLE}})\boldsymbol{b}. \tag{24}$$

It is noted that the IRPLE is essentially the direct application of the IRLS algorithm to the pseudolinear l_p-norm objective function (19), thus it follows the convergence of the IRLS algorithm. A detailed proof of the IRLS convergence can be found in [26].

Substituting $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{s} - \boldsymbol{U}^{-1}\boldsymbol{\epsilon} = \boldsymbol{A}\boldsymbol{s} - \boldsymbol{\eta}$ into (24) yields

$$\begin{aligned}\Delta\boldsymbol{s}_{\mathrm{IRPLE}} &= \hat{\boldsymbol{s}}^*_{\mathrm{IRPLE}} - \boldsymbol{s} \\ &= -\left(\boldsymbol{A}^T\boldsymbol{W}(\hat{\boldsymbol{s}}^*_{\mathrm{IRPLE}})\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{W}(\hat{\boldsymbol{s}}^*_{\mathrm{IRPLE}})\boldsymbol{\eta},\end{aligned} \tag{25}$$

and thus

$$\left(\boldsymbol{A}^T\boldsymbol{W}(\hat{\boldsymbol{s}}^*_{\mathrm{IRPLE}})\boldsymbol{A}\right)\Delta\boldsymbol{s}_{\mathrm{IRPLE}} = -\boldsymbol{A}^T\boldsymbol{W}(\hat{\boldsymbol{s}}^*_{\mathrm{IRPLE}})\boldsymbol{\eta}. \tag{26}$$

The weighting matrix $\boldsymbol{W}(\hat{\boldsymbol{s}}^*_{\mathrm{IRPLE}})$ can be written as

$$\boldsymbol{W}(\hat{\boldsymbol{s}}^*_{\mathrm{IRPLE}}) = \boldsymbol{W}^o + \Delta\boldsymbol{W} \tag{27}$$

where

$$\boldsymbol{W}^o = \boldsymbol{W}(\boldsymbol{s}) \tag{28a}$$

$$\Delta\boldsymbol{W} = \mathrm{diag}(\Delta\boldsymbol{W}_\theta, \Delta\boldsymbol{W}_\phi) \tag{28b}$$

$$\Delta\boldsymbol{W}_\theta = \mathrm{diag}(\Delta w_{\theta,1}, \ldots, \Delta w_{\theta,N}) \tag{28c}$$

$$\Delta\boldsymbol{W}_\phi = \mathrm{diag}(\Delta w_{\phi,1}, \ldots, \Delta w_{\phi,N}) \tag{28d}$$

$$\Delta w_{\theta,n} \approx \mathrm{sign}(u_{\theta,n}^{-1}\epsilon_{\theta,n})(p-2)|u_{\theta,n}|^3|\epsilon_{\theta,n}|^{p-3}\boldsymbol{A}_{\theta,n}\Delta\boldsymbol{s}_{\mathrm{IRPLE}} \tag{28e}$$

$$\Delta w_{\phi,n} \approx \mathrm{sign}(u_{\phi,n}^{-1}\epsilon_{\phi,n})(p-2)|u_{\phi,n}|^3|\epsilon_{\phi,n}|^{p-3}\boldsymbol{A}_{\phi,n}\Delta\boldsymbol{s}_{\mathrm{IRPLE}}. \tag{28f}$$

Using (28), we obtain

$$\Delta\boldsymbol{W}_\theta\boldsymbol{\eta}_\theta \approx (p-2)\boldsymbol{W}_\theta^o\boldsymbol{A}_\theta\Delta\boldsymbol{s}_{\mathrm{IRPLE}} \tag{29a}$$

$$\Delta\boldsymbol{W}_\phi\boldsymbol{\eta}_\phi \approx (p-2)\boldsymbol{W}_\phi^o\boldsymbol{A}_\phi\Delta\boldsymbol{s}_{\mathrm{IRPLE}} \tag{29b}$$

where $\boldsymbol{W}_\theta^o = \boldsymbol{W}_\theta(\boldsymbol{s})$ and $\boldsymbol{W}_\phi^o = \boldsymbol{W}_\phi(\boldsymbol{s})$. Following (27)–(29), the term $\boldsymbol{A}^T\boldsymbol{W}(\hat{\boldsymbol{s}}^*_{\mathrm{IRPLE}})\boldsymbol{\eta}$ on the right-hand side of (26) is given by

$$\begin{aligned}&\boldsymbol{A}^T\boldsymbol{W}(\hat{\boldsymbol{s}}^*_{\mathrm{IRPLE}})\boldsymbol{\eta} \\ &\quad\approx \boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{\eta} + (p-2)\left(\boldsymbol{A}_\theta^T\boldsymbol{W}_\theta^o\boldsymbol{A}_\theta\Delta\boldsymbol{s}_{\mathrm{IRPLE}}\right) \\ &\qquad + (p-2)\left(\boldsymbol{A}_\phi^T\boldsymbol{W}_\phi^o\boldsymbol{A}_\phi\Delta\boldsymbol{s}_{\mathrm{IRPLE}}\right) \\ &\quad\approx \boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{\eta} + (p-2)\left(\boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{A}\right)\Delta\boldsymbol{s}_{\mathrm{IRPLE}}.\end{aligned} \tag{30}$$

From (26) and (30), we obtain

$$\Delta\boldsymbol{s}_{\mathrm{IRPLE}} \approx -\frac{1}{p-1}\left(\boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{\eta}. \tag{31}$$

As a result, the bias of the IRLS estimate $\hat{\boldsymbol{s}}^*_{\mathrm{IRPLE}}$ is

$$\boldsymbol{\delta}_{\mathrm{IRPLE}} = \mathbb{E}\{\Delta\boldsymbol{s}_{\mathrm{IRPLE}}\} \tag{32a}$$

$$\approx -\frac{1}{p-1}\mathbb{E}\left\{\left(\boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{\eta}\right\} \tag{32b}$$

$$\approx -\frac{1}{p-1}\mathbb{E}\left\{\frac{\boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{A}}{N}\right\}^{-1}\mathbb{E}\left\{\frac{\boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{\eta}}{N}\right\}. \tag{32c}$$

Here, the approximation in (32c) follows from Slutsky's theorem [37] under the tacit assumption of sufficiently large N for the purpose of bias analysis. The second expectation term in (32c) is given by

$$\begin{aligned}\mathbb{E}\left\{\frac{\boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{\eta}}{N}\right\} &= \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left\{\boldsymbol{A}_{\theta,n}^T w_{\theta,n}^o\eta_{\theta,n}\right\} \\ &\quad + \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left\{\boldsymbol{A}_{\phi,n}^T w_{\phi,n}^o\eta_{\phi,n}\right\}\end{aligned} \tag{33}$$

where $w_{\theta,n}^o = w_{\theta,n}(\boldsymbol{s})$ and $w_{\phi,n}^o = w_{\phi,n}(\boldsymbol{s})$. It is important to note that $\boldsymbol{A}_{\theta,n}$ is correlated with $\eta_{\theta,n}$ because both are functions of $e_{\theta,n}$. Similarly, there also exists correlation between $\boldsymbol{A}_{\phi,n}$ and $\eta_{\phi,n}$. As a result, the expectation (33) is nonzero, i.e.,

$$\mathbb{E}\left\{\frac{\boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{\eta}}{N}\right\} \neq \boldsymbol{0}. \tag{34}$$

Consequently, we have

$$\boldsymbol{\delta}_{\text{IRPLE}} \neq \mathbf{0}, \tag{35}$$

thus implying the biasedness of the IRPLE. □

An explicit expression for (33) is in general difficult to derive due to the nonlinear functions in $\boldsymbol{A}_{\theta,n}$ and $\boldsymbol{A}_{\phi,n}$. For sufficiently small measurement noise, we can approximate (33) as follows. The matrix $\boldsymbol{A}_{\theta,n}^T$ can be decomposed into $\boldsymbol{A}_{\theta,n}^T \approx \boldsymbol{A}_{\theta,n}^{oT} + \boldsymbol{D}_{\theta,n}\gamma_{\theta,n}^{1/\alpha}\epsilon_{\theta,n}$, where $\boldsymbol{A}_{\theta,n}^o$ is the noise-free versions of $\boldsymbol{A}_{\theta,n}$ by replacing $\tilde{\theta}_n$ with θ_n in (13a), and $\boldsymbol{D}_{\theta,n} = [\cos\theta_n, \sin\theta_n, 0]^T$. As a result, we have

$$\begin{aligned}
&\boldsymbol{A}_{\theta,n}^T w_{\theta,n}^o \eta_{\theta,n} \\
&\quad\approx \left(\boldsymbol{A}_{\theta,n}^{oT} + \boldsymbol{D}_{\theta,n}\gamma_{\theta,n}^{1/\alpha}\epsilon_{\theta,n}\right) w_{\theta,n}^o \eta_{\theta,n} \\
&\quad\approx \boldsymbol{A}_{\theta,n}^{oT} u_{\theta,n}\,\text{sign}(\epsilon_{\theta,n})|\epsilon_{\theta,n}|^{p-1} + \boldsymbol{D}_{\theta,n}\gamma_{\theta,n}^{1/\alpha} u_{\theta,n}|\epsilon_{\theta,n}|^p.
\end{aligned} \tag{36}$$

By exploiting (3) and (5), the expectation of (36) reduces to

$$\mathbb{E}\left\{\boldsymbol{A}_{\theta,n}^T w_{\theta,n}^o \eta_{\theta,n}\right\} \approx \boldsymbol{D}_{\theta,n}\gamma_{\theta,n}^{1/\alpha} u_{\theta,n} C_\alpha(p,\alpha). \tag{37}$$

Similarly, we can decompose $\boldsymbol{A}_{\phi,n}^T$ into $\boldsymbol{A}_{\phi,n}^T = \boldsymbol{A}_{\phi,n}^{oT} + \boldsymbol{D}_{\phi,n}^{(1)}\gamma_{\phi,n}^{1/\alpha}\epsilon_{\phi,n} + \boldsymbol{D}_{\phi,n}^{(2)}\gamma_{\theta,n}^{1/\alpha}\epsilon_{\theta,n}$, where $\boldsymbol{A}_{\phi,n}^o$ is the noise-free version of $\boldsymbol{A}_{\phi,n}$ by replacing $\tilde{\theta}_n$ and $\tilde{\phi}_n$ with θ_n and ϕ_n in (15a), and $\boldsymbol{D}_{\phi,n}^{(1)}$ and $\boldsymbol{D}_{\phi,n}^{(2)}$ are defined as $\boldsymbol{D}_{\phi,n}^{(1)} = [\cos\phi_n\cos\theta_n, \cos\phi_n\sin\theta_n, \sin\phi_n]^T$ and $\boldsymbol{D}_{\phi,n}^{(2)} = [-\sin\phi_n\sin\theta_n, \sin\phi_n\cos\theta_n, 0]^T$. Consequently, we obtain

$$\begin{aligned}
&\boldsymbol{A}_{\phi,n}^T w_{\phi,n}^o \eta_{\phi,n} \\
&\quad\approx \left(\boldsymbol{A}_{\phi,n}^{oT} + \boldsymbol{D}_{\phi,n}^{(1)}\gamma_{\phi,n}^{1/\alpha}\epsilon_{\phi,n} + \boldsymbol{D}_{\phi,n}^{(2)}\gamma_{\theta,n}^{1/\alpha}\epsilon_{\theta,n}\right) w_{\phi,n}^o \eta_{\phi,n} \\
&\quad\approx \boldsymbol{A}_{\phi,n}^{oT} u_{\phi,n}\,\text{sign}(\epsilon_{\phi,n})|\epsilon_{\phi,n}|^{p-1} + \boldsymbol{D}_{\phi,n}^{(1)}\gamma_{\phi,n}^{1/\alpha} u_{\phi,n}|\epsilon_{\phi,n}|^p \\
&\qquad + \boldsymbol{D}_{\phi,n}^{(2)}\gamma_{\theta,n}^{1/\alpha} u_{\phi,n}\epsilon_{\theta,n}\,\text{sign}(\epsilon_{\phi,n})|\epsilon_{\phi,n}|^{p-1}.
\end{aligned} \tag{38}$$

Using (3) and (5), the expectation of (38) reduces to

$$\mathbb{E}\left\{\boldsymbol{A}_{\phi,n}^T w_{\phi,n}^o \eta_{\phi,n}\right\} = \boldsymbol{D}_{\phi,n}^{(1)}\gamma_{\phi,n}^{1/\alpha} u_{\phi,n} C_\alpha(p,\alpha). \tag{39}$$

Finally, by substituting (37) and (39) into (33), we have

$$\begin{aligned}
\mathbb{E}\left\{\frac{\boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{\eta}}{N}\right\} \approx \frac{C_\alpha(p,\alpha)}{N}\sum_{n=1}^{N}\Big(&\boldsymbol{D}_{\theta,n}\gamma_{\theta,n}^{1/\alpha} u_{\theta,n} \\
&+\boldsymbol{D}_{\phi,n}^{(1)}\gamma_{\phi,n}^{1/\alpha} u_{\phi,n}\Big).
\end{aligned} \tag{40}$$

Consequently, we obtain the following corollary.

Corollary 1. *Under small noise, the expression of the IRPLE bias is reduced to*

$$\begin{aligned}
\boldsymbol{\delta}_{\text{IRPLE}} \approx -\frac{C_\alpha(p,\alpha)}{p-1}\mathbb{E}\left\{\boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{A}\right\}^{-1}\sum_{n=1}^{N}\Big(&\boldsymbol{D}_{\theta,n}\gamma_{\theta,n}^{1/\alpha} u_{\theta,n} \\
&+\boldsymbol{D}_{\phi,n}^{(1)}\gamma_{\phi,n}^{1/\alpha} u_{\phi,n}\Big).
\end{aligned} \tag{41}$$

Proof. A simple substitution of (40) into (23) leads to (41). □

V. Iteratively Reweighted Instrumental-Variable Estimator

To overcome the bias problem of the IRPLE, we now propose a new IRIVE method exploiting the use of instrumental variables. Motivated by the fact that the main source for the IRPLE bias problem is the correlation between the measurement matrix $\boldsymbol{A}$ and the pseudolinear noise vector $\boldsymbol{\eta}$, the proposed IRIVE aims to eliminate this undesirable correlation by replacing $\boldsymbol{A}$ with a new matrix $\boldsymbol{G}$ in (22). Thus, the IRIVE estimate $\hat{\boldsymbol{s}}_k^{\text{IV}}$ at iteration k is given by

$$\hat{\boldsymbol{s}}_k^{\text{IV}} = (\boldsymbol{G}^T\boldsymbol{W}\boldsymbol{A})^{-1}\boldsymbol{G}^T\boldsymbol{W}\boldsymbol{b}. \tag{42}$$

Here, $\boldsymbol{G}$ is called the IV matrix [27], [28] which is ideally chosen such that $\boldsymbol{G}^T\boldsymbol{A}$ is full-rank and $\boldsymbol{E}\{\boldsymbol{G}^T\boldsymbol{\eta}\} = \mathbf{0}$ [28]. Note that the bias propagation through $\boldsymbol{W}$ from previous iterations is not the actual cause for the IRPLE bias and will naturally vanish once the bias problem due to the correlation between $\boldsymbol{A}$ and $\boldsymbol{\eta}$ is resolved.

Since the noise-free version $\boldsymbol{A}_\circ$ of $\boldsymbol{A}$ (the optimal choice for the IV matrix $\boldsymbol{G}$ [38]) is not available, a suboptimal IV matrix $\boldsymbol{G}$ at iteration k (that yields $\boldsymbol{E}\{\boldsymbol{G}^T\boldsymbol{\eta}\} \approx \mathbf{0}$) is constructed by calculating $\boldsymbol{A}_\circ$ using the estimated azimuth and elevation angles computed from the position estimate $\hat{\boldsymbol{s}}_{k-1}^{\text{IV}}$ available in the previous iteration $k-1$:

$$\boldsymbol{G}(\hat{\boldsymbol{s}}_{k-1}^{\text{IV}}) = [\boldsymbol{G}_\theta^T, \boldsymbol{G}_\phi^T]^T \tag{43}$$

where

$$\boldsymbol{G}_\theta = [\boldsymbol{G}_{\theta,1}^T, \ldots, \boldsymbol{G}_{\theta,N}^T]^T \tag{44a}$$

$$\boldsymbol{G}_{\theta,n} = [\sin\hat{\theta}_n, -\cos\hat{\theta}_n, 0] \tag{44b}$$

$$\boldsymbol{G}_\phi = [\boldsymbol{G}_{\phi,1}^T, \ldots, \boldsymbol{G}_{\phi,N}^T]^T \tag{44c}$$

$$\boldsymbol{G}_{\phi,n} = [\sin\hat{\phi}_n\cos\hat{\theta}_n, \sin\hat{\phi}_n\sin\hat{\theta}_n, -\cos\hat{\phi}_n] \tag{44d}$$

and

$$\hat{\theta}_n = \theta_n(\hat{\boldsymbol{s}}_{k-1}^{\text{IV}}), \quad \hat{\phi}_n = \phi_n(\hat{\boldsymbol{s}}_{k-1}^{\text{IV}}). \tag{45}$$

To satisfy $\boldsymbol{G}^T\boldsymbol{A}$ is full-rank and well-conditioned as required by the IV estimation method, the measurement matrix $\boldsymbol{A}$ and the IV matrix $\boldsymbol{G}$ are required to correlate well. To ensure this, the selective-angle-measurement (SAM) strategy [4], [29], [39] is incorporated into the proposed IRIVE. The main idea behind the SAM strategy is to force some rows of $\boldsymbol{G}$ to be identical to the corresponding rows of $\boldsymbol{A}$ depending on the difference between the estimated and measured angles (i.e., if $|\tilde{\theta}_n - \hat{\theta}_n| > \beta_{\theta,n}$ or $|\tilde{\phi}_n - \hat{\phi}_n| > \beta_{\phi,n}$) while reducing the contribution of these rows by dividing the corresponding weights by a large scalar κ. The SAM thresholds $\beta_{\theta,n}$ and $\beta_{\phi,n}$ should generally be chosen large enough to promote the use of instrumental variables while being not too large to maintain a strong correlation between $\boldsymbol{A}$ and $\boldsymbol{G}$. The selection of $\beta_{\theta,n}$ and $\beta_{\phi,n}$ typically depends on the measurement noise level. For the problem under consideration, the choice of these parameters does not critically affect the performance of the IRIVE (see Section VII-G for further discussion).

The computational steps of the proposed IRIVE algorithm are summarized below:

- Initialization: $\hat{\boldsymbol{s}}_0^{\text{IV}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$

- At iteration k:
  1) Compute $\hat{\theta}_n$ and $\hat{\phi}_n$ using $\hat{\boldsymbol{s}}^{\text{IV}}_{k-1}$ from the previous iteration $k-1$:
$$\hat{\theta}_n = \tan^{-1}\frac{\hat{s}^{\text{IV}}_{y,k-1} - r_{y,n}}{\hat{s}^{\text{IV}}_{x,k-1} - r_{x,n}} \tag{46a}$$
$$\hat{\phi}_n = \sin^{-1}\frac{\hat{s}^{\text{IV}}_{z,k-1} - r_{z,n}}{\|\hat{\boldsymbol{s}}^{\text{IV}}_{k-1} - \boldsymbol{r}_n\|} \tag{46b}$$
  where $\hat{\boldsymbol{s}}^{\text{IV}}_{k-1} = [\hat{s}^{\text{IV}}_{x,k-1}, \hat{s}^{\text{IV}}_{y,k-1}, \hat{s}^{\text{IV}}_{z,k-1}]^T$.
  2) Compute the weighting coefficients $w_{\theta,n}$ and $w_{\phi,n}$ using $\hat{\boldsymbol{s}}^{\text{IV}}_{k-1}$:
$$\begin{aligned} w_{\theta,n} &= \gamma_{\theta,n}^{-p/\alpha}(\|\boldsymbol{d}_n(\hat{\boldsymbol{s}}^{\text{IV}}_{k-1})\|\cos\phi_n(\hat{\boldsymbol{s}}^{\text{IV}}_{k-1}))^{-2} \\ &\quad\times|\tilde{\theta}_n - \theta_n(\hat{\boldsymbol{s}}^{\text{IV}}_{k-1})|^{p-2} \end{aligned} \tag{47a}$$
$$w_{\phi,n} = \gamma_{\phi,n}^{-p/\alpha}\|\boldsymbol{d}_n(\hat{\boldsymbol{s}}^{\text{IV}}_{k-1})\|^{-2}|\tilde{\phi}_n - \phi_n(\hat{\boldsymbol{s}}^{\text{IV}}_{k-1})|^{p-2} \tag{47b}$$
  3) Perform the SAM:
$$\begin{aligned} &\textbf{If } |\tilde{\theta}_n - \hat{\theta}_n| < \beta_{\theta,n} \text{ and } |\tilde{\phi}_n - \hat{\phi}_n| < \beta_{\phi,n} \\ &\quad\text{set } \hat{\theta}_n = \hat{\theta}_n,\ \hat{\phi}_n = \hat{\phi}_n \\ &\qquad w_{\theta,n} = w_{\theta,n},\ w_{\phi,n} = w_{\phi,n} \\ &\textbf{otherwise} \\ &\quad\text{set } \hat{\theta}_n = \tilde{\theta}_n,\ \hat{\phi}_n = \tilde{\phi}_n \\ &\qquad w_{\theta,n} = w_{\theta,n}/\kappa,\ w_{\phi,n} = w_{\phi,n}/\kappa. \end{aligned} \tag{48}$$
  4) Compute the IV matrix $\boldsymbol{G}$:
$$\boldsymbol{G}(\hat{\boldsymbol{s}}^{\text{IV}}_{k-1}) = [\boldsymbol{G}_\theta^T, \boldsymbol{G}_\phi^T]^T \tag{49}$$
  where
$$\boldsymbol{G}_\theta = [\boldsymbol{G}_{\theta,1}^T, \ldots, \boldsymbol{G}_{\theta,N}^T]^T \tag{50a}$$
$$\boldsymbol{G}_{\theta,n} = [\sin\hat{\theta}_n, -\cos\hat{\theta}_n, 0] \tag{50b}$$
$$\boldsymbol{G}_\phi = [\boldsymbol{G}_{\phi,1}^T, \ldots, \boldsymbol{G}_{\phi,N}^T]^T \tag{50c}$$
$$\boldsymbol{G}_{\phi,n} = [\sin\hat{\phi}_n\cos\hat{\theta}_n, \sin\hat{\phi}_n\sin\hat{\theta}_n, -\cos\hat{\phi}_n] \tag{50d}$$
  5) Construct the weighting matrix $\boldsymbol{W}$:
$$\boldsymbol{W}_\theta = \text{diag}(w_{\theta,1}, \ldots, w_{\theta,N}) \tag{51a}$$
$$\boldsymbol{W}_\phi = \text{diag}(w_{\phi,1}, \ldots, w_{\phi,N}) \tag{51b}$$
$$\boldsymbol{W} = \text{diag}(\boldsymbol{W}_\theta, \boldsymbol{W}_\phi) \tag{51c}$$
  6) Perform IV estimation:
$$\hat{\boldsymbol{s}}^{\text{IV}}_k = (\boldsymbol{G}^T\boldsymbol{W}\boldsymbol{A})^{-1}\boldsymbol{G}^T\boldsymbol{W}\boldsymbol{b}. \tag{52}$$

VI. Performance Analysis

In this section, we present a theoretical performance analysis for the proposed IRIVE algorithm and show that the IRIVE estimate can achieve the analytical covariance of the general least l_p-norm estimation.

Theorem 2. *Under the common assumption of sufficiently small noise and large number of measurements, the analytical covariance of the IRIVE is given by*

$$\boldsymbol{C}_{\text{IRIVE}} \approx \frac{C_\alpha(2p-2,\alpha)}{((p-1)C_\alpha(p-2,\alpha))^2}\left(\boldsymbol{G}^{oT}\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{G}^o\right)^{-1} \tag{53}$$

where $\boldsymbol{G}^o = \boldsymbol{G}(\boldsymbol{s})$.

Proof. Since the IRIVE is a variant of the IRLS, its convergence proof follows that of the IRLS as presented in [26] and is omitted here for brevity. On the convergence of the IRIVE, we have

$$\begin{aligned} \hat{\boldsymbol{s}}^*_{\text{IRIVE}} = &\left(\boldsymbol{G}(\hat{\boldsymbol{s}}^*_{\text{IRIVE}})^T\boldsymbol{W}(\hat{\boldsymbol{s}}^*_{\text{IRIVE}})\boldsymbol{A}\right)^{-1} \\ &\times\boldsymbol{G}^T(\hat{\boldsymbol{s}}^*_{\text{IRIVE}})\boldsymbol{W}(\hat{\boldsymbol{s}}^*_{\text{IRIVE}})\boldsymbol{b}. \end{aligned} \tag{54}$$

By following similar steps to (25)–(31), we can show that

$$\begin{aligned} \Delta\boldsymbol{s}_{\text{IRIVE}} &= \hat{\boldsymbol{s}}^*_{\text{IRIVE}} - \boldsymbol{s} \\ &\approx -\frac{1}{p-1}\left(\boldsymbol{G}^{oT}\boldsymbol{W}^o\boldsymbol{A}\right)^{-1}\boldsymbol{G}^{oT}\boldsymbol{W}^o\boldsymbol{\eta}. \end{aligned} \tag{55}$$

Therefore, the error covariance of $\hat{\boldsymbol{s}}^*_{\text{IRIVE}}$ is given by

$$\begin{aligned} \boldsymbol{C}_{\text{IRIVE}} &= \mathbb{E}\left\{\Delta\boldsymbol{s}_{\text{IRIVE}}\Delta\boldsymbol{s}^T_{\text{IRIVE}}\right\} \\ &\approx \frac{1}{(p-1)^2}\mathbb{E}\Big\{\left(\boldsymbol{G}^{oT}\boldsymbol{W}^o\boldsymbol{A}\right)^{-1}\boldsymbol{G}^{oT}\boldsymbol{W}^o\boldsymbol{\eta}\boldsymbol{\eta}^T\boldsymbol{W}^o\boldsymbol{G}^o \\ &\qquad\times\left(\boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{G}^o\right)^{-1}\Big\}. \end{aligned} \tag{56}$$

Applying Slutsky's theorem to (56) yields [37]

$$\begin{aligned} \boldsymbol{C}_{\text{IRIVE}} \approx &\frac{1}{(p-1)^2}\frac{1}{N}\mathbb{E}\left\{\frac{\boldsymbol{G}^{oT}\boldsymbol{W}^o\boldsymbol{A}}{N}\right\}^{-1} \\ &\times\mathbb{E}\left\{\frac{\boldsymbol{G}^{oT}\boldsymbol{W}^o\boldsymbol{\eta}\boldsymbol{\eta}^T\boldsymbol{W}^o\boldsymbol{G}^o}{N}\right\}\mathbb{E}\left\{\frac{\boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{G}^o}{N}\right\}^{-1}. \end{aligned} \tag{57}$$

The first expectation on the right-hand side of (57) can be approximated as

$$\begin{aligned} \mathbb{E}\left\{\frac{\boldsymbol{G}^{oT}\boldsymbol{W}^o\boldsymbol{A}}{N}\right\} &\approx \frac{\boldsymbol{G}^{oT}\mathbb{E}\{\boldsymbol{W}^o\}\boldsymbol{A}^o}{N} \\ &\approx \frac{\boldsymbol{G}^{oT}\mathbb{E}\{\boldsymbol{W}^o\}\boldsymbol{G}^o}{N} \end{aligned} \tag{58}$$

under the small noise assumption. Note that $\boldsymbol{A}^o$ is the noise-free version of $\boldsymbol{A}$ and in fact $\boldsymbol{A}^o = \boldsymbol{G}^o$. In addition, it is proved in Appendix B that

$$\mathbb{E}\{\boldsymbol{W}^o\} = C_\alpha(p-2,\alpha)\,\boldsymbol{U}^T\boldsymbol{U}. \tag{59}$$

Therefore, we have

$$\mathbb{E}\left\{\frac{\boldsymbol{G}^{oT}\boldsymbol{W}^o\boldsymbol{A}}{N}\right\} \approx \frac{C_\alpha(p-2,\alpha)}{N}\boldsymbol{G}^{oT}\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{G}^o. \tag{60}$$

Similarly, the third expectation of (57) is also approximated by

$$\mathbb{E}\left\{\frac{\boldsymbol{A}^T\boldsymbol{W}^o\boldsymbol{G}^o}{N}\right\} \approx \frac{C_\alpha(p-2,\alpha)}{N}\boldsymbol{G}^{oT}\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{G}^o. \tag{61}$$

By rewriting the second expectation of (57) as

$$\begin{aligned} &\mathbb{E}\left\{\frac{\boldsymbol{G}^{oT}\boldsymbol{W}^o\boldsymbol{\eta}\boldsymbol{\eta}^T\boldsymbol{W}^o\boldsymbol{G}^o}{N}\right\} \\ &\quad= \frac{1}{N}\boldsymbol{G}^{oT}\mathbb{E}\left\{(\boldsymbol{W}^o\boldsymbol{\eta})(\boldsymbol{W}^o\boldsymbol{\eta})^T\right\}\boldsymbol{G}^o \end{aligned} \tag{62}$$

and noting that (as shown in Appendix C)

$$\mathbb{E}\left\{(\boldsymbol{W}^o\boldsymbol{\eta})(\boldsymbol{W}^o\boldsymbol{\eta})^T\right\} = C_\alpha(2p-2,\alpha)\,\boldsymbol{U}^T\boldsymbol{U}, \tag{63}$$

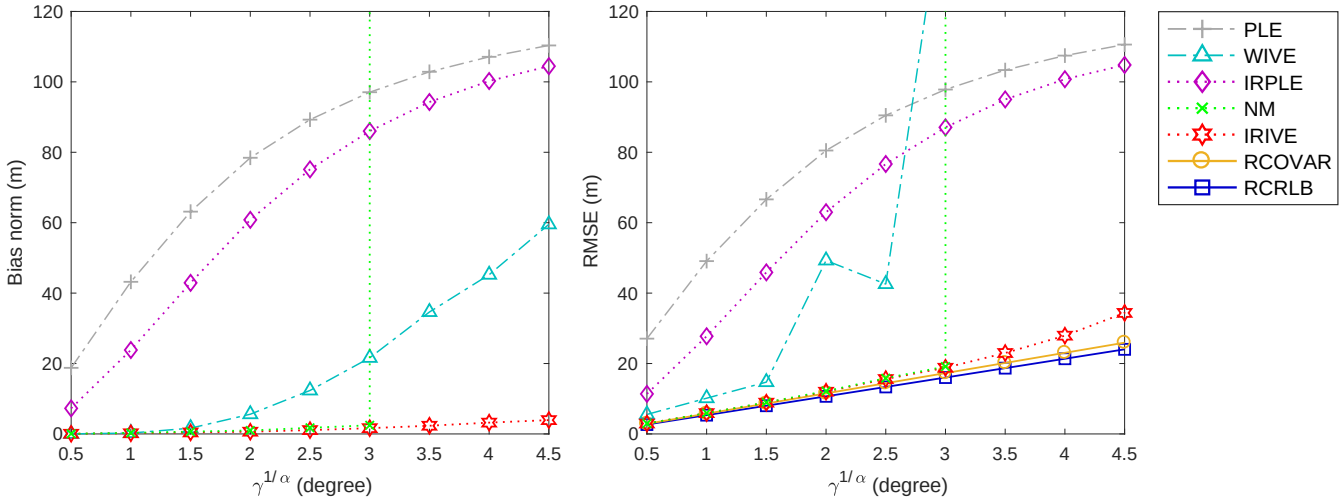


Fig. 2. Bias norm and RMSE versus noise level for the PLE, WIVE, IRPLE, NM and IRIVE algorithms.

we obtain

$$\mathbb{E}\left\{\frac{\boldsymbol{G}^{oT}\boldsymbol{W}^{o}\boldsymbol{\eta}\boldsymbol{\eta}^{T}\boldsymbol{W}^{o}\boldsymbol{G}^{o}}{N}\right\} \approx \frac{C_{\alpha}(2p-2,\alpha)}{N}\,\boldsymbol{G}^{oT}\boldsymbol{U}^{T}\boldsymbol{U}\boldsymbol{G}^{o}. \tag{64}$$

Substituting (60), (64) and (61) into (57), we obtain the analytical covariance of the IRIVE:

$$\boldsymbol{C}_{\text{IRIVE}} \approx \frac{C_{\alpha}(2p-2,\alpha)}{((p-1)C_{\alpha}(p-2,\alpha))^{2}}\left(\boldsymbol{G}^{oT}\boldsymbol{U}^{T}\boldsymbol{U}\boldsymbol{G}^{o}\right)^{-1}. \tag{65}$$

□

Corollary 2. *The IRIVE achieves the theoretical covariance of the general least l_p-norm estimation.*

Proof. The analytical covariance of estimators based on the l_p-norm minimization (11) is given by [25]

$$\boldsymbol{C}_{l_p\text{-norm}} \approx \frac{C_{\alpha}(2p-2,\alpha)}{((p-1)C_{\alpha}(p-2,\alpha))^{2}}(\boldsymbol{\nabla}^{T}\boldsymbol{\nabla})^{-1} \tag{66}$$

where $\boldsymbol{\nabla}$ is the Jacobian matrix of $\boldsymbol{\psi}$ in (8) with respect to $\boldsymbol{s}$ (see (67)–(69) in Appendix A). In fact, this covariance expression was derived in [25] assuming that the measurement noise is small and the number of measurements is sufficiently large.

After some algebraic manipulations, it is straightforward to show that $\boldsymbol{\nabla} = -\boldsymbol{U}\boldsymbol{G}^{o}$. Therefore, the terms on the right-hand side of (53) and (66) are identical. This implies that the analytical covariance of the IRIVE is equivalent to that of the general least l_p-norm estimation. □

VII. Simulation Examples

In this section, we present Monte Carlo simulations to evaluate the performance of the proposed IRPLE and IRIVE algorithms in comparison with the least l_p-norm solver based on the Nelder-Mead (NM) simplex method [40], as well as the existing pseudolinear least-squares algorithms including the pseudolinear estimator (PLE) and the weighted instrumental-variable estimator (WIVE) [2], [36]. The bias norm and root-mean-squared-error (RMSE) are used for performance comparison, which are estimated by $\left\|\frac{1}{L}\sum_{l=1}^{L}\left(\hat{\boldsymbol{s}}_{\text{X}}^{(l)}-\boldsymbol{s}\right)\right\|$ and $\left(\frac{1}{L}\sum_{l=1}^{L}\left\|\hat{\boldsymbol{s}}_{\text{X}}^{(l)}-\boldsymbol{s}\right\|^{2}\right)^{1/2}$, respectively. Here, $\hat{\boldsymbol{s}}_{\text{X}}^{(l)}$ denotes the source position estimate obtained by an algorithm, whose name is given in the subscript, at the l-th Monte Carlo run. The square root of the trace of the analytical covariance matrix of the general least l_p-norm estimation given in (66) (referred to as RCOVAR for simplicity) and the square root of the trace of the Cramér-Rao lower bound matrix [23], [25], [41] (referred to as RCRLB) are also calculated and used as the theoretical benchmark for the RMSE performance of the algorithms. Here, 'R' in RCOVAR and RCRLB is short for square root. Similar to the IRPLE and IRIVE, the NM is initialized to the PLE for ensuring a fair comparison. The IRPLE and IRIVE iterations are stopped after 20 iterations. The SAM parameters are set to: $\beta_{\theta,n} = \beta_{\phi,n} = 30^{\circ}$ and $\kappa = 10^{4}$.

A. Performance Versus Noise Dispersion

We consider a simulated 3D AOA source localization geometry with a target located at $\boldsymbol{s} = [80, 80, 60]^{T}$m and a sensor network with $N = 40$ sensors arranged uniformly in a circular configuration centered around the origin $[0, 0]^{T}$m with the radius of 15 m. The characteristic exponent parameter of impulsive noise is set to $\alpha = 1.5$, for which $p = 1.225$ is the optimum value [25] and is used in the simulations. The AOA measurement noise is i.i.d. with constant noise dispersion $\gamma_{\theta,n} = \gamma_{\phi,n} = \gamma$. A total of $L = 50{,}000$ Monte Carlo runs are carried out.

Fig. 2 shows the bias norm and RMSE performance of the estimators versus $\gamma^{1/\alpha}$, i.e., the α-order root of the noise dispersion, for $\gamma^{1/\alpha} \in \{0.5^{\circ}, 1^{\circ}, 1.5^{\circ}, \ldots, 4.5^{\circ}\}$. Note that $\gamma^{1/\alpha}$ is analogous to the standard deviation for Gaussian noise. We observe that the least-squares estimators (the PLE and WIVE), as expected, perform unreliably in the presence of impulsive noise, producing much larger bias norm and RMSE than the IRIVE. Among the least l_p-norm estimators, the IRPLE exhibits the worst performance due to its biasedness.

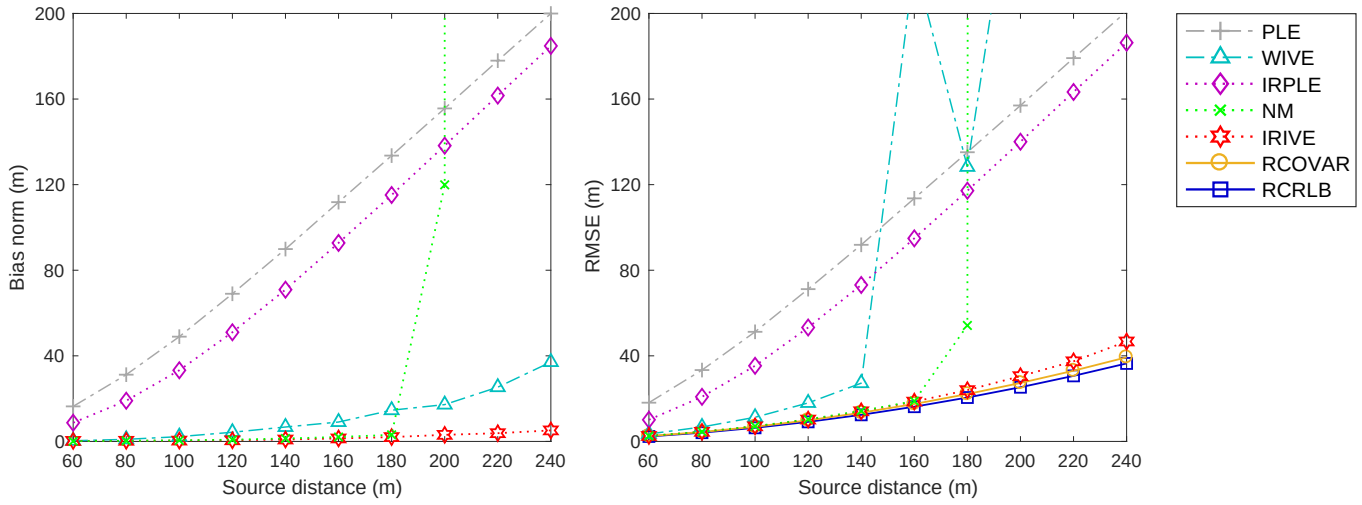


Fig. 3. Bias norm and RMSE versus sensor-source distance for the PLE, WIVE, IRPLE, NM and IRIVE algorithms.

Specifically, in agreement with the bias analysis presented in Section IV, we observe that the IRPLE yields a large estimation bias. Such a bias problem in turn leads to a poor RMSE performance for the IRPLE, far exceeding the RCRLB. On the other hand, by exploiting the use of IV estimation, the IRIVE is observed to be capable of outperforming the IRPLE by producing nearly unbiased estimates. Moreover, it is seen from Fig. 2 that the IRIVE closely attains the RCOVAR and RCRLB, thus confirming the efficiency of the IRIVE. The NM performs on par with the IRIVE for $\gamma^{1/\alpha} \leq 3°$, producing a negligible bias and achieving the RCOVAR and RCRLB. However, the NM exhibits divergence for $\gamma^{1/\alpha} > 3°$, while the IRIVE still appears to perform well at these large noise levels. This observation demonstrates the stability advantage of the IRIVE as inherited from the IRLS.

B. Performance Versus Sensor-Source Distance

We now evaluate the localization performance over the sensor-source distance using the same simulation setup as in Section VII-A except that the sensor noise level is kept at $\gamma^{1/\alpha} = 2°$ and the source position is given by $\boldsymbol{s} = d\,[\cos 30° \cos 45°, \cos 30° \sin 45°, \sin 30°]^T$. Here, d is essentially the distance between the source and the center of the sensor network. Fig. 3 shows the bias norm and RMSE performance versus the sensor-source distance d. Here, we observe a similar relative performance comparison to Fig. 2, once again confirming the performance superiority of the IRIVE over the PLE, WIVE, IRPLE and NM. We also observe from Fig. 3 that the bias and RMSE performance of the estimators degrades as the sensor-source distance d increases. This observation is consistent with the existing results for the dependence of the localization performance on the sensor-source geometry (see e.g., [42]).

C. Performance Versus Number of Sensors

Fig. 4 compares the bias norm and RMSE performance of the estimators over the number of sensors N ranging from 15 to 85 at $\gamma^{1/\alpha} = 2°$. Other simulation parameters remain the same as in Section VII-A. We observe that the IRPLE exhibits a nonvanishing bias as N increases while the biases of the IRIVE and NM tend to zero for large N. For $N \geq 35$, the RMSEs of the IRIVE and NM match the RCOVAR (i.e., the root of the analytical covariance of least l_p-norm estimation) and closely achieve the RCRLB. However, for a smaller number of sensors $N < 35$, the NM fails to converge. In contrast to the NM, the IRIVE still maintains a good estimation performance for $N < 35$, producing a negligible bias and yielding a RMSE only slightly deviating from the RCOVAR and RCRLB.

D. Performance Versus Noise Impulsiveness

We now examine the performance of the algorithms for various levels of noise impulsiveness $\alpha \in \{1.1, 1.2, \ldots, 1.9\}$. For these values of α, the corresponding optimum values of p are $1.041, 1.083, 1.127, 1.174, 1.225, 1.282, 1.348, 1.430$ and 1.546, respectively [25]. Here, we set $\gamma^{1/\alpha} = 3°$ while keeping other simulation parameters identical to those of Section VII-A. Note that $p < 1$ is required for $\alpha < 1$, and thus making robust estimation more challenging [25]. The extension of this work to the case of $\alpha < 1$ will be considered in future research.

Fig. 5 shows the bias norm and RMSE performance versus α. We observe that the performance of the PLE and WIVE degrades significantly when α decreases. This can be explained by the fact that a smaller value of α corresponds to a more impulsive noise, thereby leading to a more severe performance degradation for the least-squares estimators PLE and WIVE. The level of noise impulsiveness also affects the bias problem of the IRPLE and consequently the RMSE performance of the IRPLE, i.e.. a smaller α resulting in a poorer IRPLE performance. The NM only performs well for low levels of noise impulsiveness ($\alpha \geq 1.6$), while exhibiting an instability problem for higher levels of noise impulsiveness ($\alpha < 1.6$). The IRIVE maintains a relatively small bias while its RMSE

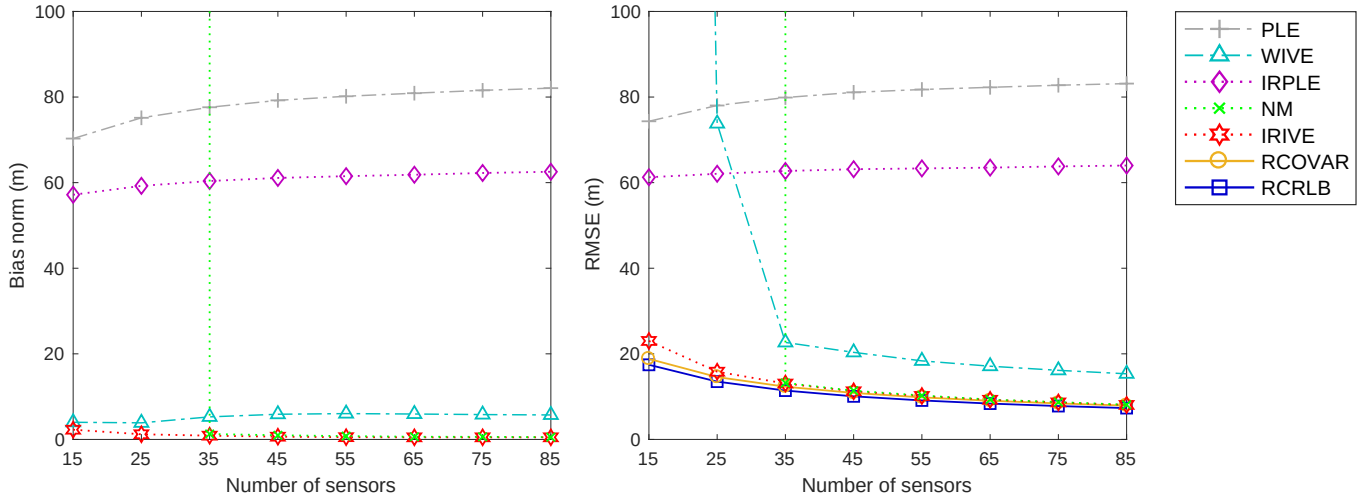


Fig. 4. Bias norm and RMSE versus number of sensors for the PLE, WIVE, IRPLE, NM and IRIVE algorithms.

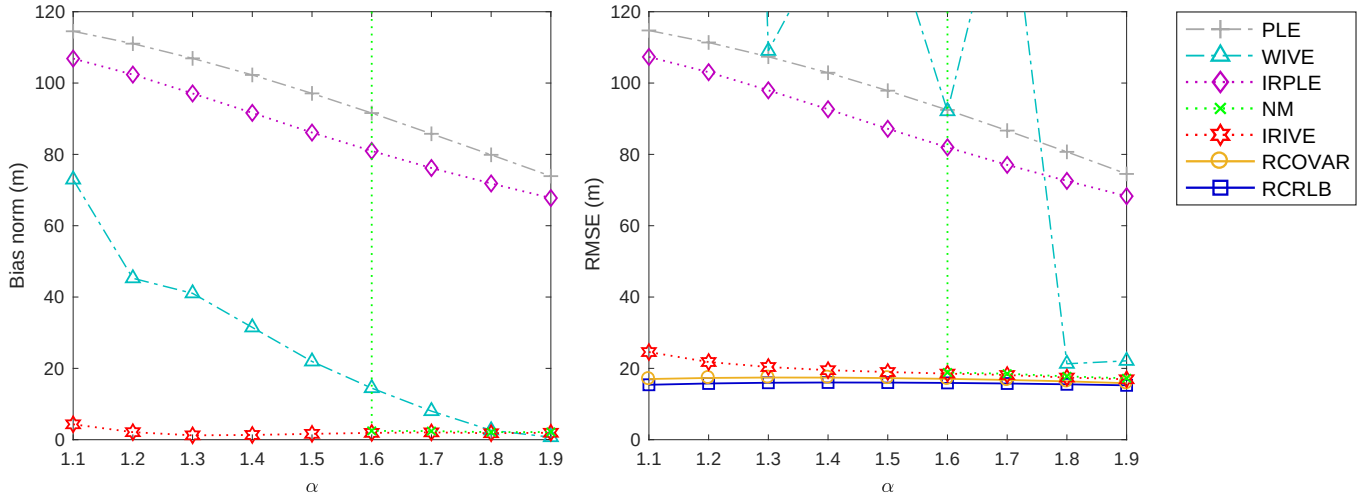


Fig. 5. Bias norm and RMSE versus noise impulsiveness for the PLE, WIVE, IRPLE, NM and IRIVE algorithms.

only slightly deviates from the RCOVAR and RCRLB as α decreases. Among the simulated estimators, the IRIVE produces the best performance and is least influenced by the noise impulsiveness level.

E. Performance Under Random Geometries

In this simulation, a total of $M = 50$ localization geometries are randomly drawn according to the uniform distribution. For each geometry realization, forty sensors are allocated randomly within 20 m from the origin, while the source is placed randomly with its distance to the origin in between 30 m and 100 m. For each geometry, the simulation is repeated for $L = 2{,}000$ Monte Carlo runs. The bias norm and RMSE are computed as $\frac{1}{M}\sum_{m=1}^{M}\left(\left\|\frac{1}{L}\sum_{l=1}^{L}\left(\hat{\boldsymbol{s}}_{\mathrm{X}}^{(m,l)}-\boldsymbol{s}^{(m)}\right)\right\|\right)$ and $\left(\frac{1}{ML}\sum_{m=1}^{M}\sum_{l=1}^{L}\left\|\hat{\boldsymbol{s}}_{\mathrm{X}}^{(m,l)}-\boldsymbol{s}^{(m)}\right\|^2\right)^{1/2}$, respectively. Here, $\boldsymbol{s}^{(m)}$ is the true source position in the m-th geometry and $\hat{\boldsymbol{s}}_{\mathrm{X}}^{(m,l)}$ is the source position estimate at the l-th Monte Carlo run in the m-th geometry. Other simulation parameters are the same as in Section VII-A. Fig. 6 compares the bias norm and RMSE of the PLE, WIVE, IRPLE, NM and IRIVE for $\gamma^{1/\alpha}$ ranging from $0.5°$ to $4.5°$. Similar observations to Section VII-A can be made here. Specifically, the results demonstrate again the unreliability of the PLE and WIVE in impulsive noise, the biasedness of the IRPLE, and the instability of the NM. On the other hand, the IRIVE enjoys the robustness of a least l_p-norm estimator, the inherent stability of the IRLS family, while at the same time overcoming the bias problem that plagues the IRPLE.

F. Performance Under Non-Identical Noise

The performance of the estimators is now evaluated under the scenario of non-identical noise. In this simulation, the measurement noises at different sensors are independent but having different noise dispersions. The noise dispersion at sensor n is modeled as $\gamma_{\theta,n} = \gamma_{\phi,n} = \gamma_{\circ}(\|\boldsymbol{d}_n\|/d_{\circ})^2$ (i.e.,

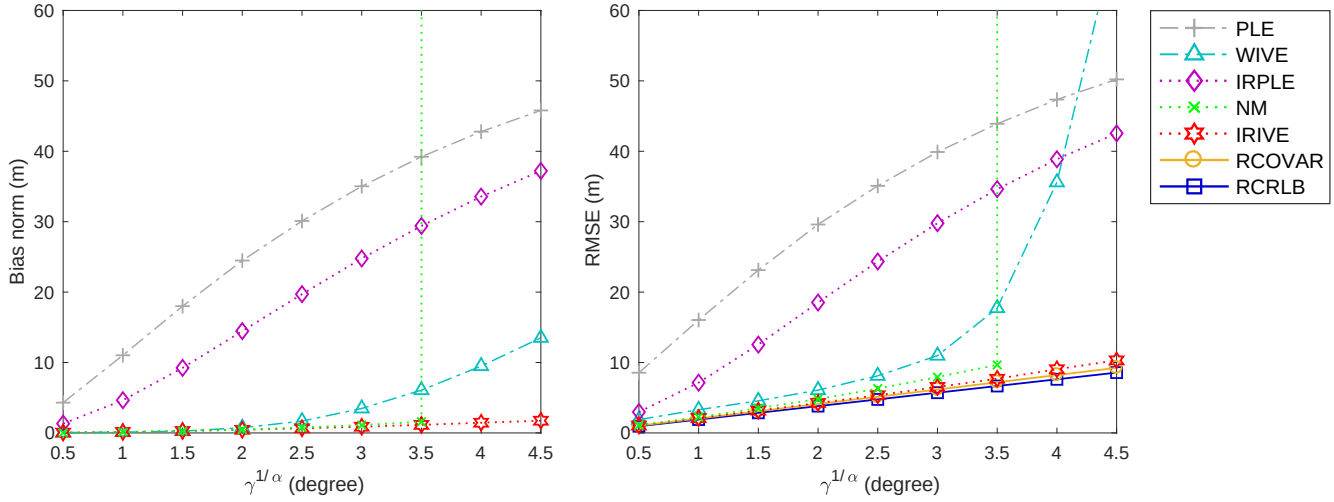


Fig. 6. Performance of the PLE, WIVE, IRPLE, NM and IRIVE algorithms under random geometries.

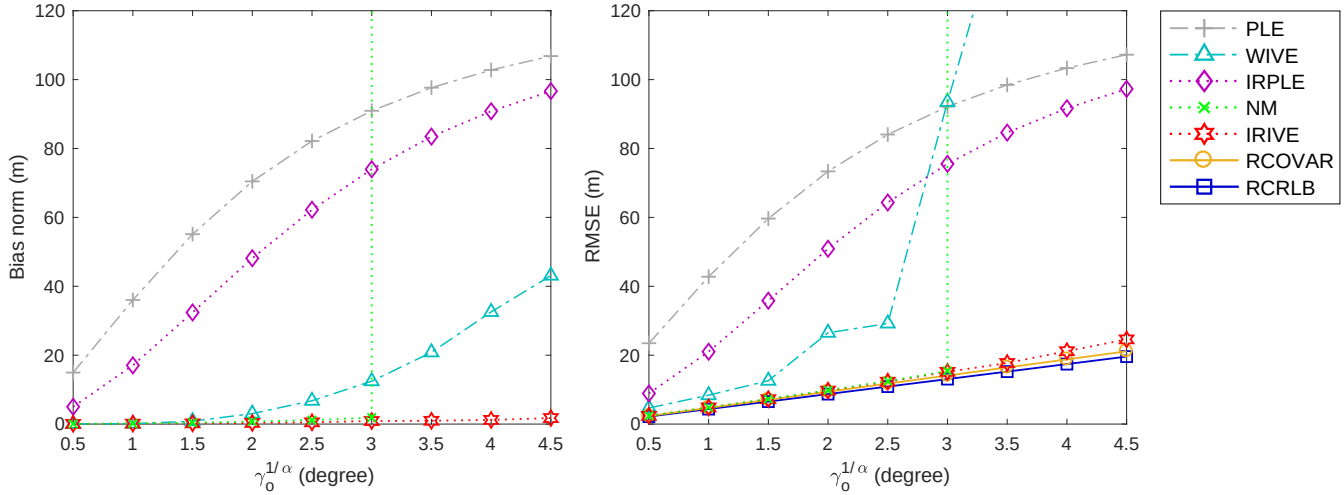


Fig. 7. Performance of the PLE, WIVE, IRPLE, NM and IRIVE algorithms under non-identical noise

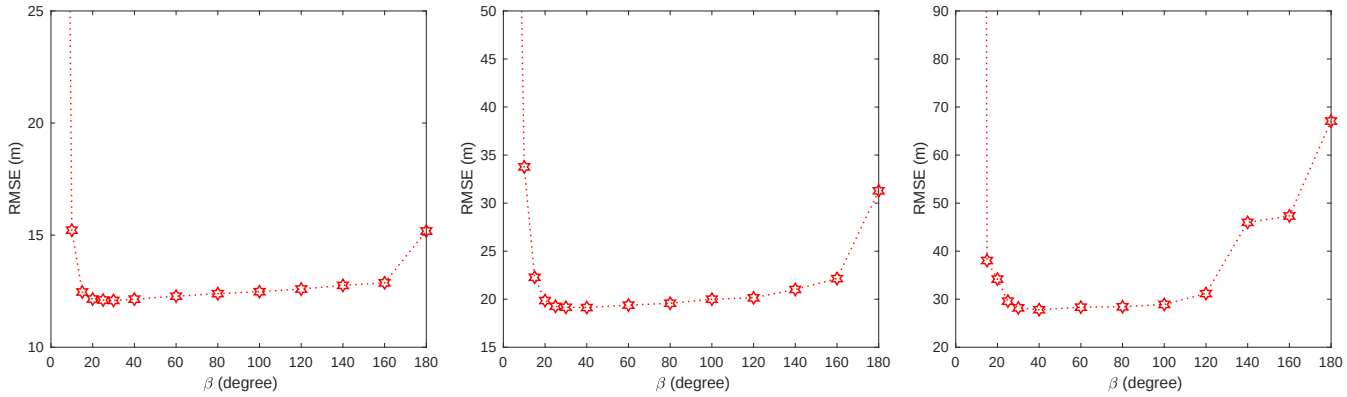
inversely proportional to the signal-to-noise ratio), where $\gamma_\circ$ is the noise dispersion value at the reference distance $d_\circ$. Other simulation parameters are kept the same as in Section VII-A. Fig. 7 shows the bias norm and RMSE performance versus various values of $\gamma_\circ^{1/\alpha}$. The results in Fig. 7 are congruent with those observed in Figs. 2 and 6, verifying the superior performance of the IRIVE.

G. Choice of SAM Threshold

Fig. 8 shows the RMSE performance of the IRIVE against various values of SAM threshold $\beta_{\theta,n} = \beta_{\phi,n} = \beta$ at noise levels $\gamma^{1/\alpha} = 2^\circ, 3^\circ$ and 4° using the same simulation scenario in Section VII-A. The RMSE are obtained via 100,000 Monte Carlo runs. We observe that the choice of β is not critical to the performance of the IRIVE as it accepts a wide range of usable values for β. The selection of β is more relaxed (i.e., a wider range of β is allowed) for smaller noise. Although the accepted range of β depends on the noise levels, i.e., $\beta \in (20^\circ, 160^\circ)$ for $\gamma^{1/\alpha} = 2^\circ$, $\beta \in (20^\circ, 120^\circ)$ for $\gamma^{1/\alpha} = 3^\circ$ and $\beta \in (25^\circ, 100^\circ)$ for $\gamma^{1/\alpha} = 4^\circ$, the optimal range of β stays roughly the same with $\beta \in (30^\circ, 60^\circ)$ for all three noise levels.

VIII. Conclusion

In this paper, we have introduced the concepts of pseudolinear estimation and IV estimation into the framework of l_p-norm minimization for robust 3D AOA source localization in the presence of impulsive α-stable noise. Two estimators were developed, namely, the IRPLE and IRIVE. Being IRLS variants, the IRPLE and IRIVE enjoy the inherent convergence property of the IRLS. The IRPLE, which is derived from direct application of the IRLS to the pseudolinear l_p-norm objective function, unfortunately produces biased source position estimates because of the correlation between the measurement matrix and pseudolinear noise vector. The IRIVE resolves this bias problem by exploiting the use of instrumental variables. A



(a) $\gamma^{1/\alpha} = 2^\circ$ (b) $\gamma^{1/\alpha} = 3^\circ$ (c) $\gamma^{1/\alpha} = 4^\circ$

Fig. 8. RMSE of the IRIVE versus SAM threshold β at $\gamma^{1/\alpha} = 2^\circ, 3^\circ$ and 4°.

theoretical analysis and numerical simulations were presented to corroborate the performance advantages of the proposed IRIVE over the IRPLE, the least-squares based PLE and WIVE, as well as the NM-based least l_p-norm solver in terms of bias, RMSE and stability.

Appendix A

The Jacobian matrix $\boldsymbol{\nabla}$ of $\boldsymbol{\psi}$ in (8) with respect to $\boldsymbol{s}$ is given by

$$\boldsymbol{\nabla} = [\boldsymbol{\nabla}_\theta^T, \boldsymbol{\nabla}_\phi^T]^T \tag{67}$$

where

$$\boldsymbol{\nabla}_\theta = [\boldsymbol{\nabla}_{\theta,1}^T, \ldots, \boldsymbol{\nabla}_{\theta,N}^T]^T \tag{68a}$$
$$\boldsymbol{\nabla}_\phi = [\boldsymbol{\nabla}_{\phi,1}^T, \ldots, \boldsymbol{\nabla}_{\phi,N}^T]^T \tag{68b}$$

and

$$\boldsymbol{\nabla}_{\theta,n} = \frac{\gamma_{\theta,n}^{-1/\alpha}}{\|\boldsymbol{d}_n\| \cos\phi_n} [-\sin\theta_n, \cos\theta_n, 0] \tag{69a}$$
$$\boldsymbol{\nabla}_{\phi,n} = \frac{\gamma_{\phi,n}^{-1/\alpha}}{\|\boldsymbol{d}_n\|} [-\sin\phi_n\cos\theta_n, -\sin\phi_n\sin\theta_n, \cos\phi_n]. \tag{69b}$$

Appendix B

The expectation of $\boldsymbol{W}^o$ is given by

$$\mathbb{E}\{\boldsymbol{W}^o\} = \mathbb{E}\{\boldsymbol{W}(\boldsymbol{s})\} = \mathrm{diag}\left(\mathbb{E}\{\boldsymbol{W}_\theta(\boldsymbol{s})\}, \mathbb{E}\{\boldsymbol{W}_\phi(\boldsymbol{s})\}\right) \tag{70}$$

where

$$\mathbb{E}\{\boldsymbol{W}_\theta(\boldsymbol{s})\} = \mathrm{diag}(\mathbb{E}\{w_{\theta,1}(\boldsymbol{s})\}, \ldots, \mathbb{E}\{w_{\theta,N}(\boldsymbol{s})\}) \tag{71a}$$
$$\mathbb{E}\{\boldsymbol{W}_\phi(\boldsymbol{s})\} = \mathrm{diag}(\mathbb{E}\{w_{\phi,1}(\boldsymbol{s})\}, \ldots, \mathbb{E}\{w_{\phi,N}(\boldsymbol{s})\}) \tag{71b}$$

and

$$\mathbb{E}\{w_{\theta,n}(\boldsymbol{s})\} = u_{\theta,n}^2 \mathbb{E}\{|\epsilon_{\theta,n}|^{p-2}\} = u_{\theta,n}^2 C_\alpha(p-2,\alpha) \tag{72a}$$
$$\mathbb{E}\{w_{\phi,n}(\boldsymbol{s})\} = u_{\phi,n}^2 \mathbb{E}\{|\epsilon_{\phi,n}|^{p-2}\} = u_{\phi,n}^2 C_\alpha(p-2,\alpha). \tag{72b}$$

From (70)–(72), we write $\mathbb{E}\{\boldsymbol{W}^o\}$ in a compact form as

$$\mathbb{E}\{\boldsymbol{W}^o\} = C_\alpha(p-2,\alpha)\, \boldsymbol{U}^T\boldsymbol{U}. \tag{73}$$

Appendix C

We have

$$\boldsymbol{W}^o\boldsymbol{\eta} = \begin{bmatrix} \boldsymbol{W}_\theta^o\boldsymbol{\eta}_\theta \\ \boldsymbol{W}_\phi^o\boldsymbol{\eta}_\phi \end{bmatrix} \tag{74}$$

where

$$\boldsymbol{W}_\theta^o\boldsymbol{\eta}_\theta = \big[u_{\theta,1}|\epsilon_{\theta,1}|^{p-1}\,\mathrm{sign}(\epsilon_{\theta,1}), \ldots, u_{\theta,N}|\epsilon_{\theta,N}|^{p-1}\,\mathrm{sign}(\epsilon_{\theta,N})\big]^T \tag{75a}$$
$$\boldsymbol{W}_\phi^o\boldsymbol{\eta}_\phi = \big[u_{\phi,1}|\epsilon_{\phi,1}|^{p-1}\,\mathrm{sign}(\epsilon_{\phi,1}), \ldots, u_{\phi,N}|\epsilon_{\phi,N}|^{p-1}\,\mathrm{sign}(\epsilon_{\phi,N})\big]^T \tag{75b}$$

By using (3) and (5), it is straightforward to show that

$$\mathbb{E}\left\{(\boldsymbol{W}^o\boldsymbol{\eta})(\boldsymbol{W}^o\boldsymbol{\eta})^T\right\} = C_\alpha(2p-2,\alpha)\, \boldsymbol{U}^T\boldsymbol{U}. \tag{76}$$

References

[1] A. G. Lingren and K. F. Gong, "Position and velocity estimation via bearing observations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 14, no. 4, pp. 564–577, Jul. 1978.

[2] K. Dogancay, "Bias compensation for the bearings-only pseudolinear target track estimator," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 59–68, Jan. 2006.

[3] Y. Wang and K. C. Ho, "An asymptotically efficient estimator in closed-form for 3-D AOA localization using a sensor network," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 6524–6535, Dec. 2015.

[4] K. Dogancay, "3D pseudolinear target motion analysis from angle measurements," *IEEE Trans. Signal Process.*, vol. 63, no. 6, pp. 1570–1580, Mar. 2015.

[5] M. Gavish and A. J. Weiss, "Performance analysis of bearing-only target location algorithms," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 817–828, Jul. 1992.

[6] C. L. Nikias and M. Shao, *Signal Processing with Alpha-stable Distributions and Applications*. New York, NY, USA: Wiley, 1995.

[7] M. Shao and C. L. Nikias, "Signal processing with fractional lower order moments: stable processes and their applications," *Proc. IEEE*, vol. 81, no. 7, pp. 986–1010, Jul. 1993.

[8] P. S. Kokoszka and M. S. Taqqu, "Fractional ARIMA with stable innovations," *Stoch. Process. Appl.*, vol. 60, no. 1, pp. 19–47, 1995.

[9] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Processes: Stochastic Models with Infinite Variance*. London, U.K.: Chapman and Hall, 1994.

[10] V. M. Zolotarev, *One-Dimensional Stable Distributions*, ser. Translations of. Mathematical Monographs. Providence, RI, USA: American Mathematical Society, 1986, vol. 65.

[11] R. Adler, R. Feldman, and M. Taqqu, Eds., *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Boston, MA: Birkhauser, 1998.

[12] P. Tsakalides and C. L. Nikias, "The robust covariation-based MUSIC (ROC-MUSIC) algorithm for bearing estimation in impulsive noise environments," *IEEE Trans. Signal Process.*, vol. 44, no. 7, pp. 1623–1633, Jul. 1996.
[13] H. Q. Liu and H. C. So, "Target tracking with line-of-sight identification in sensor networks under unknown measurement noises," *Progress Electromagn. Research*, vol. 97, pp. 373–389, 2009.
[14] E. E. Kuruoglu, "Density parameter estimation of skewed α-stable distributions," *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2192–2201, Oct. 2001.
[15] P. G. Georgiou and C. Kyriakakis, "Robust maximum likelihood source localization: the case for sub-Gaussian versus Gaussian," *IEEE Trans. Audio, Speech, Language Process.*, vol. 14, no. 4, pp. 1470–1480, Jul. 2006.
[16] M. Rupi, P. Tsakalides, E. D. Re, and C. L. Nikias, "Robust spatial filtering of coherent sources for wireless communications," *Signal Process.*, vol. 80, no. 3, pp. 381–396, 2000.
[17] P. G. Georgiou, P. Tsakalides, and C. Kyriakakis, "Alpha-stable modeling of noise and robust time-delay estimation in the presence of impulsive noise," *IEEE Trans. Multimedia*, vol. 1, no. 3, pp. 291–301, Sep. 1999.
[18] X. Jiang, W. Zeng, H. C. So, S. Rajan, and T. Kirubarajan, "Robust matched filtering in l_p-space," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6184–6199, Dec. 2015.
[19] E. E. Kuruoglu, "Nonlinear least l_p-norm filters for nonlinear autoregressive α-stable processes," *Digit. Signal Process.*, vol. 12, no. 1, pp. 119–142, 2002.
[20] G. A. Tsihrintzis and C. L. Nikias, "Fast estimation of the parameters of alpha-stable impulsive interference," *IEEE Trans. Signal Process.*, vol. 44, no. 6, pp. 1492–1503, Jun. 1996.
[21] P. Tsakalides, R. Raspanti, and C. L. Nikias, "Angle/Doppler estimation in heavy-tailed clutter backgrounds," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 35, no. 2, pp. 419–436, Apr. 1999.
[22] E. E. Kuruoglu, P. J. Rayner, and W. J. Fitzgerald, "Least l_p-norm impulsive noise cancellation with polynomial filters," *Signal Process.*, vol. 69, no. 1, pp. 1–14, 1998.
[23] W. Zeng, H. C. So, and L. Huang, "l_p-MUSIC: Robust direction-of-arrival estimator for impulsive noise environments," *IEEE Trans. Signal Process.*, vol. 61, no. 17, pp. 4296–4308, Sep. 2013.
[24] R. Gonin and A. H. Money, *Nonlinear Lp-Norm Estimation*. Boca Raton, FL, USA: CRC Press, 1989.
[25] Y. Chen, H. C. So, and E. E. Kuruoglu, "Variance analysis of unbiased least l_p-norm estimator in non-Gaussian noise," *Signal Process.*, vol. 122, pp. 190–203, 2016.
[26] R. H. Byrd and D. A. Payne, "Convergence of the iteratively reweighted least squares algorithm for robust regression," The Johns Hopkins University, Baltimore, MD, Tech. Rep. 313, 1979.
[27] M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics*. London, U.K.: Griffin, 1961, vol. 2.
[28] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 1999.
[29] N. H. Nguyen and K. Dogancay, "Improved pseudolinear Kalman filter algorithms for bearings-only target tracking," *IEEE Trans. Signal Process.*, vol. 65, no. 23, pp. 6119–6134, Dec. 2017.
[30] K. C. Ho, "Bias reduction for an explicit solution of source localization using TDOA," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2101–2114, May 2012.
[31] N. H. Nguyen and K. Dogancay, "Single-platform passive emitter localization with bearing and Doppler-shift measurements using pseudolinear estimation techniques," *Signal Process.*, vol. 125, pp. 336–348, Aug. 2016.
[32] H. C. So and L. Lin, "Linear least squares approach for accurate received signal strength based source localization," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 4035–4040, Aug. 2011.
[33] F. K. W. Chan, H. C. So, J. Zheng, and K. W. K. Lui, "Best linear unbiased estimator approach for time-of-arrival based localisation," *IET Signal Process.*, vol. 2, pp. 156–162, Jun. 2008.
[34] N. H. Nguyen and K. Dogancay, "Closed-form algebraic solutions for 3-D Doppler-only source localization," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6822–6836, Oct. 2018.
[35] G. A. Tsihrintzis and C. L. Nikias, "Fast estimation of the parameters of alpha-stable impulsive interference," *IEEE Trans. Signal Process.*, vol. 44, no. 6, pp. 1492–1503, Jun. 1996.
[36] K. Dogancay and G. Ibal, "Instrumental variable estimator for 3D bearings-only emitter localization," in *Proc. IEEE ISSNIP*, Melbourne, Australia, Dec. 2005, pp. 63–68.
[37] T. S. Ferguson, *A Course in Large Sample Theory*. London, U.K.: Chapman and Hall, 1996.
[38] K. Wong and E. Polak, "Identification of linear discrete time systems using the instrumental variable method," *IEEE Trans. Autom. Control*, vol. 12, no. 6, pp. 707–718, Dec. 1967.
[39] N. H. Nguyen and K. Dogancay, "Instrumental variable based Kalman filter algorithm for three-dimensional AOA target tracking," *IEEE Signal Process. Lett.*, vol. 25, no. 10, pp. 1605–1609, Oct. 2018.
[40] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.*, vol. 7, no. 4, pp. 308–313, 1965.
[41] B. M. Sadler, R. J. Kozick, and T. Moore, "Performance analysis for direction finding in non-Gaussian noise," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, Phoenix, AZ, USA, Mar. 1999, pp. 2857–2860.
[42] K. Dogancay and H. Hmam, "Optimal angular sensor separation for AOA localization," *Signal Process.*, vol. 88, no. 5, pp. 1248–1260, 2008.