Correspondence

An ℓ_0-norm Optimization-based Algorithm for Robust and Efficient MIMO Localization

***Abstract*—Most of the existing localization frameworks are established under the Gaussian noise assumption and thus provide unsatisfactory accuracy in the presence of outliers. This work considers the robust and efficient target localization with multiple-input multiple-output radar by adopting the idea of outlier separation and the ℓ_0-norm. Specifically, we model the outliers with an auxiliary variable and impose sparsity constraint on it. The localization task is then formulated in the form of ℓ_0-norm constrained optimization. In doing so, we integrate outlier detection and target localization into a single problem. An alternating optimization (AO) based solver is developed for the resultant optimization problem. In detail, the AO-based algorithm consists of two steps, which updates the target location and the auxiliary variable alternately. In particular, both subtasks have closed-form solutions with low computational complexity. Numerical results on both synthetic and real data verify the efficiency and accuracy of the proposed algorithm in comparison with four competing methods.**

***Index Terms*— Multiple-input multiple-output (MIMO) radar, target localization, ℓ_0-norm optimization, non-line-of-sight, outlier**

I. Introduction

Among various localization systems, the multiple-input multiple-output (MIMO) radar has demonstrated its superiority over traditional radar systems [1]–[6]. For localization task, one essential aspect that should be taken into account is the localization accuracy in noisy environment. That is, the localization algorithms should be capable of resisting noise. To this end, most of the existing methodologies assume that the noise obeys zero-mean Gaussian distribution [1]–[3]. Under this assumption, the least squares (LS) provides the optimal estimation of the

Manuscript received XXXXX 00, 0000; revised XXXXX 00, 0000; accepted XXXXX 00, 0000.

This work was supported in part by the National Natural Science Foundation of China under Grant 62306337, in part by the Youth Innovation Team Plan Project for Higher Education Institution of Shandong Province under Grant 2023KJ071, in part by the Young Innovative Talents Project of Guangdong Provincial Department of Education (Natural Science) under Grant 2023KQNCX063, in part by the Hong Kong Research Matching Grant (RMG) in the Central Pot under Project No. CP/2022/2.1, and in part by the Research and Development Fund (R&D Fund) under reference No. RD/2023/1.8. *(Corresponding author: Xiao-Peng Li)*.

Zhang-Lei Shi and Weiguo Li are with the College of Science, China University of Petroleum (East China), Qingdao, China. Wenxin Xiong and Hing Cheung So are with the Department of Electrical Engineering, City University of Hong Kong, Hong Kong, China. Xiao-Peng Li is with the State Key Laboratory of Radio Frequency Heterogeneous Integration, Shenzhen University, Shenzhen, China. Yaru Fu is with the School of Science and Technology, Hong Kong Metropolitan University, Hong Kong, China.

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

0018-9251 © 2020 IEEE

target location in the sense of maximizing the likelihood function [7]. The resultant positioning problem can be solved by adopting advanced iterative optimization algorithms [1]–[3], [7], [8].

However, the complex practical situations may incur noise of various types or their mixture, such as impulsive noise and Gaussian mixture noise [9]–[12]. In MIMO localization, electromagnetic signal transmission can be compromised by signal-blocking obstacles and reflecting surfaces, especially in complex buildings or dense urban areas, which introduces non-line-of-sight (NLOS) errors [9]–[12]. Consequently, the bistatic range (BR) measurements, i.e. total propagation distances, associated to the transmitters and receivers in the system may include outliers. Hence, mitigating the influence of outliers and improving localization accuracy are of great interest [13]–[17].

To achieve robustness against outliers, several robust estimation frameworks have been established by redesigning the optimization criterion [8]–[10], [14]–[17], aiming to reduce the impact of NLOS-induced errors. For instance, it is widely reported that the ℓ_q-norm $(0 \leq q < 2)$ with smaller q is more outlier-resistant [18]. Inspired by this, the ℓ_q-norm based robust localization models are studied in [8], [17], where $q = 1$ presents the least absolute deviation-based formulation as in [9], [14]. In addition, the average error due to outliers can be reduced by introducing a balancing parameter [15], [16]. The resultant nonconvex and nonlinear optimization problems can then be tackled by Lagrange programming neural network [9], [17], semidefinite programming (SDP) [10], majorization-minimization (MM) [8], [15], and so on [14], [16].

Apart from suppressing the influence of outliers, it would be advantageous to identify those BR's with NLOS errors [11], [19], [20]. Subsequently, the localization models can be constructed using only the BR's that are free from NLOS. In other words, NLOS identification-based localization generally consists of two stages, namely, data selection and localization. Nevertheless, identifying BR's with NLOS errors is challenging, and any missed detection or false alarm can result in a substantial accuracy loss. Additionally, such a two-stage procedure may need to be performed repeatedly in a brute force manner. Recently, there are a few works exploiting the sparsity of the outliers [21]–[23]. For localization, a sparsity-promoting regularized SDP is presented in [21]. However, this model places high demands on computational resources and requires the prior knowledge of the NLOS errors.

To conclude, though the aforementioned robust frameworks are insensitive to outliers and have achieved good localization performance, their drawbacks include computational inefficiency, reliance on the prior information of NLOS, or suppressing rather than eliminating the influence of outliers. Concerning the above, this work proposes utilizing the ℓ_0-norm optimization to achieve robust and efficient localization. In detail, we incorporate the concept of outlier separation to integrate outlier detection with localization in the form of ℓ_0-norm constrained optimization. Then, we resort to the alternating optimization (AO) [24], [25] and MM [15], [26] for developing efficient solver. Our main contributions are summarized as follows:

(i) Instead of suppressing the influence of outliers, we propose a robust localization model based on the ℓ_0-norm optimization, aiming at a one-shot framework for outlier detection and localization. Specifically, we adopt the outlier separation idea to directly model the NLOS errors in the BR measurements through introducing an auxiliary variable.

(ii) For the resultant model, an AO-based approach is developed accordingly. Specifically, AO alternately optimizes two subproblems for seeking the target location and auxiliary variable. In the subtask for target location estimation, we design a surrogate function using the MM algorithm. In doing so, both the subproblems have closed-form solutions.

(iii) Numerical results demonstrate that the proposed method outperforms four state-of-the-art (SOTA) competing algorithms in terms of both localization accuracy and computational efficiency.

The remainder of this work is arranged as follows. Section II introduces the traditional MIMO localization and several SOTA robust estimation formulations. The proposed formulation for robust MIMO target localization and the corresponding optimization algorithm are given in Section III. Section IV presents the experimental results. Finally, conclusions are drawn in Section V.

Notation: We use lower-case or upper-case letters to represent a scalars, while vectors and matrices are denoted by bold lower-case and upper-case letters, respectively. The transpose operator is signified by $(\cdot)^{\mathrm{T}}$. Other mathematical symbols are defined upon their first appearance.

II. MIMO Radar Localization

A distributed MIMO radar localization system [1], [2], [6] normally includes m transmitters and n receivers in 2- or 3-D space. Assume the positions of the transmitters and receivers are known. The aim of this system is to locate the unknown target position by utilizing the time-sum of arrival (TSOA) measurements that are the propagation distances of the signals transmitted from transmitters to receivers with reflecting by the target.

Let the positions of these transmitters, receivers, and the target in 2-D space be $\boldsymbol{t}_i = [x_i^t, y_i^t]^{\mathrm{T}}$ $(i = 1, \cdots, m)$, $\boldsymbol{r}_j = [x_j^r, y_j^r]^{\mathrm{T}}$ $(j = 1, \cdots, n)$, and $\boldsymbol{e} = [x, y]^{\mathrm{T}}$, respectively. In noise-free scenario, the TSOA-based BR measurement, aka. the propagation distance, associated with $\boldsymbol{t}_i$ and $\boldsymbol{r}_j$ is given as:

$$\hat{s}_{i,j} = \|\boldsymbol{e}-\boldsymbol{t}_i\|_2 + \|\boldsymbol{e}-\boldsymbol{r}_j\|_2,\ i = 1,\cdots,m, j = 1,\cdots,n, \quad (1)$$

where $\|\cdot\|_2$ denotes the ℓ_2-norm. Apparently, there are $m \times n$ distances in this system. However, noise exists unavoidably in practical situations. As a result, propagation

distances are contaminated and the noisy distances are expressed as

$$s_{i,j} = \hat{s}_{i,j} + \epsilon_{i,j},\ i = 1,\cdots,m, j = 1,\cdots,n, \tag{2}$$

where $\epsilon_{i,j}$ represents the noise.

Conventionally, most of the existing localization frameworks assume a zero-mean Gaussian distribution of the noise [1], [2], [6]. In such a case, optimal localization can be achieved via maximum likelihood estimation. Then, the localization problem is of the well-known least squares (LS) form, given by

$$\min_{\boldsymbol{e}} \quad \sum_{i=1}^{m}\sum_{j=1}^{n}\left(s_{i,j} - \|\boldsymbol{e}-\boldsymbol{t}_i\|_2 - \|\boldsymbol{e}-\boldsymbol{r}_j\|_2\right)^2. \tag{3}$$

In practice, the MIMO system can be corrupted by impulsive noise even outliers in practical environments. In other words, $s_{i,j}$ can be polluted by Gaussian and impulsive noise simultaneously. It is widely reported that LS technique is unable to resist the impulsive noise effectively, resulting in unsatisfactory estimate [14], [17], [18]. The reason is that the errors caused by outliers are magnified by the ℓ_2-norm and thus dominate the optimization. To achieve robustness against outliers, the ℓ_q-norm is adopted

$$\min_{\boldsymbol{e}} \quad \sum_{i=1}^{m}\sum_{j=1}^{n}\left|s_{i,j} - \|\boldsymbol{e}-\boldsymbol{t}_i\|_2 - \|\boldsymbol{e}-\boldsymbol{r}_j\|_2\right|^q, \tag{4}$$

where $0 \le q < 2$ [8], [14], [17], [23], while $q = 1$ gives the least absolute deviation based formulation [14], [23]. It is worth mentioning that ℓ_q-norm with lower order is more outlier-resistant [18], [23]. However, ℓ_q-norm ($0 < q < 1$) is nonconvex and nonsmooth, which makes the localization problem more challenging. On the other hand, those works exploring a balancing parameter [15], [16] are dedicated to dealing with the following problem:

$$\min_{\boldsymbol{e},\theta} \quad \sum_{i=1}^{m}\sum_{j=1}^{n}\left(s_{i,j} - \|\boldsymbol{e}-\boldsymbol{t}_i\|_2 - \|\boldsymbol{e}-\boldsymbol{r}_j\|_2 - \theta\right)^2, \tag{5}$$

where the objective function remains in the LS-based form. It should be noted that balancing parameter based approaches approximate bias errors in multiple transmission paths with only one estimation variable. Hence, these approaches are well-suited for scenarios, where bias errors generally exhibit even magnitudes across various transmitter-target-receiver paths. However, outlier-inducing bias errors occur at distinct scales in different paths in general, making balancing parameter-based algorithms less preferred.

III. Algorithm Development

In this section, we propose an ℓ_0-norm-based robust MIMO localization algorithm with the idea of outlier separation. We consider that impulsive noise consists of low-power dense Gaussian noise and high-power sparse impulses. To model the sparse component directly, we introduce an auxiliary variable $\boldsymbol{o}$ and then impose sparsity constraint on it. In doing so, the outlier separation-based model is:

$$\min_{\boldsymbol{e},\boldsymbol{o}} \quad \sum_{i=1}^{m}\sum_{j=1}^{n}\left(s_{i,j} - \|\boldsymbol{e}-\boldsymbol{t}_i\|_2 - \|\boldsymbol{e}-\boldsymbol{r}_j\|_2 - o_{i,j}\right)^2 \tag{6a}$$

$$\text{s.t.} \quad \|\boldsymbol{o}\|_0 \le \kappa, \tag{6b}$$

where $\|\cdot\|_0$ is the ℓ_0-norm that counts the number of nonzero elements in a vector, and $\kappa > 0$ controls the sparsity of $\boldsymbol{o}$. We term (6) as **ℓ_0-norm based Outlier Separation** (ℓ_0-OS).

Due to the highly nonconvex and nonlinear nature of positioning problem, obtaining its optimal solution is challenging [6], [7]. Besides, $\|\boldsymbol{e}-\boldsymbol{t}_i\|_2$ or $\|\boldsymbol{e}-\boldsymbol{r}_j\|_2$ can be approximately zero, when the target is near one of the transmitters or receivers [9]. Hence, the gradient-based optimizers may suffer from numerical instability or even fail to locate the target position. In order to address these difficulties, we exploit AO [24], [25] to develop an algorithm for (6). In detail, the AO-based method estimates $\boldsymbol{e}$ and $\boldsymbol{o}$ alternately until convergence, which is listed as follows:

$$\boldsymbol{e}^{k+1} = \arg\min_{\boldsymbol{e}} f(\boldsymbol{e},\boldsymbol{o}^k), \tag{7a}$$

$$\boldsymbol{o}^{k+1} = \arg\min_{\|\boldsymbol{o}\|_0 \le \kappa} f(\boldsymbol{e}^{k+1},\boldsymbol{o}), \tag{7b}$$

where

$$f(\boldsymbol{e},\boldsymbol{o}) = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(s_{i,j} - \|\boldsymbol{e}-\boldsymbol{t}_i\|_2 - \|\boldsymbol{e}-\boldsymbol{r}_j\|_2 - o_{i,j}\right)^2.$$

In practice, AO stops when the relative error of decision variables between two successive iterations is below a certain threshold or the maximum iteration number is reached.

It should be noted that (7a) is reduced to the LS problem with fixing $\boldsymbol{o}$. However, the difficulty to optimize (7a) roots from the ℓ_2-norm-based terms and the cross terms involving variable $\boldsymbol{e}$, namely, $\|\boldsymbol{e}-\boldsymbol{t}_i\|_2$, $\|\boldsymbol{e}-\boldsymbol{r}_j\|_2$, and $2\|\boldsymbol{e}-\boldsymbol{t}_i\|_2\|\boldsymbol{e}-\boldsymbol{r}_j\|_2$ [9], [15]. To address this issue, we use MM [26], [27] to devise a surrogate function for the objective function in (7a), aiming at deriving a closed-form solution of the variable $\boldsymbol{e}$. Hereby, we introduce two lemmas to construct the desired surrogate function.

Lemma 1 [27]: Given $f_1(\boldsymbol{x}) = -\|\boldsymbol{x}-\boldsymbol{c}_0\|_2$, where $\boldsymbol{c}_0$ is a constant vector. Then, $f_1(\boldsymbol{x})$ is majorized by $\tilde{f}_1(\boldsymbol{x}|\tilde{\boldsymbol{x}})$ for any $\tilde{\boldsymbol{x}}$, i.e., $f_1(\boldsymbol{x}) \le \tilde{f}_1(\boldsymbol{x}|\tilde{\boldsymbol{x}})$, where $\tilde{f}_1(\boldsymbol{x}|\tilde{\boldsymbol{x}})$ is a linear function of $\boldsymbol{x}$, given by

$$\tilde{f}_1(\boldsymbol{x}|\tilde{\boldsymbol{x}}) = -\frac{(\boldsymbol{x}-\boldsymbol{c}_0)^{\mathrm{T}}(\tilde{\boldsymbol{x}}-\boldsymbol{c}_0)}{\|\tilde{\boldsymbol{x}}-\boldsymbol{c}_0\|_2}. \tag{8}$$

Lemma 2 [27]: Given $f_2(\boldsymbol{x}) = 2\|\boldsymbol{x}-\boldsymbol{c}_1\|_2\|\boldsymbol{x}-\boldsymbol{c}_2\|_2$, where $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ are two constant vectors. Then, $f_2(\boldsymbol{x})$ is majorized by $\tilde{f}_2(\boldsymbol{x}|\tilde{\boldsymbol{x}})$ for any $\tilde{\boldsymbol{x}}$, i.e., $f_2(\boldsymbol{x}) \le \tilde{f}_2(\boldsymbol{x}|\tilde{\boldsymbol{x}})$, where $\tilde{f}_2(\boldsymbol{x}|\tilde{\boldsymbol{x}})$ is a function of $\boldsymbol{x}$ listed as follows:

$$\tilde{f}_2(\boldsymbol{x}|\tilde{\boldsymbol{x}}) = \frac{\|\tilde{\boldsymbol{x}}-\boldsymbol{c}_2\|_2}{\|\tilde{\boldsymbol{x}}-\boldsymbol{c}_1\|_2}\|\boldsymbol{x}-\boldsymbol{c}_1\|_2^2 + \frac{\|\tilde{\boldsymbol{x}}-\boldsymbol{c}_1\|_2}{\|\tilde{\boldsymbol{x}}-\boldsymbol{c}_2\|_2}\|\boldsymbol{x}-\boldsymbol{c}_2\|_2^2. \tag{9}$$

In the following, we present the update schemes for $\boldsymbol{e}^{k+1}$ and $\boldsymbol{o}^{k+1}$, respectively.

(1) update of $\boldsymbol{e}^{k+1}$

When $\boldsymbol{o}$ is fixed at the k-th iteration, the objective function of (7a) is rewritten as

$$\begin{aligned} & f(\boldsymbol{e},\boldsymbol{o}^k) \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n}\Big[(\gamma_{i,j}^k)^2 - 2\gamma_{i,j}(\|\boldsymbol{e}-\boldsymbol{t}_i\|_2 + \|\boldsymbol{e}-\boldsymbol{r}_j\|_2) \\ &\quad + \|\boldsymbol{e}-\boldsymbol{t}_i\|_2^2 + \|\boldsymbol{e}-\boldsymbol{r}_j\|_2^2 + 2\|\boldsymbol{e}-\boldsymbol{t}_i\|_2\|\boldsymbol{e}-\boldsymbol{r}_j\|_2\Big], \end{aligned} \tag{10}$$

where $\gamma_{i,j}^k = s_{i,j} - o_{i,j}^k$. Apparently, (10) still includes $\|\boldsymbol{e}-\boldsymbol{t}_i\|_2$'s and $\|\boldsymbol{e}-\boldsymbol{r}_j\|_2$'s. Besides, the cross term $2\|\boldsymbol{e}-\boldsymbol{t}_i\|_2\|\boldsymbol{x}-\boldsymbol{r}_j\|_2$ is nonconvex.

To achieve stable and efficient optimization, we seek solution from the MM framework. Specifically, we develop a surrogate function, i.e. majorizer, for $f(\boldsymbol{e},\boldsymbol{o}^k)$, such that the ℓ_2-norm based terms become smooth and linearized at $\boldsymbol{e}^k$. To this end, we first linearize $\|\boldsymbol{e}-\boldsymbol{t}_i\|_2$'s and $\|\boldsymbol{e}-\boldsymbol{r}_j\|_2$'s according to Lemma 1:

$$\begin{aligned} & -2\gamma_{i,j}(\|\boldsymbol{e}-\boldsymbol{t}_i\|_2 + \|\boldsymbol{e}-\boldsymbol{r}_j\|_2) \\ & \le -2(\boldsymbol{e}-\boldsymbol{t}_i)^{\mathrm{T}}\boldsymbol{p}_{i,j}^k - 2(\boldsymbol{e}-\boldsymbol{r}_j)^{\mathrm{T}}\boldsymbol{q}_{i,j}^k, \end{aligned} \tag{11}$$

where $\boldsymbol{p}_{i,j}^k = \frac{\gamma_{i,j}(\boldsymbol{e}^k-\boldsymbol{t}_i)}{\|\boldsymbol{e}^k-\boldsymbol{t}_i\|_2}$ and $\boldsymbol{q}_{i,j}^k = \frac{\gamma_{i,j}(\boldsymbol{e}^k-\boldsymbol{r}_j)}{\|\boldsymbol{e}^k-\boldsymbol{r}_j\|_2}$.

Secondly, applying Lemma 2 to the cross term results in

$$\begin{aligned} & 2\|\boldsymbol{e}-\boldsymbol{t}_i\|_2\|\boldsymbol{e}-\boldsymbol{r}_j\|_2 \\ & \le a_{i,j}^k\|\boldsymbol{e}-\boldsymbol{t}_i\|_2^2 + b_{i,j}^k\|\boldsymbol{e}-\boldsymbol{r}_j\|_2^2, \end{aligned} \tag{12}$$

with $a_{i,j}^k = \frac{\|\boldsymbol{e}^k-\boldsymbol{r}_j\|_2}{\|\boldsymbol{e}^k-\boldsymbol{t}_i\|_2}$ and $b_{i,j}^k = \frac{\|\boldsymbol{e}^k-\boldsymbol{t}_i\|_2}{\|\boldsymbol{e}^k-\boldsymbol{r}_j\|_2}$, leading to a convex envelope of the cross term.

Finally, combining (11) and (12) yields a surrogate function $\tilde{f}(\boldsymbol{e},\boldsymbol{o}^k|\boldsymbol{e}^k)$ for $f(\boldsymbol{e},\boldsymbol{o}^k)$ at $\boldsymbol{e}^k$, given by

$$\begin{aligned} & \tilde{f}(\boldsymbol{e},\boldsymbol{o}^k|\boldsymbol{e}^k) \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n}\Big[(\gamma_{i,j}^k)^2 - 2(\boldsymbol{e}-\boldsymbol{t}_i)^{\mathrm{T}}\boldsymbol{p}_{i,j}^k - 2(\boldsymbol{e}-\boldsymbol{r}_j)^{\mathrm{T}}\boldsymbol{q}_{i,j}^k \\ &\quad + \|\boldsymbol{e}-\boldsymbol{t}_i\|_2^2 + \|\boldsymbol{e}-\boldsymbol{r}_j\|_2^2 + a_{i,j}^k\|\boldsymbol{e}-\boldsymbol{t}_i\|_2^2 + b_{i,j}^k\|\boldsymbol{e}-\boldsymbol{r}_j\|_2^2\Big], \end{aligned} \tag{13}$$

satisfying $f(\boldsymbol{e},\boldsymbol{o}^k) \le \tilde{f}(\boldsymbol{e},\boldsymbol{o}^k|\boldsymbol{e}^k)$.

With $\tilde{f}(\boldsymbol{e},\boldsymbol{w}^k|\boldsymbol{e}^k)$, the optimization problem subject to $\boldsymbol{e}$ becomes the squared ℓ_2-norm based problem, which is a convex and smooth. Thus, setting the derivative with respect to $\boldsymbol{e}$ to zero, i.e., $\nabla_{\boldsymbol{e}}\tilde{f}(\boldsymbol{e},\boldsymbol{o}^k|\boldsymbol{e}^k) = \boldsymbol{0}$, leads to the following closed-form solution:

$$\boldsymbol{e}^{k+1} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\Big[(1+a_{i,j}^k)\boldsymbol{t}_i + (1+b_{i,j}^k)\boldsymbol{r}_j + \boldsymbol{p}_{i,j}^k + \boldsymbol{q}_{i,j}^k\Big]}{\sum_{i=1}^{m}\sum_{j=1}^{n}\big(2+a_{i,j}^k+b_{i,j}^k\big)}. \tag{14}$$

(2) update of $\boldsymbol{o}^{k+1}$

Denote $\boldsymbol{\beta} = [\beta_{1,1},,\cdots,\beta_{1,n},,\cdots,\beta_{m,1},\cdots,\beta_{m,n}]^{\mathrm{T}}$ with $\beta_{i,j} = s_{i,j} - \|\boldsymbol{e}^{k+1}-\boldsymbol{t}_i\|_2 - \|\boldsymbol{e}^{k+1}-\boldsymbol{r}_j\|_2$. Hence, the subproblem (7b) is simplified as

$$\boldsymbol{o}^{k+1} = \arg\min_{\boldsymbol{o}} \|\boldsymbol{\beta}-\boldsymbol{o}\|_2^2 \quad \text{s.t.} \quad \|\boldsymbol{o}\|_0 \le \kappa. \tag{15}$$

Note that when (15) is optimized, $\boldsymbol{o}^{k+1}$ will contain at most κ nonzero components. Thus, the objective value can be divided into two parts regarding the nonzero elements and zero elements. Inspired by this, we first define Ψ as the index set consisting of the indices of the nonzero elements in $\boldsymbol{o}$ and Ψ^C as the index set comprising the indices of zeros in $\boldsymbol{o}$. In other words, $o_{i',j'} = 0$ for $(i',j') \in \Psi^C$, while $o_{i,j} \ne 0$ for $(i,j) \in \Psi$. Then, we can divide the objective function (15) into two parts:

$$\mathcal{F}(\boldsymbol{o}) = \sum_{(i,j)\in\Psi}(\beta_{i,j}-o_{i,j})^2 + \sum_{(i',j')\in\Psi^C}(\beta_{i',j'}-o_{i',j'})^2. \tag{16}$$

Since $o_{i',j'} = 0$ for (i',j'), (16) is recast as

$$\mathcal{F}(\boldsymbol{o}) = \sum_{(i,j)\in\Psi}(\beta_{i,j}-o_{i,j})^2 + \sum_{(i',j')\in\Psi^C}\beta_{i',j'}^2. \tag{17}$$

To minimize $\mathcal{F}(\boldsymbol{o})$, $o_{i,j}$ should be equal to $\beta_{i,j}$ for all $(i,j) \in \Psi$. Otherwise, the residual $(o_{i,j}-\beta_{i,j})$ will always lead to the growth of $\mathcal{F}(\boldsymbol{o})$.

Combining the above analysis leads to

$$\mathcal{F}(\boldsymbol{o}) = \sum_{(i',j')\in\Psi^C}\beta_{i',j'}^2, \tag{18}$$

which implies that $\mathcal{F}(\boldsymbol{o})$ is proportional to the loss induced by zero elements in $\boldsymbol{o}$. Hence, to minimize $\mathcal{F}(\boldsymbol{o})$ subject to $\|\boldsymbol{o}\|_0 \le \kappa$, Ψ^C should contain the indices of the $mn-\kappa$ smallest (in absolute value) components of $\boldsymbol{\beta}$, while Ψ should contain the remaining κ indices corresponding to the κ largest (in absolute value) components of $\boldsymbol{\beta}$. It is worth mentioning that though $\|\boldsymbol{o}\|_0 \le \kappa$ is fulfilled if Ψ^C contains more than $mn-\kappa$ indices of the smallest (in absolute value) components of $\boldsymbol{\beta}$, $\mathcal{F}(\boldsymbol{o})$ becomes larger than that of Ψ^C with exactly $mn-\kappa$ indices.

In summary, the update schedule for $\boldsymbol{o}^{k+1}$ is

$$\boldsymbol{o}^{k+1} = \mathcal{H}_\kappa(\boldsymbol{\beta}), \tag{19}$$

where $\mathcal{H}_\kappa$ is an element-wise hard thresholding operator:

$$\mathcal{H}_\kappa(\beta_{i,j}) = \begin{cases} \beta_{i,j}, & \text{if } |\beta_{i,j}| \ge \mu, \\ 0, & \text{otherwise}, \end{cases} \tag{20}$$

where $|\cdot|$ is the absolute function of its argument. Herein, μ is the κ-th largest element of $|\boldsymbol{\beta}|$. If some components in $|\boldsymbol{\beta}|$ are equal to μ, one can randomly keep several of these equal elements and set others to zero such that $\|\boldsymbol{o}\|_0 = \kappa$ is satisfied. That is, $\mathcal{H}_\kappa(\boldsymbol{\beta})$ should keep only κ components of $\boldsymbol{\beta}$ when some of the elements are equal to μ. Otherwise, the sparsity constraint cannot be satisfied. It should be noted that such a random selection will not change the objective function value. In addition, if there are less than κ nonzero elements in $\boldsymbol{\beta}$, then μ is the smallest nonzero element of $|\boldsymbol{\beta}|$.

Regarding the computational complexity, the subproblem for estimating $\boldsymbol{e}$ requires the complexity of $\mathcal{O}(mn)$. In addition, the hard thresholding operation on $\boldsymbol{\beta}$ requires $\mathcal{O}(\kappa\log(mn))$. This is because the κ-sparse vector $\mathcal{H}_\kappa(\boldsymbol{\beta})$ can be efficiently computed in two steps. The first is to sort the elements of $\boldsymbol{\beta}$ in their magnitude. The second imposes sparsity on $\boldsymbol{\beta}$ by retaining the top κ of them. Thus, the total complexity at one iteration is $\mathcal{O}(mn)$ given the assumption that mn is greater than κ in outlier-resistant elliptic positioning scenario. It is worth mentioning that $\boldsymbol{e}$-update only iterates once analytically to reduce the computational complexity. Algorithm 1 summarizes the steps for the proposed ℓ_0-OS.

Algorithm 1 AO for ℓ_0-OS

Input: $\boldsymbol{s}$, κ, $\boldsymbol{t}_i$, and $\boldsymbol{r}_j$
  Initialize: $\boldsymbol{e}^0$ and $\boldsymbol{o}^0 = \boldsymbol{0}$.
  while *not converged* **do**
    Calculate $\boldsymbol{e}^{k+1}$ according to (14).
    Calculate $\boldsymbol{o}^{k+1}$ according to (19).
  end
Output: Optimal location $\boldsymbol{e}$

As for the convergence, we have $f(\boldsymbol{e}^{k+1}, \boldsymbol{o}^{k+1}) \leq f(\boldsymbol{e}^{k+1}, \boldsymbol{o}^k) \leq \tilde{f}(\boldsymbol{e}^{k+1}, \boldsymbol{o}^k | \boldsymbol{e}^k) \leq \tilde{f}(\boldsymbol{e}^k, \boldsymbol{o}^k | \boldsymbol{e}^k) = f(\boldsymbol{e}^k, \boldsymbol{o}^k)$. That is, the objective is monotonically nonincreasing. In addition, the $f(\boldsymbol{e}, \boldsymbol{o})$ is lower bounded by zero. Hence, the sequence of objective value converges to a limit point.

IV. Experiments

This section evaluates the performance of ℓ_0-OS based on both synthetic and real-world experimental data. The proposed algorithm is compared to four SOTA methods, namely, ℓ_1-norm Lagrange programming neural network (ℓ_1-LPNN) [9], message passing (MP) [14], ℓ_q-norm improved iterative reweighting (ℓ_q-IIRW) ($0 \leq q < 1$) [17], and ℓ_q-norm iteratively reweighted least squares (ℓ_q-IRLS) ($1 < q < 2$) [8].

A. Experimental Settings

In our experiments, the synthetic experimental data is obtained by developing an MIMO system with $m = n = 8$ in 2-D space. The locations of transmitters are $\boldsymbol{t}_1 = [-350, -200]^{\mathrm{T}}$m, $\boldsymbol{t}_2 = [-350, 200]^{\mathrm{T}}$m, $\boldsymbol{t}_3 = [-200, -350]^{\mathrm{T}}$m, $\boldsymbol{t}_4 = [-200, 350]^{\mathrm{T}}$m, $\boldsymbol{t}_5 = [200, -350]^{\mathrm{T}}$m, $\boldsymbol{t}_6 = [200, 350]^{\mathrm{T}}$m, $\boldsymbol{t}_7 = [350, 200]^{\mathrm{T}}$m, $\boldsymbol{t}_8 = [350, -200]^{\mathrm{T}}$m, respectively. The receivers are located at $\boldsymbol{r}_1 = [-500, 500]^{\mathrm{T}}$m, $\boldsymbol{r}_2 = [500, -500]^{\mathrm{T}}$m, $\boldsymbol{r}_3 = [550, 0]^{\mathrm{T}}$m, $\boldsymbol{r}_4 = [0, 550]^{\mathrm{T}}$m, $\boldsymbol{r}_5 = [500, 500]^{\mathrm{T}}$m, $\boldsymbol{r}_6 = [0, -600]^{\mathrm{T}}$m, $\boldsymbol{r}_7 = [-600, 0]^{\mathrm{T}}$m, $\boldsymbol{r}_8 = [0, 0]^{\mathrm{T}}$m, respectively. The true target location is $\boldsymbol{e}^* = [400, 200]^{\mathrm{T}}$m.

The localization performance is measured by root mean square error (RMSE), defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \|\boldsymbol{e}^i - \boldsymbol{e}^*\|_2^2}, \tag{21}$$

where $\boldsymbol{e}^i$ is the estimated target location in the i-th trial. In our simulations, RMSE is calculated with 1000 trials, namely, $N = 1000$. In each trial, the proposed algorithm will stop when $\frac{\|\boldsymbol{e}^k - \boldsymbol{e}^{k-1}\|_2}{\|\boldsymbol{e}^{k-1}\|_2} < 10^{-4}$. For ℓ_q-IIRW and ℓ_q-IRLS, $q = 0$ and $q = 1.5$, respectively. Note that κ can be very large in practice. Hence, we introduce a hyperparameter $p \in [0, 1]$ for evaluating the percentage of outliers instead of evaluating κ. That is, $\kappa = pmn$ for the proposed ℓ_0-OS.

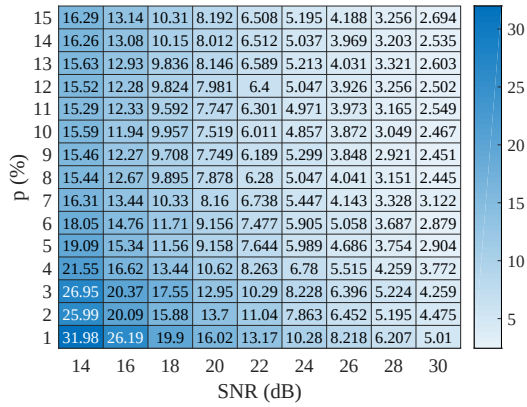

Fig. 1: Localization results of ℓ_0-OS versus p and SNR under GMM noise.

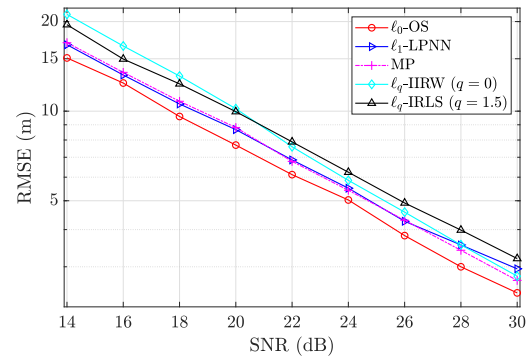

Fig. 2: Localization results of different algorithms under GMM noise when SNR ranges from 14dB to 30dB.

B. Comparison of Different Methods under GMM Noise

In this subsection, we utilize GMM to generate independent impulsive noise. The noise consists of two Gaussian variables with different variances, modeling the dense and sparse noise samples. The probability density function (PDF) of GMM is given by:

$$\varphi(y) = \frac{c_1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{y^2}{2\sigma_1^2}\right) + \frac{c_2}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{y^2}{2\sigma_2^2}\right), \tag{22}$$

where $c_1 = 1 - c_2$ with $c_2 \in [0, 1]$, and σ_1^2 and σ_2^2 are two variances. To model impulsive noise, GMM requires $\sigma_1^2 \ll \sigma_2^2$. In doing so, sparse and high power samples generated by Gaussian with large variance σ_2^2 are mixed in the dense Gaussian noise with small variance σ_1^2. In our experiments, we set $c_2 = 0.1$ and $\sigma_2^2 = 100\sigma_1^2$.

Firstly, we investigate the influence of hyperparameter p of our algorithm on performance. Fig. 1 presents the results, where the y-axis denotes p value and x-axis is the SNR of GMM noise in dB. From Fig. 1, at a certain SNR level, though there are some fluctuations, RMSE decreases first and then gradually increases along with the growth of p value. We take SNR = 20dB as an example. When $p = 1\%$, ℓ_0-OS presents RMSE of 13.17m. By increasing p from 1% to 10%, RMSE drops by 7m to around 6.011m. If we keep increasing p to 15%, RMSE

grows gradually to over 6.5m. Besides, the RMSE values are comparable, when $p \in [8\%, 12\%]$. To be specific, RMSEs are all smaller than 6.4m with minor differences.

Possible reason for such trend is that, for a small p value, say $p < 8\%$, there are still some outliers that are not modeled by $\boldsymbol{o}$. Thus, these undetected outliers will obviously lead to large RMSEs. On the contrary, when p keeps growing and finally is larger than the true percentage of outliers, a part of normal entries will be marked as outliers. As a result, the information loss makes the increase of RMSE.

On the other hand, for a fixed p, RMSE decreases as the SNR level varies from 14dB to 30dB. This trend is straightforward since a larger SNR means weaker noisy conditions. For instance, as we can observe from Fig. 1, with $p = 1\%$, the RMSE of ℓ_0-OS is improved from 31.98m to around 5m when SNR gradually approaches 30dB.

Secondly, we compare our algorithm with four SOTA methods. According to the above RMSE results and discussions related to Fig. 1, we set $p = 10\%$ for our ℓ_0-OS. Fig. 2 shows the RMSE of various algorithms with a wide range of SNR from 14dB to 30dB. Apparently, the proposed algorithm leads their counterparts with a clear margin. From Fig. 2, for most of the SNR levels, ℓ_0-OS improves the accuracy with around 1m compared to ℓ_1-LPNN and MP, while they have even more than 1m gain compared to ℓ_q-IIRW and ℓ_q-IRLS. Though the accuracy gain drops gradually when SNR value increases, the superiority of our algorithm over the competing methods is still observed clearly at all SNR levels. For example, when SNR is 20dB, the RMSE of the proposed ℓ_0-OS is 7.691m. As a comparison, the RMSE values of ℓ_1-LPNN, MP, ℓ_q-IIRW, and ℓ_q-IRLS are 8.67m, 8.812m, 10.19m, and 9.963m, respectively. At 30dB, the accuracy of ℓ_0-OS, ℓ_1-LPNN, MP, ℓ_q-IIRW, and ℓ_q-IRLS are 2.451m, 2.953m, 2.7m, 2.79m, and 3.197m, respectively, where the difference between our algorithm and other approaches are reduced. But our method still provides better estimates.

C. Comparison of Different Methods under Exponential Noise

This subsection evaluates the performance of various algorithms when one of the transmitters or receivers corresponds to NLOS propagation. In such a case, 12.5% of the BR measurements (8 out of 64) are polluted by NLOS errors. Hence, we set $p = 15\%$ for our ℓ_0-OS. In the experiments, we generate NLOS errors using exponential distribution first. Then, NLOS errors are added to BR measurements associated to one randomly selected transmitter or receiver. The standard deviation of exponential distribution varies from 10^2m to 10^5m. Besides, we add Gaussian noise with variance 100m^2 to the BR measurements as well.

As shown in Fig. 3, ℓ_0-OS has the smallest RMSEs at all noise levels, while ℓ_q-IRLS has the worst

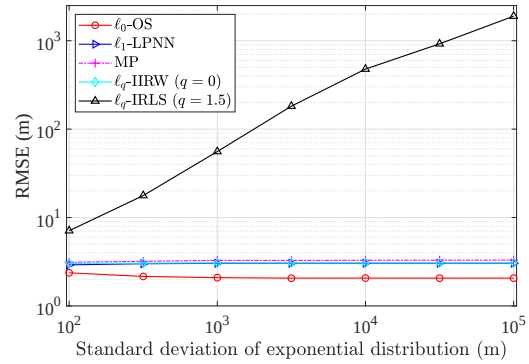

Fig. 3: Localization results of various algorithms when one transmitter or receiver corresponds to NLOS propagation.

localization performance among all algorithms. Besides, ℓ_1-LPNN, MP, and ℓ_q-IIRW are comparable, providing RMSE around 3m. In particular, the localization accuracy of our algorithm is around 2m, but lower than 2.5m at all noise levels.

D. Comparison of Different Methods under Laplacian Noise

In this subsection, we investigate the performance of various algorithms in the presence of Laplacian noise. The experimental settings are similar to those of subsection C, except that the outliers are generated by Laplace distribution. In the experiments, we fix the variance of Gaussian noise to 100m^2 and then introduce outliers randomly to one of the transmitters or receivers. Again, 12.5% of the BR measurements (8 out of 64) are contaminated by the Laplacian noise. Hence, we let $p = 15\%$ for our ℓ_0-OS.

The positioning RMSEs are plotted in Fig. 4. It is observed that the localization accuracy of ℓ_1-LPNN, MP, and ℓ_q-IIRW are around 2.9m at all noise levels. The proposed ℓ_0-OS outperforms the competing algorithms with a clear margin, whose RMSE values are around 2m.

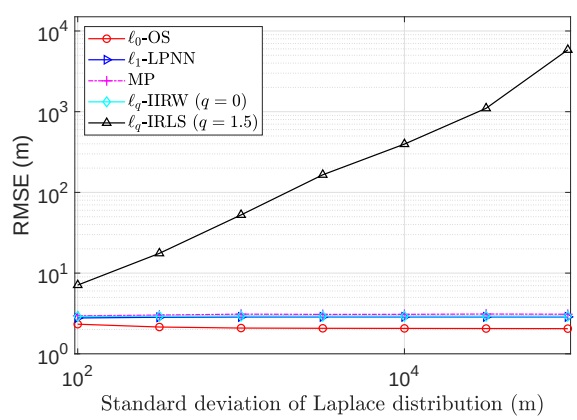

Fig. 4: RMSEs of various algorithms when Laplacian noise is randomly introduced to one transmitter or receiver.

TABLE I: Performance of different algorithms in real experiments.

Method	ℓ_0-OS	ℓ_1-LPNN	MP	ℓ_q-IIRW	ℓ_q-IRLS
RMSE (m)	0.1007	0.3701	0.1643	0.2045	0.1766

E. Comparison of Different Methods using Real Data

Apart from computer simulations with the above configuration, an acoustic localization system is implemented for evaluating the proposed algorithm. The set-up includes multiple spatially separated sound-making speakers and a Huawei Mate 20 signal-receiving device with Android 10.0 system, equipped with a Hisilicon Kirin 980 CPU and 4 GB of memory. Estimation of the speaker-smartphone distances is achieved by the generalized cross-correlation [28], where the chirp signals are modulated to the range from 19kHz to 21kHz. We conduct 60 ranging trials ($N = 60$) for each path. To add NLOS errors in specific communication channels, the common pipeline is that an experimenter takes the role of barrier standing along the transmission path. Based on the obtained speaker-smartphone distance, we construct the BR measurements for estimating the target location.

For simplicity, the smartphone is located at the frame's origin, i.e., $\boldsymbol{e}^* = [0, 0]^{\mathrm{T}}$m. The acoustic localization system consists of four transmitters and four receivers ($m = n = 4$). The locations of the transmitters are $\boldsymbol{t}_1 = [0, 1]^{\mathrm{T}}$m, $\boldsymbol{t}_2 = [2, 0]^{\mathrm{T}}$m, $\boldsymbol{t}_3 = [0, -3]^{\mathrm{T}}$m, and $\boldsymbol{t}_4 = [-4, 0]^{\mathrm{T}}$m, respectively, while the receivers are located at $\boldsymbol{r}_1 = [1.4142, 1.4142]^{\mathrm{T}}$m, $\boldsymbol{r}_2 = [-2, 2.2361]^{\mathrm{T}}$m, $\boldsymbol{r}_3 = [2.8284, -2.8284]^{\mathrm{T}}$m, and $\boldsymbol{r}_4 = [3, 4]^{\mathrm{T}}$m, respectively. To introduce NLOS errors, we obstruct the signal transmission between the smartphone and $\boldsymbol{r}_4$ by an experimenter. Thus, 4 out of 16 BR measurements are polluted and p is fixed to 25% for the ℓ_0-OS. The localization results are presented in Table I. We observe that the ℓ_0-OS achieves more accurate location estimates in the real situation compared to its four counterparts.

F. Discussion on Computational Complexity

In this subsection, we compare the five algorithms in terms of computational complexity. As discussed before, the complexity of our ℓ_0-OS is $\mathcal{O}(mn)$. In general, convergence of our algorithm can be achieved in few tens of iterations. The ℓ_1-LPNN method also costs $\mathcal{O}(mn)$ complexity per iteration. However, it usually requires thousands of iterations to converge when LPNN is discretely realized. MP requires $\mathcal{O}(mn)$ complexity per iteration and several tens of iterations to converge. For ℓ_q-IIRW and ℓ_q-IRLS, they are in the framework of iteratively reweighted least squares [29]. As a result, though their complexity increases linearly with mn, the iteratively reweighted procedure degrades their computational efficiency. Hence, our proposed algorithm and MP outperform the remaining methods in the view of computational efficiency.

TABLE II: Runtime (s) of various algorithms under different scenarios.

Method	ℓ_0-OS	ℓ_1-LPNN	MP	ℓ_q-IIRW	ℓ_q-IRLS
GMM noise	0.0078	0.0115	0.0029	0.0468	0.3418
Exponential noise	0.0094	0.0307	0.0026	0.0605	0.8957
Laplacian noise	0.0092	0.0302	0.0024	0.0595	1.7078
Real data	0.0027	0.0097	0.0018	0.0820	0.1617

To quantitatively evaluate the computational efficiency of various methods, we tabulate the average CPU time for the above experiments in Table II. It is observed that, among all algorithms, ℓ_0-OS is less efficient than MP only, which confirms our discussion.

V. Conclusion

We present an effective method called ℓ_0-OS for robust MIMO localization in the presence of impulsive noise. Specifically, it explores the outlier separation technique with an auxiliary variable. To characterize the sparse nature of outliers, we impose sparsity constraint on the auxiliary variable using ℓ_0-norm. The resultant problem is then solved by AO, where the target location and auxiliary variable are updated in closed-form. By doing so, the proposed model conducts outlier detection and target localization in a one-shot framework. Experimental results on both synthetic and real data have demonstrated that ℓ_0-OS is superior to four SOTA competing algorithms with high localization accuracy and computational efficiency.

ACKNOWLEDGMENT

The authors are sincerely grateful to Yuwei Wang and Zhi Wang from State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, China, for kindly helping us conduct real experiments. In addition, the authors would like to express their deep gratitude to the reviewers and editors for their valuable suggestions and professional work.

REFERENCES

[1] R. Amiri, F. Behnia, and M. A. M. Sadr, "Exact solution for elliptic localization in distributed MIMO radar systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1075–1086, Feb. 2018.

[2] B. K. Chalise, Y. D. Zhang, M. G. Amin, and B. Himed, "Target localization in a multi-static passive radar system through convex optimization," *Signal Process.*, vol. 102, pp. 207–215, Sep. 2014.

[3] M. Einemo and H. C. So, "Weighted least squares algorithm for target localization in distributed MIMO radar," *Signal Process.*, vol. 115, pp. 144–150, Apr. 2015.

[4] Y. Fang, S. Zhu, B. Liao, X. Li, and G. Liao, "Target localization with bistatic MIMO and FDA-MIMO dual-mode radar," *IEEE Trans. Aerosp. Electron. Syst.*, early access, Nov. 2023.

[5] K. Xiong, G. Cui, W. Yi, S. Wang, and L. Kong, "Distributed localization of target for MIMO radar with widely separated directional transmitters and omnidirectional receivers," *IEEE*

Trans. Aerosp. Electron. Syst., vol. 59, no. 3, pp. 3171–3187, 2023.

[6] H. Godrich, A. M. Haimovich, and R. S. Blum, "Target localization accuracy gain in MIMO radar-based systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2783–2803, Jun. 2010.

[7] L. Rui and K. C. Ho, "Elliptic localization: Performance study and optimum receiver placement," *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4673–4688, Sep. 2014.

[8] W. Xiong, J. He, H. C. So, J. Liang, and C.-S. Leung, "ℓ_p-norm minimization for outlier-resistant elliptic positioning in α-stable impulsive interference," *J. Franklin Inst.*, in press, Nov. 2023.

[9] Z. Shi, H. Wang, C. S. Leung, and H. C. So, "Robust MIMO radar target localization based on Lagrange programming neural network," *Signal Process.*, vol. 174, p. 107574, Sep. 2020.

[10] J. Liang, D. Wang, L. Su, B. Chen, H. Chen, and H. C. So, "Robust MIMO radar target localization via nonconvex optimization," *Signal Process.*, vol. 122, pp. 33–38, May 2016.

[11] Z. Wu, Y. Li, X. Meng, X. Lv, and Q. Guo, "A minimum joint error entropy-based localization method in mixed LOS/NLOS environments," *IEEE Internet Things J.*, vol. 10, no. 22, pp. 19 913–19 924, Nov. 2023.

[12] G. Wang, H. Chen, Y. Li, and N. Ansari, "NLOS error mitigation for TOA-based localization via convex relaxation," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4119–4131, Aug. 2014.

[13] X. P. Li, Z. Liu, Z.-L. Shi, and H. C. So, "MUSIC with capped Frobenius norm: Efficient robust direction-of-arrival estimator," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 6, pp. 8090–8103, Aug. 2023.

[14] Z. Yu, J. Li, Q. Guo, and T. Sun, "Message passing based robust target localization in distributed MIMO radars in the presence of outliers," *IEEE Signal Process. Lett.*, vol. 27, pp. 2168–2172, Dec. 2020.

[15] K. Panwar and P. Babu, "Robust multistatic target localization in the presence of NLOS errors and outliers," *IEEE Signal Process. Lett.*, vol. 29, pp. 2632–2636, Dec. 2022.

[16] W. Xiong, C. Schindelhauer, and H. C. So, "Error-reduced elliptic positioning via joint estimation of location and a balancing parameter," *IEEE Signal Process. Lett.*, vol. 29, pp. 2447–2451, Nov. 2022.

[17] X. Zhao, J. Li, and Q. Guo, "Robust target localization in distributed MIMO radar with nonconvex ℓ_p minimization and iterative reweighting," *IEEE Commun. Lett.*, vol. 27, no. 12, pp. 3230–3234, Oct. 2023.

[18] F. Wen, P. Liu, Y. Liu, R. C. Qiu, and W. Yu, "Robust sparse recovery in impulsive noise via ℓ_p-ℓ_1 optimization," *IEEE Trans. Signal Process.*, vol. 65, no. 1, pp. 105–118, Jan. 2017.

[19] W. Xiong, J. Bordoy, C. Schindelhauer, A. Gabbrielli, G. Fischer, D. J. Schott, F. Hoeflinger, S. J. Rupitsch, and H. C. So, "Data-selective least squares methods for elliptic localization with nlos mitigation," *IEEE Sens. Lett.*, vol. 5, no. 7, pp. 1–4, 2021.

[20] Z. Zheng, J. Hua, Y. Wu, H. Wen, and L. Meng, "Time of arrival and time sum of arrival based NLOS identification and localization," in *Proc. Int. Conf. Commun. Technol. (ICCT)*, Chengdu, China, Nov. 2012, pp. 1129–1133.

[21] D. Jin, F. Yin, A. M. Zoubir, and H. C. So, "Exploiting sparsity of ranging biases for NLOS mitigation," *IEEE Trans. Signal Process.*, vol. 69, pp. 3782–3795, Jun. 2021.

[22] Y.-X. Wang, C. M. Lee, L.-F. Cheong, and K.-C. Toh, "Practical matrix completion and corruption recovery using proximal alternating robust subspace minimization," *Int. J. Comput. Vision*, vol. 111, pp. 315–344, Jul. 2015.

[23] X. P. Li, Z.-L. Shi, Q. Liu, and H. C. So, "Fast robust matrix completion via entry-wise ℓ_0-norm minimization," *IEEE Trans. Cybern.*, vol. 53, no. 11, pp. 7199–7212, Nov. 2023.

[24] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, Lake Tahoe, Nevada, United States, Dec. 2013, pp. 2796–2804.

[25] J. Wang, T. Jebara, and S.-F. Chang, "Graph transduction via alternating minimization," in *Proc. Int. Conf. Mach. Learn. (ICML)*, Helsinki, Finland, Jun. 2008, pp. 1144–1151.

[26] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-Minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, Feb. 2017.

[27] R. Jyothi and P. Babu, "SOLVIT: A reference-free source localization technique using majorization minimization," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 28, pp. 2661–2673, Sep. 2020.

[28] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 24, no. 4, pp. 320–327, Aug. 1976.

[29] D. Ba, B. Babadi, P. L. Purdon, and E. N. Brown, "Convergence and stability of iteratively re-weighted least squares algorithms," *IEEE Trans. Signal Process.*, vol. 62, no. 1, pp. 183–195, Jan. 2014.

Zhang-Lei Shi
College of Science, China University of Petroleum (East China), Qingdao, China

Wenxin Xiong
Department of Electrical Engineering, City University of Hong Kong, Hong Kong, China

Xiao-Peng Li
State Key Laboratory of Radio Frequency Heterogeneous Integration, Shenzhen University, Shenzhen, China

Weiguo Li
College of Science, China University of Petroleum (East China), Qingdao, China

Yaru Fu
School of Science and Technology, Hong Kong Metropolitan University, Hong Kong, China

Hing Cheung So, Fellow, IEEE
Department of Electrical Engineering, City University of Hong Kong, Hong Kong, China