# Adaptive Robust MIMO Radar Target Localization via Capped Frobenius Norm

Jun-Ru Yang[a], Zhang-Lei Shi[a,*], Xiao-Peng Li[b], Wenxin Xiong[c], Yaru Fu[d], Xijun Liang[a]

[a] *College of Science, China University of Petroleum (East China), Qingdao, China*
[b] *College of Electronics and Information Engineering, Shenzhen University, Shenzhen, China*
[c] *Department of Electrical Engineering, City University of Hong Kong, Hong Kong SAR, China*
[d] *School of Science and Technology, Hong Kong Metropolitan University, Hong Kong SAR, China*

**Abstract**

Most of the existing algorithms for multiple-input multiple-output radar target localization assume that the bistatic range measurements are contaminated by one certain kind of noise only, such as Gaussian noise and impulsive noise. However, when the practical noise violates the original assumed distribution, their localization performance degrades severely. Therefore, adaptive and robust localization algorithms that can achieve good localization performance under both Gaussian and impulsive noise are highly desirable. In this paper, we exploit the truncated least squares loss function called capped Frobenius norm (CFN) to resist outliers. An adaptive update scheme is developed to automatically determine the upper bound of CFN using the normalized median absolute deviation. Then, the nonconvex and nonsmooth CFN-based formulation is transformed into a regularized $\ell_2$-norm optimization problem based on the half-quadratic theory. The alternating optimization (AO) algorithm is adopted as the solver, and closed-form solutions for both subproblems are derived. We also show that the sequence of objective function value generated by the devised algorithm converges. Experimental results verify the superiority of the proposed algorithm over several existing algorithms in terms of localization accuracy under impulsive noise. Furthermore, the devised algorithm can attain comparable performance to $\ell_2$-norm based methods without tweaking hyperparameters under Gaussian noise.

*Keywords:* Multiple-input multiple-output (MIMO) radar, outlier, robustness, capped Frobenius norm, half-quadratic optimization.

## 1. Introduction

The ability to effectively resist noise lies in the core for robust multiple-input multiple-output (MIMO) radar target localization [1–7]. To achieve robustness, many efforts have been dedicated to addressing noise including Gaussian noise, impulsive noise, and outliers [7–11]. Among various localization frameworks, the $\ell_2$-norm based algorithms are frequently employed to achieve the maximum likelihood estimates when the noise is Gaussian distributed. Specifically, the localization problem is formulated to minimize the sum of the

*Corresponding author

*Email addresses:* 2209060106@s.upc.edu.cn (Jun-Ru Yang), zlshi@upc.edu.cn (Zhang-Lei Shi), x.p.li@szu.edu.cn (Xiao-Peng Li), w.x.xiong@outlook.com (Wenxin Xiong), yfu@hkmu.edu.hk (Yaru Fu), liangxijunsd@upc.edu.cn (Xijun Liang)

squared residual between the real and estimated bistatic range (BR) measurements, i.e., total propagation distances from the transmitters to receivers in the system [1, 4, 6]. Such an $\ell_2$-norm based formulation is well known as least squares (LS) problem.

Despite that Gaussian distribution is commonly adopted to model noise, non-Gaussian distributed noise is unavoidable in practice. For instance, the existence of non-line-of-sight (NLOS) propagation and signal interference [7–9, 11] in real situations can introduce outliers into the BR measurements. However, the $\ell_2$-norm based algorithms are not robust to gross errors. The reason is that the loss caused by outliers is magnified by the $\ell_2$-norm and thus the optimization is dominated by the outliers [12]. As a result, the localization performance of the $\ell_2$-norm based approaches is degraded when the BR measurements are contaminated by anomalies.

To resist outliers, a mainstream idea is to suppress or reduce their influence by adopting robust loss functions [7, 9, 10, 13–17], such as $\ell_p$-norm ($0 \leq p < 2$) and its variants [9, 14, 15, 18], and the maximum correntropy criterion [7]. It is worth mentioning that the $\ell_1$-norm based formulation [9, 14], known as least absolute deviation, corresponds to the maximum likelihood estimation under Laplacian noise. One can also develop outlier detection techniques such that the target location is determined using normal BR measurements only [8, 19, 20]. Besides, it is reported that only a small proportions of data are typically corrupted by anomalies [12, 21, 22]. Inspired by this, some researchers explore the sparsity of outliers and model them by introducing auxiliary variables [23–25]. For example, the Gaussian mixture noise is considered a sparse noise embedded into a dense Gaussian noise in [25]. Then, the sparse component is modeled by a auxiliary variable, while the small dense noise is addressed according to the LS concept. To measure the sparsity of the auxiliary variable, the $\ell_0$-norm is the accurate metric that counts the number of its nonzero entries [25–27]. However, the $\ell_0$-norm is intractable because of its discontinuous and nonconvex nature [26, 27], making the $\ell_1$-norm prevail for measuring sparsity [27, 28].

As discussed above, most existing localization algorithms require that the noise complies with the assumed distribution. That is, these methods assume that the BR measurements are contaminated by one certain kind of noise only. When the practical noise violates the original assumed distribution, their localization performance degrades severely. For example, the $\ell_2$-norm based algorithms perform well under Gaussian noise, while their localization accuracy is degraded under impulsive noise. Similarly, the frameworks built on robust loss functions can attain good performance under impulsive noise, while they become inferior to the $\ell_2$-norm based algorithms in the presence of Gaussian noise. In addition, several existing algorithms also require tweaking hyperparameters, which might be time-consuming. For example, to achieve satisfactory localization accuracy, $\ell_p$-norm based approaches need to tweak the $p$ value regarding various noise intensities. The process of hyperparameter tuning further restricts the practical applicability of the existing algorithms. Hence, adaptive and robust localization frameworks that can attain high localization accuracy under both Gaussian and impulsive noise without hyperparameter tuning are of significant interest.

In this paper, we propose an adaptive robust MIMO target localization method to achieve robustness under the Gaussian noise and/or impulsive noise without tweaking hyperparameters. In detail, the MIMO target localization task is equipped with the capped Frobenius norm (CFN) [29–31] based objective function.

The CFN is a truncated $\ell_2$-norm function, where an upper bound serves as the threshold to differentiate the normal and outlier-contaminated elements. Then, the normalized median absolute deviation is employed to adaptively determine the upper bound for CFN. Although CFN is nonsmooth and nonconvex, we transform it into a tractable problem in the form of regularized $\ell_2$-norm optimization using the half-quadratic theory [32–34]. Afterwards, we incorporate the alternating optimization (AO) [35–38] with the majorization-minimization (MM) algorithm [39–41] to solve the resultant problem. In particular, we devise closed-form solutions for both subproblems of the proposed AO-based algorithm. Moreover, the convergence of objective function value generated by the suggested algorithm is presented.

The remainder of this work is arranged as follows. Section 2 provides backgrounds for the conventional $\ell_2$-norm based methods and several robust MIMO localization formulations. The proposed robust adaptive framework along with its optimization details are outlined in Section 3. Section 4 includes experimental results. Finally, conclusions are drawn in Section 5.

**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.

## 2. MIMO Radar Target Localization

Given a MIMO radar system with $M$ transmitters and $L$ receivers, the aim of this system is to locate the unknown object at position $\boldsymbol{z} = [z_1, z_2]^{\mathrm{T}}$. Let the location of $m$-th transmitter be $\boldsymbol{t}_m = [x_m^t, y_m^t]^{\mathrm{T}}$ and the location of $l$-th receiver be $\boldsymbol{r}_l = [x_l^r, y_l^r]^{\mathrm{T}}$, respectively. To locate the target, the $m$-th transmitter sends out a signal of known pattern. Then, the signal is reflected by the target and the $l$-th receiver observes the reflected signal from the target. With the orthogonality of transmitting waveforms [2, 15], the time delays $\tau_{m,l}$'s $(m = 1, \cdots, M,\ l = 1, \cdots, L)$ of signal propagation can be obtained. In the ideal noiseless scenarios, the range measurements $\tilde{d}_{m,l}$'s can be calculated by

$$\tilde{d}_{m,l} = c\tau_{m,l} = \|\boldsymbol{z} - \boldsymbol{t}_m\|_2 + \|\boldsymbol{z} - \boldsymbol{r}_l\|_2, \tag{1}$$

where $c$ represents the signal propagation speed. However, in practical positioning scenarios, noise is inevitable due to obstacles and reflection surfaces, yielding

$$d_{m,l} = \tilde{d}_{m,l} + e_{m,l} = \|\boldsymbol{z} - \boldsymbol{t}_m\|_2 + \|\boldsymbol{z} - \boldsymbol{r}_l\|_2 + e_{m,l}, \tag{2}$$

where $d_{m,l}$ is the noisy BR measurements including noise and $e_{m,l}$ denotes the noise obeying a certain distribution, such as Gaussian distribution [1, 4, 6], Laplace distribution [9, 14] , and so on [7, 15].

When the noise is zero-mean Gaussian-distributed, the target localization problem can be formulated as the following $\ell_2$-norm based unconstrained optimization problem [1, 4, 6], viz., the least squares (LS) problem:

$$\min_{\boldsymbol{z}} \sum_{m=1}^{M} \sum_{l=1}^{L} (d_{m,l} - \|\boldsymbol{z} - \boldsymbol{t}_m\|_2 \ - \|\boldsymbol{z} - \boldsymbol{r}_l\|_2)^2. \tag{3}$$

However, non-Gaussian noise such as impulsive noise or outliers is frequently encountered in real-world scenarios. In such cases, if the LS-based model (3) is adopted, the localization accuracy will be degraded severely. To achieve robustness against outliers, we can reformulate the target localization problem with a robust objective function $\psi(\cdot)$, resulting in

$$\min_{\boldsymbol{z}} \sum_{m=1}^{M} \sum_{l=1}^{L} \psi(d_{m,l} - \|\boldsymbol{z} - \boldsymbol{t}_m\|_2 \ - \|\boldsymbol{z} - \boldsymbol{r}_l\|_2), \tag{4}$$

where $\psi(\cdot)$ is less sensitive to outliers compared to the $\ell_2$-norm. For instance, selecting $\ell_1$-norm as the objective function [9, 14] presents

$$\min_{\boldsymbol{z}} \sum_{m=1}^{M} \sum_{l=1}^{L} |d_{m,l} - \|\boldsymbol{z} - \boldsymbol{t}_m\|_2 \ - \|\boldsymbol{z} - \boldsymbol{r}_l\|_2|, \tag{5}$$

which can lead to the maximum likelihood estimation for the Laplacian noise.

In addition, for general non-Gaussian noise scenarios where the noise distribution is unknown, the $\ell_p$-norm with $0 \leq p < 2$ [15] or the maximum correntropy criterion [7] can be employed for outlier rejection. Taking the $\ell_p$-norm as an example, the target localization task is reformulated as the following $\ell_p$-norm minimization problem [15]:

$$\min_{\boldsymbol{z}} \sum_{m=1}^{M} \sum_{l=1}^{L} |d_{m,l} - \|\boldsymbol{z} - \boldsymbol{t}_m\|_2 \ - \|\boldsymbol{z} - \boldsymbol{r}_l\|_2|^p, \tag{6}$$

where $0 \leq p < 2$. It is worth mentioning that the idea behind the robust function-based formulation is to suppress or reduce the influence of impulsive noise and outliers.

As previously discussed, both formulations, i.e., models (3) and (4), require that the noise follows the assumed distribution; otherwise, their performance degrades greatly. In detail, the $\ell_2$-norm-based formulation (3) is superior to formulation (4) under Gaussian noise, while its performance is degraded severely in the presence of impulsive noise or outliers. Conversely, though formulation (4) improves localization accuracy due to its robustness against outliers in non-Gaussian noise cases, it becomes inferior to formulation (3) under Gaussian noise. In addition, tuning hyperparameters to address different noise intensities is required for $\psi(\cdot)$ in general, which can be time-consuming. For instance, when $\psi(\cdot)$ is the $\ell_p$-norm $(0 \leq p < 2)$, selecting appropriate $p$ remains a challenge across various noisy scenarios.

On the other hand, one alternative framework is to achieve error reduction in the presence of anomalies by introducing a single scalar parameter $\varepsilon$, known as the balancing parameter [13, 17]. In general, the balancing parameter based algorithm aims at solving the following unconstrained optimization problem:

$$\min_{\boldsymbol{z},\varepsilon} \quad \sum_{m=1}^{M} \sum_{l=1}^{L} \left(d_{m,l} - \|\boldsymbol{z} - \boldsymbol{t}_m\|_2 - \|\boldsymbol{z} - \boldsymbol{r}_l\|_2 - \varepsilon\right)^2, \tag{7}$$

where the loss function is smooth in the form of $\ell_2$-norm. Note that (7) approximates bias errors of outliers across multiple transmission paths with a single estimation variable. Hence, balancing parameter based algorithms can achieve excellent localization results when the magnitudes of bias errors are even across various transmitter-target-receiver paths. However, outlier-inducing bias errors typically vary in scale across different transmission paths in practice, which diminishes the preference for balancing parameter based algorithms.

## 3. Algorithm Development

### 3.1. Problem Formulation

In order to attain good localization performance under both Gaussian and impulsive noise, we suggest utilizing the capped Frobenius norm [29–31], termed as CFN, given by:

$$\|\boldsymbol{y}\|_{\mathrm{CF}} = \sqrt{\sum_{i=1}^{K} \varphi_\lambda(y_i)}, \tag{8}$$

where $\varphi_\lambda(y_i) = \min(y_i^2, \lambda^2)$, and $\lambda > 0$ is the upper bound serving as the threshold to differentiate the normal and anomaly-polluted entries. It is observed that CFN performs the same as the $\ell_2$-norm for the components lower than $\lambda$ in absolute value. Otherwise, CFN assigns an equal loss value to entries larger than $\lambda$ in absolute value. In this way, CFN-based formulation is able to resist both Gaussian noise and impulsive noise once $\lambda$ is selected properly. Besides, $\lambda$ is crucial to our CFN-based formulation in terms of effectiveness and adaptivity. In the sequel, we will introduce an automatic selection strategy for $\lambda$ based on robust statistics [42].

Combining CFN with the MIMO target localization yields the following

$$\min_{\boldsymbol{z}} \sum_{m=1}^{M} \sum_{l=1}^{L} \varphi_\lambda(d_{m,l} - \|\boldsymbol{z} - \boldsymbol{t}_m\|_2 - \|\boldsymbol{z} - \boldsymbol{r}_l\|_2) \tag{9}$$

$$= \min_{\boldsymbol{z}} \sum_{m=1}^{M} \sum_{l=1}^{L} \min\left((d_{m,l} - \|\boldsymbol{z} - \boldsymbol{t}_m\|_2 - \|\boldsymbol{z} - \boldsymbol{r}_l\|_2)^2, \lambda^2\right). \tag{10}$$

Since MIMO target localization is highly nonconvex and nonlinear, introducing the nonconvex and nonsmooth CFN makes (10) more challenging and intractable. To convert (10) into a tractable problem, we leverage the half-quadratic theory [32–34].

**Theorem 1.** *[31] Let $\varphi_\lambda(y) = \min(y^2, \lambda^2)$. There is a function $\phi_\lambda(n)$, such that*

$$\min_{y} \ \varphi_\lambda(y) = \min_{y,n} \ (y - n)^2 + \phi_\lambda(n), \tag{11}$$

*where*

$$\phi_\lambda(n) = \begin{cases} -(\lambda - |n|)^2 + \lambda^2, & |n| < \lambda \\ \lambda^2, & |n| \geq \lambda \end{cases}. \tag{12}$$

*In addition, $(y - n)^2 + \phi_\lambda(n)$ is a convex function of $n$.*

Applying Theorem 1 to (10), the CFN-based robust MIMO radar target localization can be reformulated as

$$\min_{\boldsymbol{z},\boldsymbol{n}} \sum_{m=1}^{M} \sum_{l=1}^{L} (d_{m,l} - \|\boldsymbol{z} - \boldsymbol{t}_m\|_2 - \|\boldsymbol{z} - \boldsymbol{r}_l\|_2 - n_{m,l})^2 + \phi_\lambda(n_{m,l}) \tag{13}$$

where $\boldsymbol{n} = [n_{1,1}, \cdots, n_{1,L}, \cdots, n_{M,1}, \cdots, n_{M,L}]^{\mathrm{T}}$ is the auxiliary variable. Compared to (10), the objective function of (13) is a trackable $\ell_2$-norm based function with regularization. Then, the optimization is performed in the enlarged parameter space of $\boldsymbol{z}$ and $\boldsymbol{n}$.

### 3.2. Proposed Algorithm

This subsection elaborates on the proposed algorithm for (13). Denote the objective function of (13) as $\mathcal{L}(\boldsymbol{z}, \boldsymbol{n})$. Then, we adopt AO [35–38] as the solver to tackle (13), yielding the following alternatingly iterative scheme

$$\boldsymbol{z}^{k+1} = \arg \min_{\boldsymbol{z}} \mathcal{L}(\boldsymbol{z}, \boldsymbol{n}^k), \tag{14a}$$

$$\boldsymbol{n}^{k+1} = \arg \min_{\boldsymbol{n}} \mathcal{L}(\boldsymbol{z}^{k+1}, \boldsymbol{n}). \tag{14b}$$

It is seen that AO updates only one of the variables, while the other one is fixed. Since the localization problem is of high nonlinearity and nonconvexity, updating only a subset of variables at each step can facilitate mitigating optimization difficulties and reducing computational complexity. In general, the AO continues until the preset convergence condition is met. For instance, the AO stops iterating when the relative error of variables between two successive iterations is under a given threshold. It should also be noted that once the subproblems stated in (14a) and (14b) have closed-form solutions, the AO method becomes highly efficient. However, if iterative procedures are required to solve these subproblems, efficiency may significantly decrease. We term the proposed algorithm (14) as **CFN-AO**.

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

We first tackle the subtask (14a) for updating $\boldsymbol{z}$. Since $\boldsymbol{n}$ is fixed as the solution obtained from the $(k-1)$-th iteration, namely, $\boldsymbol{n}^k$, the objective of (14a) can be expanded as

$$\begin{aligned}\mathcal{L}(\boldsymbol{z}, \boldsymbol{n}^k) = &\sum_{m=1}^{M} \sum_{l=1}^{L} \left[\delta_{m,l}^2 - 2\delta_{m,l}(\|\boldsymbol{z} - \boldsymbol{t}_m\|_2 + \|\boldsymbol{z} - \boldsymbol{r}_l\|_2)\right. \\ &\left. + 2\|\boldsymbol{z} - \boldsymbol{t}_m\|_2 \|\boldsymbol{z} - \boldsymbol{r}_l\|_2 + \|\boldsymbol{z} - \boldsymbol{t}_m\|_2^2 + \|\boldsymbol{z} - \boldsymbol{r}_l\|_2^2\right],\end{aligned} \tag{15}$$

with $\delta_{m,l} = d_{m,l} - n_{m,l}^k$. Note that conventional gradient-based algorithms cannot be adopted to solve this problem due to $\|\boldsymbol{z} - \boldsymbol{t}_m\|_2$'s, $\|\boldsymbol{z} - \boldsymbol{r}_l\|_2$'s, and $2\|\boldsymbol{z} - \boldsymbol{t}_m\|_2 \|\boldsymbol{z} - \boldsymbol{r}_l\|_2$'s. Because when we take derivative with respect to $\boldsymbol{z}$, the $\ell_2$-norm based terms $\|\boldsymbol{z} - \boldsymbol{t}_m\|_2$'s and $\|\boldsymbol{z} - \boldsymbol{r}_l\|_2$'s become the denominators. Once the target is near one of the transmitters or receivers, the corresponding denominator can be approximately zero, leading to numerical instability.

To achieve stable and efficient optimization, we solve the problem in (14a) via the MM algorithm [39–41]. The MM minimizes a majorizer of the original objective function, i.e., surrogate function, rather than the original objective function in each iteration. Generally speaking, the surrogate function is designed to tightly upper bound the original objective function. Note that the challenges in solving (14a) arise from the $\ell_2$-norm based nonsmooth terms and nonconvex cross terms. Hence, we propose linearizing and convexifying these terms according to the following two lemmas:

**Lemma 1.** *[41] With a constant vector $\boldsymbol{x}_0$, the nonsmooth function $f_1(\boldsymbol{y}) = -\|\boldsymbol{y} - \boldsymbol{x}_0\|_2$ is upper bounded by $\tilde{f}_1(\boldsymbol{y}|\tilde{\boldsymbol{y}})$ for any $\tilde{\boldsymbol{y}}$, i.e., $f_1(\boldsymbol{y}) \leq \tilde{f}_1(\boldsymbol{y}|\tilde{\boldsymbol{y}})$, where $\tilde{f}_1(\boldsymbol{y}|\tilde{\boldsymbol{y}})$ is a linear function of $\boldsymbol{y}$ listed as follows:*

$$\tilde{f}_1(\boldsymbol{y}|\tilde{\boldsymbol{y}}) = -\frac{(\boldsymbol{y} - \boldsymbol{x}_0)^{\mathrm{T}}(\tilde{\boldsymbol{y}} - \boldsymbol{x}_0)}{\|\tilde{\boldsymbol{y}} - \boldsymbol{x}_0\|_2}. \tag{16}$$

**Lemma 2.** *[41] Given nonconvex function $f_2(\boldsymbol{y}) = 2\|\boldsymbol{y}-\boldsymbol{x}_1\|_2\|\boldsymbol{y}-\boldsymbol{x}_2\|_2$, where $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are constant vectors. Then, $f_2(\boldsymbol{y})$ is majorized by $\tilde{f}_2(\boldsymbol{y}|\tilde{\boldsymbol{y}})$ for any $\tilde{\boldsymbol{y}}$, i.e., $f_2(\boldsymbol{y}) \leq \tilde{f}_2(\boldsymbol{y}|\tilde{\boldsymbol{y}})$, where $\tilde{f}_2(\boldsymbol{y}|\tilde{\boldsymbol{y}})$ is a convex function of $\boldsymbol{y}$ in the following form:*

$$\tilde{f}_2(\boldsymbol{y}|\tilde{\boldsymbol{y}}) = \frac{\|\tilde{\boldsymbol{y}}-\boldsymbol{x}_2\|_2}{\|\tilde{\boldsymbol{y}}-\boldsymbol{x}_1\|_2}\|\boldsymbol{y}-\boldsymbol{x}_1\|_2^2 + \frac{\|\tilde{\boldsymbol{y}}-\boldsymbol{x}_1\|_2}{\|\tilde{\boldsymbol{y}}-\boldsymbol{x}_2\|_2}\|\boldsymbol{y}-\boldsymbol{x}_2\|_2^2. \tag{17}$$

Now we elaborate on the details for addressing (14a) according to the MM principle. Given $\boldsymbol{z}^k$ that is obtained from $(k-1)$-th iteration. Then, according to Lemma 1, the nonsmooth terms $\|\boldsymbol{z}-\boldsymbol{t}_m\|_2$'s and $\|\boldsymbol{z}-\boldsymbol{r}_l\|_2$'s can be linearized respectively as:

$$-2\delta_{m,l}\|\boldsymbol{z}-\boldsymbol{t}_m\|_2 \leq -2(\boldsymbol{z}-\boldsymbol{t}_m)^{\mathrm{T}}\boldsymbol{\gamma}_{m,l}^k, \tag{18}$$

$$-2\delta_{m,l}\|\boldsymbol{z}-\boldsymbol{r}_l\|_2 \leq -2(\boldsymbol{z}-\boldsymbol{r}_l)^{\mathrm{T}}\boldsymbol{\beta}_{m,l}^k, \tag{19}$$

with

$$\boldsymbol{\gamma}_{m,l}^k = \frac{\delta_{m,l}(\boldsymbol{z}^k-\boldsymbol{t}_m)}{\|\boldsymbol{z}^k-\boldsymbol{t}_m\|_2} \text{ and } \boldsymbol{\beta}_{m,l}^k = \frac{\delta_{m,l}(\boldsymbol{z}^k-\boldsymbol{r}_l)}{\|\boldsymbol{z}^k-\boldsymbol{r}_l\|_2}. \tag{20}$$

As for the cross terms, applying Lemma 2 leads to the following inequality:

$$2\|\boldsymbol{z}-\boldsymbol{t}_m\|_2\|\boldsymbol{z}-\boldsymbol{r}_l\|_2 \leq P_{m,l}^k\|\boldsymbol{z}-\boldsymbol{t}_m\|_2^2 + C_{m,l}^k\|\boldsymbol{z}-\boldsymbol{r}_l\|_2^2, \tag{21}$$

where

$$P_{m,l}^k = \frac{\|\boldsymbol{z}^k-\boldsymbol{r}_l\|_2}{\|\boldsymbol{z}^k-\boldsymbol{t}_m\|_2} \text{ and } C_{m,l}^k = \frac{\|\boldsymbol{z}^k-\boldsymbol{t}_m\|_2}{\|\boldsymbol{z}^k-\boldsymbol{r}_l\|_2}. \tag{22}$$

Finally, combining (15) and (18)-(22) presents us the surrogate function $\widetilde{\mathcal{L}}(\boldsymbol{z},\boldsymbol{n}^k|\boldsymbol{z}^k)$ of $\mathcal{L}(\boldsymbol{z},\boldsymbol{n}^k)$, given as:

$$\begin{aligned}\widetilde{\mathcal{L}}(\boldsymbol{z},\boldsymbol{n}^k|\boldsymbol{z}^k) =& \sum_{m=1}^{M}\sum_{l=1}^{L}\left[\delta_{m,l}^2 - 2(\boldsymbol{z}-\boldsymbol{t}_m)^{\mathrm{T}}\boldsymbol{\gamma}_{m,l}^k - 2(\boldsymbol{z}-\boldsymbol{r}_l)^{\mathrm{T}}\boldsymbol{\beta}_{m,l}^k\right.\\ &\left.+P_{m,l}^k\|\boldsymbol{z}-\boldsymbol{t}_m\|_2^2 + C_{m,l}^k\|\boldsymbol{z}-\boldsymbol{r}_l\|_2^2 + \|\boldsymbol{z}-\boldsymbol{t}_m\|_2^2 + \|\boldsymbol{z}-\boldsymbol{r}_l\|_2^2\right],\end{aligned} \tag{23}$$

such that $\mathcal{L}(\boldsymbol{z},\boldsymbol{n}^k) \leq \widetilde{\mathcal{L}}(\boldsymbol{z},\boldsymbol{n}^k|\boldsymbol{z}^k)$ is satisfied and the equality holds only when $\boldsymbol{z} = \boldsymbol{z}^k$.

It is clear that the objective function in (23) is linear and differentiable with respect to $\boldsymbol{z}$. Hence, by setting the gradient with respect to $\boldsymbol{z}$ to zero, namely, $\nabla_{\boldsymbol{z}}\widetilde{\mathcal{L}}(\boldsymbol{z},\boldsymbol{n}^k|\boldsymbol{z}^k) = \mathbf{0}$, we can achieve the closed-form solution to $\boldsymbol{z}$ as

$$\boldsymbol{z}^{k+1} = \frac{\sum\limits_{m=1}^{M}\sum\limits_{l=1}^{L}\left[(1+P_{m,l}^k)\boldsymbol{t}_m + (1+C_{m,l}^k)\boldsymbol{r}_l + \boldsymbol{\gamma}_{m,l}^k + \boldsymbol{\beta}_{m,l}^k\right]}{\sum\limits_{m=1}^{M}\sum\limits_{l=1}^{L}(2+P_{m,l}^k+C_{m,l}^k)}. \tag{24}$$

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

Once $\boldsymbol{z}^{k+1}$ is obtained by solving (14a), it is utilized to update $\boldsymbol{n}$ by resolving (14b). Specifically, the objective function of (14b) for updating $\boldsymbol{n}$ can be written as

$$\mathcal{L}(\boldsymbol{z}^{k+1},\boldsymbol{n}) = \sum_{m=1}^{M}\sum_{l=1}^{L}(s_{m,l}-n_{m,l})^2 + \phi_\lambda(n_{m,l})$$

$$= \|\boldsymbol{s} - \boldsymbol{n}\|_2^2 + \phi_\lambda(\boldsymbol{n}) \tag{25}$$

where $s_{m,l} = d_{m,l} - \|\boldsymbol{z}^{k+1} - \boldsymbol{t}_m\|_2 - \|\boldsymbol{z}^{k+1} - \boldsymbol{r}_l\|_2$ is the distance residual between $d_{m,l}$ and $\|\boldsymbol{z}^{k+1} - \boldsymbol{t}_m\|_2 + \|\boldsymbol{z}^{k+1} - \boldsymbol{r}_l\|_2$, and $\boldsymbol{s}$ is the collection of all the distance residuals, i.e., $\boldsymbol{s} = [s_{1,1}, \cdots, s_{1,L}, \cdots, s_{M,1}, \cdots, s_{M,L}]^{\mathrm{T}}$. That is, the updating of $\boldsymbol{n}$ is simplified as the regularized $\ell_2$-norm based problem with the following form:

$$\boldsymbol{n}^{k+1} = \arg\min_{\boldsymbol{n}} \|\boldsymbol{s} - \boldsymbol{n}\|_2^2 + \phi_\lambda(\boldsymbol{n}). \tag{26}$$

Note that (26) can be seen as a proximal problem. We derive a close-form solution for (26) as shown in Lemma 3.

**Lemma 3.** *For the following optimization problem*

$$n^{k+1} = \arg\min_n \psi_\lambda(n) = \arg\min_n \ (s-n)^2 + \phi_\lambda(n). \tag{27}$$

*Its optimal solution is $\mathcal{P}_\lambda(s)$, defined as*

$$n^{k+1} = \mathcal{P}_\lambda(s) = \begin{cases} 0, & |s| < \lambda, \\ s, & |s| \geq \lambda. \end{cases} \tag{28}$$

*Proof:* Plugging (12), namely, $\phi_\lambda(n)$, into (27) yields

$$\begin{aligned} n^{k+1} &= \begin{cases} \arg\min\limits_n \ (s-n)^2 - (\lambda - |n|)^2 + \lambda^2, & |n| < \lambda, \\ \arg\min\limits_n \ (s-n)^2 + \lambda^2 & |n| \geq \lambda, \end{cases} \\ &= \begin{cases} \arg\min\limits_n \ s^2 - 2n(s-\lambda), & 0 \leq n < \lambda, \\ \arg\min\limits_n \ s^2 - 2n(s+\lambda), & -\lambda < n < 0, \\ \arg\min\limits_n \ (s-n)^2 + \lambda^2, & |n| \geq \lambda, \end{cases} \\ &= \begin{cases} 0, & |s| < \lambda, \\ s, & |s| \geq \lambda. \end{cases} \end{aligned} \tag{29}$$

Specifically, when $|s| \geq \lambda$, $\psi_\lambda(n)$ is quadratic and its subgradient is $2(n - s^{k+1}) = 0$, such that $n^{k+1} = s$ is the minimizer. When $-\lambda < n^{k+1} < \lambda$, $\psi_\lambda(n)$ is linear on two intervals. And it is easy to know that $n^{k+1} = 0$ is the minimizer in the feasible region. Since $\psi_\lambda(n)$ is convex according to Theorem 1, (29) is an optimal solution to (27). The proof is complete. ■

In (26), $n_{m,l}^{k+1}$ only depends on $s_{m,l}$, that is, $\boldsymbol{n}$ is separable. Therefore, according to Lemma 3, an optimal solution to (26) is

$$\boldsymbol{n}^{k+1} = \mathcal{P}_\lambda(\boldsymbol{s}), \tag{30}$$

where $\boldsymbol{n}^{k+1}$ can be achieved via performing hard thresholding in an entry-wise manner.

It is clear that $\boldsymbol{n}^{k+1}$ is influenced by the hyperparameter $\lambda$. We suggest updating $\lambda$ adaptively along iterations. In doing so, the time-consuming procedure of hyperparameter search is avoided. Besides, the adaptive updating of $\lambda$ could also improve the adaptivity of our algorithm in various noisy scenarios. To

**Algorithm 1** CFN-AO

**Input:** $\zeta$, $\boldsymbol{t}_m$, and $\boldsymbol{r}_l$

**Initialize:** $\boldsymbol{z}^0$, $\boldsymbol{n}^0$, $\lambda^0$, and $k = 0$

**while** *not converged* **do**

Calculate $\boldsymbol{\gamma}_{m,l}^k$, $\boldsymbol{\beta}_{m,l}^k$ according to (20)

Calculate $P_{m,l}^k$, $C_{m,l}^k$ according to (22)

Update $\boldsymbol{z}^{k+1}$ according to (24)

Calculate $\lambda'$ according to (32)

Calculate $\lambda^{k+1}$ according to (31)

Update $\boldsymbol{n}^{k+1}$ according to (30)

$k = k + 1$

**end**

**Output:** Optimal location $\hat{\boldsymbol{z}}$

guarantee convergence, $\lambda$ is required to be nonincreasing along iterations. Hence, we exploit the following adaptive updating scheme for $\lambda$:

$$\lambda^{k+1} = \min(\lambda', \lambda^k), \tag{31}$$

where $\lambda'$ is determined by a robust measure for standard deviation, namely, the normalized median absolute deviation method [42]:

$$\lambda' = \zeta \times 1.4826 \times \text{Med}(|\boldsymbol{s} - \text{Med}(\boldsymbol{s})|). \tag{32}$$

Here, $\zeta > 0$ controls the confidence interval range, and $\text{Med}(\cdot)$ is the sample median operator. Since $\boldsymbol{s}$ is the collection of all the distance residual at $k$-th iteration, if the mean of $\boldsymbol{s}$ is assumed 0, $-\lambda < s_{m,l} < \lambda$ is considered as a confidence interval to identify anomalies. In the next section, the impact of $\zeta$ on the localization performance will be studied.

Moreover, since our CFN-AO optimizes $\boldsymbol{z}$ and $\boldsymbol{n}$ in an alternating manner, we suggest computing $\lambda^{k+1}$ prior to updating $\boldsymbol{n}^{k+1}$ for better localization accuracy. That is, once $\boldsymbol{z}^{k+1}$ is obtained, we use $\boldsymbol{z}^{k+1}$ to compute $\boldsymbol{s}$ and then determine $\lambda^{k+1}$. Finally, the update of $\boldsymbol{n}^{k+1}$ is achieved using $\lambda^{k+1}$. We summarize the proposed approach in Algorithm 1.

### *3.3. Convergence Analysis and Computational Complexity*

For the convergence, we have that the sequence of objective value generated by the devised CFN-AO converges to a limit point.

Specifically, we have $\mathcal{L}(\boldsymbol{z}^{k+1}, \boldsymbol{n}^{k+1}) \leq \mathcal{L}(\boldsymbol{z}^{k+1}, \boldsymbol{n}^k)$ due to the optimality of $\boldsymbol{n}^{k+1}$ to (14b). Since $\mathcal{L}(\boldsymbol{z}, \boldsymbol{n}^k)$ is majorized by $\widetilde{\mathcal{L}}(\boldsymbol{z}, \boldsymbol{n}^k|\boldsymbol{z}^k)$, the inequality $\mathcal{L}(\boldsymbol{z}^{k+1}, \boldsymbol{n}^k) \leq \widetilde{\mathcal{L}}(\boldsymbol{z}^{k+1}, \boldsymbol{n}^k|\boldsymbol{z}^k)$ and the equality $\widetilde{\mathcal{L}}(\boldsymbol{z}^k, \boldsymbol{n}^k|\boldsymbol{z}^k) = \mathcal{L}(\boldsymbol{z}^k, \boldsymbol{n}^k)$ are valid according to MM framework. Finally, $\boldsymbol{z}^{k+1}$ minimizes the surrogate function $\widetilde{\mathcal{L}}(\boldsymbol{z}, \boldsymbol{n}^k|\boldsymbol{z}^k)$, ensuring the validity of the inequality $\widetilde{\mathcal{L}}(\boldsymbol{z}^{k+1}, \boldsymbol{n}^k|\boldsymbol{z}^k) \leq \widetilde{\mathcal{L}}(\boldsymbol{z}^k, \boldsymbol{n}^k|\boldsymbol{z}^k)$. Based on the above facts, we achieve that $\mathcal{L}(\boldsymbol{z}^{k+1}, \boldsymbol{n}^{k+1}) \leq \mathcal{L}(\boldsymbol{z}^{k+1}, \boldsymbol{n}^{k}) \leq \widetilde{\mathcal{L}}(\boldsymbol{z}^{k+1}, \boldsymbol{n}^{k}|\boldsymbol{z}^{k}) \leq \widetilde{\mathcal{L}}(\boldsymbol{z}^{k}, \boldsymbol{n}^{k}|\boldsymbol{z}^{k}) = \mathcal{L}(\boldsymbol{z}^{k}, \boldsymbol{n}^{k})$. Therefore, the objective value generated by CFN-AO is monotonically nonincreasing.

Besides, as given in (13), $\mathcal{L}(\boldsymbol{z}, \boldsymbol{n})$ is a linear combination of the $\ell_2$-norm based terms $(d_{m,l} - \|\boldsymbol{z} - \boldsymbol{t}_m\|_2 - \|\boldsymbol{z} - \boldsymbol{r}_l\|_2 - n_{m,l})^2$ and the regularization terms $\phi_\lambda(n_{m,l})$. Therefore, $\mathcal{L}(\boldsymbol{z}, \boldsymbol{n})$ is lower bounded by zero. As a result, combining the monotonically nonincreasing and boundedness property of objective value leads to the convergence conclusion.

At each iteration, the complexity for updating $\boldsymbol{z}^{k+1}$ is $\mathcal{O}(ML)$. The computational complexity for updating $\lambda^{k+1}$ is dominated by the sample median operator, with the complexity of $ML\log(ML)$. The update of $\boldsymbol{n}^{k+1}$ is achieved by conducting element-wise comparison between $|s_{m,l}|$ and $\lambda^{k+1}$, which requires a complexity of $\mathcal{O}(ML)$. In summary, the total complexity of CFN-AO is $\mathcal{O}\left(ML\text{log}(ML)\right)$ per iteration.

## 4. Numerical Results

This section evaluates the localization performance of the proposed CFN-AO. The competing algorithms are $\ell_1$-norm Lagrange programming neural network ($\ell_1$-LPNN) [9], iterative message passing (IMP) [14], $\ell_p$-norm improved iterative reweighting ($\ell_p$-IIRW with $p = 1.5$) [15], balancing parameter based difference-of-convex programming (BP-DCP) [13], and balancing parameter based Lagrange programming neural network (BP-LPNN) [16].

### *4.1. Experimental Settings*

We consider a distributed MIMO radar system with 8 transmitters and 8 receivers, viz., $M = L = 8$. The locations of the transmitters and receivers are $\boldsymbol{t}_1 = [-650, -300]^{\mathrm{T}}$, $\boldsymbol{t}_2 = [-650, 300]^{\mathrm{T}}$, $\boldsymbol{t}_3 = [-150, -450]^{\mathrm{T}}$, $\boldsymbol{t}_4 = [-150, 450]^{\mathrm{T}}$, $\boldsymbol{t}_5 = [400, -150]^{\mathrm{T}}$, $\boldsymbol{t}_6 = [400, 150]^{\mathrm{T}}$, $\boldsymbol{t}_7 = [250, 550]^{\mathrm{T}}$, $\boldsymbol{t}_8 = [250, -550]^{\mathrm{T}}$, $\boldsymbol{r}_1 = [0, 0]^{\mathrm{T}}$, $\boldsymbol{r}_2 = [-300, 300]^{\mathrm{T}}$, $\boldsymbol{r}_3 = [300, -300]^{\mathrm{T}}$, $\boldsymbol{r}_4 = [250, 0]^{\mathrm{T}}$, $\boldsymbol{r}_5 = [0, 250]^{\mathrm{T}}$, $\boldsymbol{r}_6 = [550, 550]^{\mathrm{T}}$, $\boldsymbol{r}_7 = [0, -600]^{\mathrm{T}}$, $\boldsymbol{r}_8 = [-600, 0]^{\mathrm{T}}$, respectively. The target is placed at $\boldsymbol{z}^* = [300, 200]^{\mathrm{T}}$.

Both Gaussian and non-Gaussian noisy scenarios are considered, where non-Gaussian noise is modeled by Gaussian mixture model (GMM), Exponential distribution, and Laplace distribution, respectively. The GMM represents a distribution as a combination of multiple Gaussian distributions with different mean and variance. To model impulsive noise, the GMM is generally composed of two Gaussian distributions, of which the probability density function is

$$p(x) = c_1\mathcal{N}(x|\mu_1, \sigma_1^2) + c_2\mathcal{N}(x|\mu_2, \sigma_2^2)$$

where $\mathcal{N}(x|\mu_i, \sigma_i^2)$ represents the $i$-th ($i = 1, 2$) Gaussian distribution with mean $\mu_i$ and variance $\sigma_i^2$, $c_1$ and $c_2$ are the weights of each Gaussian component satisfying $c_1 + c_2 = 1$, and $\sigma_1^2 \gg \sigma_2^2$. In our experiments, we choose $c_1 = 0.1$, $c_2 = 0.9$, $\mu_1 = \mu_2 = 0$, and $\sigma_1 = 10\sigma_2$. By doing so, the noise generated by GMM can be considered a sparse impulsive noise embedded into a dense noise.

As for Laplace and Exponential distribution, the probability density function are determined by the variance/standard deviation only. Hence, we will clarify the parameters used to generate Laplacian and Exponential noise in the following subsections, respectively.

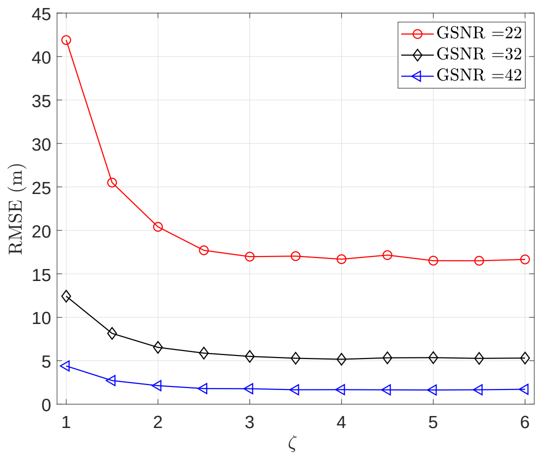


Figure 1: Localization results of CFN-AO versus $\zeta$ under Gaussian noise.

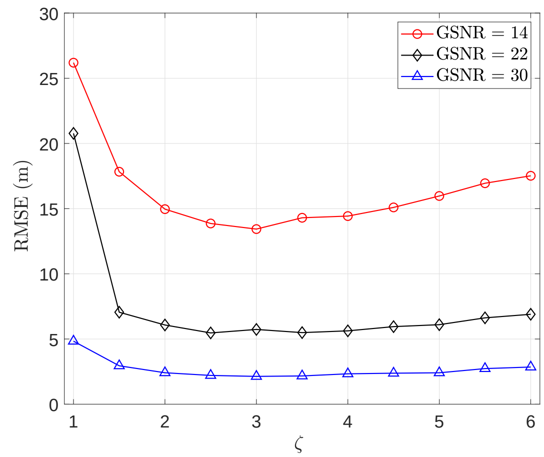


Figure 2: Localization results of CFN-AO versus $\zeta$ under GMM noise.

In the experiments, the noise is generated and added to the BR measurements $d_{m,l}$'s ($m = 1, \cdots, M, l = 1, \cdots, L$) according to the following procedure. We randomly sample points from a specific distribution such as Gaussian and Laplace distribution as the noise first, and then the noise is added to BR measurements directly. To measure the level of the noise in comparison to that of the signal, the generalized signal-to-noise ratio (GSNR) is adopted [16, 43], defined as

$$\mathrm{GSNR} = 10\log_{10}\left(\frac{\sum_{m=1}^{M}\sum_{l=1}^{L}\left(\|\boldsymbol{z}-\boldsymbol{t}_m\|_2+\|\boldsymbol{z}-\boldsymbol{r}_l\|_2\right)^2}{ML\sigma^2}\right)$$

where $\sigma^2$ is the variance of the noise.

The root mean square error (RMSE) is employed as the metric for performance comparison, given by:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\|\hat{\boldsymbol{z}}^{(i)} - \boldsymbol{z}^*\|_2^2},$$

where $N$ is the number of Monte Carlo runs, and $\hat{\boldsymbol{z}}^{(i)}$ is the estimated target position in the $i$-th Monte Carlo run. In our experiments, we set $N$ to 1000. In each trial, our CFN-AO is terminated when the relative error of the estimated target positions between two successive iterations is smaller than or equal to $\epsilon = 10^{-4}$, namely,

$$\frac{\|\boldsymbol{z}^k - \boldsymbol{z}^{k-1}\|_2}{\|\boldsymbol{z}^{k-1}\|_2} \leq \epsilon.$$

### *4.2. Impact of $\zeta$ on Localization Performance*

This subsection investigates the impact of $\zeta$ in (32) on the localization accuracy. We take Gaussian noise and GMM noise as examples, with different noise intensities and $\zeta$ varies from 1 to 6 with a step size of 0.5. The results are given in Figs. 1 and 2. It is observed that RMSE, under Gaussian noise, decreases with increasing $\zeta$ and then remain stable when $\zeta \geq 3$. Under GMM noise, RMSE shows a decrease trend first and subsequently grows with enlarging $\zeta$. As mentioned before, $\zeta$ controls the confidence interval. Thus, a smaller $\zeta$ results in a narrower confidence interval, and more entries are considered as outlier-contaminated

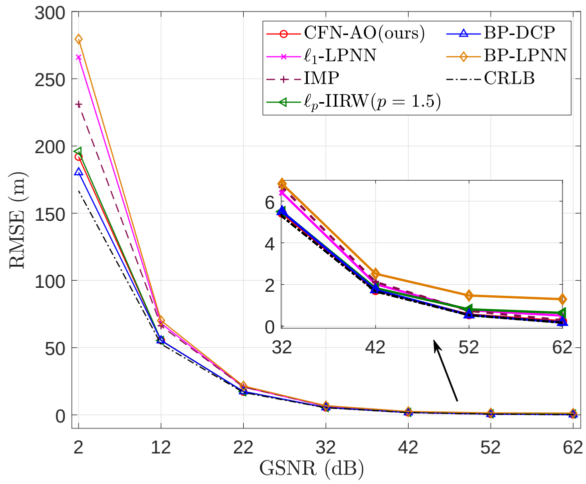


Figure 3: Localization results comparison of various algorithms under Gaussian noise.

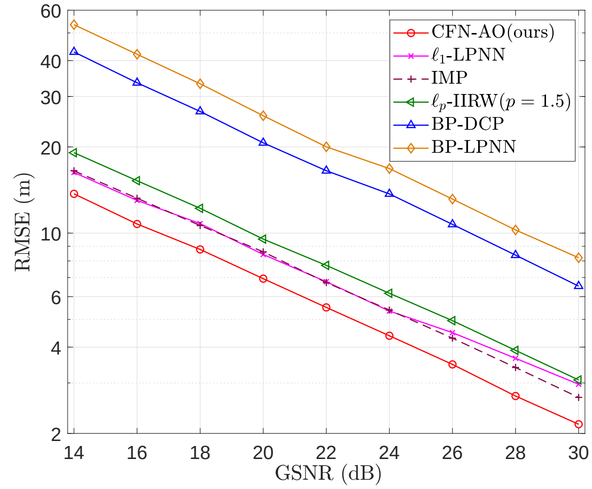


Figure 4: Localization results comparison of various algorithms under GMM noise.

data. On the contrary, a larger $\zeta$ leads to a wider confidence interval and thus more entries are classified as normal data. For Gaussian noise, since there are no anomaly-contaminated entries, a larger $\zeta$ can lead to better localization performance. For impulsive noise, a small $\zeta$ causes some normal data to be classified as anomaly-contaminated entries, while a very large $\zeta$ leads to misclassification of outlier-polluted entries as normal data. Similar behaviors can be observed under Exponential and Laplacian noise scenarios as well.

To conclude, we set $\zeta = 3$ throughout the following experiments. Besides, the noise is unknown in general and can be any kind of Gaussian or non-Gaussian noise in practice. According to our simulation results under different noisy scenarios, $\zeta = 3$ is sufficient to achieve satisfactory localization accuracy. That is, we can select $\zeta$ as 3 in practice as well.

### *4.3. Performance Comparison*

This subsection assesses the performance of our proposed algorithm under various noise situations.

*1) Performance Comparison under Gaussian Noise:* We first compare the proposed CFN-AO with the competing algorithms as well as the Cramér-Rao Lower Bound (CRLB) under Gaussian noise. The standard deviation $\sigma$ varies from $10^0$ to $10^3$, namely, $\sigma \in \{10^0, 10^{0.5}, 10^1, 10^{1.5}, 10^2, 10^{2.5}, 10^3\}$ m, corresponding to GSNR $\in \{62, 52, 42, 32, 22, 12, 2\}$ dB. The RMSE results versus various GSNRs are shown in Fig. 3. It is observed that BP-DCP, CFN-AO, and $\ell_p$-IIRW are comparable at all noise levels, which is prior to other methods. The two $\ell_1$-norm based algorithms, namely, $\ell_1$-LPNN and IMP, have comparable RMSE values when GSNR $\geq 12$, while IMP is better than $\ell_1$-LPNN when GSNR $= 2$. In addition, our CFN-AO has close RMSE to CRLB at all noise intensities, except that it is only slightly inferior to BP-DCP at GSNR $= 2$.

*2) Performance Comparison under Impulsive Noise:* Apart from Gaussian noise, we compare the performance of different algorithms under impulsive noise, including GMM noise, Exponential noise, and Laplacian noise. It is worth mentioning that Exponential noise is commonly exploited to generate NLOS errors in the presence of NLOS propagation.

For GMM, the GSNR ranges from 14dB to 30dB with an increment of 2dB, while the standard deviation

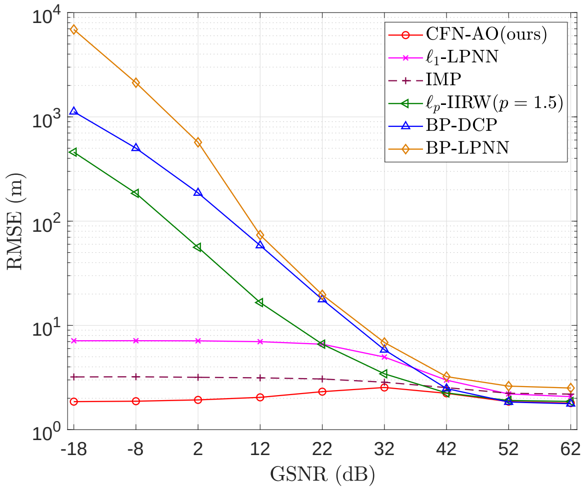


Figure 5: Localization results comparison of various algorithms under Exponential noise.

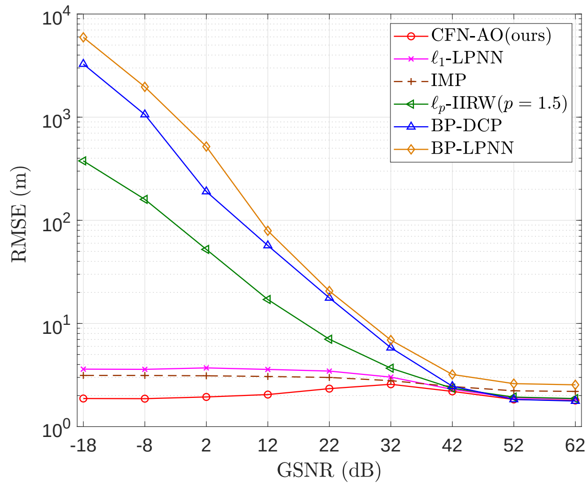


Figure 6: Localization results comparison of various algorithms under Laplacian noise.

are $\sigma \in \{10^0, 10^{0.5}, 10^1, 10^{1.5}, 10^2, 10^{2.5}, 10^3, 10^{3.5}, 10^4\}$ m for Exponential and Laplacian noise. That is, the GSNR's are $\{62, 52, 42, 32, 22, 12, 2, -8, -18\}$ dB for Exponential and Laplacian noise. Under Exponential or Laplacian noise, one of the transmitters or receivers is assumed to be influenced by outliers in our experiments. Specifically, we generate outliers using Exponential or Laplace distribution first. Then, the outliers are added to BR measurements associated to one randomly selected transmitter or receiver. Besides, Gaussian noise with standard deviation 10 m is added to the BR measurements as well. The RMSE results under GMM noise, Exponential noise, and Laplacian noise are shown in Figs. 4, 5, and 6, respectively.

From Fig. 4, namely, under the GMM noise, we observe that CFN-AO leads its competitors with a clear margin at all noise levels. That is, our CFN-AO gives out the lowest RMSE values at all noise intensities. The IMP and $\ell_1$-LPNN have comparable localization accuracies, while the $\ell_p$-IIRW is inferior to the two $\ell_1$-norm based algorithms. The BP-DCP is better than BP-LPNN, whereas the BP-LPNN presents the worst RMSE values at all GSNR levels. In addition, there is a huge gap between the BP-based algorithms, viz., BP-DCP and BP-LPNN, and other comparison algorithms.

From Figs. 5 and 6, it is seen that each algorithm exhibits similar behavior in the presence of Exponential and Laplacian noise. Thus, we take the Exponential noise as an example to analyze. When GSNR $> 42$, all the algorithms perform very similarly with comparable RMSE around 2 m. The reason might be that the standard deviation of Exponential noise is even smaller than that of Gaussian noise. As a result, the exponential noise can be considered as Gaussian. Thus, the behavior of all algorithms is similar to that under the Gaussian noise scenario.

However, the RMSE values of BP-DCP, BP-LPNN, and $\ell_p$-IIRW all are over 10 m when GSNR $\leq 42$, increasing rapidly along with the decrease of GSNR. Though the RMSEs of IMP and $\ell_1$-LPNN also increase as GSNR decreases, their RMSE values remain smaller than 10 m across all noise intensities. For our CFN-AO, its RMSE shows an increasing trend and is superior to its counterparts when GSNR $\leq 42$. A possible reason for such a trend is that, with the decrease of GSNR, i.e., the growth of $\sigma$, the magnitudes of outliers are enlarged greatly, making the outliers easy to be classified.

Table 1: Runtime comparison in different noise.

| Algorithms | Runtime (s) in various noise scenarios | | | |
|---|---|---|---|---|
| | Gaussian | GMM | Exponential | Laplacian |
| CFN-AO | 0.0071 | 0.0072 | 0.0099 | 0.0115 |
| $\ell_1$-LPNN | 0.0170 | 0.0099 | 0.0180 | 0.0240 |
| IMP | 0.0022 | 0.0016 | 0.0021 | 0.0022 |
| $\ell_p$-IIRW$(p = 1.5)$ | 0.3502 | 0.3213 | 0.4811 | 0.5774 |
| BP-DCP | 0.4501 | 0.3651 | 0.6668 | 0.7593 |
| BP-LPNN | 0.0110 | 0.0092 | 0.0404 | 0.0325 |

*3) Runtime Comparison:* To quantitatively assess the computational efficiency of various algorithms, we tabulate the average running time for all noisy scenarios in Table 1. From the table, it is observed that the efficiency of our CFN-AO is only inferior to that of IMP among all competing methods.

## 5. Conclusion

In this article, we develop a framework for robust adaptive MIMO target localization based the capped Frobenius norm. The normalized median absolute deviation strategy is exploited to adaptively determine the upper bound for the capped Frobenius norm. Based on the half-quadratic theory, we convert the minimization of the nonconvex and nonsmooth capped Frobenius norm into the tractable form of regularized least squares. Then, the AO method is adopted to address the resultant problem, yielding an efficient algorithm termed as CFN-AO. In particular, both subtasks of the suggested CFN-AO have closed-form solutions. We show that the objective value of CFN-AO converges to a limit point. Experimental results demonstrate that the CFN-AO achieves higher localization accuracy in comparison with five popular algorithms under impulsive noise. Besides, its performance is comparable to the $\ell_2$-norm based method without tuning parameter in the presence Gaussian noise.

**Acknowledgment**

The work described in this paper was supported in part by the Shandong Provincial Natural Science Foundation under Grant ZR2024QF071, in part by the National Natural Science Foundation of China under Grant 62401373, 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, in part by the Team-based Research Fund under Project No. TBRF/2024/1.10, and in part by the Shandong Provincial Natural Science Foundation under Grant ZR2023MF002.

## References


[1] M. Einemo, H. C. So, Weighted least squares algorithm for target localization in distributed MIMO radar, Signal Process. 115 (2015) 144–150.

[2] Q. He, R. S. Blum, A. M. Haimovich, Noncoherent MIMO radar for location and velocity estimation: More antennas means better performance, IEEE Trans. Signal Process. 58 (7) (2010) 3661–3680.

[3] X. Zeng, F. Zhang, B. Wang, K. J. R. Liu, Massive MIMO for high-accuracy target localization and tracking, IEEE Internet Things J. 8 (12) (2021) 10131–10145.

[4] R. Amiri, F. Behnia, M. A. M. Sadr, Exact solution for elliptic localization in distributed MIMO radar systems, IEEE Trans. Veh. Technol. 67 (2) (2018) 1075–1086.

[5] B. K. Chalise, Y. D. Zhang, M. G. Amin, B. Himed, Target localization in a multi-static passive radar system through convex optimization, Signal Process. 102 (2014) 207–215.

[6] L. Rui, K. C. Ho, Elliptic localization: Performance study and optimum receiver placement, IEEE Trans. Signal Process. 62 (18) (2014) 4673–4688.

[7] J. Liang, D. Wang, L. Su, B. Chen, H. Chen, H. C. So, Robust MIMO radar target localization via nonconvex optimization, Signal Process. 122 (2016) 33–38.

[8] Z. Wu, Y. Li, X. Meng, X. Lv, Q. Guo, A minimum joint error entropy-based localization method in mixed LOS/NLOS environments, IEEE Internet Things J. 10 (22) (2023) 19913–19924.

[9] Z. Shi, H. Wang, C. S. Leung, H. C. So, Robust MIMO radar target localization based on Lagrange programming neural network, Signal Process. 174 (2020) 107574.

[10] G. Wang, H. Chen, Y. Li, N. Ansari, NLOS error mitigation for TOA-based localization via convex relaxation, IEEE Trans. Wireless Commun. 13 (8) (2014) 4119–4131.

[11] S. S. Al-Samahi, Y. Zhang, K. C. Ho, Elliptic and hyperbolic localizations using minimum measurement solutions, Signal Process. 167 (2020) 107273.

[12] X. P. Li, Z.-L. Shi, Q. Liu, H. C. So, Fast robust matrix completion via entry-wise $\ell_0$-norm minimization, IEEE Trans. Cybern. 53 (11) (2023) 7199–7212.

[13] W. Xiong, C. Schindelhauer, H. C. So, Error-reduced elliptic positioning via joint estimation of location and a balancing parameter, IEEE Signal Process. Lett. 29 (2022) 2447–2451.

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

[15] X. Zhao, J. Li, Q. Guo, Robust target localization in distributed MIMO radar with nonconvex $\ell_p$ minimization and iterative reweighting, IEEE Commun. Lett. 27 (12) (2023) 3230–3234.

[16] W. Xiong, J. Liang, Z. Wang, H. C. So, Elliptic target positioning based on balancing parameter estimation and augmented Lagrange programming neural network, Digital Signal Process. 136 (2023) 104004.

[17] K. Panwar, P. Babu, Robust multistatic target localization in the presence of NLOS errors and outliers, IEEE Signal Process. Lett. 29 (2022) 2632–2636.

[18] Z. Liu, H. C. So, L. Zhang, X. P. Li, Robust receiver for OFDM-DCSK modulation via rank-1 modeling and $\ell_p$-minimization, Signal Process. 188 (2021) 108219.

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

[20] W. Xiong, J. Bordoy, C. Schindelhauer, A. Gabbrielli, G. Fischer, D. J. Schott, F. Hoeflinger, S. J. Rupitsch, H. C. So, Data-selective least squares methods for elliptic localization with NLOS mitigation, IEEE Sens. Lett. 5 (7) (2021) 1–4.

[21] E. J. Candès, X. Li, Y. Ma, J. Wright, Robust principal component analysis?, J. ACM 58 (3) (2011) 1–37.

[22] H. Xu, C. Caramanis, S. Mannor, Outlier-robust PCA: The high-dimensional case, IEEE Trans. Inf. Theory 59 (1) (2012) 546–572.

[23] D. Jin, F. Yin, A. M. Zoubir, H. C. So, Exploiting sparsity of ranging biases for NLOS mitigation, IEEE Trans. Signal Process. 69 (2021) 3782–3795.

[24] Z. Liu, H. C. So, X. P. Li, L. Zhang, Z.-Y. Wang, Robust and energy efficient sparse-coded OFDM-DCSK system via matrix recovery, IEEE Trans. Commun. 71 (8) (2023) 4839–4850.

[25] Z. Liu, X. P. Li, H. C. So, $\ell_0$-norm minimization-based robust matrix completion approach for MIMO radar target localization, IEEE Trans. Aerosp. Electron. Syst. 59 (5) (2023) 6759–6770.

[26] D. L. Donoho, Compressed sensing, IEEE Trans. Inf. Theory 52 (4) (2006) 1289–1306.

[27] D. L. Donoho, For most large underdetermined systems of linear equations the minimal $\ell_1$-norm solution is also the sparsest solution, Commun. Pure Appl. Math. 59 (6) (2006) 797–829.

[28] Z. Liu, C.-S. Leung, H. C. So, Formal convergence analysis on deterministic $\ell_1$-regularization based mini-batch learning for RBF networks, Neurocomputing 532 (2023) 77–93.

[29] X. Liu, H. Zhao, Robust hierarchical feature selection with a capped $\ell_2$-norm, Neurocomputing 443 (2021) 131–146.

[30] H. Yang, P. Antonante, V. Tzoumas, L. Carlone, Graduated non-convexity for robust spatial perception: From non-minimal solvers to global outlier rejection, IEEE Robot. Autom. Lett. 5 (2) (2020) 1127–1134.

[31] X. P. Li, Z.-Y. Wang, Z.-L. Shi, H. C. So, N. D. Sidiropoulos, Robust tensor completion via capped Frobenius norm, IEEE Trans. Neural Netw. Learn. Syst. Early Access (2023) (Jan. 2023).

[32] D. Geman, C. Yang, Nonlinear image recovery with half-quadratic regularization, IEEE Trans. Image Process. 4 (7) (1995) 932–946.

[33] R. He, W.-S. Zheng, T. Tan, Z. Sun, Half-quadratic-based iterative minimization for robust sparse representation, IEEE Trans. Pattern Anal. Mach. Intell. 36 (2) (2013) 261–275.

[34] M. Nikolova, M. K. Ng, Analysis of half-quadratic minimization methods for signal and image recovery, SIAM J. Sci. Comput. 27 (3) (2005) 937–966.

[35] Y. Wang, J. Yang, W. Yin, Y. Zhang, A new alternating minimization algorithm for total variation image reconstruction, SIAM J. Imag. Sci. 1 (3) (2008) 248–272.

[36] X. Yu, J.-C. Shen, J. Zhang, K. B. Letaief, Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems, IEEE J. Sel. Topics Signal Process. 10 (3) (2016) 485–500.

[37] V. Abolghasemi, S. Ferdowsi, S. Sanei, A gradient-based alternating minimization approach for optimization of the measurement matrix in compressive sensing, Signal Process. 92 (4) (2012) 999–1009.

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

[39] Y. Sun, P. Babu, D. P. Palomar, Majorization-Minimization algorithms in signal processing, communications, and machine learning, IEEE Trans. Signal Process. 65 (3) (2017) 794–816.

[40] J. P. Oliveira, J. M. Bioucas-Dias, M. A. Figueiredo, Adaptive total variation image deblurring: A majorization-minimization approach, Signal Process. 89 (9) (2009) 1683–1693.

[41] R. Jyothi, P. Babu, SOLVIT: A reference-free source localization technique using majorization minimization, IEEE/ACM Trans. Audio, Speech, Lang. Process. 28 (2020) 2661–2673.

[42] A. M. Zoubir, V. Koivunen, E. Ollila, M. Muma, Robust Statistics for Signal Processing, Cambridge University Press, 2018.

[43] W.-J. Zeng, H.-C. So, L. Huang, $\ell_p$-MUSIC: Robust direction-of-arrival estimator for impulsive noise environments, IEEE Trans. Signal Process. 61 (17) (2013) 4296–4308.