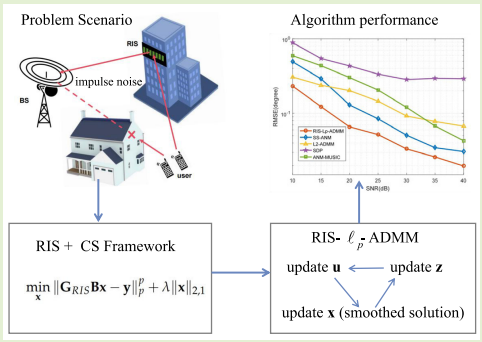
Robust RIS-Assisted DOA Estimation Under Impulsive Noise

Qingli Yan, *Member, IEEE*, and Baiting Zhou

***Abstract*—Reconfigurable intelligent surfaces (RISs) offer significant application flexibility and technical potential for 5G/6G communication and sensing due to their capability to programmably manipulate electromagnetic signal propagation paths. Addressing the challenge of direction-of-arrival (DOA) estimation in nonline-of-sight (NLOS) scenarios under impulsive noise interference, this article proposes an RIS-assisted DOA estimation scheme. The proposed DOA estimation algorithm formulates the problem within the compressed sensing (CS) framework by employing the ℓ_p norm as the residual term to suppress impulsive noise interference, and the $\ell_{2,1}$ norm as the regularization term to achieve joint sparse signal representation. Subsequently, we develop an iterative optimization algorithm using the alternating direction method of multipliers (ADMMs) to efficiently solve the nonconvex and nonsmooth sparse signal recovery problem. We further provide a theoretical analysis of the proposed algorithm for both nonconvex ($0 \leq p < 1$) and convex ($1 \leq p < 2$) cases. Simulation results demonstrate that, compared to existing DOA estimation algorithms, the proposed method achieves superior estimation accuracy in impulsive noise environments, validating the effectiveness of combining RIS-assisted transmission with the CS framework.**



***Index Terms*— Compressive sensing, direction-of-arrival (DOA) estimation, impulsive noise, reconfigurable intelligent surface (RIS).**

I. Introduction

With the further development of 5G/6G communication technologies, direction-of-arrival (DOA) estimation has become a core technology in wireless sensing. It not only optimizes the spectral utilization efficiency of information transmission and positioning performance of communication systems, but also drives the extension of wireless technologies to intelligent scenarios such as environmental monitoring and object recognition through integration with sensing technologies [1]. However, the DOA estimation accuracy of traditional methods, such as multiple signal classification (MUSIC) algorithm [2] and rotational invariance techniques (ESPRIT) [3], is decreased under nonline-of-sight (NLOS) environment.

To address the DOA estimation problem under NLOS environment, machine-learning algorithms were proposed to identify NLOS signals and mitigate them by analyzing signal characteristics [4], [5], [6], [7]. Kalman filtering (KF) [8], [9] can also incorporate NLOS errors into state variables to achieve dynamic recursive correction of ranging errors. In [10] and [11], NLOS constraints are transformed into convex optimization problems through mathematical modeling, and the optimal solution is obtained with the help of relaxation variables. Deep learning networks are used to directly extract NLOS features from signal waveforms to handle dynamic NLOS variations in time-series signals [12], [13]. In addition, hardware-level optimization strategies, such as increasing signal transmission power, selecting lower frequency bands to enhance penetration, and optimizing antenna layouts, also provide effective approaches to tackle NLOS challenges.

In recent years, reconfigurable intelligent surface (RIS) technology has shown significant advantages in expanding signal coverage and optimizing network structure [14], [15], [16]. RIS is essentially a planar array that contains a large number of reconfigurable passive units. Each unit, controlled by an intelligent controller, can independently introduce a specific phase shift to incident signals, thus collaboratively altering the propagation of reflected signals [17]. It is equivalent to setting up relay nodes between the obstacle and the receiver to forward the signal and bypass the obstacle. Intelligent control of incident signal and data transmission is realized. Therefore, the potential of RIS technology in the field of communications

Received 26 September 2025; revised 17 November 2025; accepted 18 November 2025. Date of publication 27 November 2025; date of current version 14 January 2026. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 62101443 and in part by Xi'an Science and Technology Plan Project under Grant 2025JH-GXKJRC-0134. The associate editor coordinating the review of this article and approving it for publication was Prof. Engin Masazade. *(Corresponding author: Qingli Yan.)*

The authors are with the School of Computer Science and Technology, Xi'an University of Posts and Telecommunications, Xi'an 710121, China, and also with Shaanxi Key Laboratory of Network Data Analysis and Intelligent Processing, Xi'an, Shaanxi 710121, China (e-mail: yql@xupt.edu.cn; 13892618597@163.com).

Digital Object Identifier 10.1109/JSEN.2025.3635768

has attracted extensive research attention [18], [19], and RIS-assisted schemes have also been applied to DOA estimation. Song et al. [20] utilized RIS to reconstruct the NLOS wireless sensing link and jointly optimized the transmission and RIS reflection beamforming to minimize the Cramer–Rao lower bound (CRLB) of DoA estimation. In [21], researchers achieved multiple measurements by adjusting the phase of the reflected signal from RIS units, and constructed a low-cost DOA estimation system that only requires a single fully functional receiver. Chen et al. [22] proposed two RIS phase optimization strategies, namely maximizing signal coherence and minimizing the Cramer–Rao bound (CRB), to maximize the performance of DOA estimation. In [23], a method based on multidimensional DOA estimation is proposed to solve the RIS-assisted millimeter-wave MIMO channel estimation problem.

However, most of the above studies on DOA estimation are based on the Gaussian noise assumption. In practical scenarios, the data collected by the receiving array is often corrupted not only by conventional Gaussian thermal noise but also by unpredictable strong transient interferences, such as electromagnetic pulse disturbances, signal aliasing, or hardware nonlinear effects. These interferences give rise to outliers in the measured data that, despite their small amplitude, exert a significant impact, which also named such non-Gaussian noise as impulse noise [24], [25]. It brings new challenges in wireless communication and sensing systems as traditional methods based on Gaussian assumptions will fail. The DOA estimation problem for traditional antenna arrays has received extensive attention. Kozick and Sadler [26] propose an expectation–maximization (EM)-based DOA estimation method by modeling the impulsive noise as Gaussian mixture model (GMM). Tsakalides and Nikias [27] put forward a maximum likelihood scheme for Cauchy-distributed noise. Robust statistics-based DOA estimation, such as M-estimator, S-estimator, or MM-estimator [28], [29] provides another representative approach for impulse noise environments. In addition, some robust Bayesian sparse recovery algorithms have also been proposed to address the challenges of impulsive noise, including robust sparse Bayesian learning (RSBL) [30] and robust Bayesian compressive sensing (RBCS) [31].

The DOA estimation under non-Gaussian noise for RIS-based wireless communication systems began to attract attention in recent years. He et al. [32] integrated RIS's dynamic regulation capability, robust statistical theory, and continuous sparse optimization to propose a Lorentz boundary function-atomic norm joint framework (LBF-AN), which can simultaneously suppress system noise and nontarget signal interference. Li et al. [33] constructed a logarithmic function constraint and atomic norm optimization model. In [34] and [35], the squared-sine error criterion was further introduced to enhance the adaptive impulsive noise suppression capability. However, the aforementioned methods limit their adaptability to the diverse and unpredictable characteristics of real-world impulsive noise, resulting in restricted generalization ability.

Compressed sensing (CS) provides a powerful framework for robust DOA estimation under harsh conditions. Its core advantage lies in breaking through the Nyquist limit, enabling high-resolution reconstruction with low sampling rates. Its inherent sparsity offers stronger anti-noise capability and lower signal-to-noise ratio (SNR) dependence. For CS-based methods, the off-grid problem is inevitable [36], [37]. The dictionary learning optimization method [38] compensates for grid discretization deviations by dynamically adjusting the dictionary to match the real DOA characteristics. Block sparse optimization [39] covers continuous angle intervals through block structures, refines deviations with intra-block parameters, and iteratively optimizes to adapt to real DOA, especially achieving higher estimation accuracy in scenarios with low SNR or large grid intervals.

In this article, we address the DOA estimation problem for RIS-assisted sensing systems under impulsive noise in the CS framework. The ℓ_p norm minimization method is adopted herein to address impulsive noise, and existing studies have confirmed its effectiveness in suppressing impulsive noise. Zeng et al. [40] proposed a subspace decomposition method based on the ℓ_p norm (ℓ_p-MUSIC) to improve the resolution and accuracy of DOA estimation under such noise conditions. In [41], a joint sparse recovery framework employing ℓ_p-$\ell_{2,0}$ minimization was introduced to address the robust sparse recovery problem within the multiple measurement vector (MMV) model under impulsive noise. Additionally, an ℓ_p Capon filter [42] was proposed to enhance the performance of DOA estimation in environments characterized by abnormal noise. However, the exponent p of the ℓ_p-norm in the above methods is limited to the interval $1 < p < 2$. Although the objective function within this range is convex, which is convenient for calculation and has a certain ability to suppress impulsive noise, the sparsity induced by it is limited, and its performance is restricted in strong impulsive noise environments. Therefore, we consider extending the range of the ℓ_p-norm to $0 \leq p < 2$. This range can more closely approximate the ideal ℓ_0 norm sparsity metric and is expected to achieve more accurate and robust sparse signal recovery in impulsive noise environments. To efficiently solve the resulting nonconvex optimization problem, we employ the alternating direction method of multiplier (ADMM) framework, which features excellent decomposability, parallelizability, and robust convergence properties. This framework is particularly suitable for solving large-scale problems inherent in RIS configurations. The main contributions of this study can be summarized as follows.

1) *Robust Residual Metric Based on ℓ_p Norm:* Employing the ℓ_p-norm ($0 \leq p < 2$) as the residual metric criterion, replacing the traditional ℓ_2 norm or ℓ_1 norm loss function, can achieve better sparsity and stronger flexibility. Leveraging its subquadratic growth property, the ℓ_p loss function exhibits lower sensitivity to large noise outliers.
2) *ADMM-Based Joint ℓ_p-$\ell_{2,1}$ Sparse Recovery Framework:* The proximal operator of the ℓ_p norm function is introduced into the augmented Lagrangian method

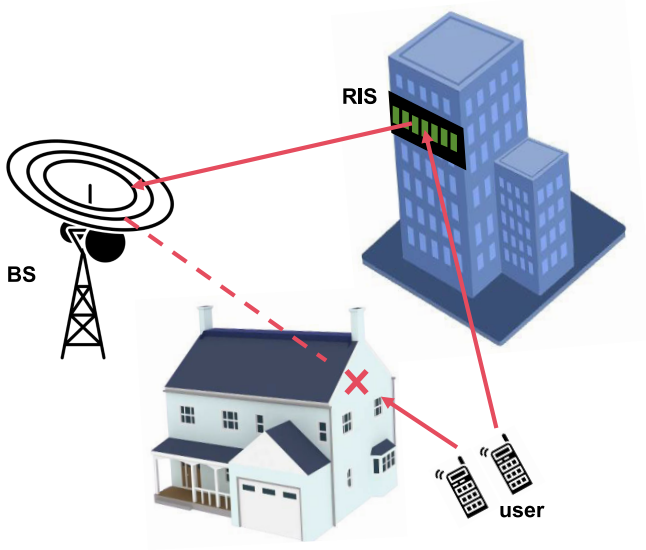


Fig. 1. RIS-assisted DOA estimation system model.

(ALM), and a smoothing strategy is applied to address challenges arising from local nonconvexity. Ultimately, an efficient ADMM-based iterative algorithm is designed to achieve sparse signal recovery.

3) *Algorithm Convergence Proof:* The convergence of the proposed RIS-ℓ_p-ADMM algorithm is rigorously proven. A detailed analysis is provided specifically for the nonconvex case where $0 \leq p < 1$, along with the corresponding convergence conditions.

The structure of this article is arranged as follows. In Section II, the RIS-assisted uplink communication scenario with impulsive noise is described in detail, and an optimization model is established using the ℓ_p norm. In Section III, the restricted isometry property (RIP) of the sensing matrix is analyzed, and the proximity operator of the ℓ_p norm is introduced to prepare for subsequent solutions. In Section IV, the steps to solve the constructed problem using the ADMM algorithm are elaborated in detail. In Section V, the convergence conditions of the proposed algorithm under nonconvex scenarios are analyzed. In Section VI, a series of detailed numerical simulations are conducted for the characteristics of the proposed algorithm, demonstrating that the method outperforms other existing methods in DOA estimation under impulsive noise. Finally, Section VII concludes this article, analyzes, and points out future research directions.

II. System Model

In this article, we consider a RIS-assisted uplink DOA estimation scenario, as illustrated in Fig. 1, in which the LOS path between the user equipment (UE) and the base station (BS) is blocked. Therefore, one RIS with passive reflective elements is deployed to establish a virtual LOS cascade link for UE–BS communication. Both the BS and the RIS are supposed to be predetermined. For simplicity, the RIS is configured with a uniform linear array (ULA) structure, which contains N reflective elements. The distance d between adjacent array elements is set to $\lambda/2$, where λ denotes the wavelength. The BS is equipped with a single antenna. Assuming there are K independent source targets in the system, the signal received by the BS at the mth snapshot can be expressed as

$$y(m) = \sum_{k=0}^{K-1} \boldsymbol{h}_{B,R}^{H} \boldsymbol{\Phi}(m) \boldsymbol{h}_R(\theta_k) s_k(m) + \omega(m) \tag{1}$$

where $\boldsymbol{\Phi}(m) = \mathrm{diag}(\eta_0(m), \eta_1(m), \ldots, \eta_{N-1}(m))$ represents the reflection coefficients matrix of the RIS at the mth snapshot, and $\eta_n(m) = \beta_n(m)e^{j\theta_n(m)}$. Here, $\beta_n(m)$ represents the amplitude of the reflection coefficient of the nth element of the RIS at the mth snapshot, and $\theta_n(m)$ denotes its phase. $s_k(m)$ is the transmitted signal by kth target. $\boldsymbol{h}_R(\theta_k)$ represents the channel between the kth UE and the RIS, and $\boldsymbol{h}_{B,R}$ is the channel from RIS to BS, which can be respectively expressed as

$$\boldsymbol{h}_R(\theta_k) = e^{j2\pi \frac{\boldsymbol{d}}{\lambda} \sin\theta_k} \tag{2}$$

and

$$\boldsymbol{h}_{B,R} = e^{j2\pi \frac{\boldsymbol{d}}{\lambda} \sin\phi} \tag{3}$$

where θ_k are the DOA of the RIS, ϕ denotes the angle of departure (AOD) from the RIS to the BS, and $\mathbf{d} \triangleq [0, d, \ldots, (N-1)d]^H$. $\omega(m)$ denotes the mth measurement noise which would display heavy-tailed characteristics due to unknown scatters and attacks.

Let $\mathbf{g}(m) = \boldsymbol{h}_{B,R} \odot \boldsymbol{\varphi}(m)$ with $\varphi(m) = [\eta_0(m), \eta_1(m), \ldots, \eta_{N-1}(m)]^H$ and $\odot$ denotes the Hadamard product. Since the BS is single-antenna, the existing DOA estimation methods cannot be used directly. We propose to achieve multiple measurements by adjusting the reflected signals from the RIS. For M time slots, we have

$$\mathbf{y} = \mathbf{G}_{\mathrm{RIS}} \mathbf{A}(\boldsymbol{\theta}) \mathbf{s} + \boldsymbol{\omega} \tag{4}$$

where

$$\mathbf{G}_{\mathrm{RIS}} \triangleq \left[\mathbf{g}(0), \mathbf{g}(1), \ldots, \mathbf{g}(M-1)\right]^{\mathrm{H}} \in \mathbb{C}^{M\times N}$$
$$\mathbf{A}(\boldsymbol{\theta}) \triangleq \left[\mathbf{h}_R(\theta_0), \mathbf{h}_R(\theta_1), \ldots, \mathbf{h}_R(\theta_{K-1})\right] \in \mathbb{C}^{N\times K}$$
$$\mathbf{s} \triangleq \left[s_0, s_1, \ldots, s_{K-1}\right]^{\mathrm{H}} \in \mathbb{C}^{K\times 1}.$$

To offer a more generalized statistical characterization of $\boldsymbol{\omega}$, we employ the symmetric alpha-stable ($S\alpha S$) distribution [43], [44], [45] to model the impulse noise. The characteristic function of the $S\alpha S$ noise distribution model is given by

$$\varphi(\omega) = \exp\left(i\alpha\omega - \gamma^{\alpha}|\omega|^{\alpha}\right) \tag{5}$$

where α is the characteristic parameter with range $0 < \alpha < 2$. The smaller the α is, the thicker the tail of the distribution model is. This continuously adjustable decay property allows the $S\alpha S$ distribution to adapt to varying impulse intensity requirements across different scenarios. γ is the scale parameter of the distribution, which takes a positive value and indicates the degree of stretching of the distribution. The infinite variance property of $S\alpha S$ distribution strictly corresponds to the unpredictable large-amplitude statistical behavior observed in impulse noise, overcoming the theoretical limitations of traditional Gaussian models.

To estimate the DOA from K targets to the RIS, the CS framework is applied in this article. Typically, the continuous angular domain is partitioned into equiangular bins to formulate an overcomplete dictionary. Received signals are then represented as sparse linear combinations of dictionary atoms. This model is ultimately solved through sparse reconstruction algorithms. However, this framework implicitly imposes a strict constraint: the true DOA must align precisely with the predefined grid points. When the actual angle deviates from the grid, the off-grid problem then arises.

Let $\tilde{\boldsymbol{\theta}} = [\tilde{\theta}_1, \tilde{\theta}_2, \ldots, \tilde{\theta}_{N_\theta}]$ be the fixed angular search grid, where N_θ is the number of search grids. Similar to the approach in [46], to solve the off-grid problem, the steering vector admits a first-order Taylor series approximation as follows:

$$a(\theta_k) = a\left(\tilde{\theta}_{n_k}\right) + a_\theta\left(\tilde{\theta}_{n_k}\right)\delta\tilde{\theta}_{n_k} \tag{6}$$

where n_k denotes the index of the closest grid point to the true DOA θ_k of the kth target in the predefined angular search grid $\tilde{\boldsymbol{\theta}}$, $\theta_k = \tilde{\theta}_{n_k} + \delta\tilde{\theta}_{n_k}$, and $a_\theta(\tilde{\theta}_{n_k}) = (\partial a(\tilde{\theta}_{n_k}))/(\partial\tilde{\theta}_{n_k})$.

Let $\Delta\boldsymbol{\theta} = \mathrm{diag}(\delta\tilde{\theta}_{n_1}, \delta\tilde{\theta}_{n_2}, \ldots, \delta\tilde{\theta}_{n_k})$, we have

$$\boldsymbol{A}(\theta) = \boldsymbol{A}\left(\tilde{\boldsymbol{\theta}}\right) + \boldsymbol{A}_{\boldsymbol{\theta}}\left(\tilde{\boldsymbol{\theta}}\right)\Delta\boldsymbol{\theta}. \tag{7}$$

The measurement model in (4) becomes

$$\boldsymbol{y} = \boldsymbol{G}_{\mathrm{RIS}}\left(\boldsymbol{A}\left(\tilde{\boldsymbol{\theta}}\right) + \boldsymbol{A}_{\boldsymbol{\theta}}\left(\tilde{\boldsymbol{\theta}}\right)\Delta\boldsymbol{\theta}\right)\boldsymbol{s} + \boldsymbol{\omega}. \tag{8}$$

Let $\boldsymbol{B} = [\boldsymbol{A}(\tilde{\boldsymbol{\theta}}) \;\; \boldsymbol{A}_{\boldsymbol{\theta}}(\tilde{\boldsymbol{\theta}})] \in \mathbb{C}^{N\times 2N_\theta}$, and let $\boldsymbol{x}^H = [\boldsymbol{s}, \Delta\boldsymbol{\theta}\boldsymbol{s}]$, we finally obtain

$$\mathbf{y} = \mathbf{G}_{\mathrm{RIS}}\mathbf{B}\mathbf{x} + \boldsymbol{\omega}. \tag{9}$$

It is clear that $\mathbf{s}$ and $\Delta\boldsymbol{\theta}\boldsymbol{s}$ have the same sparse structure, exhibiting typical block-sparsity. Thus, the problem (9) is a typical block-sparse recovery problem. Considering the impulsive property of noise $\boldsymbol{\omega}$, we propose to use the generalized ℓ_p-norm as the loss function [47]

$$\min_{\mathbf{x}} \|\mathbf{G}_{\mathrm{RIS}}\mathbf{B}\mathbf{x} - \mathbf{y}\|_p^p + \lambda\|\mathbf{x}\|_{2,1} \tag{10}$$

where p is in the range of $0 \le p < 2$. The ℓ_p norm can effectively suppress the outliers caused by impulse noise. The $\ell_{2,1}$ norm enhances structural sparsity for accurate signal recovery.

However, when $0 \le p < 1$, although the ℓ_p norm would suppress the strong impulsive noise, the optimization problem is nonconvex, which may lead to multiple local minima. Traditional convex optimization methods cannot be used to solve this optimization problem in this case. To solve this nonconvex optimization problem, we introduce a smoothing factor into the $\ell_{2,1}$ norm regularization term. The ADMM framework is then applied to solve this nonconvex optimal problem. The convergence condition is further derived in Section V.

III. Some Analyses and Preliminaries

A. Analysis on ℓ_p-Norm Model

Section III-A demonstrates that the proposed method achieves exact sparse signal recovery under limited measurements through RIP analysis.

Definition 1: We say that the sensing matrix $\mathbf{G}_{\mathrm{RIS}}\mathbf{B} \in \mathbb{C}^{M\times 2N_\theta}$ obeys the joint restricted isometry property with constant δ_s if

$$\left(1-\delta_s\right)\|\boldsymbol{x}\|_2^2 \le \|\mathbf{G}_{\mathrm{RIS}}\mathbf{B}\boldsymbol{x}\|_2^2 \le \left(1+\delta_s\right)\|\boldsymbol{x}\|_2^2 \tag{11}$$

for all K joint-sparse vectors $\mathbf{x} \in \mathbb{C}^{2N_\theta}$.

Under the guidance of the theory in [48], [49], and [50], we set $\tilde{\mathbf{x}}$ as a feasible point of the optimization problem, the noise satisfies $\|\boldsymbol{\omega}\|_p \le \zeta$, and let $\mathbf{h} = \mathbf{x} - \tilde{\mathbf{x}}$, we can get

$$\begin{aligned}\|\mathbf{G}_{\mathrm{RIS}}\mathbf{B}\mathbf{h}\|_2 &\le \|\mathbf{G}_{\mathrm{RIS}}\mathbf{B}\mathbf{x} - \mathbf{y}\|_2 + \left\|\mathbf{G}_{\mathrm{RIS}}\mathbf{B}\tilde{\mathbf{x}} - \mathbf{y}\right\|_2 \\ &\le \|\mathbf{G}_{\mathrm{RIS}}\mathbf{B}\mathbf{x} - \mathbf{y}\|_p + \left\|\mathbf{G}_{\mathrm{RIS}}\mathbf{B}\tilde{\mathbf{x}} - \mathbf{y}\right\|_p \\ &\le 2\zeta\end{aligned} \tag{12}$$

where the second inequality follows from that $\|\mathbf{x}\|_2 \le \|\mathbf{x}\|_p$ holds for arbitrary vectors, if $0 \leqslant \mathrm{p} < 2$ and $\delta_{2s} < (\sqrt{2})/2 \approx 0.707$, it has been shown in:

$$\|\mathbf{h}\|_2 \le \frac{4\sqrt{1+\delta_{2s}}}{\sqrt{2}-2\delta_{2s}}\|\mathbf{G}_{\mathrm{RIS}}\mathbf{B}\mathbf{h}\|_2. \tag{13}$$

Substituting (13) into (12), it leads to

$$\left\|\boldsymbol{x} - \tilde{\boldsymbol{x}}\right\|_2 \le \frac{4\sqrt{1+\delta_{2s}}}{\sqrt{2}-2\delta_{2s}}\zeta. \tag{14}$$

In contrast to the l_2-norm constraints in other literature, we relax the noise condition to $\|\boldsymbol{\omega}\|_p \leqslant \zeta$, $0 \le p < 2$, which is equivalent to relaxing the conditions for noise recovery. That means the proposed formulation has the capability to stably recover $\mathbf{x}$ when the noise is highly impulsive with infinite variance.

B. Proximity Operator for ℓ_p-Norm Function

In this section, we introduce the proximity operator of the ℓ_p norm to facilitate the solution process in the subsequent text.

It is known that for $\boldsymbol{x} \in \mathbb{C}^{2N_\theta}$, the proximity operator of some function $f(\boldsymbol{x})$ with the regularization parameter μ is defined as follows [51]:

$$\mathrm{prox}_{f,\mu}(\boldsymbol{t}) = \arg\min_{\boldsymbol{x}}\left\{f(\boldsymbol{x}) + \frac{\mu}{2}\|\boldsymbol{x}-\boldsymbol{t}\|_2^2\right\}. \tag{15}$$

When $f(\boldsymbol{x}) = c\parallel \boldsymbol{x} \parallel_p^p$ with $c > 0$ and $0 \le p < 2$, solving the minimization problem degenerates to resolving $2N_\theta$ single variable minimization problems. According to the different value ranges of the exponent p, there are four types of calculation paradigms for the proximity operator $\mathrm{prox}_{f,\mu}(\mathbf{t})$.

Case 1: $p = 0$: The proximity operator reduces to the classic hard-thresholding operator

$$\mathrm{prox}_{f,\mu}(\boldsymbol{t})_i = \begin{cases} 0, & |t_i| \leqslant \sqrt{\dfrac{2c}{\mu}} \\ t_i, & \text{otherwise} \end{cases} \tag{16}$$

where t_i is the ith component of the vector $\mathbf{t}$.

Case 2: $0 < p < 1$: The proximity operator falls into the following specific form [52]:

$$\mathrm{prox}_{f,\mu}(\boldsymbol{t})_i = \begin{cases} 0, & |t_i| \leqslant \sigma \\ \{0, \mathrm{sign}(t_i)\,\phi\}, & |t_i| = \sigma \\ \mathrm{sign}(t_i)\,w_i, & |t_i| > \sigma \end{cases} \tag{17}$$

where sign$(*)$ is the sign function, $\phi = \left[2c(1-p)/\mu\right]^{1/(2-p)}$, $\sigma = \phi + cp\phi^{p-1}/\mu$, and w_i is the solution of $h_1(w) = cpz^{p-1} + \mu w - \mu|t_i| = 0$ in the region $(\phi, |t_i|)$. Given the convexity of $h_1(w)$, w_i can be efficiently found via Newton's method when $|t_i| > \sigma$ [53].

Case 3: $p = 1$: Under this case, the proximity operator admits a closed-form solution as follows [54], [55], [56]:

$$\begin{aligned}\text{prox}_{f,\mu}(\boldsymbol{t})_i &= \max\left\{|t_i| - \frac{c}{\mu}, 0\right\} \cdot \text{sign}(t_i) \\ &= \text{Shrink}\left(t_i, \frac{c}{\mu}\right)\end{aligned} \tag{18}$$

where Shrink$(*)$ denotes the widely adopted soft-thresholding operator, also known as the shrinkage operator.

Case 4: $1 < p < 2$: When $f(x)$ is convex and smooth, the proximity operator admits an explicit expression derived in [53] as follows:

$$\text{prox}_{f,\mu}(\boldsymbol{t})_i = \text{sign}(t_i)\, w_i \tag{19}$$

with $w_i \geqslant 0$ and w_i is the solution of the equality

$$h_2(w) = pcw^{p-1} + \mu w + \mu|t_i| = 0. \tag{20}$$

For $t_i \neq 0$ with $h_2(0) < 0 < h_2(|t_i|)$, the function $h_2(w)$ exhibits concavity and increases monotonically with w. Consequently, for $t_i \neq 0$, the solution to (20) satisfies $0 < w_i < |t_i|$ and is computable via Newton's method.

The negativity of $h_2(0)$ under the condition $t_i \neq 0$ demonstrates that $w^0 = 0$ acts as a lower bound for the solution. However, this value cannot be utilized as an initial iteration point in numerical implementations, owing to the singularity of $h_2'(w)$ as $w \to 0$, where the derivative diverges to infinity. Given that $h_2(w)$ is a monotone concave function on $w \geqslant 0$, the solution w_0 to $h_2(w) = 0$ satisfies the following properties:

$$\begin{cases} 0 < w_0 < 1, & h_2(1) > 0 \\ w_0 \geqslant 1, & h_2(1) \leqslant 0. \end{cases} \tag{21}$$

Moreover, when $1 < p < 2$, it follows that:

$$\begin{cases} pcw_0^{p-1} + \mu w_0^{p-1} - \mu|t_i| > 0, & 0 < w_0 < 1 \\ pcw_0 + \mu w_0 - \mu|t_i| \geqslant 0, & w_0 \geqslant 1. \end{cases} \tag{22}$$

Therefore, according to (21) and (22) above, we can obtain

$$\begin{cases} \left(\dfrac{\mu|t_i|}{pc+\mu}\right)^{\frac{1}{p-1}} < w_0 < 1, & h_2(1) > 0 \\ 1 \leqslant \dfrac{\mu|t_i|}{pc+\mu} \leqslant w_0, & h_2(1) \leqslant 0. \end{cases} \tag{23}$$

We can set $\zeta = \mu|t_i|/(pc+\mu)$. One could pick a positive lower bound of the solution as the initial value as follows:

$$w_i^0 = \begin{cases} \zeta^{1/(p-1)}, & \zeta < 1 \\ \zeta, & \text{otherwise.} \end{cases} \tag{24}$$

In practical implementations, however, the computed value of $\zeta^{1/(p-1)}$ may become vanishingly small when $\zeta < 1$ and $p \to 1$. In order to resolve this problem, when $\zeta < 1$, the corresponding initial value is provided by

$$w_i^0 = \begin{cases} \chi, & h_2(\chi) \leqslant 0 \\ 0, & h_2(\chi) > 0 \end{cases} \tag{25}$$

where χ is a small positive constant.

IV. Proposed DOA Estimation Algorithm

This section employs the ADMM [57], [58] to solve the minimization problem in (10), integrating the proximity operator for ℓ_p-norm functions into the Lagrangian framework. However, directly extending the ADMM algorithm does not guarantee convergence. In the ADMM framework, with the $1 \leq p < 2$ convex case, the ℓ_p-norm loss term and the nonsmooth $\ell_{2,1}$-regularization term are naturally separated, making the problem easy to tackle. However, for $0 \leq p < 1$, both the loss term and the regularization term are nonsmooth while the loss term is nonconvex. To address the nonsmoothness, we adopt a smoothing strategy for this nonconvex-nonsmooth problem. By introducing a positive constant $\epsilon > 0$, (10) becomes

$$\min_{\boldsymbol{x},\boldsymbol{u}} \lambda \|\boldsymbol{x}\|_{2,1}^{\epsilon} + \|\boldsymbol{u}\|_p^p \quad \text{s.t } \boldsymbol{u} = \boldsymbol{G}_{\text{RIS}}\boldsymbol{B}\boldsymbol{x} - \boldsymbol{y} \tag{26}$$

where $\|\boldsymbol{x}\|_{2,1}^{\epsilon} = \sum_{i=1}^{2N_\theta}(x_i^2 + \varepsilon^2)^{1/2}$. Then, the corresponding augmented Lagrangian function is given by

$$\begin{aligned}\mathcal{L}_p(\boldsymbol{u},\boldsymbol{x},\boldsymbol{z}) = \lambda\|\boldsymbol{x}\|_{2,1}^{\epsilon} + \|\boldsymbol{u}\|_p^p - \langle \boldsymbol{z}, \boldsymbol{G}_{\text{RIS}}\boldsymbol{B}\boldsymbol{x} - \boldsymbol{y} - \boldsymbol{u}\rangle \\ + \frac{b}{2}\|\boldsymbol{G}_{\text{RIS}}\boldsymbol{B}\boldsymbol{x} - \boldsymbol{y} - \boldsymbol{u}\|_2^2\end{aligned} \tag{27}$$

where $\boldsymbol{z} \in R$ is the dual variable, $b > 0$ is a penalty parameter associated with the augmentation. Now, applying ADMM to (27) produces the following iterations:

$$\boldsymbol{u}^{t+1} = \arg\min_{\boldsymbol{u}\in\mathbb{C}}\left\{\|\boldsymbol{u}\|_p^p + \frac{b}{2}\left\|\boldsymbol{G}_{\text{RIS}}\boldsymbol{B}\boldsymbol{x}^t - \boldsymbol{y} - \boldsymbol{u} - \frac{\boldsymbol{z}^t}{b}\right\|_2^2\right\} \tag{28}$$

$$\boldsymbol{x}^{t+1} = \arg\min_{\boldsymbol{x}\in\mathbb{C}}\left\{\lambda\|\boldsymbol{x}\|_{2,1}^{\epsilon} + \frac{b}{2}\left\|\boldsymbol{G}_{\text{RIS}}\boldsymbol{B}\boldsymbol{x} - \boldsymbol{y} - \boldsymbol{u}^{t+1} - \frac{\boldsymbol{z}^t}{b}\right\|_2^2\right\} \tag{29}$$

$$\boldsymbol{z}^{t+1} = \boldsymbol{z}^t - b\left(\boldsymbol{G}_{\text{RIS}}\boldsymbol{B}\boldsymbol{x}^{t+1} - \boldsymbol{y} - \boldsymbol{u}^{t+1}\right). \tag{30}$$

The u-subproblem (28) takes the form of a proximity operator and admits efficient solution as follows:

$$\boldsymbol{u}^{t+1} = \text{prox}_{\|\boldsymbol{u}\|_p^p, b}(\boldsymbol{\varsigma}^t) = \begin{cases} \text{solved as (16)}, & p = 0 \\ \text{solved as (17)}, & 0 < p < 1 \\ \text{Shrink}\left(\boldsymbol{\varsigma}^t, \dfrac{1}{b}\right), & p = 1 \\ \text{solved as (19)}, & 1 < p < 2 \\ \dfrac{b\boldsymbol{\varsigma}^t}{b+2}, & p = 2 \end{cases} \tag{31}$$

where $\boldsymbol{\varsigma}^t = \boldsymbol{G}_{\text{RIS}}\boldsymbol{B}\boldsymbol{x}^t - \boldsymbol{y} - \boldsymbol{z}^t/b$.

Algorithm 1 RIS-ℓ_p-ADMM Algorithm

1: **Initialization:** $\boldsymbol{u}^0, \boldsymbol{x}^0, \boldsymbol{z}^0 \in \mathbb{C}$; λ, b, ϵ, and ρ are constants. Set the iteration termination criteria. Set $t = 0$.
2: **repeat**
3: Update $\boldsymbol{u}$ using:

$$\boldsymbol{u}^{t+1} = \operatorname{prox}_{\|\boldsymbol{u}\|_p^p, b}\left(\boldsymbol{\varsigma}^t\right) = \begin{cases} \text{solved as (16)}, & p = 0 \\ \text{solved as (17)}, & 0 < p < 1 \\ \operatorname{Shrink}\left(\boldsymbol{\varsigma}^t, \frac{1}{b}\right), & p = 1 \\ \text{solved as (19)}, & 1 < p < 2 \\ \frac{b\boldsymbol{\varsigma}^t}{b+2}, & p = 2. \end{cases}$$

4: Update $\boldsymbol{x}$ using:

$$\boldsymbol{x}^{t+1} = \arg\min_{\boldsymbol{x}\in\mathbb{C}} \frac{1}{2}\boldsymbol{x}^H \boldsymbol{C}\boldsymbol{x} + \boldsymbol{D}^H\boldsymbol{x} = -\boldsymbol{C}^{-1}\boldsymbol{D}.$$

5: Update $\boldsymbol{z}$ using:

$$\boldsymbol{z}^{t+1} = \boldsymbol{z}^t - b\left(\boldsymbol{G}_{\mathrm{RIS}}\boldsymbol{B}\boldsymbol{x}^{t+1} - \boldsymbol{y} - \boldsymbol{u}^{t+1}\right).$$

6: Set $t = t + 1$.
7: **until** a stopping criterion is reached

The x-subproblem (29) is now smooth and can use iterative approaches to address this minimization problem. For a fixed $\boldsymbol{x}^t$, the regularization term $\|\boldsymbol{x}\|_{2,1}^{\epsilon}$ is linearized as

$$\|\boldsymbol{x}\|_{2,1}^{\epsilon} \approx \left\|\boldsymbol{x}^t\right\|_{2,1}^{\epsilon} + \left(\nabla\left\|\boldsymbol{x}^t\right\|_{2,1}^{\epsilon}\right)^H\left(\boldsymbol{x} - \boldsymbol{x}^t\right) + \frac{\rho}{2}\left\|\boldsymbol{x} - \boldsymbol{x}^t\right\|_2^2 \tag{32}$$

and ρ is a positive constant. Let $\boldsymbol{\psi}^t = \boldsymbol{y} + \boldsymbol{u}^{t+1} + (\boldsymbol{z}^t)/b$, we then obtain

$$\begin{aligned} \boldsymbol{x}^{t+1} \approx & \arg\min_{\boldsymbol{x}\in\mathbb{C}}\left\{\lambda\left(\nabla\left\|\boldsymbol{x}^t\right\|_{2,1}^{\epsilon}\right)^H\left(\boldsymbol{x} - \boldsymbol{x}^t\right)\right. \\ & \left. + \frac{\lambda\rho}{2}\left\|\boldsymbol{x} - \boldsymbol{x}^t\right\|_2^2 + \frac{b}{2}\left\|\left(\boldsymbol{G}_{\mathrm{RIS}}\boldsymbol{B}\right)\boldsymbol{x} - \boldsymbol{\psi}^t\right\|_2^2\right\} \\ = & \arg\min_{\boldsymbol{x}\in\mathbb{C}}\left\{\frac{1}{2}\boldsymbol{x}^H\left(\lambda\rho I_{2N_\theta} + b\left(\boldsymbol{G}_{\mathrm{RIS}}\boldsymbol{B}\right)^H\boldsymbol{G}_{\mathrm{RIS}}\boldsymbol{B}\right)\boldsymbol{x}\right. \\ & \left. + \left(\lambda\left(\nabla\|\boldsymbol{x}^t\|_{2,1}^{\epsilon}\right) - \lambda\rho\boldsymbol{x}^t - b\left(\boldsymbol{G}_{\mathrm{RIS}}\boldsymbol{B}\right)^H\boldsymbol{\psi}^t\right)^H\boldsymbol{x}\right\}. \end{aligned} \tag{33}$$

Thus, it is easy to derive

$$\boldsymbol{x}^{t+1} = -\boldsymbol{C}^{-1}\boldsymbol{D} \tag{34}$$

where $\boldsymbol{C} = \lambda\rho I_{2N_\theta} + b(\boldsymbol{G}_{\mathrm{RIS}}\boldsymbol{B})^H(\boldsymbol{G}_{\mathrm{RIS}}\boldsymbol{B})$ and $\boldsymbol{D} = \lambda(\nabla\|\boldsymbol{x}\|_{2,1}) - \lambda\rho\boldsymbol{x}^t - b(\boldsymbol{G}_{\mathrm{RIS}}\boldsymbol{B})^H\boldsymbol{\psi}^t$. The proposed DOA estimation method can be summarized as Algorithm 1.

V. Convergence Analysis

As discussed in Section IV, the problem (10) is nonconvex when $0 \le p < 1$ and convex only when $1 \le p < 2$. The convergence of the ADMM solution in the convex case is relatively easy to prove which can refer to [59] and [60]. For any $b > 0$, under the assumption of $\rho > \lambda_{\max}((\mathbf{G}_{\mathrm{RIS}}\mathbf{B})^H(\mathbf{G}_{\mathrm{RIS}}\mathbf{B}))$, the sequence $(\mathbf{x}^t, \mathbf{u}^t, \mathbf{z}^t)$ generated by RIS-ℓ_p-ADMM Algorithm from any initiated value $(\mathbf{x}^0, \mathbf{u}^0, \mathbf{z}^0)$ converges. In the following sections, we will provide the convergence analysis in for $0 \le p < 1$.

We first demonstrate that changes in the dual iterates are bounded by those in the primal iterates

$$\begin{aligned} \left\|\mathbf{z}^{t+1} - \mathbf{z}^t\right\|_2^2 \le & \operatorname{slant}\frac{2\rho^2}{\vartheta}\left\|\mathbf{x}^{t+1} - \mathbf{x}^t\right\|_2^2 \\ & + \frac{2(\rho + \mathcal{C})^2}{\vartheta}\left\|\mathbf{x}^t - \mathbf{x}^{t-1}\right\|_2^2 \end{aligned} \tag{35}$$

where $\epsilon > 0$, $\mathcal{C} = 1/\epsilon$, and $\vartheta > 0$. Applying Fermat's rule to (29) yields

$$\begin{aligned} 0 = & \nabla\left\|\mathbf{x}^t\right\|_{2,1}^{\epsilon} + \rho\left(\mathbf{x}^{t+1} - \mathbf{x}^t\right) \\ & + b\left(\mathbf{G}_{\mathrm{RIS}}\mathbf{B}\right)^H\left(\mathbf{G}_{\mathrm{RIS}}\mathbf{B}\mathbf{x}^{t+1} - \mathbf{y} - \mathbf{u}^{t+1} - \frac{\mathbf{z}^t}{b}\right) \end{aligned} \tag{36}$$

which together with (30) implies that

$$\left(\mathbf{G}_{\mathrm{RIS}}\mathbf{B}\right)^H\mathbf{z}^{t+1} = \nabla\left\|\mathbf{x}^t\right\|_{2,1}^{\epsilon} + \rho\left(\mathbf{x}^{t+1} - \mathbf{x}^t\right). \tag{37}$$

Then,

$$\begin{aligned} & \left\|\left(\mathbf{G}_{\mathrm{RIS}}\mathbf{B}\right)^H\left(\mathbf{z}^{t+1} - \mathbf{z}^t\right)\right\|_2^2 \\ & = \|\nabla\|\mathbf{x}^t\|_{2,1}^{\epsilon} - \nabla\|\mathbf{x}^{t-1}\|_{2,1}^{\epsilon} + \rho\mathbf{x}^{t+1} - \rho\mathbf{x}^t + \rho\mathbf{x}^t - \rho\mathbf{x}^{t-1}\|_2^2 \\ & \le \left(\left\|\nabla\left\|\mathbf{x}^t\right\|_{2,1}^{\epsilon} - \nabla\left\|\mathbf{x}^{t-1}\right\|_{2,1}^{\epsilon}\right\|_2 + \rho\left\|\mathbf{x}^{t+1} - \mathbf{x}^t\right\|_2 + \rho\left\|\mathbf{x}^t - \mathbf{x}^{t-1}\right\|_2\right) \\ & \le \left(\mathcal{C}\left\|\mathbf{x}^t - \mathbf{x}^{t-1}\right\|_2 + \rho\left\|\mathbf{x}^{t+1} - \mathbf{x}^t\right\|_2 + \rho\left\|\mathbf{x}^t - \mathbf{x}^{t-1}\right\|_2\right)^2 \\ & \le 2\rho^2\left\|\mathbf{x}^{t+1} - \mathbf{x}^t\right\|_2^2 + 2(\rho + \mathcal{C})^2\left\|\mathbf{x}^t - \mathbf{x}^{t-1}\right\|_2^2. \end{aligned} \tag{38}$$

Further, suppose that $(\mathbf{G}_{\mathrm{RIS}}\mathbf{B})(\mathbf{G}_{\mathrm{RIS}}\mathbf{B})^H \succcurlyeq \vartheta I_{2N_\theta}$ for some $\vartheta > 0$, we get

$$\left\|\left(\mathbf{G}_{\mathrm{RIS}}\mathbf{B}\right)^H\left(\mathbf{z}^{t+1} - \mathbf{z}^t\right)\right\|_2^2 \geqslant \vartheta\left\|\mathbf{z}^{t+1} - \mathbf{z}^t\right\|_2^2 \tag{39}$$

which at once implies (35) as desired.

By the definition of the minimizer $\mathbf{u}^{t+1}$, we get

$$\mathcal{L}_{p,\epsilon}\left(\mathbf{u}^{t+1}, \mathbf{x}^t, \mathbf{z}^t\right) - \mathcal{L}_{p,\epsilon}\left(\mathbf{u}^t, \mathbf{x}^t, \mathbf{z}^t\right) \leqslant 0. \tag{40}$$

Then by (39) and (40), we have

$$\begin{aligned} & \mathcal{L}_{p,\epsilon}\left(\mathbf{u}^{t+1}, \mathbf{x}^{t+1}, \mathbf{z}^t\right) - \mathcal{L}_{p,\epsilon}\left(\mathbf{u}^{t+1}, \mathbf{x}^t, \mathbf{z}^t\right) \\ & \qquad \leqslant \left(\frac{\mathcal{C}}{2} - \rho\right)\left\|\mathbf{x}^{t+1} - \mathbf{x}^t\right\|_2^2. \end{aligned} \tag{41}$$

By the definition of $\mathcal{L}_{p,\epsilon}(\mathbf{u}, \mathbf{x}, \mathbf{z})$ and (30), we get

$$\begin{aligned} & \mathcal{L}_{p,\epsilon}\left(\mathbf{u}^{t+1}, \mathbf{x}^{t+1}, \mathbf{z}^{t+1}\right) - \mathcal{L}_{p,\epsilon}\left(\mathbf{u}^{t+1}, \mathbf{x}^{t+1}, \mathbf{z}^t\right) \\ & \qquad \le \operatorname{slant}\frac{1}{b}\left\|\mathbf{z}^{t+1} - \mathbf{z}^t\right\|_2^2. \end{aligned} \tag{42}$$

Combining with (35), (40), (41), and (42), it follows that:

$$\begin{aligned} & \mathcal{L}_{p,\epsilon}\left(\mathbf{u}^{t+1}, \mathbf{x}^{t+1}, \mathbf{z}^{t+1}\right) - \mathcal{L}_{p,\epsilon}\left(\mathbf{u}^t, \mathbf{x}^t, \mathbf{z}^t\right) \\ & \le \left(\frac{2\rho^2}{b\vartheta} + \frac{\mathcal{C}}{2} - \rho\right)\left\|\mathbf{x}^{t+1} - \mathbf{x}^t\right\|_2^2 + \frac{2(\mathcal{C} + \rho)^2}{b\vartheta}\left\|\mathbf{x}^t - \mathbf{x}^{t-1}\right\|_2^2; \end{aligned} \tag{43}$$

Then, if constant b obeys the following inequality:

$$b > \frac{2\rho^2 + 2(\rho + \mathcal{C})^2}{\vartheta\left(\rho - \frac{\mathcal{C}}{2}\right)} \tag{44}$$

we can get

$$\begin{aligned}&\tilde{\mathcal{L}}_{p,\epsilon}\left(\mathbf{x}^{t+1}, \mathbf{u}^{t+1}, \mathbf{z}^{t+1}, \mathbf{x}^{t}\right)\\ &\leq \text{slant}\tilde{\mathcal{L}}_{p,\epsilon}\left(\mathbf{x}^{t}, \mathbf{u}^{t}, \mathbf{z}^{t}, \mathbf{x}^{t-1}\right) - a_2\left\|\mathbf{x}^{t+1} - \mathbf{x}^{t}\right\|_2^2\end{aligned} \tag{45}$$

where $\tilde{\mathcal{L}}_{p,\epsilon}(\mathbf{x}, \mathbf{u}, \mathbf{z}, \tilde{\mathbf{x}}) \leqslant \mathcal{L}_{p,\epsilon}(\mathbf{x}, \mathbf{u}, \mathbf{z}) - a_1\|\mathbf{x} - \tilde{\mathbf{x}}\|_2^2$ with $a_1, a_2 > 0$ are given by

$$a_1 = \frac{2(\rho + \mathcal{C})^2}{b\vartheta}, \quad a_2 = \rho - \frac{2\rho^2 + 2(\rho + \mathcal{C})^2}{b\vartheta} - \frac{\mathcal{C}}{2}. \tag{46}$$

Next, due to (37) and the fact that $\|\nabla\|\mathbf{x}\|_{2,1}^{\epsilon}\|_2^2 \leqslant 2N_\theta$ for any $\mathbf{x}$, we get

$$\begin{aligned}&\left\|(\mathbf{G}_{\text{RIS}}\mathbf{B})^H\mathbf{z}^t\right\|_2^2\\ &\leqslant \left(\left\|\nabla\left\|\mathbf{x}^{t-1}\right\|_{2,1}^{\epsilon}\right\|_2 + \rho\left\|\mathbf{x}^t - \mathbf{x}^{t-1}\right\|_2\right)^2\\ &\leqslant 2\left\|\nabla\left\|\mathbf{x}^{t-1}\right\|_{2,1}^{\epsilon}\right\|_2^2 + 2\rho^2\left\|\mathbf{x}^t - \mathbf{x}^{t-1}\right\|_2^2\\ &\leqslant 4N_\theta + 2\rho^2\left\|\mathbf{x}^t - \mathbf{x}^{t-1}\right\|_2^2.\end{aligned} \tag{47}$$

Let $\boldsymbol{\nu}^t = (\mathbf{u}^t, \mathbf{x}^t, \mathbf{z}^t, \mathbf{x}^{t-1})$. Then, $\tilde{\mathcal{L}}_{p,\epsilon}(\tilde{\boldsymbol{\nu}}^t)$ is lower semi-continuous and lower bounded, implying it is bounded below. Moreover, since $\tilde{\mathcal{L}}_{p,\epsilon}(\tilde{\boldsymbol{\nu}}^t)$ is nonincreasing (as established previously), it follows that:

$$\begin{aligned}\tilde{\mathcal{L}}_{p,\epsilon}\left(\boldsymbol{\nu}^1\right) &\geqslant \tilde{\mathcal{L}}_{p,\epsilon}\left(\boldsymbol{\nu}^t\right)\\ &= \lambda\left\|\mathbf{x}^t\right\|_{2,1}^{\epsilon} + \left\|\mathbf{u}^t\right\|_p^p\\ &\quad + \frac{b}{2}\left\|\mathbf{G}_{\text{RIS}}\mathbf{B}\mathbf{x}^t - \mathbf{y} - \mathbf{u}^t - \frac{\mathbf{z}^t}{b}\right\|_2^2\\ &\quad - \frac{1}{2b}\left\|\mathbf{z}^t\right\|_2^2 + a_1\left\|\mathbf{x}^t - \mathbf{x}^{t-1}\right\|_2^2\\ &\geqslant \lambda\left\|\mathbf{x}^t\right\|_{2,1}^{\epsilon} + \left\|\mathbf{u}^t\right\|_p^p\\ &\quad + \frac{b}{2}\left\|\mathbf{G}_{\text{RIS}}\mathbf{B}\mathbf{x}^t - \mathbf{y} - \mathbf{u}^t - \frac{\mathbf{z}^t}{b}\right\|_2^2 - \frac{2N_\theta}{b\vartheta}\\ &\quad - \left(\frac{\rho^2}{b\vartheta} - \frac{2(\rho + \mathcal{C})^2}{b\vartheta}\right)\left\|\mathbf{x}^t - \mathbf{x}^{t-1}\right\|_2^2.\end{aligned} \tag{48}$$

The coercivity of $\|\mathbf{x}\|_{2,1}^{\epsilon}$ and $\|\mathbf{u}^t\|_p^p$, combined with (47), ensures that $\mathbf{x}^t$, $\mathbf{u}^t$, and $\mathbf{z}^t$ are bounded. Consequently, $\mathbf{z}^t$ admits a convergent subsequence with limit point $\mathbf{z}^*$. By (45), we get

$$\begin{aligned}&a_2\sum_{t=1}^{N_{\text{iter}}}\left\|\mathbf{x}^{t+1} - \mathbf{x}^t\right\|_2^2\\ &\leqslant \sum_{t=1}^{N_{\text{iter}}}\left(\tilde{\mathcal{L}}_{p,\epsilon}\left(\boldsymbol{\nu}^t\right) - \tilde{\mathcal{L}}_{p,\epsilon}\left(\boldsymbol{\nu}^{t+1}\right)\right)\\ &= \tilde{\mathcal{L}}_{p,\epsilon}\left(\boldsymbol{\nu}^1\right) - \tilde{\mathcal{L}}_{p,\epsilon}\left(\boldsymbol{\nu}^{t+1}\right)\\ &\leqslant \tilde{\mathcal{L}}_{p,\epsilon}\left(\boldsymbol{\nu}^1\right) - \tilde{\mathcal{L}}_{p,\epsilon}\left(\boldsymbol{\nu}^*\right) < \infty\end{aligned} \tag{49}$$

where the third inequality holds because $\tilde{\mathcal{L}}_{p,\epsilon}(\boldsymbol{\nu}^t)$ converges and $\tilde{\mathcal{L}}_{p,\epsilon}(\boldsymbol{\nu}^t) > \tilde{\mathcal{L}}_{p,\epsilon}(\boldsymbol{\nu}^*)$ for all $t > 0$. Taking $N_{\text{iter}} \to \infty$ yields, we have

$$\sum_{t=1}^{\infty}\left\|\boldsymbol{x}^{t+1} - \boldsymbol{x}^t\right\|_2^2 < \infty. \tag{50}$$

By the same reasoning, we can get

$$\sum_{t=1}^{\infty}\left\|\mathbf{u}^{t+1} - \mathbf{u}^t\right\|_2^2 < \infty \tag{51}$$

and

$$\sum_{t=1}^{\infty}\left\|\mathbf{z}^{t+1} - \mathbf{z}^t\right\|_2^2 < \infty. \tag{52}$$

Therefore, we derive

$$\sum_{t=1}^{\infty}\left\|\boldsymbol{\nu}^{t+1} - \boldsymbol{\nu}^t\right\|_2^2 < \infty \tag{53}$$

and as $t \to \infty$

$$\left\|\tilde{\boldsymbol{\nu}}^{t+1} - \tilde{\boldsymbol{\nu}}^t\right\|_2^2 \to 0. \tag{54}$$

Then, by the optimality theory, we get

$$\begin{cases}0 \in \partial\left\|\mathbf{u}^{t+1}\right\|_p^p + b\boldsymbol{\nu}\left(\mathbf{G}_{\text{RIS}}\mathbf{B}\right)\left(\mathbf{x}^{t+1} - \mathbf{x}^t\right) + \psi\mathbf{z}^{t+1}\\ 0 = \nabla\left\|\mathbf{x}^t\right\|_{2,1}^{\epsilon} + \rho\left(\mathbf{x}^{t+1} - \mathbf{x}^t\right) - (\mathbf{G}_{\text{RIS}}\mathbf{B})^H\mathbf{z}^{t+1}\\ \mathbf{z}^{t+1} = \mathbf{z}^t - b\left((\mathbf{G}_{\text{RIS}}\mathbf{B})\mathbf{x}^{t+1} - \mathbf{y} - \mathbf{u}^{t+1}\right).\end{cases} \tag{55}$$

For a convergent subsequence $\boldsymbol{\nu}^{t_i}$, because $\lim_{t\to+\infty}\|\tilde{\boldsymbol{\nu}}^{t+1} - \tilde{\boldsymbol{\nu}}^t\|_2^2 = 0$, $\tilde{\mathbf{z}}^{t_i}$ and $\tilde{\nu}^{t_i+1}$ have the same limit point ν^*: $= (\mathbf{u}^*, \mathbf{x}^*, \mathbf{z}^*)$.

Moreover, since $\hat{\mathcal{L}}(\tilde{\boldsymbol{\nu}}^t)$ is convergent, $\|\mathbf{u}^t\|_p^p$ is also convergent. Then, passing to the limit in (55) along the subsequence $\tilde{\mathbf{z}}^{t_i}$ yields $0 \in \partial\|\mathbf{u}^*\|_p^p + \psi\mathbf{z}^*$, $(\mathbf{G}_{\text{RIS}}\mathbf{B})^H\mathbf{z}^* = \nabla\|\mathbf{x}^*\|_{2,1}^{\epsilon}$ and $(\mathbf{G}_{\text{RIS}}\mathbf{B})\mathbf{x}^* - \mathbf{u}^* = \mathbf{y}$, since z^* is a critical point of $\mathcal{L}_{p,\epsilon}$.

Let $\tilde{\boldsymbol{\nu}}^t = \left(\mathbf{u}^t, \mathbf{x}^t, \mathbf{z}^t\right)$, the generated sequence $\{\tilde{\boldsymbol{\nu}}^t\}$ has finite length

$$\sum_{t=0}^{\infty}\left\|\tilde{\boldsymbol{\nu}}^{t+1} - \tilde{\boldsymbol{\nu}}^t\right\|_2 < \infty \tag{56}$$

which implies that $\{\tilde{\boldsymbol{\nu}}^t\}$ is a Cauchy sequence and hence convergent.

As mentioned above, similar to the proof in [61], when we let $\tilde{\boldsymbol{\nu}}^{t+1} = (\mathbf{u}^{t+1}, \mathbf{x}^{t+1}, \mathbf{z}^{t+1}, \mathbf{x}^t)$ and $\mathcal{E} > 0$, for each t there exists

$$\begin{aligned}&\text{dist}\left(\partial\tilde{\mathcal{L}}\left(\tilde{\boldsymbol{\nu}}^{t+1}\right), 0\right)\\ &\leqslant \mathcal{E}\left(\left\|\mathbf{x}^t - \mathbf{x}^{t+1}\right\| + \left\|\mathbf{x}^t - \mathbf{x}^{t+1}\right\| + \left\|\mathbf{x}^t - \mathbf{x}^{t-1}\right\|\right)\end{aligned} \tag{57}$$

it follows that $\lim_{t\to\infty}\text{dist}(\partial\tilde{\mathcal{L}}(\tilde{\boldsymbol{\nu}}^{t+1}), 0) = 0$.

Based on the above discussions, we can prove that when $\epsilon > 0$, $\mathcal{C} = 1/\epsilon$, and $(\mathbf{G}_{\text{RIS}}\mathbf{B})(\mathbf{G}_{\text{RIS}}\mathbf{B})^H \succcurlyeq \vartheta I_M$ with some $\vartheta > 0$, then for any $0 \leq \text{p} < 2$, if $\rho = \alpha/\epsilon > 1/2\epsilon$ and $b > D/\epsilon$, where $D = (4(2\alpha^2 + 2\alpha + 1))/(\vartheta(2\alpha - 1))$, the sequence $\{(\mathbf{u}^t, \mathbf{x}^t, \mathbf{z}^t)\}$ generated by Algorithm 1 converges to a stationary point.

TABLE I
SETTING OF SIMULATION PARAMETERS

Parameter	Value
Number of targets	$K = 3$
Number of RIS elements	$N = 32$
Number of measurements	$M = 40$
Range of spatial angle	$(-45^\circ, 45^\circ)$
Directions of targets	$[-26.34^\circ, 9.45^\circ, 30.78^\circ]$
AOD from RIS to BS	$\phi = 45^\circ$
Maximum iterations	$N_{iter} = 100$
Grid interval	0.5°
Proximal parameter	$\rho = 10^3$
Penalty parameter	$b = 10^4$
Smoothing parameter	$\epsilon = 10^{-3}$

VI. NUMERICAL SIMULATIONS

To comprehensively verify the performance advantages of the proposed RIS-ℓ_p-ADMM algorithm, we have conducted a large number of simulation experiments. The main parameters are listed in Table I. Furthermore, we adopt a stopping tolerance of 10^{-3} (applicable to both primal and dual residuals) and specify the maximum number of iterations for the algorithm as 300.

For impulsive noise, we introduce the SNR to quantify the relative intensity between signal power and noise dispersion, and it is defined as

$$\mathrm{SNR} = 10\log_{10}\left(\frac{\mathbb{E}\left\{|s(t)|^2\right\}}{\gamma^{\alpha}}\right). \tag{58}$$

We assume that all targets has the same transmitted signal power. The characteristic parameter α and scale parameter γ of the $S\alpha S$ distribution are set to 1.2 and 0.2, respectively. unless otherwise specified.

The estimation accuracy of the proposed RIS-ℓ_p-ADMM is measured using root-mean-square error (RMSE) expressed as

$$\mathrm{RMSE} = \sqrt{\frac{1}{K N_{\mathrm{mc}}}\sum_{i_{\mathrm{mc}}=0}^{N_{\mathrm{mc}}-1}\left\|\hat{\theta}_{i_{\mathrm{mc}}} - \theta_{i_{\mathrm{mc}}}\right\|_2^2} \tag{59}$$

where N_{mc} is the number of Monte Carlo simulations, $\hat{\theta}_{i_{\mathrm{mc}}}$ is the estimated value of DOA in the i_{mc}th simulation, and $\theta_{i_{\mathrm{mc}}}$ is the true value of DOA. In this article, $N_{\mathrm{mc}} = 100$.

Fig. 2 shows the convergence performance of the proposed method under different SNR conditions. It can be clearly observed that the algorithm achieves stable convergence within 100 iterations across various noise levels. Based on this, in all subsequent experiments, we set the number of iterations of the algorithm to 100, so as to ensure estimation accuracy while balancing computational efficiency.

Fig. 3 shows the impact of grid interval. We can see that the smaller the sampling grid partition, the lower the estimation error will be. Although a finer grid leads to higher accuracy, the computational complexity also increases. We select 0.5° to make a tradeoff between computational complexity and accuracy.

Considering that different values of the characteristic parameter and scale parameter will make the $S\alpha S$ distribution present different pulse characteristics, we plot Fig. 4 to display the influence of characteristic parameter α and scale parameter γ on the RMSE. We can see that the proposed method has a lower RMSE when $\alpha = 1.2$ under the same value of γ. And the estimation performance is relatively stable when $p < 1$, and the RMSE has a tendency to increase when $p > 1$. This is because the proposed method has a stronger ability to suppress outliers when $p < 1$.

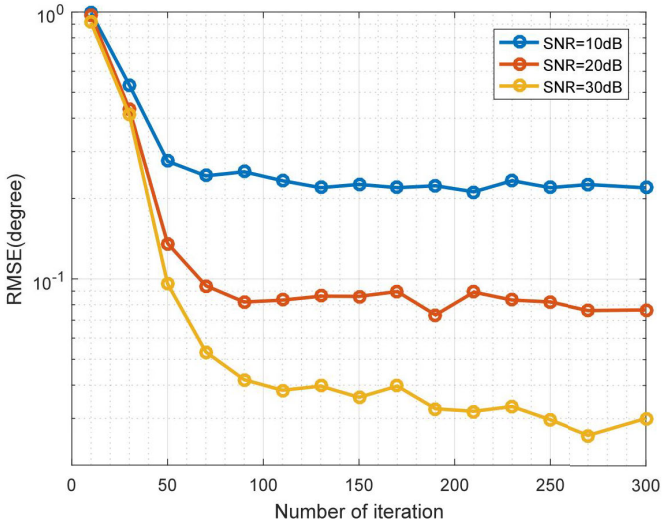


Fig. 2. RMSE of DOA estimation with different number of iterations.

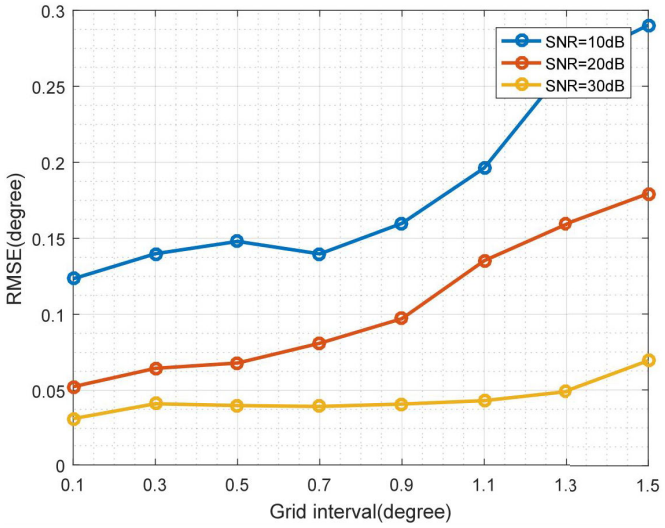


Fig. 3. RMSE of DOA estimation with different grid intervals.

Next, we investigate the performance of the number of RIS elements on DOA estimation as shown in Fig. 5. As the number of RIS elements increases, the DOA estimation accuracy also improves.

To comprehensively evaluate the performance of the proposed method, we compare it with the following four advanced methods.

1) *ANM-MUSIC [63]:* A novel atomic norm-based method was proposed to remove the interference signals and reconstruct the sparse signal. Additionally, a novel

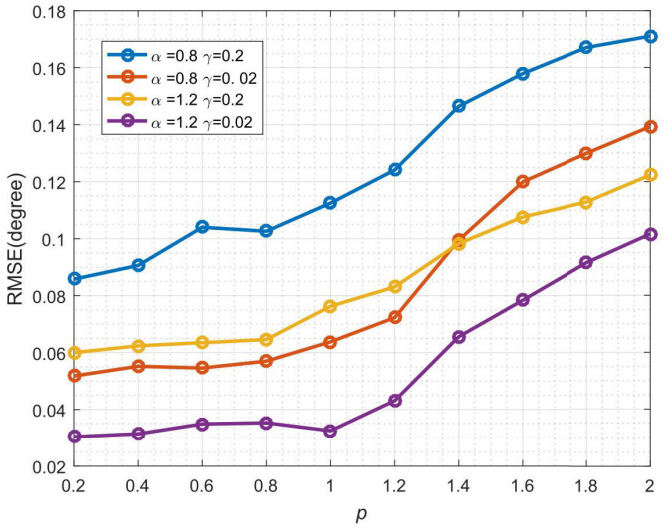


Fig. 4. RMSE of DOA estimation with different p values.

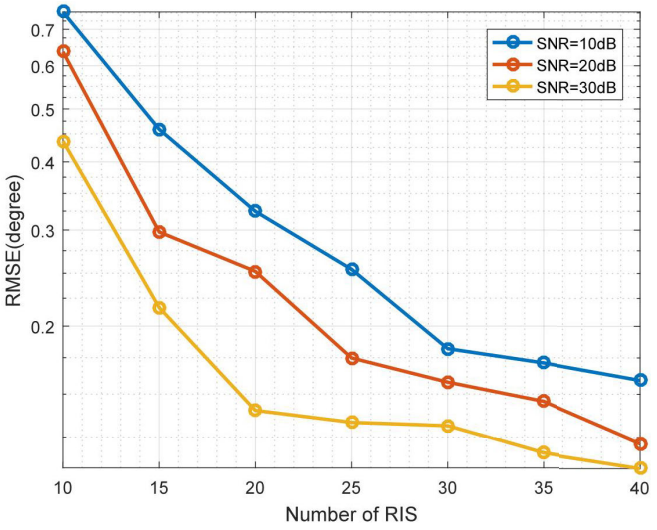


Fig. 5. RMSE of the DOA estimation with different number of RIS elements.

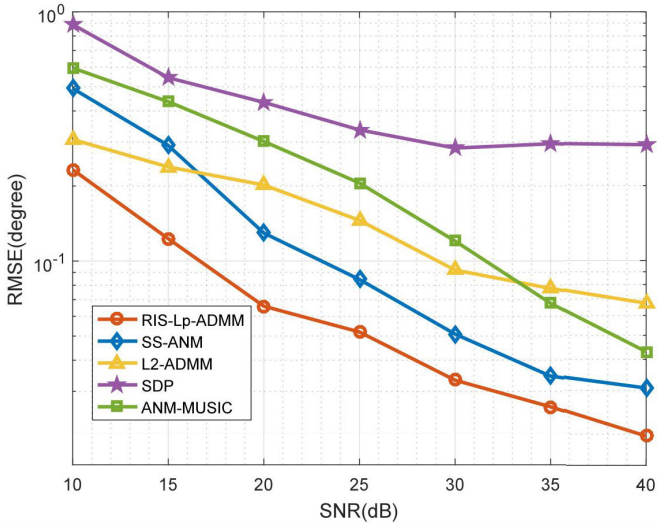


Fig. 6. Comparison of methods with different SNR.

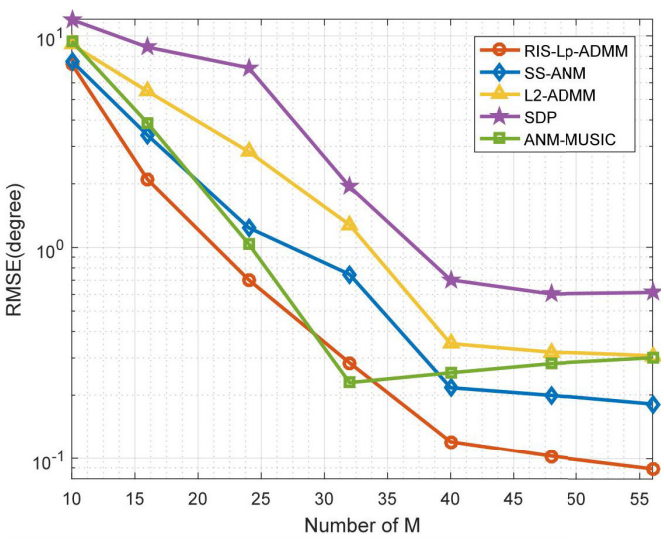


Fig. 7. Comparison of methods with different number of measurements.

Hankel-based MUSIC method is formulated to obtain the DOA information after the interference removal.

2) *Semidefinite Programming (SDP) [21]:* SDP was applied to solve the RIS-aided sparse DOA estimation problem that constructs an optimization framework based on atomic norm.
3) ℓ_2*-ADMM [62]:* The ℓ_2 norm was proposed as the residual error and atomic norm as sparsity constraint to achieve sparse DOA estimation for RIS-aided passive sensing scenarios, which are finally solved by iterative optimization using the ADMM.
4) *SS-ANM [34]:* The squared sine (SS) function was proposed as the cost function to suppress large outliers in impulse noise.

As shown in Fig. 6, the proposed method achieves a significant improvement in DOA estimation accuracy. By contrast, the methods of SDP, ℓ_2-ADMM, and ANM-MUSIC fail to account for the impulsive noise issue, resulting in relatively large RMSEs. It is worth noting that despite their inability to address outliers, all three methods effectively avoid off-grid errors by leveraging the atomic norm framework. Additionally, the SS-ANM can suppress the impact of impulsive noise to a certain extent, however, its overall performance remains inferior to that of the proposed method.

Fig. 7 shows the performance of these algorithms with different number of measurements. We can observe that an increase in the number of measurements brings a significant positive gain to algorithm performance. Meanwhile, compared with other competing algorithms, the proposed algorithm achieves higher DOA estimation accuracy with fewer measurements.

In addition, we compare the influence of the number of RIS elements on different algorithms in Fig. 8. We can see that in a certain range, increasing the number of RIS elements can reduce the estimation error, and the RMSE of our method is

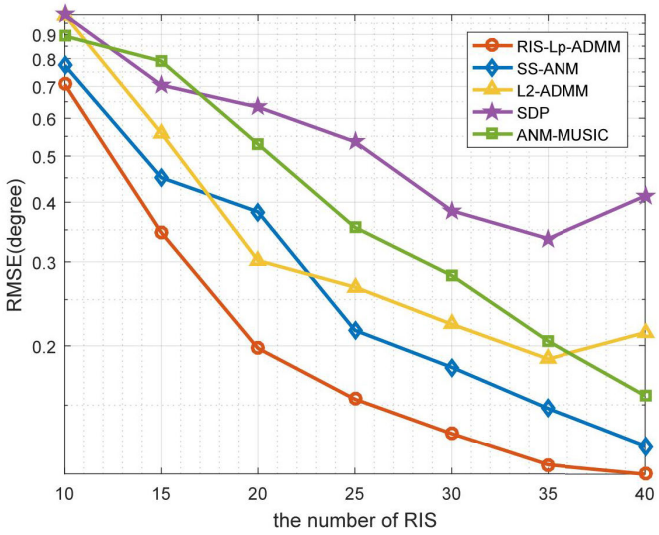


Fig. 8. Comparison of methods with different number of RIS elements.

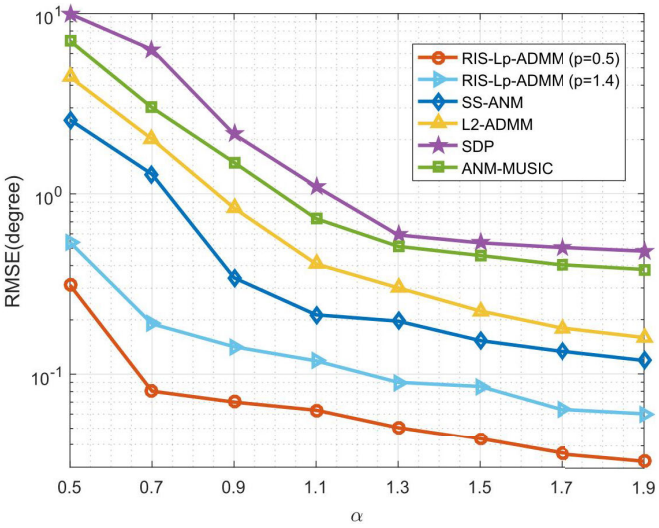


Fig. 9. Comparison of methods with different α (the characteristic parameter of $S\alpha S$ impulsive noise).

smaller than that of other methods under the same number of RIS elements.

Fig. 9 further illustrates the RMSEs of various methods under different characteristic parameters α. It is clear to observe that our method maintains excellent accuracy even when the system is subjected to high-intensity impulsive noise interference. This feature profoundly highlights the core advantage of our algorithm: it demonstrates stronger robustness compared to other methods in more extreme impulsive environments.

VII. Conclusion

In this article, the robust DOA estimation approach is developed by introducing the ℓ_p norm loss and $\ell_{2,1}$ regularization model to deal with the impulsive noise. The proximity operator of the ℓ_p norm function is introduced into the framework of the augmented Lagrange method, and ADMM is used to solve the iterative problem. For the nonconvex case of partial p values, a smoothing strategy is adopted for the regularization term. We further analyze and prove the convergence conditions of our proposed algorithm in the nonconvex ($0 \leq p < 1$) and convex ($1 \leq p < 2$) cases. Simulation experiments demonstrate that our proposed RIS-ℓ_p-ADMM algorithm shows better performance compared with other well-known algorithms under impulsive noise.

References

[1] M. P. Wylie and J. Holtzman, "The non-line of sight problem in mobile location estimation," *Proc. 5th Int. Conf. Universal Pers. Commun. (ICUPC)*, vol. 2, Cambridge, MA, USA, Oct. 1996, pp. 827–831.

[2] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.

[3] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process*, vol. ASSP-37, no. 7, pp. 984–995, Jul. 1989.

[4] X. Wang, Z. Wang, and B. O'Dea, "A TOA-based location algorithm reducing the errors due to non-line-of-sight (NLOS) propagation," *IEEE Trans. Veh. Technol.*, vol. 52, no. 1, pp. 112–116, Jan. 2003.

[5] N. Khajehnouri and A. H. Sayed, "A non-line-of-sight equalization scheme for wireless cellular location," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 1, Hong Kong, China, Apr. 2003, pp. 549–552.

[6] L. Cong and W. Zhuang, "Nonline-of-sight error mitigation in mobile location," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 560–573, Mar. 2005.

[7] Y. T. Chan, W. Y. Tsui, H. C. So, and P.-C. Ching, "Time-of-arrival based localization under NLOS conditions," *IEEE Trans. Veh. Technol.*, vol. 55, no. 1, pp. 17–24, Jan. 2006.

[8] M. Nájar and J. Vidal, "Kalman tracking for mobile location in NLOS situations," in *Proc. IEEE Pers.*, 2004, pp. 2203–2207.

[9] E. García, P. Poudereux, Á. Hernández, J. Ureña, and D. Gualda, "A robust UWB indoor positioning system for highly complex environments," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, Seville, Spain, Mar. 2015, pp. 3386–3391.

[10] H. Wu, L. Liang, X. Mei, and Y. Zhang, "A convex optimization approach for NLOS error mitigation in TOA-based localization," *IEEE Signal Process. Lett.*, vol. 29, pp. 677–681, 2022.

[11] G. Wang, A. M.-C. So, and Y. Li, "Robust convex approximation methods for TDOA-based localization under NLOS conditions," *IEEE Trans. Signal Process.*, vol. 64, no. 13, pp. 3281–3296, Jul. 2016.

[12] C. Wu et al., "Learning to localize: A 3D CNN approach to user positioning in massive MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4556–4570, Jul. 2021.

[13] K. Bregar and M. Mohorcic, "Improving indoor localization using convolutional neural networks on computationally restricted devices," *IEEE Access*, vol. 6, pp. 17429–17441, 2018.

[14] W. Yan, X. Yuan, Z.-Q. He, and X. Kuai, "Passive beamforming and information transfer design for reconfigurable intelligent surfaces aided multiuser MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1793–1808, Aug. 2020.

[15] W. Zhao, G. Wang, S. Atapattu, T. A. Tsiftsis, and C. Tellambura, "Is backscatter link stronger than direct link in reconfigurable intelligent surface-assisted system?" *IEEE Commun. Lett.*, vol. 24, no. 6, pp. 1342–1346, Jun. 2020.

[16] Z. Chen et al., "Reconfigurable-intelligent-surface-assisted B5G/6G wireless communications: Challenges, solution, and future opportunities," *IEEE Commun. Mag.*, vol. 61, no. 1, pp. 16–22, Jan. 2023.

[17] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2020.

[18] M. Di Renzo et al., "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.

[19] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, May 2021.

[20] X. Song, J. Xu, F. Liu, T. X. Han, and Y. C. Eldar, "Intelligent reflecting surface enabled sensing: Cramêr–Rao bound optimization," *IEEE Trans. Signal Process.*, vol. 71, pp. 2011–2026, 2023.

[21] P. Chen, Z. Yang, Z. Chen, and Z. Guo, "Reconfigurable intelligent surface aided sparse DOA estimation method with non-ULA," *IEEE Signal Process. Lett.*, vol. 28, pp. 2023–2027, 2021.

[22] H. Chen, Y. Bai, Q. Wang, H. Chen, L. Tang, and P. Han, "DOA estimation assisted by reconfigurable intelligent surfaces," *IEEE Sensors J.*, vol. 23, no. 12, pp. 13433–13442, Jun. 2023.

[23] Z. Chen, J. Tang, L. Huang, Z.-Q. He, K.-K. Wong, and J. Wang, "Robust target positioning for reconfigurable intelligent surface assisted MIMO radar systems," *IEEE Trans. Veh. Technol.*, vol. 72, no. 11, pp. 15098–15102, Nov. 2023.

[24] M. Guo, Y. Sun, J. Dai, and C. Chang, "Robust DOA estimation for burst impulsive noise," *Digit. Signal Process.*, vol. 114, Jul. 2021, Art. no. 103059.

[25] Q. Liu, Y. Gu, and H. C. So, "DOA estimation in impulsive noise via low-rank matrix approximation and weakly convex optimization," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 6, pp. 3603–3616, Dec. 2019.

[26] R. J. Kozick and B. M. Sadler, "Robust subspace estimation in non-Gaussian noise," in *Proc. IEEE Int. Conf. Acoust.*, vol. 6, Aug. 2002, pp. 3818–3821.

[27] P. Tsakalides and C. L. Nikias, "Maximum likelihood localization of sources in noise modeled as a stable process," *IEEE Trans. Signal Process.*, vol. 43, no. 11, pp. 2700–2713, Nov. 1995.

[28] A. M. Zoubir, V. Koivunen, Y. Chakhchoukh, and M. Muma, "Robust estimation in signal processing: A tutorial-style treatment of fundamental concepts," *IEEE Signal Process. Mag.*, vol. 29, no. 4, pp. 61–80, Jul. 2012.

[29] M. Muma, Y. Cheng, F. Roemer, M. Haardt, and A. M. Zoubir, "Robust source number enumeration for R-dimensional arrays in case of brief sensor failures," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Kyoto, Japan, Mar. 2012, pp. 3709–3712.

[30] J. Dai and H. C. So, "Sparse Bayesian learning approach for outlier-resistant direction-of-arrival estimation," *IEEE Trans. Signal Process.*, vol. 66, no. 3, pp. 744–756, Feb. 2018.

[31] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2346–2356, Jun. 2008.

[32] L. He, L. Li, R. Ying, C. Yu, and Q. Li, "A RIS-supported DOA estimation using combined norm constraint in non-Gaussian noise environments," in *Proc. IEEE 6th Int. Conf. Electron. Inf. Commun. Technol. (ICEICT)*, Qingdao, China, Jul. 2023, pp. 1314–1317.

[33] L. Li, L. He, Y. Li, and P. S. R. Diniz, "Robust RIS-based DOA estimation with mixed constraints," *IEEE Signal Process. Lett.*, vol. 31, pp. 2260–2264, 2024.

[34] L. Li, C. Yu, Y. Li, Z. Huang, and Q. Wu, "Joint squared-sine function and ANM-based DOA estimation with RIS," *IEEE Trans. Veh. Technol.*, vol. 72, no. 12, pp. 16856–16860, Dec. 2023.

[35] C. Yu et al., "DOA estimation for RIS assisted multi-user wireless communication system with squared-sine error criterion and Lawson norm constraint," in *Proc. IEEE 6th Int. Conf. Electron. Inf. Commun. Technol. (ICEICT)*, Qingdao, China, Jul. 2023, pp. 850–853.

[36] Z. Wang, G. Sun, J. Tong, and Y. Ji, "Pattern synthesis for sparse linear arrays via atomic norm minimization," *IEEE Antennas Wireless Propag. Lett.*, vol. 20, pp. 2215–2219, 2021.

[37] M. A. Abdelhay, N. O. Korany, and S. E. El-Khamy, "Synthesis of uniformly weighted sparse concentric ring arrays based on off-grid compressive sensing framework," *IEEE Antennas Wireless Propag. Lett.*, vol. 20, pp. 448–452, 2021.

[38] H. Zamani, H. Zayyani, and F. Marvasti, "An iterative dictionary learning-based algorithm for DOA estimation," *IEEE Commun. Lett.*, vol. 20, no. 9, pp. 1784–1787, Sep. 2016.

[39] J. Huang, J. Wang, W. Wang, and F. Zhang, "Sharp sufficient condition of block signal recovery via ℓ_2/ℓ_1 minimisation," *IET Signal Process.*, vol. 13, no. 5, pp. 495–505, Jul. 2019.

[40] W.-J. Zeng, H. C. So, and L. Huang, "ℓ_p-MUSIC: Robust direction-of-arrival estimator for impulsive noise environments," *IEEE Trans. Signal Process.*, vol. 61, no. 17, pp. 4296–4308, Sep. 2013.

[41] Z.-Q. He, Z.-P. Shi, L. Huang, H. Li, and H. C. So, "Sparse recovery of multiple measurement vectors in impulsive noise: A smooth block successive minimization algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Shanghai, China, Mar. 2016, pp. 4543–4547.

[42] Y. Yang and Y. Chen, "ℓ_p-norm based capon filter for robust DOA estimation," in *Proc. IEEE Int. Conf. Service Operations Logistics, Informat. (SOLI)*, Macau, China, Jun. 2024, pp. 1–6.

[43] F. Wen, P. Liu, Y. Liu, R. C. Qiu, and W. Yu, "Robust sparse recovery for compressive sensing in impulsive noise using ℓ_p-norm model fitting," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Shanghai, China, Mar. 2016, pp. 4643–4647.

[44] P. Tsakalides and C. L. Nikias, "Robust adaptive beamforming in alpha-stable noise environments," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. Conf.*, May 1996, pp. 2884–2887.

[45] 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.

[46] Y. Guo, B. Sun, N. Li, and D. Fang, "Variational Bayesian inference-based counting and localization for off-grid targets with faulty prior information in wireless sensor networks," *IEEE Trans. Commun.*, vol. 66, no. 3, pp. 1273–1283, Mar. 2018.

[47] X.-Z. Kong et al., "Joint ℓ_p-norm and $\ell_{2,1}$-norm constrained graph Laplacian PCA for robust tumor sample clustering and gene network module discovery," *Frontiers Genet.*, vol. 12, Feb. 2021, Art. no. 621317.

[48] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comp. Rendus. Mathématique*, vol. 346, nos. 9–10, pp. 589–592, Apr. 2008.

[49] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5302–5316, Nov. 2009.

[50] Z. Tan, Y. C. Eldar, A. Beck, and A. Nehorai, "Smoothing and decomposition for analysis sparse recovery," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1762–1774, Apr. 2014.

[51] P. L. Combettes and J. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, vol. 49. New York, NY, USA: Springer, Oct. 2011, pp. 185–212.

[52] G. Marjanovic and V. Solo, "On l_q optimization and matrix completion," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5714–5724, Nov. 2012.

[53] Z. Xu, X. Chang, F. Xu, and H. Zhang, "$L_{1/2}$ regularization: A thresholding representation theory and a fast solver," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 7, pp. 1013–1027, Jul. 2012.

[54] J. Yang and Y. Zhang, "Alternating direction algorithms for ℓ_1-problems in compressive sensing," *SIAM J. Scientific Comput.*, vol. 33, no. 1, pp. 250–278, Jan. 2011.

[55] Y. Xiao, H. Zhu, and S. Y. Wu, "Primal and dual alternating direction algorithms for ℓ_1-ℓ_1-norm minimization problems in compressive sensing," *Comput. Optim. Appl.*, vol. 54, pp. 441–459, May 2013.

[56] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, Jan. 2009.

[57] F. Bian, J. Liang, and X. Zhang, "A stochastic alternating direction method of multipliers for non-smooth and non-convex optimization," *Inverse Problems*, vol. 37, no. 7, Jul. 2021, Art. no. 075009.

[58] L. W. Zhong and J. T. Kwok, "Fast stochastic alternating direction method of multipliers," in *Proc. Int. Conf. Mach. Learn. (ICML)*, Beijing, China, Jan. 2014, pp. 46–54.

[59] M. Hong, Z.-Q. Luo, and M. Razaviyayn, "Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, South Brisbane, QLD, Australia, Apr. 2015, pp. 3836–3840.

[60] Y. Ouyang, Y. Chen, G. Lan, and E. Pasiliao, "An accelerated linearized alternating direction method of multipliers," *SIAM J. Imag. Sci.*, vol. 8, no. 1, pp. 644–681, Jan. 2015.

[61] F. Wang, Z. Xu, and H.-K. Xu, "Convergence of Bregman alternating direction method with multipliers for nonconvex composite problems," 2014, *arXiv:1410.8625*.

[62] P. Chen, Z. Chen, P. Miao, and Y. Chen, "RIS-ADMM: A RIS and ADMM-based passive and sparse sensing method with interference removal," *IEEE Commun. Lett.*, vol. 28, no. 4, pp. 867–871, Apr. 2024.

[63] Z. Chen, P. Chen, Z. Guo, Y. Zhang, and X. Wang, "A RIS-based vehicle DOA estimation method with integrated sensing and communication system," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 6, pp. 5554–5566, Jun. 2024.

Qingli Yan (Member, IEEE) received the Ph.D. degree in electrical engineering from Northwestern Polytechnical University, Xi'an, China, in 2019.

She is currently an Associate Professor with the School of Computer Science and Technology, Xi'an University of Posts and Telecommunications, Xi'an. Her research interests include source localization, secure intelligent communications and sensing, and artificial intelligent communications.

Baiting Zhou received the B.S. degree in information and computing science from Shaanxi University of Science and Technology, Xi'an, China, in 2023. She is currently pursuing the master's degree with Xi'an University of Posts and Telecommunications, Xi'an.

Her research interests include target localization in complex environments and intelligent communications and sensing.