

Article

Joint Smoothed l_0 -Norm DOA Estimation Algorithm for Multiple Measurement Vectors in MIMO Radar

Jing Liu ^{1,2}, Weidong Zhou ^{1,*} and Filbert H. Juwono ³¹ College of Automation, Harbin Engineering University, Harbin 150001, China; liujing@hrbeu.edu.cn² School of Electrical, Electronic and Computer Engineering, The University of Western Australia, Perth 6009, Australia³ Department of Electrical Engineering, Universitas Indonesia, Depok 16424, Indonesia; filbert@ieee.org

* Correspondence: zhouweidong@hrbeu.edu.cn; Tel.: +86-451-8256-8107

Academic Editor: Vittorio M. N. Passaro

Received: 13 March 2017; Accepted: 2 May 2017; Published: 8 May 2017

Abstract: Direction-of-arrival (DOA) estimation is usually confronted with a multiple measurement vector (MMV) case. In this paper, a novel fast sparse DOA estimation algorithm, named the joint smoothed l_0 -norm algorithm, is proposed for multiple measurement vectors in multiple-input multiple-output (MIMO) radar. To eliminate the white or colored Gaussian noises, the new method first obtains a low-complexity high-order cumulants based data matrix. Then, the proposed algorithm designs a joint smoothed function tailored for the MMV case, based on which joint smoothed l_0 -norm sparse representation framework is constructed. Finally, for the MMV-based joint smoothed function, the corresponding gradient-based sparse signal reconstruction is designed, thus the DOA estimation can be achieved. The proposed method is a fast sparse representation algorithm, which can solve the MMV problem and perform well for both white and colored Gaussian noises. The proposed joint algorithm is about two orders of magnitude faster than the l_1 -norm minimization based methods, such as l_1 -SVD (singular value decomposition), RV (real-valued) l_1 -SVD and RV l_1 -SRACV (sparse representation array covariance vectors), and achieves better DOA estimation performance.

Keywords: direction-of-arrival estimation; joint smoothed l_0 -norm; multiple measurement vectors; sparse signal reconstruction; MIMO radar

1. Introduction

Colocated multiple-input multiple-output (MIMO) radar has attracted a growing interest recently because it can achieve higher resolution and better parameter identification compared with conventional phased-array radar [1]. As one type of colocated MIMO radar, the monostatic MIMO radar is equipped with closely-located transmit and receive arrays, which result in the same angle for direction-of-departure (DOD) and direction-of-arrival (DOA) [2]. Aiming at DOA estimation, a large quantity of methods have been proposed, most of them are based on the signal and noise subspaces [3–6], such as the representative MUSIC (multiple signal classification) algorithm [3] and ESPRIT (estimation of signal parameters via rotational invariance techniques) algorithm [4].

In the area of sensor array signal processing, sparse representation is an important technique to estimate the parameters. When reconstructing the sparse signal, to avoid the NP-hard (non-deterministic polynomial-time hard) l_0 -norm minimization, different methods such as those based on the relaxed constraint l_1 -norm minimization [7], the focal underdetermined system solution (FOCUSS) [8] and the sparse Bayesian learning (SBL) [9] have been proposed. These algorithms were originally designed for the single measurement vector (SMV) problem. Then, to extend them to the multiple measurement vector (MMV) case, some algorithms, e.g., M-FOCUSS and M-SBL, were proposed [10,11]. Recently, the emerging application that exploits the sparse representation technique

to achieve the DOA estimation has been increasingly attractive. Compared with the alternative methods, sparse representation-based DOA estimation methods can achieve higher angle resolution and better adapt to some challenging scenarios such as low signal-to-noise ratio (SNR), as discussed in [7]. Among the abovementioned sparse signal reconstruction methods, l_1 -norm minimization is widely used for sparse DOA estimation because of its desirable solution accuracy and reasonable computing speed [7,10,12]. The DOA estimation algorithms using l_1 -norm minimization include l_1 -SRACV (sparse representation array covariance vectors) [13], real-valued (RV) l_1 -SRACV [14], and the iterative algorithm with reweighting l_1 -norm minimization [15]. All of these methods involve the SMV problem when recovering the signal. Aiming at the MMV case, the l_1 -SVD (singular value decomposition) [7], RV l_1 -SVD [16], reweighted l_1 -SVD [17] and CMSR (covariance matrix sparse representation) [10] were proposed for DOA estimation. However, the computational complexity of the l_1 -norm minimization is relatively large compared with that of the fast sparse representation method named smoothed l_0 -norm algorithm, which was proposed in [18,19] under the SMV circumstance, and applied in [20,21] for the imaging and the power-line communications. To achieve fast sparse DOA estimation, by designing a reweighted continuous function in the SMV case, the reweighted smoothed l_0 -norm (RSL0) algorithm was proposed in [22] based on the covariance vector. With enlarged array aperture and better angle estimation performance, the RSL0 algorithm [22] is about two orders of magnitude faster than the l_1 -norm minimization based methods.

The previously mentioned DOA estimation methods are all proposed under the ideal circumstance with Gaussian white noise. However, the noise is often correlated in practical, thus the additive noise is spatially modeled as Gaussian colored noise rather than white noise. It has been verified that, with colored noise, the performance of the conventional algorithms including the methods mentioned above, is seriously degraded, except for the specially designed methods such as those based on fourth-order cumulants (FOC) [12,23,24]. These FOC-based methods are computationally expensive when they are applied to estimate the DOAs for MIMO radar. In [25], the reduced-dimension (RD) FOC-based sparse representation method (RD l_1 -SRFOC) was proposed for DOA estimation in MIMO radar with array errors. RD l_1 -SRFOC is an l_1 -norm minimization based sparse representation method, which uses the high-order cumulants to deal with the colored noise and simultaneously solves the problem of mutual coupling.

It can be concluded that the existing sparse DOA estimation should improve the following problems: (I) for the l_1 -norm minimization based methods, the computation complexities need to be lowered down and the computing time needs speeding up; (II) the fast sparse DOA estimation algorithm, i.e., the reweighted smoothed l_0 -norm algorithm [22], is based on the covariance vector that is deemed the SMV problem. The fast sparse algorithm tailored for the MMV case needs to be studied and put forward; and (III) fast sparse algorithms for solving the colored Gaussian noise have not been available so far.

To solve the three problems stated above, a joint smoothed l_0 -norm algorithm for DOA estimation in the MIMO radar is proposed in this paper. The new sparse algorithm first obtains a low-complexity FOC-based data matrix to eliminate the white or colored Gaussian noises. Secondly, the proposed algorithm designs a joint smoothed function tailored for the MMV case. Thirdly, with the MMV-based joint smoothed function, a joint smoothed l_0 -norm sparse representation framework is constructed. Finally, the corresponding gradient-based sparse signal reconstruction is designed, and then DOA estimation is achieved. The proposed algorithm is a fast sparse DOA estimation algorithm, which can solve the multiple measurement vector problem, and perform well for both of the white and colored Gaussian noise environments. The proposed algorithm is about two orders of magnitude faster than the l_1 -norm minimization based methods. This is because it approximates the l_0 -norm by a joint smoothed function and performs the gradient-based steepest ascent scheme, which avoids the convex optimization problem involved in the l_1 -norm minimization. The proposed algorithm can provide better DOA estimation performance than l_1 -SVD, RV l_1 -SVD and RV l_1 -SRACV methods.

The rest of this paper is organized as follows. Section 2 presents the MIMO radar system and array signal models. In Section 3, the implementation process of the proposed algorithm is described in detail. Section 4 gives some related remarks and discussions regarding the advantages, the extended applications, the computational complexity and the noise elimination of the proposed algorithm. The experimental results of different methods are compared in Section 5. Finally, Section 6 concludes this paper.

Throughout the paper, the notations $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^{-1}$ and $(\cdot)^+$ denote conjugation, transpose, conjugate transpose, inverse and pseudo-inverse, respectively. \mathbf{I}_K represents a $K \times K$ -dimensional unit matrix. $\|\cdot\|_0$ and $\|\cdot\|_2$ denote the l_0 -norm and the l_2 -norm, respectively. In addition, we use \otimes , $E(\cdot)$ and $\text{vec}(\cdot)$ to separately indicate the Kronecker product, the expectation and the vectorization operators.

2. Problem Formulation

A narrowband monostatic MIMO radar system is considered, and it is shown in Figure 1. The transmit and the receive arrays of this system are equipped with M transmit and N receive antennas, respectively. The arrays are both half-wavelength spaced uniform linear arrays (ULAs), whose effects of the array errors, including mutual coupling and gain-phase errors, can be ignored. The transmitting antennas transmit M orthogonal narrowband waveforms, such as BPSK (binary phase shift keying) modulated signal waveforms. In the far field, there are P uncorrelated targets regarded as point scatterers at the same range. In addition, it is assumed that $P \leq M + N - 2$ [22,25,26]. Thus, at the receive array, the N antennas are impinged by the echo signals reflected by the P targets. Because of the closely located arrays in the monostatic MIMO radar, for the p th target, DOA and DOD are the same and can be denoted by θ_p . Then, by matched filtering operation, the $N \times 1$ dimensional complex envelop of the output of the m th carrier matched filter is expressed as [14,27]

$$\mathbf{x}_m(t) = \sum_{p=1}^P \mathbf{a}_r(\theta_p) \mathbf{a}_{tm}^T(\theta_p) s_p(t) + \mathbf{n}_m(t), \quad (1)$$

where $\mathbf{a}_r(\theta_p) = [1, e^{j\pi \sin(\theta_p)}, e^{j\pi 2 \sin(\theta_p)}, \dots, e^{j\pi(N-1) \sin(\theta_p)}]^T$ is the receive steering vector, \mathbf{a}_{tm} is the m th element of the transmit steering vector $\mathbf{a}_t(\theta_p) = [1, e^{j\pi \sin(\theta_p)}, e^{j\pi 2 \sin(\theta_p)}, \dots, e^{j\pi(M-1) \sin(\theta_p)}]^T$, $s_p(t)$ contains the target reflection coefficient and the transmitted baseband signal such as non-circular signal, and $\mathbf{n}_m(t)$ is the noise vector after the m th matched filter. After all the matched filters, the received data vector, i.e., the vector composed of the outputs of the M matched filters, is given by [2,25,28]

$$\begin{aligned} \mathbf{x}(t) &= \text{vec}([\mathbf{x}_1^T(t), \dots, \mathbf{x}_M^T(t)]^T) \\ &= \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \end{aligned} \quad (2)$$

where $\mathbf{x}(t) \in \mathbb{C}^{MN \times 1}$ and $\mathbf{n}(t) = [\mathbf{n}_1^T(t), \dots, \mathbf{n}_M^T(t)]^T \in \mathbb{C}^{MN \times 1}$ is the zero-mean Gaussian white or colored noise vector. $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_P(t)]^T \in \mathbb{C}^{P \times 1}$ is the reflected signal vector after the M carrier matched filters, and it is assumed to be statistically independent and non-Gaussian with zero-mean. The target model is considered as the classical Swerling case II, namely, the radar cross section (RCS) fluctuations are constant during a snapshot period and vary independently from snapshot to snapshot [29]. $\mathbf{n}(t)$ and $\mathbf{s}(t)$ are independent of each other. Moreover, $\mathbf{A} \in \mathbb{C}^{MN \times P}$ is the transmit–receive steering matrix, and its detailed expression is

$$\begin{aligned} \mathbf{A} &= [\mathbf{a}_t(\theta_1) \otimes \mathbf{a}_r(\theta_1), \dots, \mathbf{a}_t(\theta_P) \otimes \mathbf{a}_r(\theta_P)] \\ &= [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_P)], \end{aligned} \quad (3)$$

where $\mathbf{a}(\theta_p) = \mathbf{a}_t(\theta_p) \otimes \mathbf{a}_r(\theta_p) \in \mathbb{C}^{MN \times 1}$ is the transmit–receive steering vector for $p = 1, 2, \dots, P$. Therefore, by collecting J snapshots, the $MN \times J$ dimensional received data matrix in monostatic MIMO radar is represented as

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N}, \tag{4}$$

where $\mathbf{S} = [\mathbf{s}(t_1), \dots, \mathbf{s}(t_J)] \in \mathbb{C}^{P \times J}$ and $\mathbf{N} = [\mathbf{n}(t_1), \dots, \mathbf{n}(t_J)] \in \mathbb{C}^{MN \times J}$ are the signal matrix and complex Gaussian noise matrix, respectively. Based on the received data \mathbf{X} , sparse DOA estimation can be viewed as the signal reconstruction that subjects to [14]

$$\mathbf{X} = \mathbf{A}_{\hat{\theta}}\mathbf{S}_{\hat{\theta}} + \mathbf{N}, \tag{5}$$

where $\mathbf{A}_{\hat{\theta}} = [\mathbf{a}(\hat{\theta}_1), \dots, \mathbf{a}(\hat{\theta}_L)]$ is the complete dictionary with the discrete sample grid $\{\hat{\theta}_i\}_{i=1}^L, L \gg P$. $\mathbf{S}_{\hat{\theta}}$ has the same row support with \mathbf{S} . Let $\mathbf{s}_{\hat{\theta}} = [\mathbf{s}(\hat{\theta}_1), \dots, \mathbf{s}(\hat{\theta}_L)]$ be a sparse vector whose i th element can be equal to the l_2 -norm of the i th row in $\mathbf{S}_{\hat{\theta}}$. To obtain the sparsest solution of Equation (5), an ideal constraint is the l_0 -norm method by minimizing the nonzero number of $\mathbf{s}_{\hat{\theta}}$, which can be expressed as

$$\min \|\mathbf{s}_{\hat{\theta}}\|_0, \quad \text{s.t.} \quad \mathbf{X} = \mathbf{A}_{\hat{\theta}}\mathbf{S}_{\hat{\theta}} + \mathbf{N}. \tag{6}$$

Unfortunately, the l_0 -norm minimization method in Equation (6) is NP-hard. To solve this problem, conventional sparse DOA estimation methods [7,10,13–15,17] can obtain the DOAs with the l_1 -norm minimization. However, the computational complexity of the l_1 -norm minimization is relatively large. In [22], by solving the array covariance vector based SMV sparse signal reconstruction problem, the reweighted smoothed l_0 -norm algorithm greatly improves the computation speed. To develop a fast sparse DOA estimation algorithm tailored for the MMV problem existed in the MIMO radar systems, a joint smoothed l_0 -norm algorithm is proposed in the following.

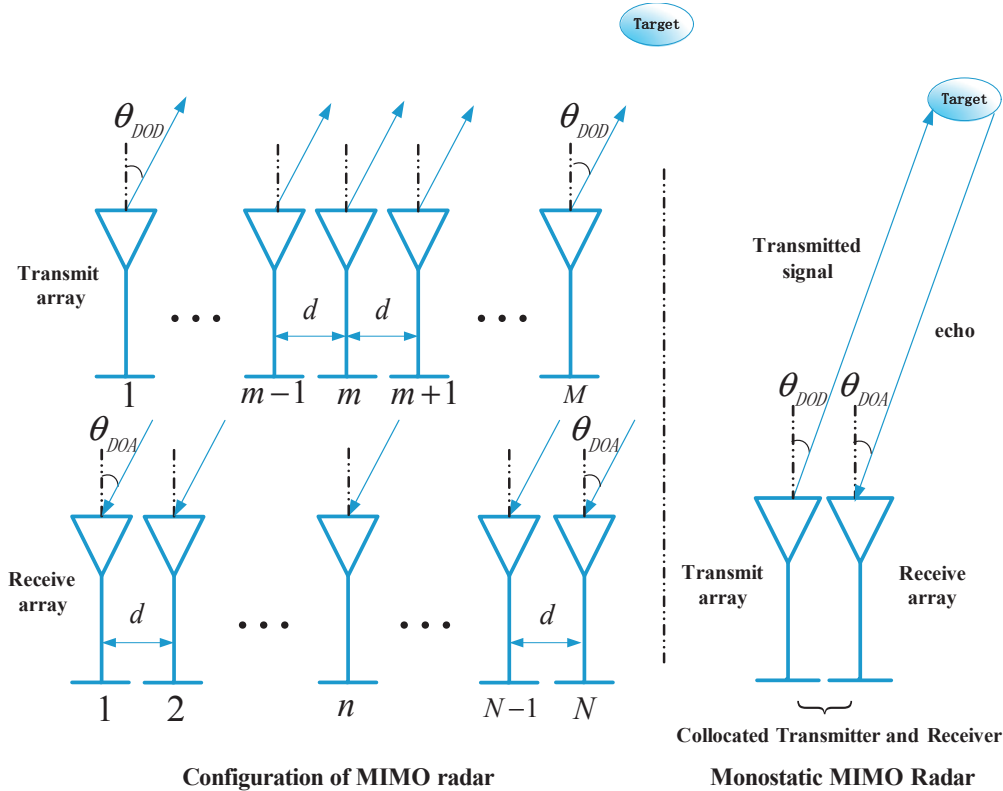


Figure 1. Configuration of monostatic multiple-input multiple-output radar.

3. The Proposed Algorithm

3.1. High-Order Cumulants Based Data Matrix Construction

Based on Equation (3), the detailed expression of the steering vector $\mathbf{a}(\theta_p)$ is

$$\begin{aligned}\mathbf{a}(\theta_p) &= [\mathbf{a}_r^T(\theta_p), e^{j\pi \sin(\theta_p)} \mathbf{a}_r^T(\theta_p), e^{j\pi(M-1) \sin(\theta_p)} \mathbf{a}_r^T(\theta_p)]^T \\ &= [1, \dots, z^{N-1}, z, \dots, z^N, \dots, z^{M-1}, \dots, z^{M+N-2}]^T,\end{aligned}\quad (7)$$

with $z = e^{j\pi \sin(\theta_p)}$. It can be observed that there are only $M + N - 1$ distinct elements in $\mathbf{a}(\theta_p)$; thus, the dimension of the received data vector $\mathbf{x}(t)$ can be reduced. Let $\mathbf{b}(\theta_p)$ be a new steering vector composed of the distinct elements. It is written as

$$\mathbf{b}(\theta_p) = [1, e^{j\pi \sin(\theta_p)}, e^{j\pi 2 \sin(\theta_p)}, \dots, e^{j\pi(M+N-2) \sin(\theta_p)}]^T, \quad (8)$$

where $\mathbf{b}(\theta_p) \in \mathbb{C}^{(M+N-1) \times 1}$. According to the element structures of $\mathbf{a}(\theta_p)$ and $\mathbf{b}(\theta_p)$, their relationship can be expressed as [25]

$$\mathbf{a}(\theta_p) = \mathbf{G}\mathbf{b}(\theta_p), \quad (9)$$

where $\mathbf{G} = [\mathbf{L}_0^T, \mathbf{L}_1^T, \dots, \mathbf{L}_{M-1}^T]^T \in \mathbb{C}^{MN \times (M+N-1)}$ and $\mathbf{L}_m = [\mathbf{0}_{N \times m}, \mathbf{I}_N, \mathbf{0}_{N \times (M-m-1)}] \in \mathbb{C}^{N \times (M+N-1)}$ for $m = 0, 1, \dots, M-1$. Based on the relationship in Equation (9), a reduced-dimensional matrix can be defined as [25]

$$\mathbf{R} = (\mathbf{G}^H \mathbf{G})^{(-\frac{1}{2})} \mathbf{G}^H \in \mathbb{C}^{(M+N-1) \times MN}. \quad (10)$$

Therefore, the dimensional reduction for the received data is carried out as follows [25]:

$$\begin{aligned}\tilde{\mathbf{x}}(t) &= \mathbf{R}\mathbf{G}\mathbf{B}\mathbf{s}(t) + \mathbf{R}\mathbf{n}(t) \\ &= (\mathbf{G}^H \mathbf{G})^{(-\frac{1}{2})} \mathbf{G}^H \mathbf{G}\mathbf{B}\mathbf{s}(t) + \tilde{\mathbf{n}}(t) \\ &= \mathbf{F}\mathbf{B}\mathbf{s}(t) + \tilde{\mathbf{n}}(t),\end{aligned}\quad (11)$$

where $\tilde{\mathbf{x}}(t) = \mathbf{R}\mathbf{x}(t) \in \mathbb{C}^{(M+N-1) \times 1}$, $\mathbf{B} = [\mathbf{b}(\theta_1), \dots, \mathbf{b}(\theta_P)] \in \mathbb{C}^{(M+N-1) \times P}$ and $\tilde{\mathbf{n}}(t) \in \mathbb{C}^{(M+N-1) \times 1}$ are the new low-dimensional received data vector, transmit–receive steering matrix and noise vector, respectively. In addition, $\mathbf{F} = (\mathbf{G}^H \mathbf{G})^{\frac{1}{2}}$ is a known diagonal matrix, and its (i, i) th element can be expressed as

$$\mathbf{F}(i, i) = \begin{cases} \sqrt{i} & i = 1, 2, \dots, \beta \\ \sqrt{\beta} & i = \beta + 1, \beta + 2, \dots, \beta + \tau - 1 \\ \sqrt{M + N - i} & i = \beta + \tau, \dots, M + N - 1, \end{cases} \quad (12)$$

with $\beta = \min(M, N)$ and $\tau = |M - N| + 2$. Note that the noise vector $\tilde{\mathbf{n}}(t)$ remains Gaussian after the dimensional reduction, for asymptotic normal distribution has the invariance speciality of linear transformation. To eliminate the Gaussian white noise or colored noise in $\tilde{\mathbf{x}}(t)$, fourth-order cumulant is exploited under the circumstance of collecting J snapshots. The definition of the FOC is given as follows [12,24,25]

$$\begin{aligned}C_{4\tilde{\mathbf{x}}}(k_1, k_2, k_3, k_4) &= \text{cum}\{\tilde{\mathbf{x}}_{k_1}, \tilde{\mathbf{x}}_{k_2}^*, \tilde{\mathbf{x}}_{k_3}, \tilde{\mathbf{x}}_{k_4}^*\} = E(\tilde{\mathbf{x}}_{k_1} \tilde{\mathbf{x}}_{k_2}^* \tilde{\mathbf{x}}_{k_3} \tilde{\mathbf{x}}_{k_4}^*) - E(\tilde{\mathbf{x}}_{k_1} \tilde{\mathbf{x}}_{k_2}^*)E(\tilde{\mathbf{x}}_{k_3} \tilde{\mathbf{x}}_{k_4}^*) \\ &\quad - E(\tilde{\mathbf{x}}_{k_1} \tilde{\mathbf{x}}_{k_3})E(\tilde{\mathbf{x}}_{k_2}^* \tilde{\mathbf{x}}_{k_4}^*) - E(\tilde{\mathbf{x}}_{k_1} \tilde{\mathbf{x}}_{k_4}^*)E(\tilde{\mathbf{x}}_{k_2}^* \tilde{\mathbf{x}}_{k_3}),\end{aligned}\quad (13)$$

where $E(\tilde{\mathbf{x}}_{k_1} \tilde{\mathbf{x}}_{k_2}^* \tilde{\mathbf{x}}_{k_3} \tilde{\mathbf{x}}_{k_4}^*) \approx \frac{1}{J} \sum_{t=1}^J \tilde{\mathbf{x}}_{k_1}(t) \tilde{\mathbf{x}}_{k_2}^*(t) \tilde{\mathbf{x}}_{k_3}(t) \tilde{\mathbf{x}}_{k_4}^*(t)$, $E(\tilde{\mathbf{x}}_{k_1} \tilde{\mathbf{x}}_{k_2}^*) \approx \frac{1}{J} \sum_{t=1}^J \tilde{\mathbf{x}}_{k_1}(t) \tilde{\mathbf{x}}_{k_2}^*(t)$, $\tilde{\mathbf{x}}_{k_i}$ stands for the k_i th element in $\tilde{\mathbf{x}}$, and $1 \leq k_i \leq M + N - 1$ for $i = 1, 2, 3, 4$. With the assumptions about the noise in Equation (2), the fourth-order cumulant of Gaussian noise has the following property:

$$C_{4\tilde{\mathbf{n}}}(k_1, k_2, k_3, k_4) = \text{cum}\{\tilde{\mathbf{n}}_{k_1}, \tilde{\mathbf{n}}_{k_2}^*, \tilde{\mathbf{n}}_{k_3}, \tilde{\mathbf{n}}_{k_4}^*\} = 0, \tag{14}$$

where $C_{4\tilde{\mathbf{n}}}(k_1, k_2, k_3, k_4)$ represents the fourth-order cumulant of $\tilde{\mathbf{n}}$ corresponding to the indices (k_1, k_2, k_3, k_4) . For the signal assumed in Equation (2), the fourth-order cumulant satisfies

$$C_{4s}(p_1, p_2, p_3, p_4) = \begin{cases} c_p & p_1 = p_2 = p_3 = p_4 = p \\ 0 & \text{otherwise,} \end{cases} \tag{15}$$

where $c_p = \text{cum}\{\mathbf{s}_p, \mathbf{s}_p^*, \mathbf{s}_p, \mathbf{s}_p^*\}$ is the FOC of the signal for the p th target, and $1 \leq p_i \leq P$ for $i = 1, 2, 3, 4$. Based on Equations (14) and (15), an FOC-based data matrix can be constructed with its (k_1, k_2) th element being obtained from

$$\begin{aligned} \mathbf{Y}(k_1, k_2) &= C_{4\tilde{\mathbf{x}}}(k_1, k_2, k_2, k_2) \\ &= \text{cum}\{\tilde{\mathbf{x}}_{k_1}, \tilde{\mathbf{x}}_{k_2}^*, \tilde{\mathbf{x}}_{k_2}, \tilde{\mathbf{x}}_{k_2}^*\} \\ &= \sum_{p=1}^P [\mathbf{Fb}(\theta_p)]_{k_1} [\mathbf{Fb}^*(\theta_p)]_{k_2} |\mathbf{F}(k_2, k_2)|^2 c_p, \end{aligned} \tag{16}$$

where $[\mathbf{Fb}(\theta_p)]_{k_1}$ and $[\mathbf{Fb}^*(\theta_p)]_{k_2}$ are the k_1 th and the k_2 th elements in $\mathbf{Fb}(\theta_p)$ and $\mathbf{Fb}^*(\theta_p)$. As a result, the data matrix \mathbf{Y} can be expressed as

$$\mathbf{Y} = \mathbf{FBD}_s(\mathbf{F}^3\mathbf{B})^H, \tag{17}$$

where $\mathbf{D}_s = \text{diag}(c_1, c_2, \dots, c_P)$ is a diagonal matrix composed of the fourth-order cumulants for \mathbf{s}_p , $p = 1, 2, \dots, P$. Then, the singular value decomposition of \mathbf{Y} can be performed as follows:

$$\mathbf{Y} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H, \tag{18}$$

where \mathbf{U} is the left singular vector, \mathbf{V} is the right singular vector, and $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_{M+N-1})$ is the singular value matrix with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{M+N-1}$. Therefore, the signal subspace \mathbf{V}_s is obtained by extracting the first P right singular vectors that correspond to $(\lambda_1, \lambda_2, \dots, \lambda_P)$. In addition, the last $M + N - 1 - P$ left singular vectors corresponding to $(\lambda_{P+1}, \lambda_{P+2}, \dots, \lambda_{M+N-1})$ are used to make up the noise subspace $\mathbf{U}_n \in \mathbb{C}^{(M+N-1) \times (M+N-1-P)}$. With $\mathbf{V}_s \in \mathbb{C}^{(M+N-1) \times P}$, \mathbf{Y} can be transformed into a dimension-reduced data matrix, that is

$$\mathbf{Y}_s = \mathbf{YV}_s = \mathbf{FBD}_s(\mathbf{F}^3\mathbf{B})^H\mathbf{V}_s, \tag{19}$$

where $\mathbf{Y}_s \in \mathbb{C}^{(M+N-1) \times P}$.

3.2. Designs of Joint Smoothed Function and Joint Smoothed l_0 -Norm Framework for MMV Case

In the sparse representation framework for DOA estimation, a complete dictionary containing all interest DOAs is required. Thus, let $\{\hat{\theta}_i\}_{i=1}^L$ be the discrete sample grid. According to Equation (19), the complete dictionary can be constructed as $\mathbf{B}_{\hat{\theta}} = [\mathbf{b}(\hat{\theta}_1), \mathbf{b}(\hat{\theta}_2), \dots, \mathbf{b}(\hat{\theta}_L)] \in \mathbb{C}^{(M+N-1) \times L}$. The DOA estimation is considered as a sparse signal reconstruction problem that is subject to

$$\mathbf{Y}_s = \mathbf{FB}_{\hat{\theta}}\mathbf{T}_{\hat{\theta}}, \tag{20}$$

where $\mathbf{T}_{\hat{\theta}} \in \mathbb{C}^{L \times P}$ is a sparse matrix and has the same row support with $\mathbf{T} = \mathbf{D}_s(\mathbf{F}^3\mathbf{B})^H\mathbf{V}_s$. Then, DOA estimation can be achieved by measuring the sparsity of $\mathbf{T}_{\hat{\theta}}$. Note that, in the sparse signal recovery problem in Equation (20), the low-dimensional data \mathbf{Y}_s is a matrix rather than a vector. Therefore, recovering $\mathbf{T}_{\hat{\theta}}$ involves the MMV problem. To develop a fast sparse DOA estimation algorithm that can solve the MMV problem, in the following, the proposed joint smoothed l_0 -norm algorithm designs

a new joint smoothed function and then derives its new steepest ascent scheme. For obtaining the smoothed estimation of the sparsity of $\mathbf{T}_{\hat{\theta}}$, we first design a joint continuous function as follows

$$f_{\sigma}[\mathbf{T}_{\hat{\theta}}(i, :)] = \exp\{-[|\mathbf{T}_{\hat{\theta}}(i, 1)|^2 + |\mathbf{T}_{\hat{\theta}}(i, 2)|^2 + \dots + |\mathbf{T}_{\hat{\theta}}(i, P)|^2]/2P\sigma^2\}, \quad (21)$$

where $\mathbf{T}_{\hat{\theta}}(i, p)$ is the (i, p) th element of $\mathbf{T}_{\hat{\theta}}$ for $p = 1, 2, \dots, P$ and $i = 1, 2, \dots, L$. Thus, the result of the designed joint continuous function $f_{\sigma}[\mathbf{T}_{\hat{\theta}}(i, :)]$ can be expressed as follows:

$$\lim_{\sigma \rightarrow 0} f_{\sigma}[\mathbf{T}_{\hat{\theta}}(i, :)] = \begin{cases} 1 & \mathbf{T}_{\hat{\theta}}^{l_2}(i) = 0 \\ 0, & \mathbf{T}_{\hat{\theta}}^{l_2}(i) \neq 0, \end{cases} \quad (22)$$

where $\mathbf{T}_{\hat{\theta}}^{(l_2)} \in \mathbb{C}^{L \times 1}$ is the $L \times 1$ dimensional vector, and $[\mathbf{T}_{\hat{\theta}}^{(l_2)}](i) = \|\mathbf{T}_{\hat{\theta}}(i, :)\|_2$. Namely, the i th element of $\mathbf{T}_{\hat{\theta}}^{(l_2)}$ is the l_2 -norm of the i th row in $\mathbf{T}_{\hat{\theta}}$. Then, a series of weights $[r_{w1}, r_{w2}, \dots, r_{wL}]$ are introduced into the joint continuous function $f_{\sigma}[\mathbf{T}_{\hat{\theta}}(i, :)]$. Therefore, in the proposed joint smoothed l_0 -norm algorithm, the final design of the joint smoothed function is

$$\begin{aligned} f_{w\sigma}[\mathbf{T}_{\hat{\theta}}(i, :)] &= \exp\left\{-\frac{r_{wi}}{2P\sigma^2} \sum_{p=1}^P [|\mathbf{T}_{\hat{\theta}}(i, p)|^2]\right\} \\ &= e^{-\frac{r_{wi}}{2P\sigma^2} \sum_{p=1}^P [|\mathbf{T}_{\hat{\theta}}(i, p)|^2]}, \end{aligned} \quad (23)$$

where r_{wi} is the weight coefficient, which is obtained from $r_{wi} = \|\mathbf{Fb}(\hat{\theta}_i)\|^H \mathbf{U}_n / r_{max}$ with $r_{max} = \max\{\|\mathbf{Fb}(\hat{\theta}_1)\|^H \mathbf{U}_n, \|\mathbf{Fb}(\hat{\theta}_2)\|^H \mathbf{U}_n, \dots, \|\mathbf{Fb}(\hat{\theta}_L)\|^H \mathbf{U}_n\}$. In Equation (23), the preset parameter σ is known and adjusts the smoothness of $f_{w\sigma}[\mathbf{T}_{\hat{\theta}}(i, :)]$. In addition, P stands for the total number of the columns in the data matrix \mathbf{Y}_s , namely, the column number of the sparse matrix $\mathbf{T}_{\hat{\theta}}$. For the true target DOA θ_p , $0 < r_{wp} \ll 1$ because of the orthogonality between the signal subspace and the noise subspace [5]. Therefore, the approximate value of the joint smoothed function $f_{w\sigma}[\mathbf{T}_{\hat{\theta}}(i, :)]$ corresponding to $\hat{\theta}_i$ can be calculated as follows:

$$f_{w\sigma}[\mathbf{T}_{\hat{\theta}}(i, :)] \approx \begin{cases} 1 & r_{wi} \rightarrow 1, \mathbf{T}_{\hat{\theta}}^{l_2}(i) \ll \sigma \\ 0 & r_{wi} \ll 1, \mathbf{T}_{\hat{\theta}}^{l_2}(i) \gg \sigma, \end{cases} \quad (24)$$

where a small σ is required to guarantee the approximation. Based on Equation (24), the sparsity of $\mathbf{T}_{\hat{\theta}}$ can be measured by

$$\|\mathbf{T}_{\hat{\theta}}^{l_2}\|_0 \approx L - \sum_{i=1}^L f_{w\sigma}[\mathbf{T}_{\hat{\theta}}(i, :)] = L - F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2}), \quad (25)$$

where $F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2}) = \sum_{i=1}^L f_{w\sigma}[\mathbf{T}_{\hat{\theta}}(i, :)]$. According to Equation (25), to recover the sparse matrix $\mathbf{T}_{\hat{\theta}}$, the minimization of $\|\mathbf{T}_{\hat{\theta}}^{l_2}\|_0$ can be turned into the maximization of $F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})$. As a consequence, the MMV-based joint smoothed l_0 -norm sparse representation framework is represented as

$$\max_{\mathbf{T}_{\hat{\theta}}} F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2}), \quad \text{s.t.} \quad \mathbf{Y}_s = \mathbf{FB}_{\hat{\theta}}\mathbf{T}_{\hat{\theta}}. \quad (26)$$

3.3. Design of the MMV-Based Signal Reconstruction in the Joint Smoothed l_0 -Norm Algorithm

For the joint smoothed function $F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})$, the steepest ascent based on the gradient is designed to recover the sparse signal. Aiming at the sparse signal reconstruction in Equation (26) that involves the MMV problem, an initial matrix is needed for the joint smoothed l_0 -norm algorithm to start the

steepest ascent process. The initial matrix $\mathbf{U}_0^{\hat{\theta}}$ can be chosen as the solution of Equation (26) with $\sigma \rightarrow \infty$, i.e., the minimum l_2 -norm solution calculated by

$$\mathbf{U}_0^{\hat{\theta}} = (\mathbf{FB}_{\hat{\theta}})^+ \mathbf{Y}_s \in \mathbb{C}^{(M+N-1) \times P}, \quad (27)$$

where $(\mathbf{FB}_{\hat{\theta}})^+$ is the pseudo-inverse of $\mathbf{FB}_{\hat{\theta}}$. When recovering the sparse signal for Equation (26), the selected value of σ has an effect on the computation time and the solution accuracy. The larger σ is, the smoother $F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})$ will be. Larger σ can result in less local maxima and makes it more easy to maximize $F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})$. In this case, the computation process can be speeded up. However, large σ causes the deviation of the approximation in Equation (25), and it will further decrease the precision of the signal reconstruction. To guarantee both of the speed and the accuracy, a decreasing sequence for σ , i.e., $[\sigma_1, \sigma_2, \dots, \sigma_K]$ with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_K$, which is known as the graduated non-convexity (GNC) approach [18,22], can be exploited. Even though large σ is desired to avoid a lot of local maxima and improve the speed, the value of σ_1 in the decreasing σ sequence may be set as one to four times of the maximum absolute value of $\mathbf{U}_0^{\hat{\theta}}$. This is because in $f_{w\sigma}[\mathbf{T}_{\hat{\theta}}(i, :)]$, when $\sigma > 4 \max\{|\mathbf{U}_0^{\hat{\theta}}|\}$, σ acts virtually like infinity for all elements with $e^{-\frac{1}{2p\sigma^2} \sum_{p=1}^P |\mathbf{T}_{\hat{\theta}}(i,p)|^2} > 0.96 \approx 1$. Then, the other values in the σ sequence can be chosen as $\sigma_{k+1} = \alpha \sigma_k$ for $k = 1, 2, \dots, K-1$. To make the steepest ascent algorithm be not far from the actual maximum, and be less likely to get trapped into local maxima, σ is required to change slowly. This guarantees that the GNC approach can work. In general, the decreasing factor α satisfies $0.5 \leq \alpha < 1$. σ_K should be a small value determined by the desired accuracy, and a parameter σ_{off} can be used to set the lower limit of the decreasing σ sequence.

For each σ_k in the decreasing σ sequence, graduated non-convexity based Q iterations are carried out to maximize $F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})$. In each iteration, $\mathbf{T}_{\hat{\theta}}$ is first updated as the following form:

$$\mathbf{T}_{\hat{\theta}} \leftarrow \mathbf{T}_{\hat{\theta}} + \tilde{\mu} \nabla F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2}) = \mathbf{T}_{\hat{\theta}} - \mu \Delta \mathbf{T}_{\hat{\theta}}, \quad (28)$$

where $\tilde{\mu} = \mu \sigma^2$ is a step-size parameter, and μ should be $\mu \geq 1$. In addition, $\nabla F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})$ is the gradient of $F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})$ corresponding to $\mathbf{T}_{\hat{\theta}}$. In the following, $\Delta \mathbf{T}_{\hat{\theta}}$ will be derived in detail. Let the detailed expression of $\mathbf{T}_{\hat{\theta}}$ be

$$\mathbf{T}_{\hat{\theta}} = [\mathbf{t}_1^T, \mathbf{t}_2^T, \dots, \mathbf{t}_L^T]^T, \quad (29)$$

where $\mathbf{t}_i \in \mathbb{C}^{1 \times P}$ is the i th row of $\mathbf{T}_{\hat{\theta}}$. According to Equation (28), \mathbf{t}_i can be written as

$$\mathbf{t}_i \leftarrow \mathbf{t}_i + \tilde{\mu} \nabla_{\mathbf{t}_i} F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2}) = \mathbf{t}_i - \mu \Delta \mathbf{t}_i, \quad (30)$$

where $\nabla_{\mathbf{t}_i} F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})$ stands for the $1 \times P$ dimensional gradient vector of $F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})$ in regard to $\mathbf{t}_i = [\mathbf{T}_{\hat{\theta}}(i, 1), \mathbf{T}_{\hat{\theta}}(i, 2), \dots, \mathbf{T}_{\hat{\theta}}(i, P)]$. The p th element of $\nabla_{\mathbf{t}_i} F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})$ is detailedly expressed as

$$\begin{aligned} \nabla_{\mathbf{t}_i} F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})(p) &= \partial F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2}) / \partial \mathbf{t}_{ip} \\ &= \partial \sum_{i=1}^L e^{-\frac{r_{wi}}{2p\sigma^2} \sum_{p=1}^P |\mathbf{T}_{\hat{\theta}}(i,p)|^2} / \partial \mathbf{T}_{\hat{\theta}}(i, p), \end{aligned} \quad (31)$$

where $\mathbf{t}_{ip} = \tilde{a} + \tilde{b}j$ is the p th parameter of \mathbf{t}_i with the real component \tilde{a} and imaginary component \tilde{b} . $\partial F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2}) / \partial \mathbf{t}_{ip} = [\partial F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2}) / \partial \tilde{a}] + [\partial F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2}) / \partial \tilde{b}]j$ denotes the 1st-order derivative of $F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})$. As a result, $\nabla_{\mathbf{t}_i} F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})(p)$ is derived as follows:

$$\begin{aligned}\nabla_{\mathbf{t}_i} F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})(p) &= e^{-\frac{r_{wi}}{2P\sigma^2} \sum_{p=1}^P \|\mathbf{T}_{\hat{\theta}}(i,p)\|^2} \partial \left\{ -\frac{r_{wi}}{2P\sigma^2} \sum_{p=1}^P \|\mathbf{T}_{\hat{\theta}}(i,p)\|^2 \right\} / \partial \mathbf{T}_{\hat{\theta}}(i,p) \\ &= -\frac{r_{wi} \mathbf{T}_{\hat{\theta}}(i,p)}{P\sigma^2} \exp \left\{ -\frac{r_{wi}}{2P\sigma^2} \sum_{p=1}^P \|\mathbf{T}_{\hat{\theta}}(i,p)\|^2 \right\}.\end{aligned}\quad (32)$$

Combining Equation (30) with Equation (32), the p th element of $\Delta \mathbf{t}_i$ can be obtained, that is,

$$\begin{aligned}\Delta \mathbf{t}_{ip} &= -\sigma^2 \nabla_{\mathbf{t}_i} F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})(p) \\ &= \frac{r_{wi} \mathbf{T}_{\hat{\theta}}(i,p)}{P} \exp \left\{ -\frac{r_{wi}}{2P\sigma^2} \sum_{p=1}^P \|\mathbf{T}_{\hat{\theta}}(i,p)\|^2 \right\}.\end{aligned}\quad (33)$$

With $\Delta \mathbf{t}_{ip}$, the $1 \times P$ dimensional vector $\Delta \mathbf{t}_i$ is written as

$$\begin{aligned}\Delta \mathbf{t}_i &= -\sigma^2 [\partial F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2}) / \partial \mathbf{T}_{\hat{\theta}}(i,1), \partial F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2}) / \partial \mathbf{T}_{\hat{\theta}}(i,2), \dots, \partial F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2}) / \partial \mathbf{T}_{\hat{\theta}}(i,P)] \\ &= [\Delta \mathbf{t}_{i1}, \Delta \mathbf{t}_{i2}, \dots, \Delta \mathbf{t}_{iP}]\end{aligned}\quad (34)$$

for $i = 1, 2, \dots, L$. As a result, based on Equations (33) and (34), the sparse matrix to be recovered in each iteration is first updated as

$$\mathbf{T}_{\hat{\theta}} \leftarrow \mathbf{T}_{\hat{\theta}} - \mu \Delta \mathbf{T}_{\hat{\theta}} \in \mathbb{C}^{L \times P}, \quad (35)$$

with

$$\Delta \mathbf{T}_{\hat{\theta}} = [\Delta \mathbf{t}_1^T, \Delta \mathbf{t}_2^T, \dots, \Delta \mathbf{t}_L^T]^T, \quad (36)$$

where $\Delta \mathbf{T}_{\hat{\theta}} \in \mathbb{C}^{L \times P}$. Then, this iteration projects $\mathbf{T}_{\hat{\theta}}$ back to the feasible set $\mathcal{T}_{\hat{\theta}} = \{\mathbf{T}_{\hat{\theta}} \mid \mathbf{Y}_s = \mathbf{F}\mathbf{B}_{\hat{\theta}}\mathbf{T}_{\hat{\theta}}\}$, which satisfies Equation (20), that is,

$$\mathbf{T}_{\hat{\theta}} \leftarrow \mathbf{T}_{\hat{\theta}} - (\mathbf{F}\mathbf{B}_{\hat{\theta}})^+ [(\mathbf{F}\mathbf{B}_{\hat{\theta}})\mathbf{T}_{\hat{\theta}} - \mathbf{Y}_s]. \quad (37)$$

Based on Equations (27)–(37), after Q iterations for $\sigma = \sigma_k$, the MMV based sparse signal reconstruction in Equation (26) starts the maximization process of the joint smoothed function $F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})$ for $\sigma = \sigma_{k+1}$. Let $\mathbf{T}_f^{\hat{\theta}}$ be the output of $\mathbf{T}_{\hat{\theta}}$ by maximizing $F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})$ for $\sigma = \sigma_k$. Consequently, the final sparse matrix is recovered as $\mathbf{T}_f^{\hat{\theta}}$. In order to evaluate the reconstruction results of each row in $\mathbf{T}_f^{\hat{\theta}}$, and then achieve the DOA estimation, let

$$\bar{l}_p = (\mathbf{T}_f^{\hat{\theta}})^{l_2}, \quad (38)$$

with $[(\mathbf{T}_f^{\hat{\theta}})^{l_2}](i) = \|\mathbf{T}_f^{\hat{\theta}}(i, :)\|_2$. Therefore, by searching the spectrum of \bar{l}_p , the target DOAs are obtained. The summary of the proposed joint smoothed l_0 -norm algorithm is listed below:

- Step 1: Compute the fourth-order cumulants based matrix \mathbf{Y}_s from Equations (11), (16) and (19).
- Step 2: Design the joint smoothed function $f_{w\sigma}[\mathbf{T}_{\hat{\theta}}(i, :)]$ tailored for the MMV case as Equation (23).
- Step 3: Construct the joint smoothed l_0 -norm sparse representation framework in Equation (26).
- Step 4: Execute the fast MMV-based sparse signal reconstruction with Equations (27)–(37).
- Step 5: Attain the DOA estimation based on Equation (38).

4. Related Remarks

Remark 1. With the joint smoothed function $F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})$, the proposed joint smoothed l_0 -norm minimization based sparse signal reconstruction algorithm is a new fast sparse DOA estimation algorithm, which is tailored for the MMV problem. By designing a specific joint continuous function that exploits all data information in the data matrix \mathbf{Y}_s , and then, deriving its steepest ascent process to achieve the maximization, the sparse solutions can be obtained. In the extended and improved smoothed l_0 -norm algorithms, the proposed joint continuous function $f_{\sigma}[\mathbf{T}_{\hat{\theta}}(i, :)]$ in Equation (21) is applicable for all MMV cases with the form $\mathbf{Y}_s = \hat{\mathbf{B}}_{\hat{\theta}}\mathbf{T}_{\hat{\theta}}$, where $\hat{\mathbf{B}}_{\hat{\theta}}$ is the

complete dictionary and $\mathbf{T}_{\hat{\theta}}$ is the $L \times P$ dimensional sparse matrix that needs to be reconstructed. Under this circumstance, gradient-based maximizing $\sum_{i=1}^L f_{\sigma}[\mathbf{T}_{\hat{\theta}}(i, :)]$ is similar to the derivation of Equations (27)–(37), and its $\Delta \mathbf{t}_{ip}$ corresponding to $\sum_{i=1}^L f_{\sigma}[\mathbf{T}_{\hat{\theta}}(i, :)]$ is concluded as

$$\Delta \mathbf{t}_{ip} = \frac{\mathbf{T}_{\hat{\theta}}(i, p)}{P} \exp\left\{-\frac{1}{2P\sigma^2} \sum_{p=1}^P [|\mathbf{T}_{\hat{\theta}}(i, p)|^2]\right\} \quad (39)$$

for $i = 1, 2, \dots, L$. With $\Delta \mathbf{t}_{ip}$ in Equation (39), the proposed joint smoothed l_0 -norm algorithm regarding the joint continuous function $f_{\sigma}[\mathbf{T}_{\hat{\theta}}(i, :)]$ can be applied to other fast sparse signal reconstruction problems that involve multiple measurement vectors.

Remark 2. The proposed method has low computational complexity because its sparse signal recovery is a graduated non-convexity procedure, rather than the convex optimization problem involved in the l_1 -norm minimization. When recovering the sparse signal matrix $\mathbf{T}_{\hat{\theta}}$, the main computational burden is caused by the QK iterations, which require about $O[PLQK + (M + N - 1)PLQK]$, where L is the number of the sample grid cells in the complete dictionary $\mathbf{B}_{\hat{\theta}}$. For the SMV based reweighted smoothed l_0 -norm algorithm in [22], recovering the sparse signal vector requires about $O[(M + N - 1)^2LQK]$. However, l_1 -norm minimization based DOA estimation methods, such as l_1 -SVD and RV l_1 -SVD, require $O[(PL)^3]$ and $O[\frac{1}{4}(PL)^3]$, respectively. In addition, the RV l_1 -SRACV needs $O[\frac{1}{4}L^3]$. As L is much larger than P , $M + N - 1$, Q and K , the computation speed of the proposed joint smoothed l_0 -norm algorithm is much faster than the l_1 -norm minimization based methods.

Remark 3. The proposed method can deal with both of the white and colored Gaussian noises, thereby making it desirable for use in practical applications. This is because, in the proposed joint smoothed l_0 -norm algorithm, high-order cumulants are utilized to process the reduced-dimensional data $\mathbf{x}(t)$. Note that, for the dimension reduction in Equation (11), the reduced-dimensional matrix \mathbf{R} satisfies $\mathbf{R}\mathbf{R}^H = \mathbf{I}_{M+N-1}$. Hence, if the noise vector $\tilde{\mathbf{n}}(t)$ is white Gaussian with covariance matrix $\rho^2\mathbf{I}_{MN}$, it remains white Gaussian with covariance matrix $\mathbf{R}\rho^2\mathbf{I}_{MN}\mathbf{R}^H = \rho^2\mathbf{I}_{M+N-1}$ after the transformation [2,25]. On the other hand, if the noise vector $\mathbf{n}(t)$ is colored Gaussian with covariance matrix \mathbf{V}_n , its covariance matrix turns into $\mathbf{R}\mathbf{V}_n\mathbf{R}^H$ after the dimensional reduction. However, in this case, $\tilde{\mathbf{n}}(t)$ is still Gaussian according to the invariance speciality of linear transformation in asymptotic normal distribution. Therefore, even though the dimensional reduction makes the change of the noise happen, the high-order cumulant process in Equation (16) can eliminate it.

5. Simulation Results

In this part, some simulations are implemented to evaluate the performance and the computation speed of the proposed method. The simulation results of the proposed method are compared with those of the l_1 -SVD [7], RV l_1 -SVD [16], RV l_1 -SRACV [14] and reweighted smoothed l_0 -norm (RSL0) [22] algorithms. In all of the simulations, a narrowband monostatic MIMO radar system with ULAs is considered. The antennas of the transmit array and the receive array are both half-wavelength spaced, and their numbers are $M = 6$ and $N = 6$, respectively. The transmitted signals are assumed to be BPSK modulated [27,30], moreover, they are mutually orthogonal and different from each other. The noise is white Gaussian with the covariance matrix $\mathbf{R} = \mathbf{I}_{MN}$, or colored Gaussian with the elements in the covariance matrix being $\mathbf{R}(k_1, k_2) = 0.75^{|k_1 - k_2|} e^{j\pi(k_1 - k_2)/2}$. By means of the minimum description length (MDL) principle or the Akaike information criterion (AIC) principle [31], the prior knowledge of the target number is assumed to be known as P .

When constructing the complete dictionary, the discrete sample grid is set from -90° to 90° with the uniform interval distance $\Delta\hat{\theta} = 0.05^\circ$ for the proposed algorithm, as well as the l_1 -SVD, RV l_1 -SVD, RV l_1 -SRACV and RSL0 algorithms. The parameters σ_1 , μ and σ_{off} are fixedly set at $\sigma_1 = 4\max\{|\mathbf{U}_0^{\hat{\theta}}|\}$, $\mu = 2.5$ and $\sigma_{off} = 0.0007$ for the proposed algorithm and the RSL0 algorithm. The effects of the other parameters, i.e., the decreasing factor α and the iteration number Q , will be evaluated and given in

detail in the following experiments. The definitions for the signal-to-noise ratio (SNR) and the root mean square error (RMSE) are shown, respectively, as follows:

$$\text{SNR} = 10\log_{10}(\|\mathbf{AS}\|_F^2/\|\mathbf{N}\|_F^2),$$

$$\text{RMSE} = \sqrt{\frac{1}{500P} \sum_{i=1}^{500} \sum_{p=1}^P (\hat{\theta}_{p,i} - \theta_p)^2}, \quad (40)$$

where $\hat{\theta}_{p,i}$ is the angle estimation of the p th target DOA θ_p for the i th Monte Carlo trial.

Figure 2 depicts the spatial spectrum of the proposed method for white noise and colored noise, where SNR = 5 dB and SNR = −10 dB are considered. There are three targets located at $\theta_1 = -8.2^\circ$, $\theta_2 = 0^\circ$ and $\theta_3 = 21^\circ$. The spatial spectrum is plotted by computing $10\log_{10}[|\bar{\ell}_p|/\max(|\bar{\ell}_p|)]$ with $\bar{\ell}_p$ being the final sparse solution. Figure 2 shows that the proposed joint smoothed l_0 -norm minimization algorithm is effective for the MMV case of DOA estimation in monostatic MIMO radar. Moreover, it is suitable for the practical radar systems with colored Gaussian noise.

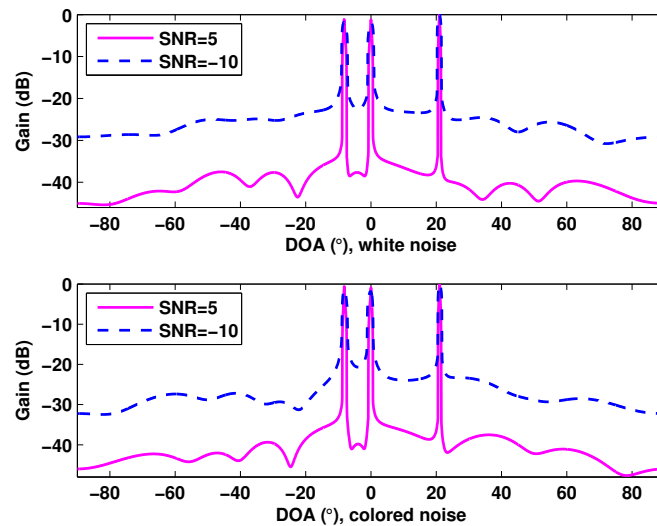


Figure 2. Spatial spectrum of the proposed method for white noise and colored noise.

Figure 3a,b evaluates the effects of the parameters α and Q on the DOA estimation performance of the proposed method, with $Q = 3$ and $\alpha = 0.5$, respectively. Moreover, $J = 1000$. Three uncorrelated targets with the DOAs $\theta_1 = -8.2^\circ$, $\theta_2 = 0^\circ$ and $\theta_3 = 21^\circ$ are considered. In Figure 3, it can be verified that the proposed method keeps the same angle estimation performance when the parameters α and Q vary from $\alpha = 0.5$ to $\alpha = 0.9$ and $Q = 3$ to $Q = 35$. In Figure 3b, the RMSE decreases with the increase of Q when $Q \leq 3$, and it reaches an asymptotic value at $Q = 3$. For the iterations of the steepest ascent algorithm, we just need a region near the maximum value of $F_{w\sigma}(\mathbf{T}_{\hat{\theta}}^{l_2})$; thus, a fixed small positive integer $Q = 3$ is applicable.

With respect to Figure 3, the effects of the parameters α and Q on the average computation time of the proposed method are shown in Tables 1 and 2. The simulation conditions keep the same with those of Figure 3. As the implementing processes of the analyzed methods are not affected by the type of the noise, the computation time is the same for the white and the colored Gaussian noises. It can be observed that the computation speed becomes slower when α and Q increase. However, larger α and Q are not able to improve the estimation performance of the proposed method, as shown in Figure 3a,b. As a result, small values of α and Q are enough to guarantee the estimation accuracy and the computation speed. In the following experiments, the parameters α and Q are set at $\alpha = 0.5$ and $Q = 3$.

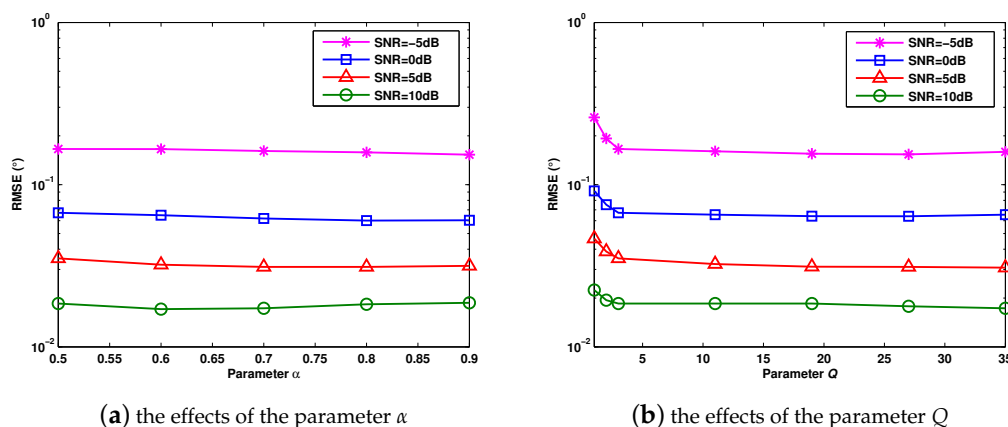


Figure 3. RMSE (root mean square error) versus different values of the parameters α and Q .

Table 1. Average computation time for different values of the parameter α .

| α | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|------------------------------|--------|--------|--------|--------|--------|
| Average Computation Time (s) | 0.0401 | 0.0498 | 0.0752 | 0.1137 | 0.2511 |

Table 2. Average computation time for different values of the parameter Q .

| Q | 3 | 11 | 19 | 27 | 35 |
|------------------------------|--------|--------|--------|--------|--------|
| Average Computation Time (s) | 0.0399 | 0.1382 | 0.2576 | 0.3696 | 0.4738 |

Figure 4a,b verify the RMSE of DOA estimation versus SNR for white noise and colored noise, where $J = 1000$. Consider three targets whose DOAs are $\theta_1 = -8.2^\circ$, $\theta_2 = 0^\circ$ and $\theta_3 = 21^\circ$. According to Figure 4a,b, we can conclude that the RSL0 algorithm achieves the best DOA estimation for white noise. For colored noise, the proposed algorithm and the RSL0 algorithm provide better angle estimation performance than the l_1 -SVD, RV l_1 -SVD and RV l_1 -SRACV methods.

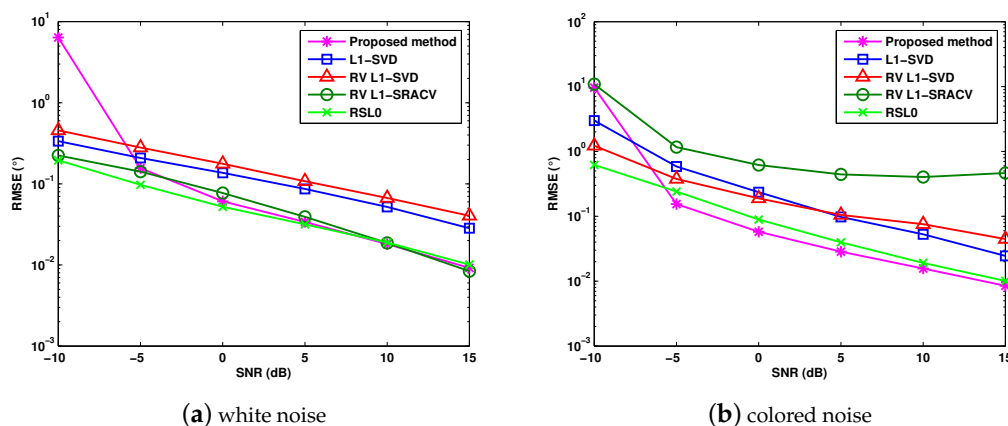


Figure 4. RMSE versus SNR (signal-to-noise ratio) for different sparse DOA (direction-of-arrival) estimation algorithms.

The processing time of signal reconstruction is compared in Table 3, where $J = 1000$ and SNR = 0 dB. Different antenna numbers and target numbers are considered. For $p = 2$, the DOAs

are $\theta_2 = 0^\circ$ and $\theta_3 = 21^\circ$. For $p = 3$, they are $\theta_1 = -8.2^\circ$, $\theta_2 = 0^\circ$ and $\theta_3 = 21^\circ$. Table 3 verifies that, compared with the l_1 -norm minimization based methods, the proposed algorithm has much lower computational complexity and is about two orders of magnitude faster.

Table 3. Average computation time of the signal reconstruction for different algorithms.

| (M,N,P) | Average Computation Time (s) | | | | |
|---------|------------------------------|------------|---------------|-----------------|--------|
| | Proposed Method | l_1 -SVD | RV l_1 -SVD | RV l_1 -SRACV | RSL0 |
| (4,4,2) | 0.0269 | 2.1067 | 1.5499 | 1.3213 | 0.0135 |
| (4,4,3) | 0.0323 | 2.5139 | 2.0591 | 1.3368 | 0.0139 |
| (5,5,2) | 0.0273 | 2.4984 | 2.1071 | 2.1277 | 0.0252 |
| (5,5,3) | 0.0373 | 3.1170 | 2.7076 | 2.3842 | 0.0254 |
| (6,6,2) | 0.0328 | 2.6323 | 2.2911 | 3.2908 | 0.0438 |
| (6,6,3) | 0.0395 | 3.5439 | 3.1187 | 4.7397 | 0.0441 |

SVD (singular value decomposition), RV (real-valued), SRACV (sparse representation array covariance vectors), RSL0 (reweighted smoothed l_0 -norm).

Figure 5a,b illustrates the RMSE versus snapshots for Gaussian white noise and colored noise, where SNR = 0 dB, and the DOAs of three targets are $\theta_1 = -8.2^\circ$, $\theta_2 = 0^\circ$ and $\theta_3 = 21^\circ$. It can be observed that, for white noise, the RSL0 algorithm has the minimum RMSE. For colored noise, the proposed algorithm provides the best angle estimation performance when the snapshots vary from $J = 450$ to $J = 4050$, whereas, in this case, the RV l_1 -SRACV method performs worse than the other algorithms.

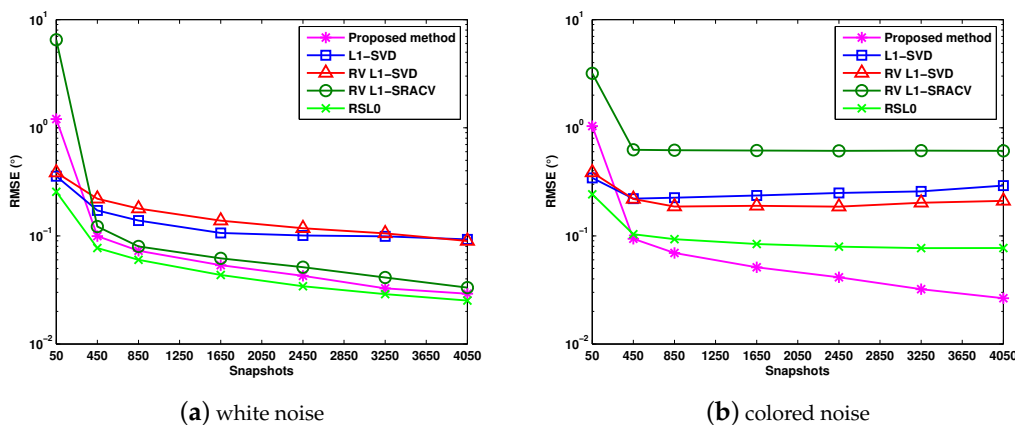


Figure 5. RMSE versus snapshots with SNR = 0 dB.

Figure 6a,b verifies the DOA estimation performance of different sparse methods versus angle separation for Gaussian white noise and colored noise, where $J = 1000$, and the DOAs of two targets are considered as $\theta_1 = 0^\circ$ and $\theta_2 = \theta_1 + \tilde{\theta}$ with $\tilde{\theta}$ varying from $\tilde{\theta} = 4^\circ$ to $\tilde{\theta} = 16^\circ$. For the white noise in Figure 6a, the RSL0 algorithm in [22] performs the best, and the proposed algorithm provides better estimation performance than the l_1 -SVD, RV l_1 -SVD and RV l_1 -SRACV methods. For the colored noise in Figure 6b, the performance of the proposed algorithm stands out among the analyzed methods, while the performance of the RV l_1 -SRACV is severely degraded.

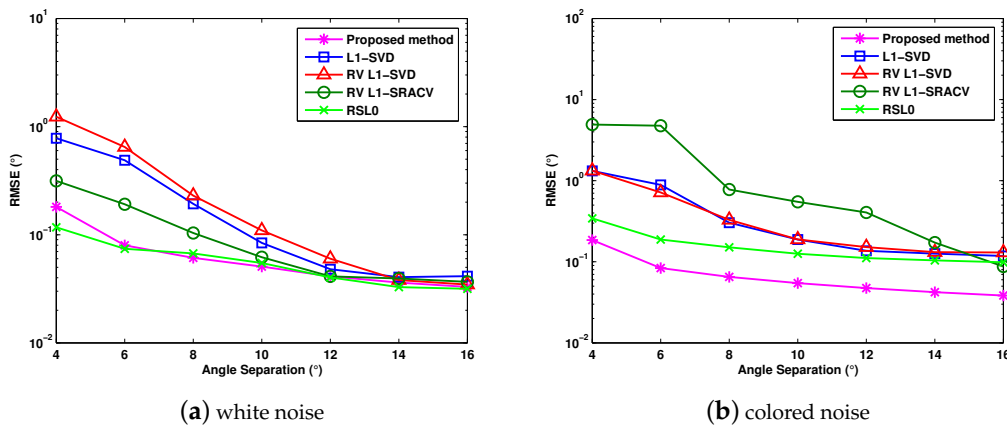


Figure 6. RMSE versus angle separation with SNR = 0 dB.

Figure 7a,b verifies the relationships between the target resolution probability and SNR for different sparse methods, where $J = 1000$, and three targets $\theta_1 = -8.2^\circ$, $\theta_2 = 0^\circ$ and $\theta_3 = 21^\circ$ are considered. The i th Monte Carlo trial is regarded as a successful one if the estimation results $\bar{\theta}_1$, $\bar{\theta}_2$ and $\bar{\theta}_3$ satisfy $\max\{|\bar{\theta}_1 - \theta_1|, |\bar{\theta}_2 - \theta_2|, |\bar{\theta}_3 - \theta_3|\} \leq 0.1^\circ$. Figure 7a,b verifies that the RSL0 algorithm performs the best for the white noise, and the proposed algorithm provides the highest target resolution probability for the colored noise. Colored noise causes the serious performance degradation for the analyzed methods except for the proposed algorithm.

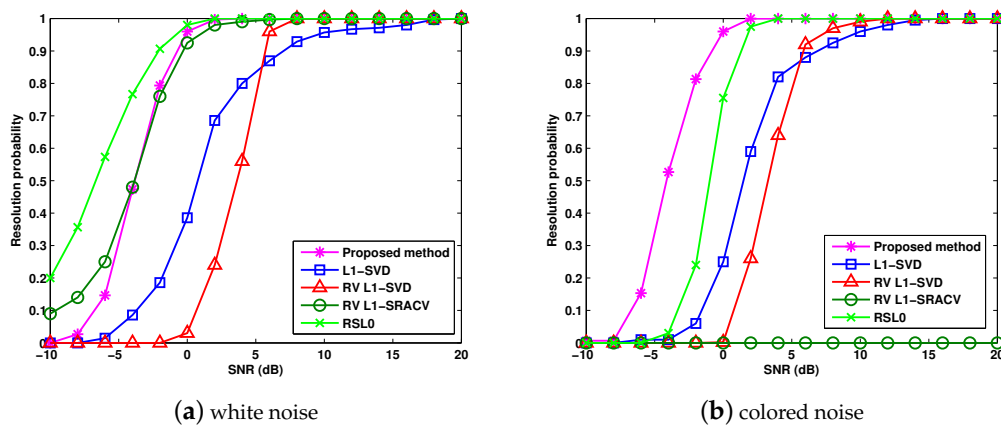


Figure 7. RMSE versus target resolution probability.

6. Conclusions

Direction-of-arrival estimation is usually confronted with the MMV case. In this paper, for DOA estimation in MIMO radar, we have proposed a joint smoothed l_0 -norm algorithm tailored for the multiple measurement vector case. By designing the new joint smoothed function and its gradient-based sparse signal reconstruction, the proposed method is a fast sparse DOA estimation algorithm that can solve the MMV problem. In addition, it is applicable to the colored Gaussian noise. The extended applications and the computational complexity of the proposed method are analyzed. The experimental results have verified that the proposed algorithm is about two orders of magnitude faster than the l_1 -norm minimization based methods, such as l_1 -SVD, RV l_1 -SVD and RV l_1 -SRACV, and performs well for both white and colored Gaussian noises.

Acknowledgments: This work was supported by the China Scholarship Council and the Fundamental Research Funds for the Central Universities (HEUCF160401).

Author Contributions: Jing Liu and Weidong Zhou provided the idea of this work and wrote the paper. Filbert H. Juwono assisted with performing the experiments and analyzing the results.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this paper:

| | |
|------|--------------------------------|
| MIMO | Multiple-Input Multiple-Output |
| DOA | Direction-of-Arrival |
| FOC | Fourth-Order Cumulants |
| MMV | Multiple Measurement Vector |
| SMV | Single Measurement Vector |
| SVD | Singular Value Decomposition |
| ACV | Array Covariance Vectors |
| SR | Sparse Representation |
| RV | Real-Valued |

References

- Li, J.; Stoica, P. MIMO radar with colocated antennas. *IEEE Signal Process. Mag.* **2007**, *24*, 106–114.
- Zhou, W.D.; Liu, J.; Zhu, P.; Gong, W.; Hou, J. Noncircular sources-based sparse representation algorithm for direction of arrival estimation in MIMO radar with mutual coupling. *Algorithms* **2016**, *9*, 61.
- Gao, X.; Zhang, X.F.; Feng, G.P.; Wang, Z.Q.; Xu, D.Z. On the MUSIC-derived approaches of angle estimation for bistatic MIMO radar. In Proceedings of the 2009 International Conference on Wireless Networks and Information Systems, Shanghai, China, 28–29 December 2009.
- Chen, D.; Chen, B.; Qin, G. Angle estimation using ESPRIT in MIMO radar. *Electron. Lett.* **2008**, *44*, 770–771.
- Nie, W.K.; Feng, D.Z.; Xie, H.; Li, J.; Xu, P.F. Improved MUSIC algorithm for high resolution angle estimation. *Signal Process.* **2016**, *122*, 87–92.
- Qian, C.; Huang, L.; So, H.C. Computationally efficient ESPRIT algorithm for direction-of-arrival estimation based on Nyström method. *Signal Process.* **2014**, *94*, 74–80.
- Malioutov, D.; Cetin, M.; Willsky, A.S. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE Trans. Signal Process.* **2005**, *53*, 3010–3022.
- Gorodnitsky, I.F.; Rao, B.D. Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm. *IEEE Trans. Signal Process.* **1997**, *45*, 600–616.
- Wipf, D.P.; Rao, B.D. Sparse Bayesian learning for basis selection. *IEEE Trans. Signal Process.* **2004**, *52*, 2153–2164.
- Liu, Z.M.; Huang, Z.T.; Zhou, Y.Y. Array signal processing via sparsity-inducing representation of the array covariance matrix. *IEEE Trans. Aerosp. Electron. Syst.* **2013**, *49*, 1710–1724.
- Cotter, S.F.; Rao, B.D.; Engan, k.; Kreutz-Delgado, K. Sparse solutions to linear inverse problems with multiple measurement vectors. *IEEE Trans. Signal Process.* **2005**, *53*, 2477–2488.
- Li, S.; Jiang, X.X.; He, W.; Wang, Y.G. Direction of arrival estimation via sparse representation of fourth order statistics. In Proceedings of the 2013 IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC), KunMing, China, 5–8 August 2013.
- Yin, J.; Chen, T. Direction-of-arrival estimation using a sparse representation of array covariance vectors. *IEEE Trans. Signal Process.* **2011**, *59*, 4489–4493.
- Wang, X.P.; Wang, W.; Li, X.; Liu, J. Real-valued covariance vector sparsity-inducing DOA estimation for monostatic MIMO radar. *Sensors* **2015**, *15*, 28271–28286.
- Cands, E.J.; Wakin, M.B.; Boyd, S.P. Enhancing sparsity by reweighted L_1 minimization. *J. Fourier Anal. Appl.* **2008**, *14*, 877–905.
- Dai, J.; Xu, X.; Zhao, D. Direction-of-arrival estimation via real-valued sparse representation. *IEEE Antennas Wirel. Propag. Lett.* **2013**, *12*, 376–379.

17. Wang, X.P.; Wang, W.; Liu, J.; Li, X.; Wang, J.X. A sparse representation scheme for angle estimation in monostatic MIMO radar. *Signal Process.* **2014**, *104*, 258–263.
18. Mohimani, G.H.; Babaie-Zadeh, M.; Jutten, C. Complex-valued sparse representation based on smoothed l^0 norm. In Proceedings of the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, NV, USA, 31 March–4 April 2008.
19. Mohimani, G.H.; Babaie-Zadeh, M.; Jutten, C. A fast approach for overcomplete sparse decomposition based on smoothed l^0 norm. *IEEE Trans. Signal Process.* **2009**, *57*, 289–301.
20. Bu, H.X.; Tao, R.; Bai, X.; Zhao, J. Regularized smoothed l_0 norm algorithm and its application to CS-based radar imaging. *Signal Process.* **2016**, *122*, 115–122.
21. Juwono, F.H.; Guo, Q.H.; Huang, D.F.; Wong, K.P.; Xu, L. Impulsive noise detection in PLC with smoothed l_0 -norm. In Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brisbane, Australia, 19–24 April 2015.
22. Liu, J.; Zhou, W.D.; Juwono, F.H.; Huang, D.F. Reweighted smoothed l_0 -norm based DOA estimation for MIMO radar. *Signal Process.* **2017**, *137*, 44–51.
23. Lie, S.; Leyman, A.R.; Chew, Y.H. Fourth-order and weighted mixed order direction-of-arrival estimators. *IEEE Signal Process. Lett.* **2006**, *13*, 691–694.
24. Liu, C.; Ye, Z.F.; Zhang, Y.F. DOA estimation based on fourth-order cumulants with unknown mutual coupling. *Signal Process.* **2009**, *89*, 1839–1843.
25. Liu, J.; Zhou, W.D.; Wang, X.P. Fourth-order cumulants-based sparse representation approach for DOA estimation in MIMO radar with unknown mutual coupling. *Signal Process.* **2016**, *128*, 123–130.
26. Liu, J.; Wang, X.P.; Zhou, W.D. Covariance vector sparsity-aware DOA estimation for monostatic MIMO radar with unknown mutual coupling. *Signal Process.* **2016**, *119*, 21–27.
27. Bencheikh, M.L.; Wang, Y. Non circular ESPRIT-RootMUSIC joint DOA-DOD estimation in bistatic MIMO radar. In Proceedings of the 7th International Workshop on Systems, Signal Processing and their Applications (WOSSPA), Tipaza, Algeria, 9–11 May 2011.
28. Wang, X.P.; Wang, W.; Li, X.; Liu, Q.; Liu, J. Sparsity-aware DOA estimation scheme for noncircular source in MIMO radar. *Sensors* **2016**, *16*, 539.
29. Yan, H.; Li, J.; Liao, G. Multitarget identification and localization using bistatic MIMO radar systems. *EURASIP J. Adv. Signal Process.* **2008**, doi:10.1155/2008/283483.
30. Zhang, M.; Liu, L.T.; Diao, M. LPI radar waveform recognition based on time-frequency distribution. *Sensors* **2016**, *16*, 1682.
31. Wax, M.; Kailath, T. Detection of signals by information theoretic criteria. *IEEE Trans. Acoust. Speech Signal Process.* **1985**, *33*, 387–392.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).