Adaptive Widely-Linear Constrained Constant Modulus Reduced-Rank Beamforming

Xiaomin Wu, Yunlong Cai, Rodrigo C. de Lamare,
Benoit Champagne, and Minjian Zhao

***Abstract*—We propose a blind adaptive reduced-rank widely linear beamforming algorithm using the generalized sidelobe canceller structure for interference suppression. A structured Krylov-subspace based approach is devised to construct the dimensionality reducing transformation matrix and a recursive least squares algorithm is developed according to the widely linear constrained constant modulus (CCM) criterion to update the reduced-rank filter. We analyze the convergence and complexity of the proposed algorithm and validate its performance gains through simulations.**

***Index Terms*—Adaptive beamforming, widely linear techniques, reduced-rank techniques, constrained constant modulus.**

I. Introduction

Large-scale antenna array systems have been receiving significant attention for future aeronautical applications and communications [1]-[3]. Aerial vehicles or platforms equipped with large-scale antenna arrays can dramatically increase the system capacity and improve the quality and reliability of wireless links. The increased computational complexity associated with large-dimensional received data vectors motivates the use of efficient adaptive beamforming techniques, which are among the most commonly used approaches to continually adapt the beamformer weights for detecting a desired signal, while coping with changes in the radio signal environment and reducing computational complexity [4]-[6]. Blind algorithms, which can work without any requirements for training symbols, can further improve the information rates and the efficiency of communication systems employing antenna arrays. The most popular design criteria for adaptive blind beamformers are the constrained minimum variance (CMV) [7]-[9] and the constrained constant modulus (CCM) [10]-[14] due to their effectiveness and simplicity. The CMV-based algorithms are designed with the aim of minimizing the filter output power while maintaining a constant response in the direction of the signal of interest. The CCM-based algorithms, which attempt to minimize the mean deviation of the squared output from constant values, exploit additional information about the underlying signal constellation, and can therefore achieve superior performance as compared with the CMV techniques.

Recently, some researchers have proposed new robust beamforming algorithms for non-Gaussian signals, where the l_p norm $(p \geq 1)$ of the output is minimized while constraining the magnitude response of any steering vector within a specific uncertainty set [15], [16]. In many situations of interest, the received data vector $\mathbf{r}$ at the array output[1] is assumed to be second-order circular with rotation invariant probability distribution. Consequently, the complementary covariance matrix $\mathbf{R}_c = \mathbb{E}\{\mathbf{r}\mathbf{r}^T\}$ equals zero, and for this reason, only the covariance matrix $\mathbf{R} = \mathbb{E}\{\mathbf{r}\mathbf{r}^H\}$ is utilized in conventional schemes. However, this ideal and general assumption may not be satisfied in practice since non-circularities in the observed data may arise from many sources, such as: the difference in signal powers between the real and imaginary parts, correlation between the real and imaginary parts, or the particular structure of the signal constellation used for digital transmission [21]. Typically, when the received data vector $\boldsymbol{r}$ contains noncircular modulated signals, such as binary phase shift keying (BPSK) modulation, $\mathbf{R}_c$ is no longer a zero matrix. Under such circumstances, a more general estimation scheme, which takes into consideration both the received vector $\mathbf{r}$ and its complex conjugate $\mathbf{r}^*$, is needed to fully exploit the second-order statistics of the data. Referred to as widely linear (WL) beamformer, this more general scheme can lead to higher signal-to-interference-plus noise ratio (SINR) or smaller mean square error (MSE) in the estimation of a desired signal [17]-[23].

However, one problem for the standard, i.e., full-rank adaptive algorithms is that their convergence performance deteriorates rapidly with an increase in the eigenvalue spread of the received data covariance matrix, as measured by its condition number [24]-[27]. This situation is usually worse in a large-scale antenna array system with numerous filter coefficients to be estimated. In this context, reduced-rank signal processing has become a key technique to provide faster convergence and increased robustness against interference as compared to standard methods. In this approach, the large-dimensional received data vector is projected onto a lower dimensional subspace with the aid of a transformation matrix

[1]For one data snapshot after demodulation and sampling.

X. Wu, Y. Cai, and M. Zhao are with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: wxm075910@zju.edu.cn; ylcai@zju.edu.cn; mjzhao@zju.edu.cn).

R. C. de Lamare is with CETUC-PUC-Rio, 22453-900 Rio de Janeiro, Brazil, and also with the Communications Research Group, Department of Electronics, University of York, YO10 5DD York, U.K. (e-mail: r-cdl500@ohm.york.ac.uk).

B. Champagne is with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada, H3A 0E9 (e-mail: benoit.champagne@mcgill.ca).

This work was supported in part by the National Natural Science Foundation of China under Grant 61471319, Zhejiang Provincial Natural Science Foundation of China under Grant LY14F010013, the Fundamental Research Funds for the Central Universities, and the National High Technology Research and Development Program (863 Program) of China under Grant 2014AA01A707.

and a reduced-rank filter is then designed to process the low-dimensional data vector within this subspace. We note that WL processing doubles the size of the received data vector which further motivates the use of reduced-rank techniques [27]-[30]. Various WL-based reduced-rank beamforming algorithms have been introduced in previous studies, including the eigen-decomposition method [22], the multi-stage Wiener filter (MSWF) [31], and the auxiliary vector filtering (AVF) [32]. In particular, Song *et al.* [31] designed a WL-MSWF reduced-rank algorithm according to the minimum mean squared error (MMSE) criterion to suppress the interference in a high data-rate direct-sequence ultra wideband (DS-UWB) system. The MSWF-based adaptive WL processing algorithm is implemented with the aid of training symbols. In [32], a non-data-aided adaptive beamforming algorithm based on WL processing techniques and the AVF algorithm was developed for non-circular signals. The WL-AVF algorithm recursively updates the filter weights by a sequence of auxiliary vectors that are designed according to the widely-linearly constrained minimum variance (WLCMV) criterion. Both the MSWF and AVF methods involve the construction of a low rank Krylov-subspace (KS) for the purpose of rank-reduction, which has led to excellent performance in several applications and can be combined with different design criteria. Besides these recent studies investigating WL reduced-rank techniques, to the best of our knowledge, there has been no work focusing on the design of adaptive reduced-rank beamforming algorithms using the widely linearly constrained constant modulus (WLCCM) criterion.

In this work, a novel blind adaptive reduced-rank WL beamforming algorithm based on the KS technique is proposed for interference suppression in large-scale antenna array systems. The proposed beamforming algorithm operates in the generalized sidelobe canceller (GSC) structure[2] where the KS technique is employed for rank reduction [33]. To reduce the computational complexity of the conventional WL realization scheme based on the stacking of the received data and its complex conjugate, the structure of the augmented covariance matrix is taken into consideration as prior information to devise a structured KS-based transformation matrix. We then develop a recursive least squares (RLS) algorithm based on the WLCCM criterion to update the reduced-rank filter. We refer to the new blind adaptive reduced-rank beamforming algorithm so obtained as the WLCCM-KS. A theoretical performance analysis in terms of convergence behavior and achievable SINR for the proposed algorithm and its linear counterpart is provided. In addition, we investigate the computational complexity of the proposed algorithm and compare it with that of other existing reduced-rank algorithms. Simulation results verify the analytical results and show that the proposed WLCCM-KS algorithm outperforms its linear counterpart as well as the full-rank algorithms, achieving the best convergence performance and steady-state SINR among all the analyzed methods with a relatively low complexity.

The remainder of this paper is structured as follows. Section II briefly describes the system model and the problem statement, while the KS based reduced-rank scheme with the GSC structure is introduced in Section III. In Section IV, we develop the proposed WLCCM-KS algorithm and provide the computational complexity analysis. The performance analysis in terms of convergence properties and achievable SINR for the proposed algorithm is conducted in Section V. The supporting simulation results and their discussion are presented in Section VI. Finally, Section VII draws the conclusions.

II. System Model and Problem Statement

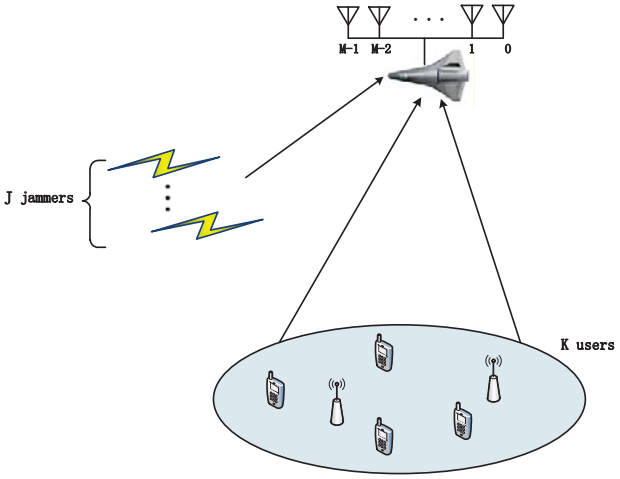

Fig. 1: The aeronautical communication system model.

We consider a wireless communication scenario as shown in Fig. 1, where K user signals and J jamming signals impinge on an aerial vehicle or platform which is equipped with a uniform linear array (ULA) comprised of M identical omnidirectional antenna elements[3], and M is a fairly large number with $K + J < M$. Our interest is focused on communication problems where a reasonably large antenna array is used to extract the signal of a desired user located along a known direction of arrival (DOA).

The $M \times 1$ sampled array output vector (or snapshot) at discrete time $i \in \{0, 1, 2, \ldots\}$, can be modeled as

$$\mathbf{r}(i) = \sum_{k=0}^{K-1} b_k(i)\mathbf{a}(\theta_k) + \sum_{j=0}^{J-1} c_j(i)\mathbf{a}(\phi_j) + \mathbf{n}(i). \quad (1)$$

In this expression, we assume that the sequence of transmitted signals by the kth user, i.e., $\{b_k(i)\}$, contains independent and identically distributed (i.i.d) random variables with zero mean drawn from a given symbol set with constant modulus. The quantity $c_j(i)$ denotes the jth jammer signal at snapshot i, which is typically assumed to be an i.i.d. sequence of

[2]The GSC structure uses a main branch along with an array of auxiliary branches. While interference may be present in both the main and auxiliary branches, the desired user signal is mostly present in the main branch due to its high directional gain and the use of a so-called signal blocking matrix along the auxiliary branches. Consequently, the latter can be used to form an estimate of the main branch interference that can be used for cancelation [33]. For these reasons, the use of the GSC structure is often preferred in applications.

[3]ULA is considered for the sake of simplifying the presentation, while generalization to other antenna configurations is straightforward.

Gaussian variables with zero mean and variance σ_j^2. The vectors $\mathbf{a}(\theta_k)$ and $\mathbf{a}(\phi_j)$ denote the $M \times 1$ steering vectors of the user and jamming signals, with respective DOAs θ_k and ϕ_j. The term $\mathbf{n}(i) \in \mathbb{C}^{M\times 1}$ is an additive noise vector, which is modeled as an i.i.d. sequence of spatially white Gaussian random vectors with zero-mean and covariance matrix $\mathbb{E}\{\mathbf{n}(i)\mathbf{n}^H(i)\} = \sigma_n^2\mathbf{I}_M$, where σ_n^2 denotes the noise variance, $\mathbf{I}_M$ is an identity matrix of order M, and $(.)^H$ stands for the Hermitian transpose operation. The random sequences $\{b_k(i)\}, \{c_j(i)\}$ and $\{\mathbf{n}(i)\}$ are mutually independent. Let λ_c denote the wavelength at the operating frequency and $d = \lambda_c/2$ be the inter-element spacing of the ULA. The corresponding $M \times 1$ steering vector is given by

$$\mathbf{a}(\theta) = [1, e^{-j2\pi\frac{d\cos\theta}{\lambda_c}}, ..., e^{-j2\pi\frac{(M-1)d\cos\theta}{\lambda_c}}]^T. \quad (2)$$

Without loss of generality, we assume that user $k = 0$ is the desired user while the remaining $K - 1$ users and the J jammers are interferers. The desired-signal-to-noise ratio (SNR) specific to each sensor element is given by SNR $= \frac{\sigma_0^2}{\sigma_n^2}$, and the desired-signal-to-jammer-j ratio (SJR) is given by SJR$_j = \frac{\sigma_0^2}{\sigma_j^2}$, where $\sigma_0^2 = \mathbb{E}\{|b_0(i)|^2\}$ denotes the desired signal's power.

The design of a linear full-rank beamformer is equivalent to forming a spatial filter $\mathbf{w}(i) \in \mathbb{C}^{M\times 1}$ that provides an estimate of the desired user symbol, as expressed by $y(i) = \mathbf{w}^H(i)\mathbf{r}(i)$, where $\mathbf{r}(i)$ denotes the sampled and demodulated output of the antenna array at the ith snapshot. For the GSC structure, the full-rank weight vector $\mathbf{w}(i) = \gamma\mathbf{a}(\theta_0) - \mathbf{B}\mathbf{w}_g(i)$ [33], where γ is a real-valued scalar introduced to guarantee the convexity of the optimization problem [29], angle θ_0 is the DOA of the desired user signal, and $\mathbf{a}(\theta_0)$ denotes the corresponding normalized steering vector. $\mathbf{B}$ is the signal blocking matrix, which spans a subspace that is orthogonal to the steering vector $\mathbf{a}(\theta_0)$ [8]. We calculate the weight vector $\mathbf{w}_g(i)$ for the GSC structure according to the CCM criterion, which is a positive measure of the deviation of the squared output from a constant value, along with a constraint on the array response to the desired signal. Considering the specific GSC structure, the CCM beamformer is converted into an unconstrained optimization problem with the following cost function:

$$J_{CM}(\mathbf{w}_g(i)) = \mathbb{E}\{(|y(i)|^2 - 1)^2\}, \quad (3)$$

where $y(i)$ is the output of the GSC beamformer denoted as

$$y(i) = (\gamma\mathbf{a}(\theta_0) - \mathbf{B}\mathbf{w}_g(i))^H\mathbf{r}(i), \quad (4)$$

and $\mathbf{w}_g(i)$ is a filter to be designed.

For a large-scale antenna array system with M sensors, the convergence speed for the full-rank blind adaptive beamformer is typically rather slow. As a result, we resort to reduced-rank techniques to overcome this problem.

III. Krylov-Subspace Based Reduced-Rank Scheme with the GSC Structure

Reduced-rank signal processing techniques have been the focus of many recent works [27]-[30]. These approaches reduce the number of adaptive filter coefficients by projecting the received signal vector onto a lower dimensional subspace and performing the weight adaptation in this subspace. In this section, we describe the reduced-rank CCM beamformer design based on the KS technique. For motivations explained earlier, our interest is centered on the GSC structure illustrated in Fig. 2.

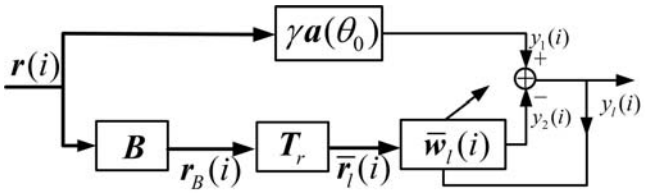

Fig. 2: Reduced-rank beamforming scheme with the GSC structure

As can be seen, similar to the full-rank GSC beamformer, the reduced-rank GSC beamformer output is composed of a constrained component and an unconstrained component. For the constrained component (top or main branch), the output is

$$y_1(i) = \gamma\mathbf{a}^H(\theta_0)\mathbf{r}(i). \quad (5)$$

We note that the desired signal, the interfering users' signals, the jamming signals and the noise component can all pass through the top branch. For the first three, the corresponding output is a weighted version of the input, where the weight equals the inner product between the corresponding normalized steering vector and that of the desired user. As for the unconstrained component (the bottom or auxiliary branch), the received data vector first passes through a signal blocking matrix $\mathbf{B}$, which can be obtained by the singular value decomposition, the QR decomposition [8], [34], or the correlation subtractive structure (CSS) [35]. In this work, we use the CSS structure of the blocking matrix described by

$$\mathbf{B} = \mathbf{I}_M - \frac{\mathbf{a}(\theta_0)\mathbf{a}^H(\theta_0)}{\mathbf{a}^H(\theta_0)\mathbf{a}(\theta_0)} \in \mathbb{C}^{M\times M}. \quad (6)$$

We note that, $\mathbf{B}$ is conjugate symmetric and idempotent, i.e., with the property that $\mathbf{B}^m = \mathbf{B}$, for any positive integer m. Furthermore, the computational complexity of the product $\mathbf{B}\mathbf{r}(i)$ is restricted to $O(M)$, instead of $O(M^2)$ for a general matrix $\mathbf{B}$. Thus, the blocked signal vector $\mathbf{r}_B(i) \in \mathbb{C}^{M\times 1}$ is given by

$$\mathbf{r}_B(i) = \mathbf{B}^H\mathbf{r}(i) = \mathbf{B}\mathbf{r}(i). \quad (7)$$

The blocked signal is processed by the transformation matrix $\mathbf{T}_r \in \mathbb{C}^{M\times D}$. In this work, for the construction of $\mathbf{T}_r$, we utilize the KS technique as exposed in [37], The standard rank-D ($1 \leq D \ll M$) KS can be represented by

$$\mathcal{K}_{\mathcal{D}} = \text{Span}\{\mathbf{a}(\theta_0), \mathbf{R}\mathbf{a}(\theta_0), ..., \mathbf{R}^{D-1}\mathbf{a}(\theta_0)\}, \quad (8)$$

where $\mathbf{R} = \mathbf{E}\{\mathbf{r}(i)\mathbf{r}^H(i)\}$ denotes the array covariance matrix [36]. The transformation matrix which directly utilizes the standard KS is suitable for the direct form processing (DFP) structure, which processes the original received data vector directly. However, our proposed scheme is based on the GSC structure, which processes the blocked signal vector $\mathbf{r}_B(i)$ with

only the interference and noise components in it. Thus, we make some modifications to the construction of the transformation matrix. For the GSC structure, the auxiliary branch is devised to recover the interference-plus-noise component[4] which has passed through the main branch and then cancel it by subtraction. Thus, we choose the first projection vector $\boldsymbol{\rho}_1$ in order to maximize the magnitude of the correlation between its output $(\mathbf{B}\boldsymbol{\rho}_1)^H\mathbf{r}(i)$ and the output of the main branch $\gamma\mathbf{a}^H(\theta_0)\mathbf{r}(i)$ under the constraint that $\boldsymbol{\rho}_1^H\boldsymbol{\rho}_1 = 1$. Hence, the constrained optimization problem can be formulated as:

$$\begin{aligned}\boldsymbol{\rho}_1 &= \arg\min_{\boldsymbol{\rho}} |\mathbb{E}\{(\mathbf{B}\boldsymbol{\rho})^H\mathbf{r}(i)(\gamma\mathbf{a}^H(\theta_0)\mathbf{r}(i))^*\}| \\ &\text{s.t.} \quad \boldsymbol{\rho}^H\boldsymbol{\rho} = 1.\end{aligned} \tag{9}$$

We note that both the cost function and the constraint are phase invariant; in other words, any vector $\boldsymbol{\rho}_1 e^{j\phi}$ where ϕ is an arbitrary phase can be chosen as the optimal solution. However, the solution in common use is the one which forces the cost function to be real. This problem can be solved through Lagrange multipliers optimization, and we obtain

$$\boldsymbol{\rho}_1 = \frac{\mathbf{B}\mathbf{R}\mathbf{a}(\theta_0)}{\|\mathbf{B}\mathbf{R}\mathbf{a}(\theta_0)\|}. \tag{10}$$

Considering that the projection vector is applied to the blocked data vector $\mathbf{r}_B(i)$, and utilizing the idempotence property of $\mathbf{B}$, we can omit $\mathbf{B}$ in the expression of $\boldsymbol{\rho}_1$. In addition, the normalization factor $1/\|\mathbf{R}\mathbf{a}(\theta_0)\|$ is denoted as ν_1, and we obtain

$$\boldsymbol{\rho}_1 = \nu_1\mathbf{R}\mathbf{a}(\theta_0). \tag{11}$$

In the same way, the d-th ($2 < d \leq D$) projection vector $\boldsymbol{\rho}_d$ which maximizes the magnitude of the correlation between its output $(\mathbf{B}\boldsymbol{\rho}_d)^H\mathbf{r}(i)$ and the output of the previous projected vector $(\mathbf{B}\boldsymbol{\rho}_{d-1})^H\mathbf{r}(i)$ can be chosen as

$$\boldsymbol{\rho}_d = \nu_d\mathbf{R}\mathbf{B}\boldsymbol{\rho}_{d-1}, \tag{12}$$

where $\nu_d = 1/\|\mathbf{R}\mathbf{B}\boldsymbol{\rho}_{d-1}\|$ denotes the normalization factor. Finally, we collect all the D projection vectors and define a modified rank-D transformation matrix that is well-suited to the GSC structure via the following expression

$$\begin{aligned}\mathbf{T}_r &= [\nu_1\mathbf{R}\mathbf{a}(\theta_0), \nu_2\mathbf{R}\mathbf{B}\mathbf{R}\mathbf{a}(\theta_0), ..., \nu_D(\mathbf{R}\mathbf{B})^{D-1}\mathbf{R}\mathbf{a}(\theta_0)] \\ &\doteq [\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, ..., \boldsymbol{\rho}_D],\end{aligned} \tag{13}$$

which can be formed iteratively with $\boldsymbol{\rho}_1 = \nu_1\mathbf{R}\mathbf{a}(\theta_0)$, and recursively applying $\boldsymbol{\rho}_k = \nu_k\mathbf{R}\mathbf{B}\boldsymbol{\rho}_{k-1}$. The reduced-rank estimation can capture most, in the maximum correlation sense, of the interference-plus-noise component that has passed through the main branch. The transformation matrix $\mathbf{T}_r$ maps the blocked signal vector $\mathbf{r}_B(i)$ into a low-dimensional reduced-rank vector, which is given by

$$\bar{\mathbf{r}}_l(i) = \mathbf{T}_r^H\mathbf{r}_B(i). \tag{14}$$

Following this transformation step, the reduced-rank vector $\bar{\mathbf{r}}_l(i) \in \mathbb{C}^{D\times 1}$ is processed by a reduced-rank filter $\bar{\mathbf{w}}_l(i) \in \mathbb{C}^{D\times 1}$, which is to be iteratively updated using an RLS type of algorithm to be developed in the next section. The resulting unconstrained output is

$$y_2(i) = \bar{\mathbf{w}}_l^H(i)\bar{\mathbf{r}}_l(i). \tag{15}$$

Finally, the beamformer output is obtained as the difference

$$\begin{aligned}y_l(i) &= y_1(i) - y_2(i) \\ &= \gamma\mathbf{a}^H(\theta_0)\mathbf{r}(i) - \bar{\mathbf{w}}_l^H(i)\bar{\mathbf{r}}_l(i) \\ &= \mathbf{w}^H(i)\mathbf{r}(i),\end{aligned} \tag{16}$$

where the equivalent full-rank weight vector for the reduced-rank GSC structure $\mathbf{w}(i)$ is given by

$$\mathbf{w}(i) = \gamma\mathbf{a}(\theta_0) - \mathbf{B}\mathbf{T}_r\bar{\mathbf{w}}_l(i) \tag{17}$$

IV. Proposed Blind Adaptive Widely linear Reduced-Rank Algorithm

For many applications with non-circular sources, the second-order statistics are fully described by both the covariance matrix $\mathbf{R} = \mathbf{E}\{\mathbf{r}(i)\mathbf{r}^H(i)\}$ and the complementary covariance matrix $\mathbf{R}_c = \mathbf{E}\{\mathbf{r}(i)\mathbf{r}^T(i)\} \neq \mathbf{0}$. In order to exploit the additional information contained in $\mathbf{R}_c$, we combine the received signal $\mathbf{r}(i)$ with its complex conjugate $\mathbf{r}^*(i)$ into an augmented vector $\tilde{\mathbf{r}}(i)$ using an injective transformation $\mathcal{T}$, as shown below

$$\mathbf{r}(i) \xrightarrow{\mathcal{T}} \tilde{\mathbf{r}}(i): \qquad \tilde{\mathbf{r}}(i) = \frac{1}{\sqrt{2}}[\mathbf{r}^T(i), \mathbf{r}^H(i)]^T \in \mathbb{C}^{2M\times 1}. \tag{18}$$

In the WL case, the size of the augmented vector obtained by (18) is twice that of the observed signal, thereby providing extra degrees of freedom to suppress interference [38],[39]. Moreover, the WL processing can take advantage of the non-circular property of the desired signal to further improve the performance. With the increase in the dimension of the data vector in large-scale arrays, it is therefore crucial to take full advantage of reduced-rank signal processing techniques to achieve a faster convergence and increased robustness to interference.

The block diagram of the WLCCM reduced-rank beamforming algorithm with GSC structure is similar to Fig. 2. The difference lies in that all the elements including $\mathbf{r}(i), \mathbf{a}(\theta_0), \mathbf{B}$ and $\mathbf{T}_r$ are extended to their WL variants, respectively denoted as $\tilde{\mathbf{r}}(i), \tilde{\mathbf{a}}(\theta_0), \tilde{\mathbf{B}}$ and $\tilde{\mathbf{T}}_r$. The augmented vectors $\tilde{\mathbf{r}}(i)$ and $\tilde{\mathbf{a}}(\theta_0)$ are obtained from (18), whereas $\tilde{\mathbf{B}} = \mathbf{I}_{2M} - \frac{\tilde{\mathbf{a}}(\theta_0)\tilde{\mathbf{a}}^H(\theta_0)}{\tilde{\mathbf{a}}^H(\theta_0)\tilde{\mathbf{a}}(\theta_0)}$ and $\tilde{\mathbf{T}}_r = [\varrho_1\tilde{\mathbf{R}}\tilde{\mathbf{a}}(\theta_0), \varrho_2\tilde{\mathbf{R}}\tilde{\mathbf{B}}\tilde{\mathbf{R}}\tilde{\mathbf{a}}(\theta_0), ..., \varrho_D(\tilde{\mathbf{R}}\tilde{\mathbf{B}})^{D-1}\tilde{\mathbf{R}}\tilde{\mathbf{a}}(\theta_0)]$, where $\tilde{\mathbf{R}} = \mathbb{E}\{\tilde{\mathbf{r}}(i)\tilde{\mathbf{r}}^H(i)\}$, while the normalization factor $\varrho_d (1 \leq d \leq D)$ ensures that the Euclidian norm of each projection vector equals 1. However, this *direct* WL scheme does not fully exploit the structure of $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{T}}_r$, and consequently, results in extra computational complexity. In this section, we propose an equivalent structured version of the WLCCM reduced-rank beamforming scheme which reduces the complexity by simplifying the construction of reduced-rank data vector $\bar{\mathbf{r}}(i)$. Furthermore, a computational complexity analysis of the proposed algorithm and the existing reduced-rank algorithms is given for comparison.

[4]In the following, we refer to the interfering users' signals plus the jamming signals as interference.

A. The Proposed WLCCM-KS Algorithm

The development of the proposed WLCCM-KS beamforming algorithm involves two steps, namely, the construction of the reduced-rank data vector $\bar{\mathbf{r}}(i)$ and the design of an RLS algorithm to iteratively update the reduced-rank filter $\bar{\mathbf{w}}(i)$.

1) Construction of the reduced-rank data vector: In the first step, we take advantage of the structure of the WL covariance matrix to reduce the computational complexity when constructing the reduced-rank data vector. Firstly, it follows from (18) that the augmented covariance matrix can be written as

$$\tilde{\mathbf{R}} = \frac{1}{2}\begin{bmatrix} \mathbf{R} & \mathbf{R}_c \\ \mathbf{R}_c^* & \mathbf{R}^* \end{bmatrix}, \tag{19}$$

which conforms to the block conjugate structure, that is, the lower halves of the augmented covariance matrix can be completely determined by its upper halves, or vice-versa. Moreover, the augmented blocking matrix $\tilde{\mathbf{B}}$ is also block conjugate and can be partitioned into four sub-matrices

$$\begin{aligned}\tilde{\mathbf{B}} &= \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 \\ \mathbf{B}_2^* & \mathbf{B}_1^* \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{I} - \frac{\mathbf{a}(\theta_0)\mathbf{a}^H(\theta_0)}{2\mathbf{a}^H(\theta_0)\mathbf{a}(\theta_0)} & -\frac{\mathbf{a}(\theta_0)\mathbf{a}^T(\theta_0)}{2\mathbf{a}^H(\theta_0)\mathbf{a}(\theta_0)} \\ (-\frac{\mathbf{a}(\theta_0)\mathbf{a}^T(\theta_0)}{2\mathbf{a}^H(\theta_0)\mathbf{a}(\theta_0)})^* & (\mathbf{I} - \frac{\mathbf{a}(\theta_0)\mathbf{a}^H(\theta_0)}{2\mathbf{a}^H(\theta_0)\mathbf{a}(\theta_0)})^* \end{bmatrix}.\end{aligned} \tag{20}$$

Let us rewrite the augmented steering vector as

$$\tilde{\mathbf{a}}(\theta_0) = \mathcal{T}\{\mathbf{a}(\theta_0)\} = \frac{1}{\sqrt{2}}[\mathbf{a}^T(\theta_0), \mathbf{a}^H(\theta_0)]^T, \tag{21}$$

then, the first column vector (projection vector) of augmented transformation matrix $\tilde{\mathbf{T}}_r$ can be expressed as

$$\begin{aligned}\tilde{\boldsymbol{\rho}}_1 &= \varrho_1 \tilde{\mathbf{R}}\tilde{\mathbf{a}}(\theta_0) \\ &= \frac{1}{\sqrt{2}}\begin{bmatrix} \varrho_1\left(\frac{1}{2}\mathbf{R}\mathbf{a}(\theta_0) + \frac{1}{2}\mathbf{R}_c\mathbf{a}^*(\theta_0)\right) \\ \varrho_1\left(\frac{1}{2}\mathbf{R}\mathbf{a}(\theta_0) + \frac{1}{2}\mathbf{R}_c\mathbf{a}^*(\theta_0)\right)^* \end{bmatrix} \\ &= \mathcal{T}\{\bar{\boldsymbol{\rho}}_1\},\end{aligned} \tag{22}$$

where $\bar{\boldsymbol{\rho}}_1 = \varrho_1\left(\frac{1}{2}\mathbf{R}\mathbf{a}(\theta_0) + \frac{1}{2}\mathbf{R}_c\mathbf{a}^*(\theta_0)\right)$, and the scalar $\varrho_1 = 1/\|\frac{1}{2}\mathbf{R}\mathbf{a}(\theta_0) + \frac{1}{2}\mathbf{R}_c\mathbf{a}^*(\theta_0)\|$ is a normalization factor. According to (22), $\bar{\boldsymbol{\rho}}_1 \in \mathbb{C}^{M\times 1}$ contains all the necessary information to construct $\tilde{\boldsymbol{\rho}}_1 \in \mathbb{C}^{2M\times 1}$.

Next, we will make some simplifications to the second projection vector $\tilde{\boldsymbol{\rho}}_2$, which is obtained by left multiplying $\tilde{\boldsymbol{\rho}}_1$ with $\tilde{\mathbf{R}}\tilde{\mathbf{B}}$, followed by a normalization operation. Using (19) and (20), we can obtain

$$\begin{aligned}\tilde{\mathbf{R}}\tilde{\mathbf{B}} &= \frac{1}{2}\begin{bmatrix} \mathbf{R} & \mathbf{R}_c \\ \mathbf{R}_c^* & \mathbf{R}^* \end{bmatrix}\begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 \\ \mathbf{B}_2^* & \mathbf{B}_1^* \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} \mathbf{R}\mathbf{B}_1 + \mathbf{R}_c\mathbf{B}_2^* & \mathbf{R}\mathbf{B}_2 + \mathbf{R}_c\mathbf{B}_1^* \\ (\mathbf{R}\mathbf{B}_2 + \mathbf{R}_c\mathbf{B}_1^*)^* & (\mathbf{R}\mathbf{B}_1 + \mathbf{R}_c\mathbf{B}_2^*)^* \end{bmatrix}.\end{aligned} \tag{23}$$

Then, $\tilde{\boldsymbol{\rho}}_2$ is given by

$$\begin{aligned}\tilde{\boldsymbol{\rho}}_2 &= \varrho_2\tilde{\mathbf{R}}\tilde{\mathbf{B}}\tilde{\boldsymbol{\rho}}_1 \\ &= \frac{1}{\sqrt{2}}\begin{bmatrix} \varrho_2\left(\frac{1}{2}\mathbf{R}(\mathbf{B}_1\bar{\boldsymbol{\rho}}_1 + \mathbf{B}_2\bar{\boldsymbol{\rho}}_1^*) + \frac{1}{2}\mathbf{R}_c(\mathbf{B}_1\bar{\boldsymbol{\rho}}_1 + \mathbf{B}_2\bar{\boldsymbol{\rho}}_1^*)^*\right) \\ \varrho_2\left(\frac{1}{2}\mathbf{R}(\mathbf{B}_1\bar{\boldsymbol{\rho}}_1 + \mathbf{B}_2\bar{\boldsymbol{\rho}}_1^*) + \frac{1}{2}\mathbf{R}_c(\mathbf{B}_1\bar{\boldsymbol{\rho}}_1 + \mathbf{B}_2\bar{\boldsymbol{\rho}}_1^*)^*\right)^* \end{bmatrix} \\ &= \mathcal{T}\{\bar{\boldsymbol{\rho}}_2\},\end{aligned} \tag{24}$$

where $\bar{\boldsymbol{\rho}}_2 = \varrho_2\left(\frac{1}{2}\mathbf{R}(\mathbf{B}_1\bar{\boldsymbol{\rho}}_1 + \mathbf{B}_2\bar{\boldsymbol{\rho}}_1^*) + \frac{1}{2}\mathbf{R}_c(\mathbf{B}_1\bar{\boldsymbol{\rho}}_1 + \mathbf{B}_2\bar{\boldsymbol{\rho}}_1^*)^*\right)$, with $\varrho_2 = 1/\|\frac{1}{2}\mathbf{R}(\mathbf{B}_1\bar{\boldsymbol{\rho}}_1 + \mathbf{B}_2\bar{\boldsymbol{\rho}}_1^*) + \frac{1}{2}\mathbf{R}_c(\mathbf{B}_1\bar{\boldsymbol{\rho}}_1 + \mathbf{B}_2\bar{\boldsymbol{\rho}}_1^*)^*\|$. In the same way, if we apply the recursive relation between $\tilde{\boldsymbol{\rho}}_d$ and $\tilde{\boldsymbol{\rho}}_{d-1}$, that is $\tilde{\boldsymbol{\rho}}_d = \varrho_d\tilde{\mathbf{R}}\tilde{\mathbf{B}}\tilde{\boldsymbol{\rho}}_{d-1}$ $(3 \le d \le D)$, we obtain that

$$\tilde{\boldsymbol{\rho}}_d = \mathcal{T}\{\bar{\boldsymbol{\rho}}_d\}, \tag{25}$$

where $\bar{\boldsymbol{\rho}}_d = \varrho_d\left(\frac{1}{2}\mathbf{R}(\mathbf{B}_1\bar{\boldsymbol{\rho}}_{d-1} + \mathbf{B}_2\bar{\boldsymbol{\rho}}_{d-1}^*) + \frac{1}{2}\mathbf{R}_c(\mathbf{B}_1\bar{\boldsymbol{\rho}}_{d-1} + \mathbf{B}_2\bar{\boldsymbol{\rho}}_{d-1}^*)^*\right)$.

In practice, $\mathbf{R}$ and $\mathbf{R}_c$ are often estimated by the time average of i received snapshots $\mathbf{r}(n), n = 1, \ldots, i$, where the snapshot index i is now re-introduced. That is,

$$\hat{\mathbf{R}}(i) = \frac{1}{i}\sum_{n=1}^{i}\mathbf{r}(n)\mathbf{r}^H(n) \quad \hat{\mathbf{R}}_c(i) = \frac{1}{i}\sum_{n=1}^{i}\mathbf{r}(n)\mathbf{r}^T(n). \tag{26}$$

Thus, the WL transformation matrix can be written as

$$\tilde{\mathbf{T}}_r(i) = \mathcal{T}\{\mathbf{P}(i)\}, \tag{27}$$

where we define

$$\mathbf{P}(i) = [\bar{\boldsymbol{\rho}}_1(i), \bar{\boldsymbol{\rho}}_2(i), ..., \bar{\boldsymbol{\rho}}_D(i)]. \tag{28}$$

In that sense, $\mathbf{P}(i) \in \mathbb{C}^{M\times D}$ contains the same information as $\tilde{\mathbf{T}}_r(i) \in \mathbb{C}^{2M\times D}$. After further matrix manipulations, the reduced-rank vector with WL processing can be rewritten as

$$\begin{aligned}\bar{\mathbf{r}}(i) &= (\tilde{\mathbf{B}}\tilde{\mathbf{T}}_r(i))^H\tilde{\mathbf{r}}(i) \\ &= \mathcal{R}\{(\mathbf{B}_1\mathbf{P}(i) + \mathbf{B}_2\mathbf{P}^*(i))^H\mathbf{r}(i)\},\end{aligned} \tag{29}$$

where $\mathcal{R}\{\cdot\}$ denotes the real part of its argument. The block diagram summarizing the above structured approach is depicted as Fig. 3.

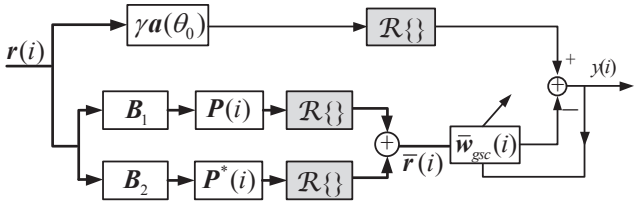

Fig. 3: Proposed WL reduced-rank scheme with the GSC structure.

2) Adaptive implementation of the reduced-rank filter: Next, we derive the structured RLS algorithm for the adaptive implementation of the reduced-rank filter. The reduced-rank weight vector $\bar{\mathbf{w}}(i)$ is obtained by minimizing the unconstrained exponentially weighted least-squares cost function

$$J_{CM}(\bar{\mathbf{w}}(i)) = \sum_{n=1}^{i}\alpha^{i-n}(|y(n)|^2 - 1)^2, \tag{30}$$

where $|y(n)|^2 = y^*(n)(\gamma\mathcal{R}\{\mathbf{a}^H(\theta_0)\mathbf{r}(n)\} - \bar{\mathbf{w}}^H(i)\bar{\mathbf{r}}(n))$, and α is a forgetting factor chosen as a positive scalar, close to, but less than 1. The range of the values of γ which ensures the convexity of the optimization problem is derived in Appendix A. Letting $\tilde{\mathbf{x}}(n) = y^*(n)\bar{\mathbf{r}}(n)$, $\tilde{d}(n) = \gamma y^*(n)\mathcal{R}\{\mathbf{a}^H(\theta_0)\mathbf{r}(n)\} - 1$, and substituting the remaining

TABLE I: The proposed WLCCM-KS algorithm

Initialization with a specified rank D:

$\tilde{\mathbf{Q}}^{-1}(0) = \delta \mathbf{I}_D, \bar{\mathbf{w}}(0) = [1, 0, ..., 0]^T$

For the ith **snapshot** $i = 1, 2, ...$

Compute $\hat{\mathbf{R}}(i)$ and $\hat{\mathbf{R}}_c(i)$ according to (26)

Calculate $\mathbf{P}(i)$ according to (28)

$\bar{\mathbf{r}}(i) = \mathcal{R}\{(\mathbf{B}_1\mathbf{P}(i) + \mathbf{B}_2\mathbf{P}^*(i))^H \mathbf{r}(i)\}$

$y(i) = \gamma\mathcal{R}\{\mathbf{a}^H(\theta_0)\mathbf{r}(i)\} - \bar{\mathbf{w}}^H(i-1)\bar{\mathbf{r}}(i)$

$\tilde{\mathbf{x}}(i) = y^*(i)\bar{\mathbf{r}}(i)$

$\tilde{d}(i) = \gamma y^*(i)\mathcal{R}\{\mathbf{a}^H(\theta_0)\mathbf{r}(i)\} - 1$

Update the reduced-rank coefficient $\bar{\mathbf{w}}$

$\tilde{\mathbf{k}}(i) = \frac{\tilde{\mathbf{Q}}^{-1}(i-1)\tilde{\mathbf{x}}(i)}{\alpha + \tilde{\mathbf{x}}^H(i)\tilde{\mathbf{Q}}^{-1}(i-1)\tilde{\mathbf{x}}(i)}$

$\tilde{\xi}(i) = \tilde{d}(i) - \bar{\mathbf{w}}^H(i-1)\tilde{\mathbf{x}}(i)$

$\tilde{\mathbf{Q}}^{-1}(i) = \alpha^{-1}(\tilde{\mathbf{Q}}^{-1}(i-1) - \tilde{\mathbf{k}}(i)\tilde{\mathbf{x}}^H(i)\tilde{\mathbf{Q}}^{-1}(i-1))$

$\bar{\mathbf{w}}(i) = \bar{\mathbf{w}}(i-1) + \tilde{\mathbf{k}}(i)\tilde{\xi}^*(i)$

$y(n)$ by $\gamma\mathcal{R}\{\mathbf{a}^H(\theta_0)\mathbf{r}(n)\} - \bar{\mathbf{w}}^H(i-1)\bar{\mathbf{r}}(n)$, (30) can be approximated as

$$J_{CM}(\bar{\mathbf{w}}(i)) \approx \sum_{n=1}^{i} \alpha^{i-n}[\tilde{d}(n) - \bar{\mathbf{w}}^H(i)\tilde{\mathbf{x}}(n)]^2, \tag{31}$$

which is now quadratic in the unknown weight vector $\bar{\mathbf{w}}(i)$. By taking the gradient of (31) with respect to $\bar{\mathbf{w}}^*(i)$ and equating it to zero, after further manipulations we obtain

$$\bar{\mathbf{w}}(i) = \tilde{\mathbf{Q}}^{-1}(i)\tilde{\mathbf{p}}(i), \tag{32}$$

where

$$\tilde{\mathbf{Q}}(i) = \alpha\tilde{\mathbf{Q}}(i-1) + \tilde{\mathbf{x}}(i)\tilde{\mathbf{x}}^H(i) \tag{33}$$

$$\tilde{\mathbf{p}}(i) = \alpha\tilde{\mathbf{p}}(i-1) + \tilde{\mathbf{x}}(i)\tilde{d}^*(i) \tag{34}$$

To avoid the matrix inversion and reduce the complexity, we apply the matrix inversion lemma [33] to (32), and obtain the following recursive expression

$$\bar{\mathbf{w}}(i) = \bar{\mathbf{w}}(i-1) + \tilde{\mathbf{k}}(i)\tilde{\xi}^*(i), \tag{35}$$

where

$$\tilde{\mathbf{k}}(i) = \frac{\tilde{\mathbf{Q}}^{-1}(i-1)\tilde{\mathbf{x}}(i)}{\alpha + \tilde{\mathbf{x}}^H(i)\tilde{\mathbf{Q}}^{-1}(i-1)\tilde{\mathbf{x}}(i)}, \tag{36}$$

$$\tilde{\xi}(i) = \tilde{d}(i) - \bar{\mathbf{w}}^H(i-1)\tilde{\mathbf{x}}(i), \tag{37}$$

$$\tilde{\mathbf{Q}}^{-1}(i) = \alpha^{-1}(\tilde{\mathbf{Q}}^{-1}(i-1) - \tilde{\mathbf{k}}(i)\tilde{\mathbf{x}}^H(i)\tilde{\mathbf{Q}}^{-1}(i-1)). \tag{38}$$

Based on (35)-(38), we obtain the reduced-rank filter updating procedure for the proposed adaptive WLCCM-KS algorithm with the GSC structure, which is summarized in Table I.

With this new structured scheme, we do not need to use the transformation (18) and all the calculations are processed with vectors of lengths less than or equal to M, thereby significantly reducing the computational complexity as compared to the conventional direct WL scheme.

TABLE II: Computational complexity of reduced-rank algorithms

Algorithms	Real multiplications	Real additions
Structured WLCCM-KS-RLS	$8DM^2 + 12M^2 + 23DM$ $+2M + 3D^2 + 9D + 1$	$8DM^2 + 8M^2 + 18DM$ $+2M + 2D^2 - 3$
Direct WLCCM-KS-RLS	$16DM^2 + 24M^2 + 10DM$ $+32M + 3D^2 + 5D + 1$	$16DM^2 + 16M^2$ $+6DM + 24M + 2D^2 - D - 3$
WL-AVF	$32DM^2 + 24M^2$ $+40DM + 8M + 4D$	$32DM^2 + 16M^2$ $+32DM - 4D$
LCCM-KS-RLS	$4DM^2 + 6M^2 + 13DM$ $+4M + 10D^2 + 20D + 2$	$4DM^2 + 4M^2 + 11DM$ $+4M + 8D^2 + 9D - 1$
L-AVF	$8DM^2 + 6M^2$ $+20DM + 4D$	$8DM^2 + 4M^2$ $+16DM - 4D$

B. Computational Complexity Analysis

We investigate the computational complexity of the proposed WLCCM-KS algorithm with the GSC structure, where the complexity is evaluated in terms of the number of real additions and real multiplications for each snapshot of length M of the received data vector. We compare the complexity of the proposed algorithms with that of the direct algorithm, i.e. as obtained by application of the conventional WL processing scheme, the existing WL-AVF reduced-rank algorithm [32], and their linear counterparts, referred to by the acronyms LCCM-KS and L-AVF, respectively. The complexity figures are listed in Table II, while Fig. 4 illustrates the total number of real operations (real multiplications plus real additions) per snapshot for each reduced-rank algorithm as a function of M, where we choose the rank $D = 3$. As can be seen, the proposed WLCCM-KS algorithm, implemented either in the direct or in the structured form, reduces the computational complexity, compared with the existing WL-AVF algorithm. It is worth emphasizing that the proposed structured WLCCM-KS algorithm is less complex than the direct form, since it fully utilizes the structure of the augmented covariance matrix and exploits it to make some key simplifications. We also observe that the WL algorithms generally have a higher computational complexity than their linear counterparts. This should not come as a surprise since in the construction of the transformation matrix, the WL algorithms make use of both the original observed vector and its complex conjugate.

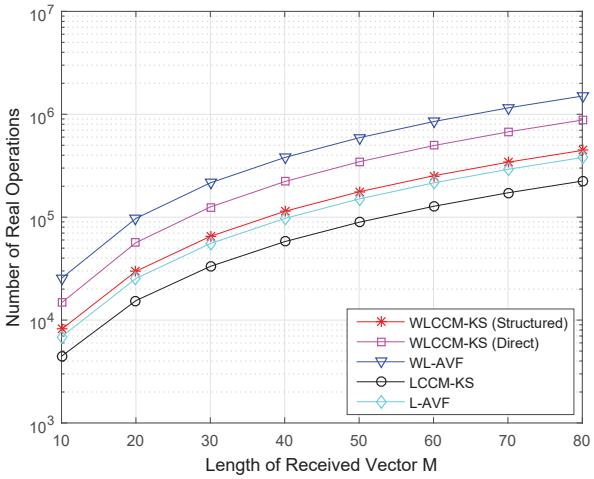

Fig. 4: Number of real operations per snapshot versus length of received vector M for various reduced-rank adaptive algorithms (the rank D is set to 3)

V. Analysis of the Proposed Algorithm

In this section, we theoretically analyze the convergence properties of the sequence of reduced-rank weight vectors produced by the proposed WLCCM-KS algorithm and derive the analytical steady-state MSE expression. Furthermore, a qualitative analysis of the SINR performance is carried out to further emphasize the superiority of the WL processing over its corresponding linear processing.

A. Convergence of the Mean Reduced-Rank Weight Vector

We consider the convergence of the mean reduced-rank weight vector $\bar{\mathbf{w}}(i)$ for the proposed WLCCM-KS algorithm, which is characterized by the reduced-rank weight error vector $\boldsymbol{\epsilon}(i)$, defined as the difference between the instantaneous and optimal values of the weight vector[5], that is,

$$\boldsymbol{\epsilon}(i) = \bar{\mathbf{w}}(i) - \bar{\mathbf{w}}_o. \tag{39}$$

Multiplying both sides of the adaptation equation (35) with $\tilde{\mathbf{Q}}(i)$, we subsequently obtain

$$\tilde{\mathbf{Q}}(i)\bar{\mathbf{w}}(i) = \tilde{\mathbf{Q}}(i)\bar{\mathbf{w}}(i-1) + \tilde{\mathbf{Q}}(i)\tilde{\mathbf{k}}(i)\tilde{\xi}^*(i). \tag{40}$$

Invoking (33) and (36), we obtain

$$\tilde{\mathbf{Q}}(i)\tilde{\mathbf{k}}(i) = \tilde{\mathbf{x}}(i). \tag{41}$$

Substituting (33),(37) and (41) into (40) yields

$$\tilde{\mathbf{Q}}(i)\bar{\mathbf{w}}(i) = \alpha\tilde{\mathbf{Q}}(i-1)\bar{\mathbf{w}}(i-1) + \tilde{\mathbf{x}}(i)\tilde{d}^*(i). \tag{42}$$

We define the optimum error at time i as

$$\tilde{e}_o(i) = \tilde{d}(i) - \bar{\mathbf{w}}_o^H\tilde{\mathbf{x}}(i). \tag{43}$$

Right multiplying both sides of (33) with the optimum reduced-rank weight vector $\bar{\mathbf{w}}_o$, and subtracting the resulting equation from (42), we obtain the following recursive relation for the reduced-rank weight error vector

$$\tilde{\mathbf{Q}}(i)\boldsymbol{\epsilon}(i) = \alpha\tilde{\mathbf{Q}}(i-1)\boldsymbol{\epsilon}(i-1) + \tilde{\mathbf{x}}(i)\tilde{e}_o^*(i). \tag{44}$$

When $i \to \infty$, the asymptotic stationarity property of the estimated augmented correlation matrix $\hat{\tilde{\mathbf{R}}}(i)$ indicates that $\hat{\tilde{\mathbf{R}}}(i) \simeq \tilde{\mathbf{R}}$, where the asymptotic variance of the error matrix is very small. Since the transformation matrix $\tilde{\mathbf{T}}_r(i)$ is directly related to $\hat{\tilde{\mathbf{R}}}(i)$, we can assume that $\tilde{\mathbf{T}}_r(i)$ is also asymptotically stationary with

$$\begin{aligned}\tilde{\mathbf{T}}_r(i) &\simeq \tilde{\mathbf{T}}_r(i-1) \simeq \tilde{\mathbf{T}}_r \\ &= [\varrho_1\tilde{\mathbf{R}}\tilde{\mathbf{a}}(\theta_0), \varrho_2\tilde{\mathbf{R}}\tilde{\mathbf{B}}\tilde{\mathbf{R}}\tilde{\mathbf{a}}(\theta_0), ..., \varrho_D(\tilde{\mathbf{R}}\tilde{\mathbf{B}})^{D-1}\tilde{\mathbf{R}}\tilde{\mathbf{a}}(\theta_0)].\end{aligned} \tag{45}$$

Accordingly, as the algorithm converges to its steady-state, $\tilde{\mathbf{x}}(i)$ only depends on the input vector $\mathbf{r}(i)$, thus we can similarly obtain that $\tilde{\mathbf{Q}}^{-1}(i)\tilde{\mathbf{Q}}(i-1) \simeq \mathbf{I}_D$ [14], [40], [41]. Multiplying both sides of (44) by $\tilde{\mathbf{Q}}^{-1}(i)$ and after some simplifications, we obtain the following recursive relation for the reduced-rank weight error vector

$$\boldsymbol{\epsilon}(i) \simeq \alpha\boldsymbol{\epsilon}(i-1) + \tilde{\mathbf{Q}}^{-1}(i)\tilde{\mathbf{x}}(i)\tilde{e}_o^*(i). \tag{46}$$

Noticing that when $i \to \infty$, the optimum error $\tilde{e}_o(i)$ is orthogonal [40] with $\tilde{\mathbf{x}}(i)$, namely

$$\mathbb{E}\{\tilde{\mathbf{x}}(i)\tilde{e}_o^*(i)\} = \mathbf{0}, \tag{47}$$

we obtain the following recursive equation for the mean reduced-rank weight error vector

$$\mathbb{E}\{\boldsymbol{\epsilon}(i)\} \simeq \alpha\mathbb{E}\{\boldsymbol{\epsilon}(i-1)\}. \tag{48}$$

Since $0 < \alpha < 1$, when $i \to \infty$, we obtain

$$\mathbb{E}\{\boldsymbol{\epsilon}(i)\} = \mathbf{0}, \tag{49}$$

which implies that the mean reduced-rank weight error vector converges to zero or equivalently, that the reduced-rank weight vector converges to its optimum value.

B. Convergence of MSE

Next, we analyze the MSE convergence for the proposed reduced-rank beamforming algorithm and develop an analytical expression of the steady-state MSE. When $i \to \infty$, the steady-state MSE can be written as

$$\begin{aligned}\lim_{i\to\infty}\tilde{\zeta}_{\text{mse}}(i) &= \lim_{i\to\infty}\mathbb{E}\{|b_0(i) - \tilde{\mathbf{w}}^H(i)\tilde{\mathbf{r}}(i)|^2\} \\ &= \lim_{i\to\infty}\mathbb{E}\{|b_0(i) - \tilde{\mathbf{w}}_o^H\tilde{\mathbf{r}}(i) - \tilde{\boldsymbol{\varepsilon}}^H(i)\tilde{\mathbf{r}}(i)|^2\} \\ &= (1-2\gamma)\sigma_0^2 + \tilde{\mathbf{w}}_o^H\tilde{\mathbf{R}}\tilde{\mathbf{w}}_o + \lim_{i\to\infty}\mathbb{E}\{tr[\tilde{\mathbf{R}}\tilde{\boldsymbol{\varepsilon}}(i)\tilde{\boldsymbol{\varepsilon}}^H(i)]\} \\ &= \tilde{\zeta}_{\text{min}} + \lim_{i\to\infty}\tilde{\zeta}_{\text{ex}}(i),\end{aligned} \tag{50}$$

where $\tilde{\mathbf{w}}(i) = \gamma\tilde{\mathbf{a}}(\theta_0) - \tilde{\mathbf{B}}\tilde{\mathbf{T}}_r(i)\bar{\mathbf{w}}(i)$ and $\tilde{\mathbf{w}}_o = \gamma\tilde{\mathbf{a}}(\theta_0) - \tilde{\mathbf{B}}\tilde{\mathbf{T}}_r\bar{\mathbf{w}}_o$ are the equivalent transient and optimum full-rank weight vector for the reduced-rank algorithm, respectively. The difference $\tilde{\boldsymbol{\varepsilon}}(i) = \tilde{\mathbf{w}}(i) - \tilde{\mathbf{w}}_o$ denotes the weight error vector. The steady-state MSE consists of the minimum MSE component $\tilde{\zeta}_{\text{min}} = (1-2\gamma)\sigma_0^2 + \tilde{\mathbf{w}}_o^H\tilde{\mathbf{R}}\tilde{\mathbf{w}}_o$ and the excess MSE component $\lim_{i\to\infty}\tilde{\zeta}_{\text{ex}}(i) = \lim_{i\to\infty}\mathbb{E}\{tr[\tilde{\mathbf{R}}\tilde{\boldsymbol{\varepsilon}}(i)\tilde{\boldsymbol{\varepsilon}}^H(i)]\}$. The unavoidable weight error vector $\tilde{\boldsymbol{\varepsilon}}(i)$ is the source of the excess MSE.

In order to calculate the excess MSE, we need to derive two equations as a preliminary step. Firstly, using the asymptotic stationary property of $\tilde{\mathbf{T}}_r(i)$ given in (45), we obtain when $i \to \infty$

$$\tilde{\boldsymbol{\varepsilon}}(i) = -\tilde{\mathbf{B}}\tilde{\mathbf{T}}_r\boldsymbol{\epsilon}(i), \tag{51}$$

which is the first equation needed. In addition, we define the correlation matrix of $\tilde{\mathbf{x}}(i)$ as $\tilde{\mathbf{R}}_x = \mathbb{E}\{\tilde{\mathbf{x}}(i)\tilde{\mathbf{x}}^H(i)\}$. Recalling that $\tilde{\mathbf{x}}(i) = y^*(i)\tilde{\mathbf{T}}_r^H(i)\tilde{\mathbf{B}}^H\tilde{\mathbf{r}}(i)$, we have

$$\begin{aligned}\lim_{i\to\infty}\tilde{\mathbf{R}}_x(i) &= \lim_{i\to\infty}\mathbb{E}\{|y(i)|^2\tilde{\mathbf{T}}_r^H(i)\tilde{\mathbf{B}}^H\tilde{\mathbf{r}}(i)\tilde{\mathbf{r}}^H(i)\tilde{\mathbf{B}}\tilde{\mathbf{T}}_r(i)\} \\ &\simeq \lim_{i\to\infty}\mathbb{E}\{|y(i)|^2\}\tilde{\mathbf{T}}_r^H\tilde{\mathbf{B}}^H\tilde{\mathbf{R}}\tilde{\mathbf{B}}\tilde{\mathbf{T}}_r,\end{aligned} \tag{52}$$

where we used the asymptotic property in (45), and this is the second needed equation.

[5]The optimal value $\bar{\mathbf{w}}_o$ is the solution to the minimization problem with cost function $J = \mathbb{E}\{(|y(i)|^2 - 1)^2\}$, where $y(i) = \gamma\tilde{\mathbf{a}}^H(\theta_0)\tilde{\mathbf{r}}(i) - \bar{\mathbf{w}}^H(i)\tilde{\mathbf{T}}_r^H\tilde{\mathbf{B}}\tilde{\mathbf{r}}(i)$. However, there is no closed form expression for $\bar{\mathbf{w}}_o$ and an iterative approach must be employed to reach a solution [14].

Subsequently, by applying (51) and (52), the expression of the excess MSE can be further modified as follows

$$\begin{aligned}\lim_{i\to\infty}\tilde{\zeta}_{\text{ex}}(i) &= \lim_{i\to\infty}\mathbb{E}\{\text{tr}[\tilde{\mathbf{R}}\tilde{\mathbf{B}}\tilde{\mathbf{T}}_r\boldsymbol{\epsilon}(i)\boldsymbol{\epsilon}^H(i)\tilde{\mathbf{T}}_r^H\tilde{\mathbf{B}}^H]\}\\ &= \lim_{i\to\infty}\mathbb{E}\{\text{tr}[\tilde{\mathbf{T}}_r^H\tilde{\mathbf{B}}^H\tilde{\mathbf{R}}\tilde{\mathbf{B}}\tilde{\mathbf{T}}_r\boldsymbol{\epsilon}(i)\boldsymbol{\epsilon}^H(i)]\}\\ &= \lim_{i\to\infty}\frac{\mathbb{E}\{\text{tr}[\tilde{\mathbf{R}}_x(i)\boldsymbol{\epsilon}(i)\boldsymbol{\epsilon}^H(i)]\}}{\mathbb{E}\{|y(i)|^2\}}.\end{aligned} \tag{53}$$

Considering the recursive relation for the reduced-rank weight error vector described in (46), we obtain

$$\begin{aligned}\boldsymbol{\Theta}(i) &= \mathbb{E}\{\boldsymbol{\epsilon}(i)\boldsymbol{\epsilon}^H(i)\}\\ &= \alpha^2\mathbb{E}\{\boldsymbol{\epsilon}(i-1)\boldsymbol{\epsilon}^H(i-1)\}\\ &\quad-\alpha\mathbb{E}\{\boldsymbol{\epsilon}(i-1)\tilde{e}_o(i)\tilde{\mathbf{x}}^H(i)\tilde{\mathbf{Q}}^{-1}(i)\}\\ &\quad-\alpha\mathbb{E}\{\tilde{\mathbf{Q}}^{-1}(i)\tilde{\mathbf{x}}(i)\tilde{e}_o^*(i)\boldsymbol{\epsilon}^H(i-1)\}\\ &\quad+\mathbb{E}\{\tilde{\mathbf{Q}}^{-1}(i)\tilde{\mathbf{x}}(i)\tilde{e}_o^*(i)\tilde{e}_o(i)\tilde{\mathbf{x}}^H(i)\tilde{\mathbf{Q}}^{-1}(i)\}.\end{aligned} \tag{54}$$

When $i\to\infty$, the algorithm has converged to its steady-state, and thus we can assume that $\tilde{e}_o(i)$, $\tilde{\mathbf{x}}(i)$ and $\tilde{\mathbf{Q}}^{-1}(i)$ are uncorrelated. Furthermore, making use again of the orthogonality between $\tilde{e}_o(i)$ and $\tilde{\mathbf{x}}(i)$ as stated in (47) yields

$$\boldsymbol{\Theta}(i) = \alpha^2\boldsymbol{\Theta}(i-1) + \sigma_e^2\mathbb{E}\{\tilde{\mathbf{Q}}^{-1}(i)\}\tilde{\mathbf{R}}_x\mathbb{E}\{\tilde{\mathbf{Q}}^{-1}(i)\}, \tag{55}$$

where $\sigma_e^2 = \mathbb{E}\{\tilde{e}_o^*(i)\tilde{e}_o(i)\}$ denotes the minimum value of the CCM reduced-rank algorithm. When $i\to\infty$, we can assume that $\tilde{\mathbf{Q}}^{-1}(i)\tilde{\mathbf{R}}_x \simeq (1-\alpha)\mathbf{I}_{DD}$ [40], thus (55) can be simplified as

$$\boldsymbol{\Theta}(i) = \alpha^2\boldsymbol{\Theta}(i-1) + (1-\alpha)^2\sigma_e^2\tilde{\mathbf{R}}_x^{-1}. \tag{56}$$

Considering that $0<\alpha<1$, the solution $\boldsymbol{\Theta}(i)$ to the above equation (56) converges to a finite limit. Hence upon equating $\boldsymbol{\Theta}(i) = \boldsymbol{\Theta}(i-1)$, we obtain

$$\boldsymbol{\Theta}(i) = \frac{(1-\alpha)^2}{1-\alpha^2}\sigma_e^2\tilde{\mathbf{R}}_x^{-1}. \tag{57}$$

We then substitute (57) into (53) and obtain the compact form of the excess MSE given by

$$\lim_{i\to\infty}\tilde{\zeta}_{\text{ex}}(i) = \lim_{i\to\infty}\frac{D(1-\alpha)^2\sigma_e^2}{(1-\alpha^2)\mathbb{E}\{|y(i)|^2\}}. \tag{58}$$

Finally, we have the steady-state MSE which is given by

$$\begin{aligned}\lim_{i\to\infty}\tilde{\zeta}_{\text{mse}}(i) &= (1-2\gamma)\sigma_0^2 + \tilde{\mathbf{w}}_o^H\tilde{\mathbf{R}}\tilde{\mathbf{w}}_o\\ &\quad+\lim_{i\to\infty}\frac{D(1-\alpha)^2\sigma_e^2}{(1-\alpha^2)\mathbb{E}\{|y(i)|^2\}}.\end{aligned} \tag{59}$$

C. Achievable SINR

From the block diagram shown in Fig. 3, we explicitly note that the WL reduced-rank vector $\bar{\mathbf{r}}(i)$ is real-valued. Thus, the filter coefficient $\bar{\mathbf{w}}(i)$ and the output of the filter $y(i)$ are also real-valued. Then the optimum augmented WL weighting vector $\tilde{\mathbf{w}}_o = \gamma\tilde{\mathbf{a}}(\theta_0) - \tilde{\mathbf{B}}\tilde{\mathbf{T}}_r\bar{\mathbf{w}}_o$ is conjugate symmetric, and can be expressed as $\tilde{\mathbf{w}}_o = \mathcal{T}\{\mathbf{w}_{\text{o,WL}}\}$, where by definition $\mathbf{w}_{\text{o,WL}}\in\mathbb{C}^{M\times 1}$ contains the same information as $\tilde{\mathbf{w}}_o\in\mathbb{C}^{2M\times 1}$. The corresponding optimal weight vector $\mathbf{w}_{\text{o,WL}}$ minimizes the cost function $E\{(|\mathcal{R}\{y(i)\}|^2-1)^2\}$, where $y(i) = \mathbf{w}^H\mathbf{r}(i)$. The optimum output SINR can be equivalently expressed as

$$\begin{aligned}\text{SINR}_{\text{WL}} &= \frac{\mathbb{E}\{|\mathcal{R}\{\mathbf{w}_{\text{o,WL}}^H\mathbf{s}(i)\}|^2\}}{\mathbb{E}\{|\mathcal{R}\{\mathbf{w}_{\text{o,WL}}^H\mathbf{v}(i)\}|^2\}}\\ &= \frac{\gamma^2 M}{\mathbb{E}\{|\mathcal{R}\{\mathbf{w}_{\text{o,WL}}^H\mathbf{v}(i)\}|^2\}},\end{aligned} \tag{60}$$

where $\mathbf{s}(i)$ and $\mathbf{v}(i)$ denote the desired signal and the interference-plus-noise components, respectively. On the one hand, the optimal weight vector $\mathbf{w}_{\text{o,L}}$ minimizes the linear cost function $E\{(|y(i)|^2-1)^2\}$, and the corresponding optimum SINR is given by

$$\text{SINR}_{\text{L}} = \frac{\gamma^2 M}{\mathbb{E}\{|\mathbf{w}_{\text{o,L}}^H\mathbf{v}(i)|^2\}}. \tag{61}$$

According to (60), if we substitute $\mathbf{w}_{\text{o,WL}}$ for $\mathbf{w}_{\text{o,L}}$, the resulting $\text{SINR}' = \frac{\gamma^2 M}{\mathbb{E}\{|\mathcal{R}\{\mathbf{w}_{\text{o,L}}^H\mathbf{v}(i)\}|^2\}} \le \text{SINR}_{\text{WL}}$. On the other hand, the operation $\mathcal{R}\{\cdot\}$ nearly reduces the interference-plus-noise power by half, that is $\text{SINR}' \approx 2\text{SINR}_L$. Consequently, the optimum SINR of the WL processing exhibits an almost 3dB gain over that of the linear one [32].

VI. SIMULATIONS

In this section, we assess the output SINR and the MSE performance of the proposed WLCCM-KS algorithm through numerical simulations. In addition, we verify the validity of the derived analytical results for the steady-state MSE. For the SINR, we evaluate the convergence and steady-state performance of the proposed algorithm and compare it with the existing WL-AVF algorithm [32], its linear counterpart LCCM-KS algorithm as well as the CCM and CMV criteria based full-rank (FR) algorithms, with the respective acronyms WLCCM-FR, LCCM-FR, WLCMV-FR, LCMV-FR. The FR scheme directly devises an adaptive algorithm to update the weighting coefficient corresponding to each component of the blocked data vector $\mathbf{r}_B(i)$ without the projection procedure. The output SINR of the WL processing is given by

$$\text{SINR}(i) = \frac{\tilde{\mathbf{w}}^H(i)\tilde{\mathbf{R}}_s\tilde{\mathbf{w}}(i)}{\tilde{\mathbf{w}}^H(i)\tilde{\mathbf{R}}_{in}\tilde{\mathbf{w}}(i)}, \tag{62}$$

where $\tilde{\mathbf{R}}_s$ and $\tilde{\mathbf{R}}_{in}$ denote the augmented covariance matrices of the desired signal and the interference-plus-noise component in the observation space, respectively.

In our simulations, we consider a fairly large ULA system where the number M of antenna elements, which has a great impact on the output SINR performance, is varied between 10 to 30.[6] For simplicity, we assume that the signal of user $k\in\{0,\ldots,K-1\}$ is taken from the BPSK set $\{\pm 1\}$ with equal probability, so that its power is normalized to unity[7]. The DOA of the desired user ($k=0$) is set to

[6] At an operational wavelength of 1.5GHz [42], this corresponds to an array aperture size between 1 meter and 3 meters.

[7] In our work, we consider the constant modulus criterion after demodulation (i.e., using match filtering with the same pulse shape) under the assumption of perfect synchronization. The beamforming is done using the optimally generated sample in each symbol duration, which is affected by the added Gaussian noise. Under these standard modeling conditions, the BPSK signal is exactly constant modulus.

$\theta_0 = 50°$, while for the other $K-1$ interference users, the DOAs are set as $(40°, 70°, 20°, 80°)$. We assume that $J = 4$ white Gaussian jamming signals impinge on the array with DOAs of $(60°, 90°, 30°, 10°)$, where for each jammer, the corresponding SJR$=$ 0 dB. In the simulation results given below, all the SINR related curves are averaged over 200 independent runs, whereas the MSE related ones are obtained by averaging 2000 runs.

A. Effects of the Number of Antenna Elements M

Firstly, we investigate the effect of the number of antenna elements M on the output SINR performance of each analyzed algorithm. Fig. 5 shows the impact of M on the output SINR performance with the reference snapshot set as $N = 1000$ to ensure that the steady-regime is achieved. The performance of the optimum minimum variance distortionless response (MVDR) filter is also given for comparison. The input SNR for each antenna element is set to -3dB. We find that as the value of M increases, the steady-state SINR of each algorithm generally becomes larger, which is especially evident for the optimum MVDR solution. Moreover, the proposed WLCCM-KS algorithm tends to yield a greater performance gain compared with the second best WL-AVF algorithm when M is larger $(M > 26)$. Besides, when M is smaller $(M < 14)$, WL processing generally provides a greater performance gain compared with the linear processing. This can be explained by noting that when M is small, WL processing doubles the dimension of the processing vector so as to obtain extra freedoms for interference suppression.

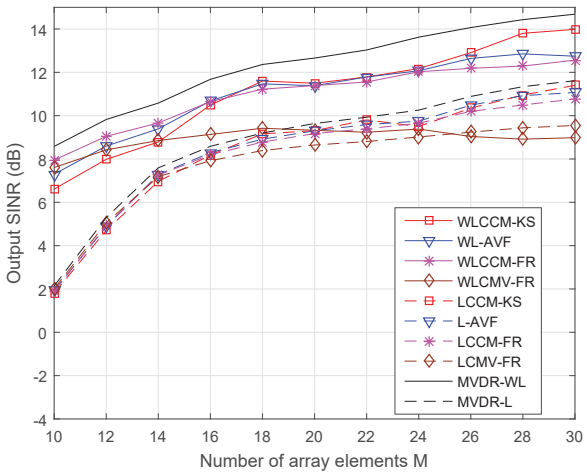

Fig. 5: The impact of M on the output SINR performance (N=1000)

Correspondingly, Fig. 6 shows the impact of M on the output SINR performance with reference snapshot set as $N = 200$, which represents a transient-state performance in the process of convergence. The simulation environment is the same as that of Fig. 5. It can be seen that our proposed WLCCM-KS algorithm obtains the best output SINR performance compared with all the other analyzed algorithms. Besides, one of the problems for the full-rank algorithms is

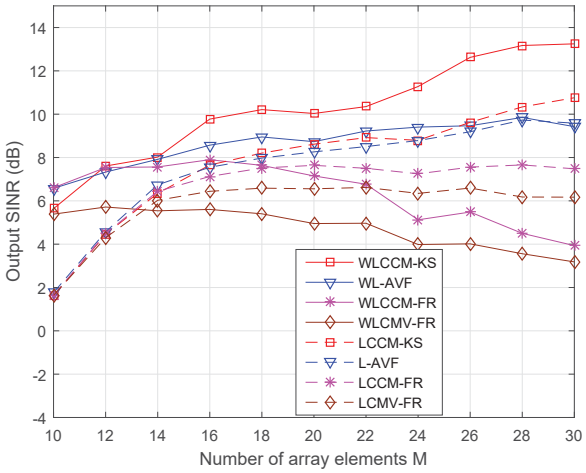

Fig. 6: The impact of M on the output SINR performance (N=200)

that the output SINR does not increase as M grows, and even decreases for the WL processing schemes, when $M \geq 16$. This shows that as M increases, the convergence speed of the full-rank algorithms deteriorates greatly [24]-[27]. Thus, in such case, the use of reduced-rank technique is of crucial importance. Considering limitations of a realistic application and the simulation results for the output SINR performance versus the value of M, we finally choose M as 20 in the remaining experiments.

B. Effects of Rank D

The output SINR performance of the reduced-rank algorithms depends on the selection of rank D. In this part, we evaluate the effects of D on the steady-state SINR performance for the proposed and existing reduced-rank algorithms, so as to find the most adequate value of D for a fair comparison. Interestingly, it has been observed that in various scenarios with the KS based reduced-rank technique, the optimal rank D does not scale significantly with the number of users K and the length of the observation vector M. For a blind algorithm, generally $D \leq 5$ can be chosen [27]. This knowledge can dramatically compress the range of D to be considered. In Fig. 7, we show the SINR performance of the reduced-rank algorithms versus D, while the other simulation parameters are the same as that of Fig. 5. The reference snapshots is set as $N = 1000$ to ensure that the steady-regime is achieved. We observe that for the WLCCM-KS algorithm and its linear counterpart, the performance is enhanced when D increases from 1 to 3, however when D continues to grow, the performance nearly remains the same. Comprehensively taking into consideration the performance and complexity, the optimal choice of rank appears to be $D = 3$. While for the WL-AVF and L-AVF algorithms, the most adequate value to reach the best performance is $D = 2$.

C. Beampatterns

Fig. 8 illustrates the beampatterns of the proposed WLCCM-KS reduced-rank beamforming algorithm and its linear version

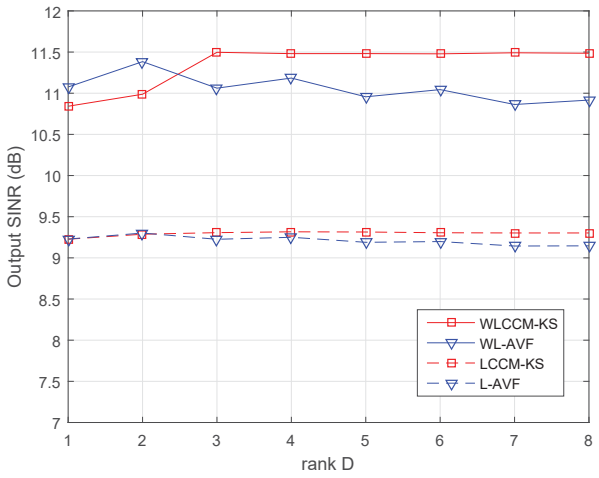

Fig. 7: SINR performance of the reduced-rank algorithms versus the rank D (SNR= -3dB, $N = 1000$)

LCCM-KS algorithm, respectively. The reference snapshot is set as $N = 1000$, where the simulation environment is just the same as that of Fig. 7. For both algorithms, the rank D is chosen to be the most appropriate one, i.e., $D = 3$. It is evident that the mainlobes of the two schemes considered are both directed towards the DOA of the desired user. However, the WLCCM-KS algorithm forms a narrower mainlobe with the same aperture, which will greatly enhance the angle resolution capability. Moreover, the proposed WLCCM-KS algorithm generally yields lower sidelobes, which allows better interference suppression.

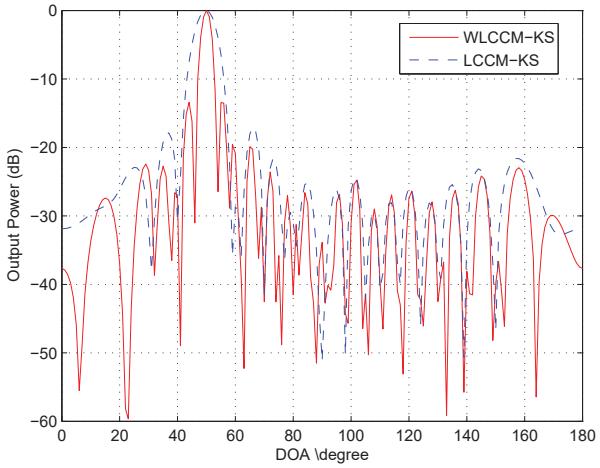

Fig. 8: Beampattern of the reduced-rank beamforming scheme

D. SINR Convergence Performance

In the experiment shown in Fig. 9, we consider the same simulation environment as given in Fig. 8, and investigate the output SINR convergence versus the number of snapshots. The proposed WLCCM-KS algorithm exhibits a faster convergence and a higher steady-state SINR compared with the WL-AVF algorithm and the full-rank schemes. It can be observed that the WL algorithms outperform their linear counterparts. This can be explained by the fact that the received data is non-circular and WL processing fully exploits the second-order statistics. Moreover, the SINR performance of the CCM-based full-rank algorithm is superior to that of the CMV-based one.

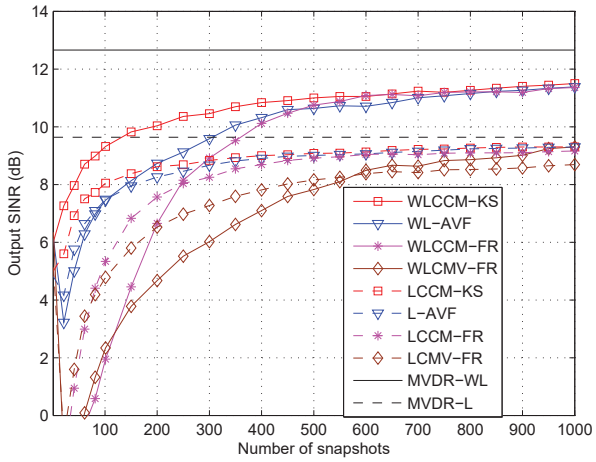

Fig. 9: Output SINR convergence performance (SNR= -3dB)

In Fig. 10, we show the steady-state SINR performance as a function of the input SNR, where the simulation scenario is the same as that in Fig. 9 and the reference snapshot is set as $N = 1000$. Generally, the SINR increases monotonically with the input SNR, but our proposed algorithm shows a better performance with a smaller gap from the optimum MVDR results. This conclusion is consistent with the results in Fig. 9. Besides, the WL algorithms yield an additional gain compared with the conventional linear algorithms.

In the next experiment, we analyze the robustness of the various algorithms under study to the minimum angular separation. Here, we assume there are two interfering users with DOAs $(50^\circ + \triangle\theta, 50^\circ - \triangle\theta)$, where $\triangle\theta$ is the minimum angular separation between the interferer and the desired user. Besides, two jammers impinge on the array with DOAs of $(10^\circ, 30^\circ)$. The input SNR is set to -3dB, and reference snapshot is set as $N = 1000$. As the results of Fig. 11 show, the WL algorithms are less sensitive to the minimum angular separation than the linear algorithms, which suffer a great performance degradation when the minimum angular separation becomes less than 7°.

E. MSE performance

In this part, we evaluate the MSE performance of the proposed WLCCM-KS algorithm and its linear counterpart. To this end, the MSE estimated by simulation is compared to the analytical steady-state MSE expression (59) derived in Section V-B. With respect to the approximation of the optimum CCM reduced-rank weight vector, we apply the steady-state reduced-rank weight vector as $\bar{\mathbf{w}}_o$, as obtained

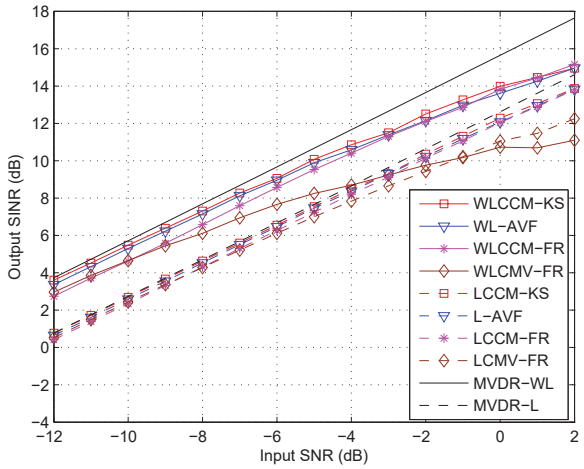

Fig. 10: Steady-state SINR versus the input SNR ($N = 1000$)

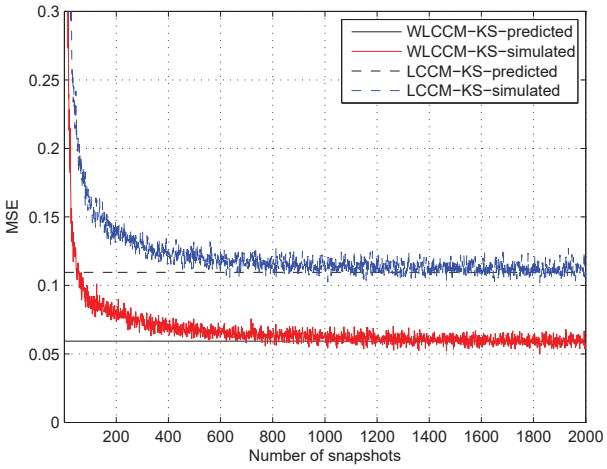

Fig. 12: MSE performance versus the number of snapshots (SNR= -3dB)

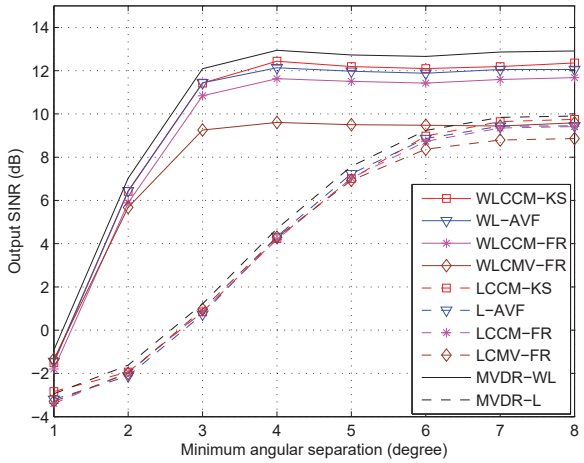

Fig. 11: Steady-state SINR versus the minimum angular separation between the interferer and the desired user ($N = 1000$, SNR= -3dB)

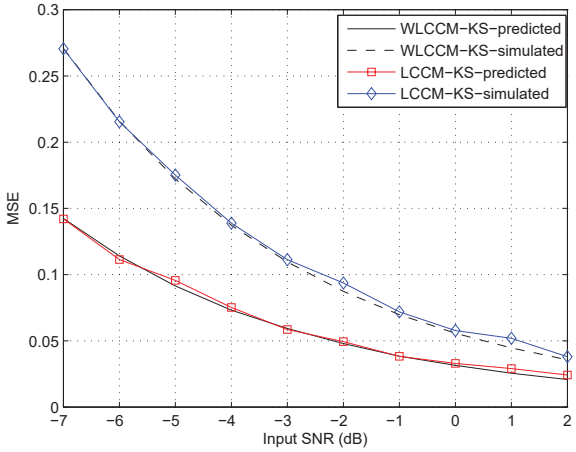

Fig. 13: Steady-state MSE performance versus the input SNR ($N = 2000$)

by averaging relevant quantities over independent simulation trials, and the corresponding optimum full-rank weight vector is given by $\tilde{\mathbf{w}}_o = \gamma\tilde{\mathbf{a}}(\theta_0) - \tilde{\mathbf{B}}\tilde{\mathbf{T}}_r\bar{\mathbf{w}}_o$. For the calculation of the analytical steady-state MSE of the linear version of the algorithm (LCCM-KS), the procedure is very similar, and we only need to replace the optimum WL weight vector with its linear counterpart. In Fig. 12, the input SNR is set to -3dB, while the interfering users and jammers are configured in the same way as in Fig. 9. We observe that as the number of snapshots increases, the simulated MSE decreases rapidly and converges to the analytical result. The steady-state MSE of WL processing is much smaller than its linear counterpart. Fig. 13 depicts the steady-state MSE performance as a function of the input SNR, where reference snapshot is set as $N = 2000$ to ensure steady-state condition. It can be seen that the analytical results and the simulation results agree well with each other.

VII. Conclusion

In this paper, we have proposed a novel blind adaptive reduced-rank WL beamforming algorithm with the GSC structure based on the KS technique for interference suppression in aeronautical communication systems. In order to reduce the computational complexity of the direct WL reduced-rank scheme, a structured KS based method has been devised to construct the transformation matrix needed for dimensionality reduction by considering the structure of the augmented covariance matrix. An RLS algorithm has been developed according to the WLCCM criterion to update the reduced-rank filter. In addition, we have analyzed the performance of the proposed WLCCM-KS-RLS algorithm in terms of computational complexity, convergence behvior and achievable SINR. The simulation results demonstrated the validity of the analytical results and showed that the proposed reduced-rank beamforming algorithm significantly outperforms existing full-rank and reduced-rank beamforming algorithms with a

relatively low complexity. Although in this work we employed BPSK signals to illustrate our ideas, in future work the proposed algorithm can be also extended to other types of non-circular modulated signals employed in aeronautical communications.

APPENDIX A
DERIVATION OF VALUE RANGE FOR γ

According to equation (1) and (18), the augmented vector $\tilde{\mathbf{r}}$ can be represented as

$$\tilde{\mathbf{r}} = \sum_{k=0}^{K-1} b_k \tilde{\mathbf{a}}(\theta_k) + \sum_{k=0}^{J-1} c_j^r \tilde{\mathbf{a}}(\phi_j) + \sum_{j=0}^{J-1} i c_j^i \breve{\mathbf{a}}(\phi_j) + \tilde{\mathbf{n}}, \quad (63)$$

where c_j^r and c_j^i denote the real and imaginary part of c_j, respectively. Here we drop the time index i for simplicity. The vector $\tilde{\mathbf{a}}(\theta_k)$ is the augmented steering vector, which satifies $\tilde{\mathbf{a}}(\theta_k) = \frac{1}{\sqrt{2}}[\tilde{\mathbf{a}}^T(\theta_k), \tilde{\mathbf{a}}^H(\theta_k)]^T \in \mathbb{C}^{2M\times 1}$, and we define $\breve{\mathbf{a}}(\theta_k) = \frac{1}{\sqrt{2}}[\tilde{\mathbf{a}}^T(\theta_k), -\tilde{\mathbf{a}}^H(\theta_k)]^T \in \mathbb{C}^{2M\times 1}$ as the quasi-augmented steering vector. Furthermore, we rewrite equation (63) in matrix form as follows for convenience:

$$\tilde{\mathbf{r}} = \tilde{\mathbf{A}}(\theta)\tilde{\mathbf{b}} + \tilde{\mathbf{n}}, \quad (64)$$

where $\tilde{\mathbf{A}}(\theta) = [\tilde{\mathbf{a}}(\theta_0), ..., \tilde{\mathbf{a}}(\theta_{K-1}), \tilde{\mathbf{a}}(\phi_0), ..., \tilde{\mathbf{a}}(\phi_0), \breve{\mathbf{a}}(\phi_0), ..., \breve{\mathbf{a}}(\phi_J - 1)]$,
and $\tilde{\mathbf{b}} = [b_0, ..., b_{K-1}, c_0^r, ..., c_{J-1}^r, ic_0^i, ..., ic_{J-1}^i]$. The CM cost function can be expressed as

$$\begin{aligned} J_{CM} &= \mathbb{E}[(|y|^2 - 1)^2] \\ &= \mathbb{E}[|y|^4] - 2\mathbb{E}[|y|^2] + 1, \end{aligned} \quad (65)$$

where the output

$$y = \tilde{\mathbf{w}}^H \mathbf{r} = \tilde{\mathbf{w}}^H \tilde{\mathbf{A}}(\theta)\tilde{\mathbf{b}} + \tilde{\mathbf{w}}^H \tilde{\mathbf{n}}, \quad (66)$$

and $\tilde{\mathbf{w}}$ denotes the equivalent full-rank weight vector for the proposed reduced-rank algorithm. For conciseness, we define $z_1 = \tilde{\mathbf{w}}^H \tilde{\mathbf{A}}(\theta)\tilde{\mathbf{b}} = \mathbf{u}^H(n)\mathbf{b}(n)$, $z_2 = \tilde{\mathbf{w}}^H \tilde{\mathbf{n}}$, where $\mathbf{u} = [u_0, ..., u_{K+2J-1}]^T$ and

$$u_k = \begin{cases} \tilde{\mathbf{a}}^H(\theta_k)\tilde{\mathbf{w}}, & 0 \le k \le K-1 \\ \tilde{\mathbf{a}}^H(\phi_{k-K})\tilde{\mathbf{w}}, & K \le k \le K+J-1 \\ \breve{\mathbf{a}}^H(\phi_{k-K-J})\tilde{\mathbf{w}}, & K+J \le k \le K+2J-1 \end{cases} \quad (67)$$

Thus, the output $y = z_1 + z_2$. Then, after some mathematical operations and simplifications, the CM cost function can be rewritten as

$$J_{CM} = J_1(\mathbf{u}) + \sigma_n^2 J_2(\tilde{\mathbf{w}}), \quad (68)$$

where

$$J_1(\mathbf{u}) = 2(\mathbf{u}^H\mathbf{u})^2 - \sum_{k=0}^{K+2J-1} u_k^4 - 2\mathbf{u}^H\mathbf{u} + 1, \quad (69)$$

$$J_2(\tilde{\mathbf{w}}) = (4\mathbf{u}^H\mathbf{u} - 2 + 3\sigma_n^2 \tilde{\mathbf{w}}^H\tilde{\mathbf{w}})\tilde{\mathbf{w}}^H\tilde{\mathbf{w}}. \quad (70)$$

Considering the constraint that the array response remains constant in the direction of the desired user, we define $E = u_0 u_0^* = \gamma^2 \|\tilde{\mathbf{a}}(\theta_0)\|^2$, and $\bar{\mathbf{u}} = [u_1, ..., u_{K+2J-1}]^T$. Then, the CM cost function can be expressed as

$$J_{CM} = J_1(\bar{\mathbf{u}}) + \sigma_n^2 J_2(\tilde{\mathbf{w}}), \quad (71)$$

where

$$J_1(\bar{\mathbf{u}}) = 2(E + \bar{\mathbf{u}}^H\mathbf{u})^2 - \left(E^2 + \sum_{k=1}^{K+2J-1} u_k^4\right) - 2\left(E + \mathbf{u}^H\mathbf{u}\right) + 1, \quad (72)$$

$$J_2(\tilde{\mathbf{w}}) = \left(4\left(E + \mathbf{u}^H\mathbf{u}\right) - 2 + 3\sigma_n^2 \tilde{\mathbf{w}}^H\tilde{\mathbf{w}}\right)\tilde{\mathbf{w}}^H\tilde{\mathbf{w}}. \quad (73)$$

To evaluate the convexity of J_{CM}, we compute its Hessian matrix using the rule $\mathbf{H} = \frac{\partial}{\partial \tilde{\mathbf{w}}^H}\frac{\partial J_{CM}}{\partial \tilde{\mathbf{w}}}$. This yields $\mathbf{H} = \mathbf{H}_1 + \sigma_n^2\mathbf{H}_2$, with

$$\mathbf{H}_1 = 4\bar{\mathbf{A}}[(E - 1/2)\mathbf{I} + \bar{\mathbf{u}}^H\bar{\mathbf{u}}\mathbf{I} + \bar{\mathbf{u}}\bar{\mathbf{u}}^H - \mathrm{diag}(|u_1|^2, ..., |u_{K+2J-1}|^2)]\bar{\mathbf{A}}^T, \quad (74)$$

$$\begin{aligned} \mathbf{H}_2 = (4E - 2)\mathbf{I} &+ 6\sigma_n^2(\tilde{\mathbf{w}}^H\tilde{\mathbf{w}}\mathbf{I} + \tilde{\mathbf{w}}\tilde{\mathbf{w}}^H) \\ &+ 4\big(\tilde{\mathbf{w}}^H\bar{\mathbf{A}}\bar{\mathbf{A}}^H\tilde{\mathbf{w}}\mathbf{I} + (\bar{\mathbf{A}}\bar{\mathbf{A}}^H)^T\tilde{\mathbf{w}}^H\tilde{\mathbf{w}} \\ &+ (\tilde{\mathbf{w}}\tilde{\mathbf{w}}^H\bar{\mathbf{A}}\bar{\mathbf{A}}^H)^T + (\tilde{\mathbf{w}}^H\bar{\mathbf{A}}\bar{\mathbf{A}}^H\tilde{\mathbf{w}})^T\big), \end{aligned} \quad (75)$$

where $\bar{\mathbf{A}}(\theta) = [\tilde{\mathbf{a}}(\theta_1), ..., \tilde{\mathbf{a}}(\theta_{K-1}), \tilde{\mathbf{a}}(\phi_0), ..., \tilde{\mathbf{a}}(\phi_0), \breve{\mathbf{a}}(\phi_0), ..., \breve{\mathbf{a}}(\phi_J - 1)]$.

We recall that $\mathbf{H}$ is positive definite if $\mathbf{c}^H\mathbf{H}\mathbf{c} > 0$ for all nonzero vector $\mathbf{c}$. For $\mathbf{H}_1$ expressed in (74), the sum of the second and fourth terms yields a positive definite matrix and the third term $\bar{\mathbf{u}}\bar{\mathbf{u}}^H$ is also positive definite. Thus, the first term provides the condition $E = \gamma^2\|\tilde{\mathbf{a}}(\theta_0)\|^2 \ge 1/2$, which ensures the convexity of $J_1(\bar{\mathbf{u}})$. As for $\mathbf{H}_2$ expressed in (75), it is easily seen that we can select a sufficiently large value of E such that $\mathbf{H}_2$ is positive definite in any bounded region.

Finally, we conclude that by properly selecting the constant γ such that $E = \gamma^2\|\tilde{\mathbf{a}}^H(\theta_0)\|^2 \ge 1/2$, which yields $\gamma \ge \frac{1}{\sqrt{2}\|\tilde{\mathbf{a}}(\theta_0)\|^2}$, the convexity of the CM cost function can be guaranteed, and the algorithm can reach the global minimum.

REFERENCES

[1] R. C. de Lamare, "Massive MIMO systems: signal processing challenges and future trends," *URSI Radio Science Bulletin*, 2013.

[2] S. K. Mohammed, A., A. Chockalingam, and B. S. Rajan, "High rate space-time coded large-MIMO systems: Low-complexity detection and channel estimation", *IEEE Journal Sel. Topics Signal Process.* , vol. 3, no. 6, pp. 958-974, Dec. 2009.

[3] J. Arnau, B. Devillers, C. Mosquera, and A. Pérez-Neira, "Performance study of multiuser interference mitigation schemes for hybrid broadband multibeam satellite architectures", *EURASIP J. on Wireless Commun. and Networking*, 2012:132, Apr. 2012.

[4] A. B. Gershman, E. Nemeth, and J. F. Bohme, "Experimental performance of adaptive beamforming in a sonar environment with a towed array and moving interfering sources," *IEEE Trans. Signal Process.* , vol. 48, no. 1, pp. 246-250, Jan. 2000.

[5] S. Anderson, M. Millnet, M. Viberg, and B. Wahlberg, "An adaptive array for mobile communication systems," *IEEE Trans. Vehi. Technol.* , vol.40, pp. 230-236, Feb. 1991.

[6] J. R. Guerci, J. S. Goldstein, and I. S. Reed, "Optimal and adaptive reduced-rank STAP," *IEEE Trans. Aerosp. Electron. Syst.* , vol. 36, no. 2, pp. 647-663, Apr. 2000.

[7] M. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 944-960, Jul. 1995.

[8] Z. Xu and K. T. Michail, "Blind adaptive algorithms for minimum variance CDMA receivers," *IEEE Trans. Commun.* , vol. 49, no. 1, pp. 180-194, 2001.

[9] R. C. de Lamare and R. Sampaio-Neto, "Low-complexity variable step-size mechanisms for stochastic gradient algorithms in minimum variance CDMA receivers," *IEEE Trans. Signal Process.* , vol. 54, no. 6, pp. 2302-2317, Jun. 2006.

[10] M. Gu and L. Tong, "Geometrical characterizations of constant mudulus receivers," *IEEE Trans. Signal Process.* , vol. 47, no. 10, pp. 2745-2756, Oct. 1999.

[11] R. C. de Lamare and R. Sampaio-Neto, "Blind adaptive code-constrained constant modulus algorithms for CDMA interference suppression in multipath channels," *IEEE Commun. Lett.* , vol. 9, no. 4, pp. 334-336, Apr. 2005.

[12] H. Zeng, L. Tong, and C. R. Johnson, "Relationships between the constant modulus and Wiener receivers," *IEEE Trans. Inf. Theory,* vol. 44, no. 4, pp. 1523-1538, Jul. 1998.

[13] Y. Cai and R. C. de Lamare, "Low-complexity variable step-size mechanism for code-constrained constant modulus stochastic gradient algorithms applied to CDMA interference suppression," *IEEE Trans. Signal Process.* , vol. 57, no. 1, pp. 313-323, Jan. 2009.

[14] Y. Cai, R. C. de Lamare, M. Zhao, and J. Zhong, "Low-complexity variable forgetting factor mechanism for blind adaptive constrained constant modulus algorithms," *IEEE Trans. Signal Process.* , vol. 60, no. 8, pp. 3988-4002, Aug. 2012.

[15] X. Jiang , W. -J. Zeng , A. Yasotharan , H. C. So and T. Kirubarajan , "Robust beamforming by linear programming," *IEEE Trans. Signal Process.* , vol. 62, no. 7, pp. 1834-1849, Mar. 2014.

[16] X. Jiang , W.-J. Zeng , A. Yasotharan , H. C. So and T. Kirubarajan, "Quadratically constrained minimum dispersion beamforming via gradient projection," *IEEE Trans. Signal Process.* , vol. 63, no. 1, pp. 192-205, Jan. 2015.

[17] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.* , vol. 43, no. 8, pp. 2030-2033, Aug. 1995.

[18] Y. Shi, L. Huang, C. Qian, and H.C. So, "Shrinkage linear and widely-linear complex-valued least mean squares algorithms for adaptive beamforming," *IEEE Trans. Signal Process.* , vol. 61, no. 1, pp. 119-131, Jan. 2015.

[19] Y. Wang, L. Huang, and Y. Shi, "Robust widely linear adaptive MVDR beamformer based on interference-plus-noise covariance matrix and steering vector estimation," *Proc. of IEEE China Summit and Int. Conf on Signal and Information Processing (ChinaSIP 2015)*, July 2015, Chengdu, China.

[20] J.-S. Lim, K.-Y. Han and J. Jeon, "Adaptive step-size widely linear linearly constrained constant modulus algorithm for DS-CDMA receivers in nonstationary interference environments," *Signal Processing*, vol. 87, no. 6, pp. 1523-1527, 2007.

[21] D. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models.* Hoboken, NJ: Wiley, 2009.

[22] P. J. Schreier and L. L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Trans. Signal Process. ,* vol. 51, no. 3, pp. 714-725, Mar. 2003.

[23] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals.* Cambridge, U.K.: Cambridge Univ, Press, 2010.

[24] A. M. Haimovich and Y. Bar-Ness, "An eigenanalysis interference canceler," *IEEE Trans. Signal Process. ,* vol. 39, no. 1, pp. 76-84, Jan. 1991.

[25] J. S. Goldstein and I. S. Reed, "Reduced-rank adaptive filtering," *IEEE Trans. Signal Process. ,* vol. 45, no. 2, pp. 492-496, Feb. 1997.

[26] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A multistage representation of the Wiener filter based on orthogonal projections," *IEEE Trans. Inf. Theory,* vol. 44, no. 11, pp. 2943-2959, Nov. 1998.

[27] M. L. Honig and J. S. Goldstein, "Adaptive reduced-rank interference suppression based on the multistage Wiener filter," *IEEE Trans. Commun.* , vol. 50, no. 6, pp. 986-994, Jun. 2002.

[28] X. Wang and H. V. Poor, "Blind multiuser detection: A subspace approach," *IEEE Trans. Inf. Theory*, vol. 44, pp. 677-690, Mar. 1998

[29] R. C. de Lamare, M. Haardt, and R. Sampaio-Neto, "Blind adaptive constrained reduced-rank parameter estimation based on constant modulus design for CDMA interference suppression," *IEEE Trans. Signal Process.* , vol. 56, no. 6, pp. 2470-2482, Jun. 2008.

[30] D. A. Pados and G. N. Karystinos, "An iterative algorithm for the computation of the MVDR filter," *IEEE Trans. Signal Process.* , vol. 49, No. 2, February, 2001.

[31] N. Song, R. C. de Lamare, M. Haardt, and M. Wolf, "Adaptive widely linear reduced-rank interference suppression based on the multi-stage wiener filter," *IEEE Trans. Signal Process.* , vol. 60, no. 8, pp. 4003-4016, Aug. 2012.

[32] N. Song, J. Steinwandt, L. Wang, R. C. de Lamare, and M. Haardt, "Non-data-aided adaptive beamforming algorithm based on the widely linear auxiliary vector filter," *Proc. Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Prague, Czech Repulic, May 2011.

[33] S. Haykin, *Adaptive Filter Theory*, 4th ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.

[34] J. S. Goldstein and I. S. Reed, "Theory of partially adaptive radar," *IEEE Trans. Aerosp. Electron. Syst.* , vol. 33, no. 4, pp. 1309-1325, Oct. 1997.

[35] S. Werner, M. With, and V. Koivunen, "Householder multistage Wiener filter for space-time navigation receivers," *IEEE Trans. Aerosp. Electron. Syst.* , vol. 43, no. 3, pp. 975-988, Jul. 2007.

[36] M. L. Honig and W. Xiao, "Adaptive reduced-rank interference suppression with adaptive rank selection," *Proc. MILCOM*, vol. 2, pp. 747-751, 2000.

[37] S. D. Somasundaram, P. Li, N. H. Parsons and R. C. de Lamare, "Reduced-dimension robust capon beamforming using Krylov-subspace techniques," *IEEE Trans. Aerosp. Electron. Syst.* , vol. 51, no. 1, pp. 270-289, Jul. 2015.

[38] P. Chevalier, J. P. Delmas, A. Oukaci, "Optimal widely linear MVDR beamformers for nocircular signals," *Proc. Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 3573-3576, Mar. 2009.

[39] T. Adal, P. J. Schreier and L. L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety" , *IEEE Trans. Signal Process.* , vol. 59, no. 11, pp. 5101-5125, Nov. 2011.

[40] E. Eleftheriou and D. D. Falconer, "Tracking properties and steady-state performance of RLS adaptive filter algorithms," *IEEE Trans. Acoustics Speech Signal Process.* , vol. 34, no. 5, pp. 1097-1110, Oct. 1986.

[41] P. H. Vincent and X. D. Wang, "Code-aided interference suppression for DS/CDMA communications - part II: Parallel blind adaptive implentations," *IEEE Trans. Commun.* , vol. 45, no. 9, pp. 1112-1122, Sep. 1997.

[42] J. J. Abularach A, C. Rodriguez R, M. P. C. de Almeida, L. da Silva Mello, G. Neto, F. Giacomini, "Coexistence of aeronautical mobile telemetry and IMT systems in the 1300–1518 MHz band," *Int. Workshop on Telecommunications (IWT)* pp. 1-5, Jun. 2015.