Reduced-Rank STAP Algorithms using Joint Iterative Optimization of Filters

RUI FA
RODRIGO C. DE LAMARE
University of York

We develop a reduced-rank space-time adaptive processing (STAP) method based on joint iterative optimization of filters (JOINT) for airborne radar applications. The proposed method consists of a bank of full-rank adaptive filters, which forms the projection matrix, and an adaptive reduced-rank filter that operates at the output of the bank of filters. We describe the proposed method for both the direct-form processor (DFP) and the generalized sidelobe canceller (GSC) structures. Adaptive algorithms including the stochastic gradient (SG), the recursive least square (RLS), and their hybrid algorithms are derived for the efficient implementation of the JOINT STAP method. The computational complexity analysis of the proposed algorithms is shown in terms of the number of multiplications and additions per snapshot. Furthermore, the convexity analysis of the proposed method is carried out. Simulations for a clutter-plus-jamming suppression application show that the proposed STAP algorithm outperforms the state-of-the-art reduced-rank schemes in convergence and tracking at significantly lower complexity.

Manuscript received November 27, 2008; revised April 16 and September 16, 2009; released for publication May 13, 2010.

IEEE Log No. T-AES/47/3/941755.

Refereeing of this contribution was handled by Y. Abramovich.

This work is funded by the Ministry of Defence (MoD), UK, Project MoD, Contract RT/COM/S/021.

Authors' address: Communications Research Group, Dept. of Electronics, University of York, Heslington, York, YO10 5DD, UK, E-mail: (rf533@ohm.york.ac.uk).

I. INTRODUCTION

A requirement of airborne surveillance radar systems is to detect moving targets in a severe and dynamic interference environment which may be composed of clutter and jamming. Space-time adaptive processing (STAP) has been motivated as a key enabling technology for advanced airborne radar applications following the landmark publication by Brennan and Reed [1]. By a joint-domain optimization of the spatial and temporal degrees-of-freedom (DOFs), STAP algorithms can improve slow-moving target detection through better mainlobe clutter suppression, provide better detection in combined clutter and jamming enviroment, and offer a significant increase in output signal-to-interference-plus-noise-ratio (SINR), as compared with traditional factored approaches [2]. However, the large computational complexity of the optimum STAP algorithm often prohibits its full-rank processing in real-time implementations. The computational complexity requirements of the optimal full-rank STAP algorithm have order $\mathcal{O}(M^3)$, where M is the dimension of the full-rank filter, due primarily to a covariance matrix inversion operation. An even more challenging problem, which is raised by conventional STAP techniques, is that when the number of elements in the filter is large, it requires a large number of samples, at least twice as many as the DOFs, to reach its steady state within 3 dB of the optimal performance. In dynamic scenarios, filters with large M usually fail or provide poor performance in tracking signals embedded in interference.

To address computational complexity and sample support issues, a diverse set of reduced-rank STAP techniques has been proposed. In the comprehensive report by Ward [2] and the book by Klemm [3], several factored reduced-dimension methods were developed to reduce the number of statistical unknowns associated with the interference based on forming "beams" in angle and/or Doppler. However, there is generally a price to be paid in interference cancellation performance. The first statistical reduced-rank method, which outperforms the conventional "beams"-type reduced-rank STAP techniques, was based on an eigenvector decomposition of the target-free covariance matrix, which is known as the principal-components (PC) method [4, 5]. By retaining the principal eigenvectors of the total target-free covariance matrix, the interference was indeed confined to a relatively low-rank subspace. Another class of eigen-decomposition methods was based on the cross-spectral metric (CSM) [6], [7]. Compared with the PC method, the CSM method yields a signal-dependent rank-ordering of the interference eigenvectors. However, the PC and the CSM methods have the problem of heavy computational load due

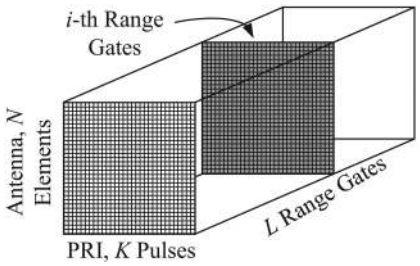

Fig. 1. Radar CPI datacube.

to the eigen-decomposition. The family of Krylov subspace methods has been investigated thoroughly in recent years. This class of reduced-rank adaptive filtering algorithms, including the multistage Wiener filter (MSWF) [8–10] and the auxiliary-vector filter (AVF) with orthogonal auxiliary vectors [11], projects the observation onto a lower dimensional Krylov subspace. Improvements to the AVF include the version with nonorthogonal auxiliary vectors [12, 13]. The joint domain localized (JDL) approach, which is a beamspace reduced-dimension algorithm, was proposed by Wang and Cai [24] and investigated in both homogeneous and nonhomogenous environments [25, 26]. Recently, the JDL algorithm was applied to the code-division multiple-access (CDMA) systems and was reported to outperform the MSWF algorithm with limited training [27].

In the present paper, we develop a reduced-rank approach to the STAP design utilizing a method based on joint iterative optimization of adaptive filters (JOINT) which was recently proposed in [14] and generalize the idea of adaptive interpolated filters with time-varying interpolators [15–17]. The proposed STAP algorithm consists of a bank of full-rank adaptive filters, which constitutes the projection matrix, and an adaptive reduced-rank filter that operates at the output of the bank of full-rank filters. The reduced-rank STAP algorithm employing the proposed JOINT scheme is investigated in both the direct-form processor (DFP) and the generalized sidelobe canceller (GSC) [18, 19] structures. We also derive the stochastic gradient (SG) and the recursive least squares (RLS) algorithms for their implementations. As a consequence, the computational complexity of the STAP with the SG algorithm is reduced substantially to $\mathcal{O}(MD)$ and the computational complexity of the STAP with the RLS algorithm is reduced to $\mathcal{O}(M^2)$ for each computing cycle of clock time, where D is the dimension of the reduced-rank filter.

This paper is organized as follows. Section II introduces the signal model, the optimum STAP algorithm and the fundamental of reduced-rank signal processing. The new reduced-rank STAP scheme is proposed and detailed in Section III, and the proposed adaptive algorithms and their complexity analysis are also presented in Section III. The convexity analysis of the proposed method and the study of its convergence properties are carried out in Section IV. Examples of the performance assessment of the proposed reduced-rank STAP are provided in Section V using both simulated and experimental radar data. Finally, conclusions are given in Section VI.

II. PROBLEM STATEMENT

The system under consideration is a pulsed Doppler radar residing on an airborne platform. The radar antenna is a uniformly spaced linear array antenna consisting of N elements. Radar returns are collected in a coherent processing interval (CPI), which is referred to as the 3-D radar datacube shown in Fig. 1, where L denotes the number of samples collected to cover the range interval. The data is then processed at one range of interest, which corresponds to a slice of the CPI datacube. This slice is a $K \times N$ matrix which consists of $N \times 1$ spatial snapshots for K pulses at the range of interest. It is convenient to stack the matrix column-wise to form the $M \times 1, M = KN$ vector $\mathbf{r}(i)$, termed the ith range gate space-time snapshot, $1 \leq i \leq L$ [1, 2].

A. Signal Model

The function of a radar is to ascertain whether targets are present in the data. Thus, given a space-time snapshot, radar detection is a binary hypothesis problem, where hypothesis $\mathbf{H}_0$ corresponds to target absence and hypothesis $\mathbf{H}_1$ corresponds to target presence. The radar space-time snapshot is then expressed for each of the two hypotheses in the following form,

$$\begin{aligned} \mathbf{H}_0 &: \quad \mathbf{r}(i) = \mathbf{v}(i) \\ \mathbf{H}_1 &: \quad \mathbf{r}(i) = a\mathbf{s} + \mathbf{v}(i) \end{aligned} \tag{1}$$

where a is a zero-mean complex Gaussian random variable with variance σ_s^2, $\mathbf{v}(i)$ denotes the input interference-plus-noise vector which consists of clutter $\mathbf{r}_c(i)$, jamming $\mathbf{r}_j(i)$, and the white noise $\mathbf{r}_n(i)$. These three components are assumed to be mutually uncorrelated. Thus, the $M \times M$ covariance matrix $\mathbf{R}$ of the undesired clutter-plus-jammer-plus-noise component can be modeled as

$$\mathbf{R} = \mathbb{E}\{\mathbf{v}(i)\mathbf{v}^{\mathrm{H}}(i)\} = \mathbf{R}_c + \mathbf{R}_j + \mathbf{R}_n \tag{2}$$

where H represents Hermitian transpose, $\mathbf{R}_c = \mathbb{E}\{\mathbf{r}_c(i)\mathbf{r}_c^{\mathrm{H}}(i)\}$, $\mathbf{R}_j = \mathbb{E}\{\mathbf{r}_j(i)\mathbf{r}_j^{\mathrm{H}}(i)\}$ and $\mathbf{R}_n = \mathbb{E}\{\mathbf{r}_n(i)\mathbf{r}_n^{\mathrm{H}}(i)\}$ denote clutter, jamming, and noise covariance matrix, respectively, and $\mathbb{E}$ denotes expectation. The vector $\mathbf{s}$, which is the $M \times 1$ normalized space-time steering vector in the space-time look-direction, can be defined as

$$\mathbf{s} = \mathbf{b}(\varpi_t) \otimes \mathbf{a}(\vartheta_t) \tag{3}$$

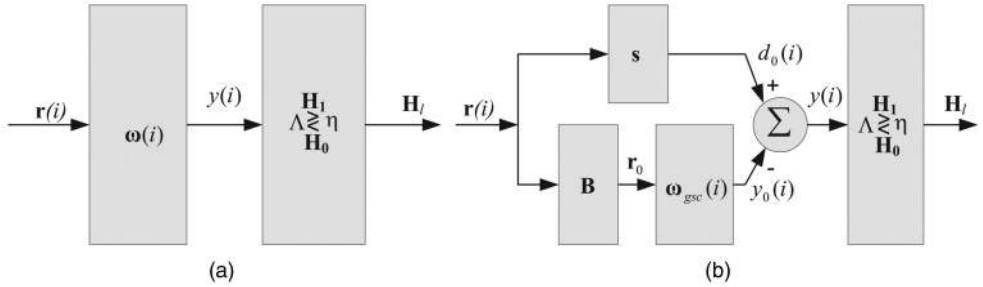

Fig. 2. (a) Full-rank direct-form processor. (b) Full-rank GSC processor.

where $\mathbf{b}(\varpi_t)$ is the $K \times 1$ normalized temporal steering vector at the target Doppler frequency ϖ_t and $\mathbf{a}(\vartheta_t)$ is the $N \times 1$ normalized spatial steering vector in the direction provided by the target spatial frequency ϑ_t. The notation $\otimes$ denotes Kronecker product.

B. Optimum Radar Signal Processing

To detect the presence of targets, each range bin is processed by an adaptive 2D beamformer (to achieve maximum output SINR) followed by a hypothesis test to determine the target presence or absence. The STAP algorithm normally can be implemented in two different structures: one is the DFP structure and the other is the GSC structure whose diagrams are depicted in Fig. 2(a) and (b) respectively. In Fig. 2(a), $\{\mathbf{r}(i)\}_{i=1}^{L}$ denotes the secondary data, which is a target-free signal. We assume that the secondary data $\{\mathbf{r}(i)\}_{i=1}^{L}$ are independent and identically distributed (IID). $y(i) = \boldsymbol{\omega}^{\mathrm{H}}(i)\mathbf{r}(i)$ is the output of the STAP algorithm, where $\boldsymbol{\omega}^{\mathrm{H}}(i)$ is the STAP filter. The optimum full-rank STAP algorithm for DFP[1] obtained by an unconstrained optimization of the SINR is given as follows:

$$\boldsymbol{\omega}_{\mathrm{opt}} = k\mathbf{R}^{-1}\mathbf{s} \tag{4}$$

where k is an arbitrary non-zero complex number. The optimal constrained weight vector for maximizing the output SINR, while maintaining a normalized response in the target spatial-Doppler look-direction was originally given in [18] by

$$\boldsymbol{\omega}_{\mathrm{opt}} = \frac{\mathbf{R}^{-1}\mathbf{s}}{\mathbf{s}^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{s}}. \tag{5}$$

The solution in (5) can also be obtained by solving the linearly constrained minimum variance (LCMV) problem as [36]

$$\boldsymbol{\omega}_{\mathrm{opt}} = \arg\min_{\boldsymbol{\omega}(i)} \boldsymbol{\omega}^{\mathrm{H}}(i)\mathbf{R}\boldsymbol{\omega}(i) \qquad \text{subject to} \quad \mathbf{s}^{\mathrm{H}}\boldsymbol{\omega}(i) = 1. \tag{6}$$

For the GSC structure as shown in Fig. 2(b), the output d_0 of the radar mainbeam is provided by

$$d_0(i) = \mathbf{s}^{\mathrm{H}}\mathbf{r}(i) \tag{7}$$

and the interference-plus-noise subspace data vector $\mathbf{r}_0(i)$ is defined to be an $(M-1) \times 1$ vector at the output of the signal blocking matrix given by

$$\mathbf{r}_0(i) = \mathbf{B}\mathbf{r}(i) \tag{8}$$

where $\mathbf{B}$ denotes the $(M-1) \times M$ signal blocking matrix which can be directly obtained by using the singular value decomposition (SVD) and the QR decomposition algorithms [20]. Thus, $\mathbf{Bs} = \mathbf{0}_{(M-1)\times 1}$, so that the matrix $\mathbf{B}$ effectively blocks any signal coming from the spatial-Doppler look-direction. Let us consider the equivalent transformation defined by the operator $\mathbf{T} = [\mathbf{s}, \mathbf{B}^{\mathrm{H}}]^{\mathrm{H}}$. The transformation of radar return $\mathbf{r}(i)$ yields a vector $\tilde{\mathbf{r}}(i)$ which has the form

$$\tilde{\mathbf{r}}(i) = \mathbf{T}\mathbf{r}(i) = \begin{bmatrix} \mathbf{s}^{\mathrm{H}}\mathbf{r}(i) \\ \mathbf{B}\mathbf{r}(i) \end{bmatrix} = \begin{bmatrix} d_0(i) \\ \mathbf{r}_0(i) \end{bmatrix} \tag{9}$$

and the associated covariance matrix $\mathbf{R}_{\tilde{r}}$ given by

$$\mathbf{R}_{\tilde{r}} = \mathbb{E}[\tilde{\mathbf{r}}(i)\tilde{\mathbf{r}}^{\mathrm{H}}(i)] = \mathbf{T}\mathbf{R}\mathbf{T}^{\mathrm{H}} = \begin{bmatrix} \sigma_0^2 & \mathbf{p}_0^{\mathrm{H}} \\ \mathbf{p}_0 & \mathbf{R}_0 \end{bmatrix} \tag{10}$$

where $\sigma_0^2 = \mathbf{s}^{\mathrm{H}}\mathbf{R}\mathbf{s}$, the $(M-1) \times 1$ cross-correlation vector between the noise subspace data vector and the beamformer output $\mathbf{p}_0 = \mathbb{E}\{\mathbf{r}_0(i)d_0^*(i)\}$ and the $(M-1) \times (M-1)$ noise subspace covariance matrix $\mathbf{R}_0 = \mathbf{B}\mathbf{R}\mathbf{B}^{\mathrm{H}}$. Due to the transformation on the steering vector yielding the unit transformed steering vector $\mathbf{e}_1$ as

$$\mathbf{e}_1 = \mathbf{T}\mathbf{s} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{11}$$

the optimal weight vector for GSC in these transformed coordinates is given by

$$\boldsymbol{\omega}'_{\mathrm{opt}} = \frac{\mathbf{R}_{\tilde{r}}^{-1}\mathbf{e}_1}{\mathbf{e}_1^{\mathrm{H}}\mathbf{R}_{\tilde{r}}^{-1}\mathbf{e}_1} = \begin{bmatrix} 1 \\ -\boldsymbol{\omega}_{\mathrm{gsc}} \end{bmatrix} \tag{12}$$

where $\boldsymbol{\omega}_{\mathrm{gsc}} = \mathbf{R}_0^{-1}\mathbf{p}_0$. Thus the optimal weight vector may be expressed as

$$\boldsymbol{\omega}_{\mathrm{opt}} = \mathbf{T}^{\mathrm{H}}\boldsymbol{\omega}'_{\mathrm{opt}} = \mathbf{s} - \mathbf{B}^{\mathrm{H}}\boldsymbol{\omega}_{\mathrm{gsc}}. \tag{13}$$

C. Reduced-Rank Signal Processing

The diagrams of reduced-rank processors with the DFP and the GSC structures are depicted in

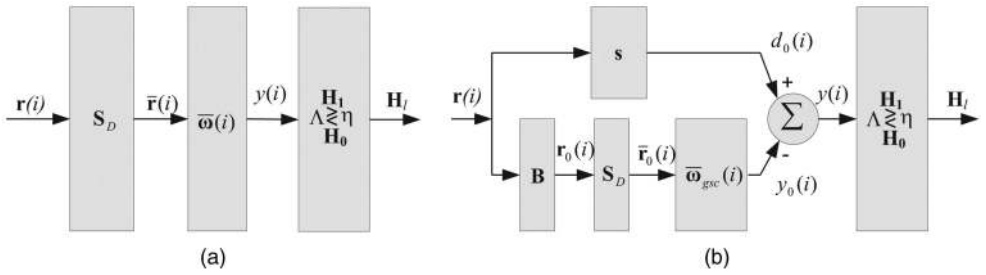

Fig. 3. (a) Reduced-rank DFP. (b) Reduced-rank GSC processor.

Fig. 3(a) and (b), respectively. The basic idea of reduced-rank algorithms is to reduce the number of adaptive coefficients by projecting the received vectors onto a lower dimensional subspace as illuminated in the figure. Considering the DFP structure, let $\mathbf{S}_D$ denote the $M \times D$ projection matrix with column vectors which are an $M \times 1$ basis for a D-dimensional subspace, where $D < M$. Thus, the received signal $\mathbf{r}(i)$ is transformed into its reduced-rank version $\bar{\mathbf{r}}(i)$ given by

$$\bar{\mathbf{r}}(i) = \mathbf{S}_D^{\mathrm{H}}\mathbf{r}(i). \tag{14}$$

The reduced-rank signal is processed by an adaptive reduced-rank filter $\bar{\omega}(i) \in \mathcal{C}^{D\times 1}$. Subsequently, the decision is made based on the filter output $y(i)$. The optimum minimum variance distortionless response (MVDR) solution for the reduced-rank weight vector $\bar{\omega}_{\mathrm{opt}}$ is obtained by [28]

$$\begin{aligned} \bar{\omega}_{\mathrm{opt}} &= \arg\min_{\bar{\omega}} \bar{\omega}^{\mathrm{H}}\bar{\mathbf{R}}\bar{\omega} \qquad \text{subject to} \quad \bar{\omega}^{\mathrm{H}}\bar{\mathbf{s}} = 1 \\ \Rightarrow \bar{\omega}_{\mathrm{opt}} &= \frac{\bar{\mathbf{R}}^{-1}\bar{\mathbf{s}}}{\bar{\mathbf{s}}^{\mathrm{H}}\bar{\mathbf{R}}^{-1}\bar{\mathbf{s}}} \end{aligned} \tag{15}$$

where $\bar{\mathbf{R}} = \mathbf{S}_D^{\mathrm{H}}\mathbf{R}\mathbf{S}_D$ denotes the reduced-rank covariance matrix and $\bar{\mathbf{s}} = \mathbf{S}_D^{\mathrm{H}}\mathbf{s}$ denotes the reduced-rank steering vector.

The reduced-rank adaptive GSC shown in Fig. 3(b) utilizes an $(M-1) \times D$ projection matrix $\mathbf{S}_D$ to form $D \times 1$ reduced-rank observation data vector

$$\bar{\mathbf{r}}_0(i) = \mathbf{S}_D^{\mathrm{H}}\mathbf{r}_0(i) \tag{16}$$

and the associated $D \times D$ reduced-rank covariance matrix is given by

$$\bar{\mathbf{R}}_0 = \mathbf{S}_D^{\mathrm{H}}\mathbf{R}_0\mathbf{S}_D. \tag{17}$$

The reduced-rank data vector $\bar{\mathbf{r}}_0(i) = \mathbf{S}_D^{\mathrm{H}}\mathbf{r}_0(i)$ is then processed by the $D \times 1$ reduced-rank filter $\bar{\omega}_{\mathrm{gsc}}(i)$, which can be optimized to minimize the squared error between the mainbeam output $d_0(i)$ and the sidelobe output $y_0(i)$ as

$$\begin{aligned} \bar{\omega}_{\mathrm{gsc}} &= \arg\min_{\bar{\omega}(i)} \mathbb{E}\{|d_0(i) - \bar{\omega}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}\mathbf{r}_0(i)|^2\} \\ \Rightarrow \bar{\omega}_{\mathrm{gsc}} &= \bar{\mathbf{R}}_0^{-1}\bar{\mathbf{p}}_d \end{aligned} \tag{18}$$

where $\bar{\mathbf{p}}_d = \mathbf{S}_D^{\mathrm{H}}\mathbf{B}\mathbf{R}\mathbf{s}$ denotes the reduced-rank cross-correlation vector. Since the equivalent transformation matrix $\bar{\mathbf{T}}$ with $(D+1) \times M$-dimension is given by

$$\bar{\mathbf{T}} = \begin{bmatrix} \mathbf{s}^{\mathrm{H}} \\ \mathbf{S}_D^{\mathrm{H}}\mathbf{B} \end{bmatrix}$$

the optimal reduced-rank weight vector for the GSC with $D \times 1$-dimension in the transformed domain is given by

$$\bar{\omega}_{\mathrm{opt}} = \begin{bmatrix} 1 \\ -\bar{\omega}_{\mathrm{gsc}} \end{bmatrix}.$$

The challenge left to us is how to efficiently design and optimize the projection matrix $\mathbf{S}_D$. The PC method which is also known as the eigencanceller method [4] suggested to form the projection matrix using the eigenvectors of the covariance matrix $\mathbf{R}$ corresponding to the eigenvalues with significant magnitude. The CSM method, a counterpart of the PC method belonging to the eigen-decomposition algorithm family, outperforms the PC method because it employs the projection matrix which contains the eigenvectors which contribute the most towards maximizing the SINR [7]. A family of closely related reduced-rank adaptive filters, such as the MSWF [8] and the AVF [11] algorithms, employs a set of basis vectors as the projection matrix which spans the same subspace, known as the Krylov subspace [29, 30]. The Krylov subspace is generated by taking the powers of the covariance matrix of observations on a cross-correlation (or steering) vector. Despite the improved convergence and tracking performance achieved with these methods, the remained problem is their high complexity and the existence of numerical problems for implementation.

III. THE PROPOSED STAP ALGORITHM

In this section, we present the principles of the proposed reduced-rank scheme based on the joint iterative optimization of adaptive filters (JOINT) scheme for both the DFP and the GSC structures. We also develop efficient and low-complexity adaptive implementations for our proposed scheme using the SG, the RLS, and the hybrid algorithms and we compare the complexity of the proposed JOINT scheme with other existing algorithms, namely, the full-rank filters, the JDL, the MSWF and the AVF

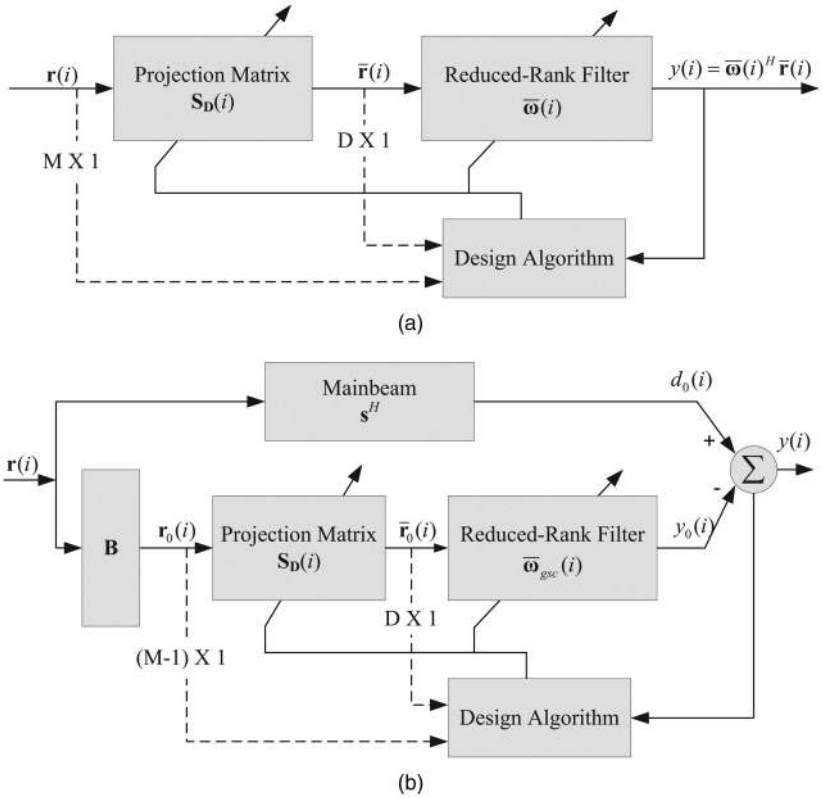

Fig. 4. Proposed JOINT STAP schematic diagrams. (a) DFP structure. (b) GSC structure.

algorithms, in terms of multiplications and additions per snapshot. Furthermore, we consider the important issue of rank selection and propose an adaptive technique to automatically determine the rank.

A. Proposed Reduced-Rank STAP Algorithm

ere we detail the principles of the proposed reduced-rank scheme using a projection operator based on adaptive filters. For the DFP structure which is depicted in Fig. 4(a), our proposed JOINT algorithm employs a projection matrix $\mathbf{S}_D(i)$ with dimensions $M \times D$, which is responsible for the dimensionality reduction, to project the $M \times 1$ data vector $\mathbf{r}(i)$ onto the rank-reduced data vector $\bar{\mathbf{r}}(i)$. The $D \times 1$-dimensional reduced-rank filter $\bar{\omega}(i)$ linearly combines the vector $\bar{\mathbf{r}}(i)$ in order to yield a scalar estimate $y(i)$. The projection matrix $\mathbf{S}_D(i)$ and the reduced-rank filter $\bar{\omega}(i)$ are jointly optimized in the proposed scheme according to the constrained minimum variance (CMV) criterion. Specifically, the projection matrix is structured as a bank of D full-rank filters $\mathbf{s}_d(i) = [s_{1,d}(i), s_{2,d}(i), \ldots, s_{M,d}(i)]^T, (d = 1, \ldots, D)$ with dimensions $M \times 1$ as given by $\mathbf{S}_D(i) = [\mathbf{s}_1(i)\ \mathbf{s}_2(i) \cdots \mathbf{s}_D(i)]$. Let us now mathematically express the output estimate $y(i)$ of the reduced-rank scheme as a function of the received data $\mathbf{r}(i)$, the projection matrix $\mathbf{S}_D(i)$ and the reduced-rank filter $\bar{\omega}(i)$:

$$y(i) = \bar{\omega}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}(i) = \bar{\omega}^{\mathrm{H}}(i)\bar{\mathbf{r}}(i). \tag{19}$$

Note that for $D = 1$, the novel scheme becomes a conventional full-rank filtering scheme with an addition weight parameter $\bar{\omega}_D$ that provides a gain. For $D > 1$, the signal processing tasks are changed and the full-rank filters compute a subspace projection and the reduced-rank filter estimates the desired signal. The CMV expression for the filters $\mathbf{S}_D(i)$ and $\bar{\omega}(i)$ can be computed via the constrained optimization problem

$$\begin{aligned} [\bar{\omega}_{\mathrm{opt}}, \mathbf{S}_{D,opt}] &= \arg\min_{\bar{\omega}(i), \mathbf{S}_D(i)} \mathbb{E}[|\bar{\omega}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}(i)|^2] \\ \text{subject to} \quad & \bar{\omega}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{s} = 1 \end{aligned} \tag{20}$$

where $\bar{\omega}(i) \in \mathcal{C}^D, \mathbf{S}_D(i) \in \mathcal{C}^{M \times D}$. The constrained optimization problem in (20) can be transformed by the method of Lagrange multipliers into an unconstrained optimization problem whose cost function is

$$\mathcal{L}_{\mathrm{MV}} = \mathbb{E}[|\bar{\omega}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}(i)|^2] + 2\Re[\lambda^*(\bar{\omega}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{s} - 1)] \tag{21}$$

where λ is a scalar Lagrange multiplier, $*$ denotes complex conjugate, and the operator $\Re[\cdot]$ selects the real part of the argument. By fixing $\bar{\omega}(i)$, minimizing (21) with respect to $\mathbf{S}_D(i)$ and solving for λ, we get

$$\mathbf{S}_D(i) = \frac{\mathbf{R}^{-1}\mathbf{s}\bar{\omega}^{\mathrm{H}}(i)\mathbf{R}_{\bar{\omega}}^{-1}}{\bar{\omega}^{\mathrm{H}}(i)\mathbf{R}_{\bar{\omega}}^{-1}\bar{\omega}(i)\mathbf{s}^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{s}} \tag{22}$$

where $\mathbf{R} = \mathbb{E}[\mathbf{r}(i)\mathbf{r}^{\mathrm{H}}(i)]$, $\mathbf{R}_{\bar{\omega}} = \mathbb{E}[\bar{\omega}(i)\bar{\omega}^{\mathrm{H}}(i)]$. By fixing $\mathbf{S}_D(i)$, minimizing (21) with respect to $\bar{\omega}(i)$, and

solving for λ, we arrive at the expression for $\bar{\omega}(i)$

$$\bar{\omega}(i) = \frac{\bar{\mathbf{R}}^{-1}(i)\bar{\mathbf{s}}}{\bar{\mathbf{s}}^{\mathrm{H}}\bar{\mathbf{R}}^{-1}(i)\bar{\mathbf{s}}} \tag{23}$$

where $\bar{\mathbf{R}}(i) = \mathbf{S}_D^{\mathrm{H}}(i)\mathbf{R}\mathbf{S}_D(i)$, $\bar{\mathbf{s}} = \mathbf{S}_D^{\mathrm{H}}(i)\mathbf{s}$.

For the GSC structure, the proposed algorithm is illustrated in Fig. 4(b). As mentioned in the previous section, the reduced-rank adaptive GSC utilizes an $(M-1)\times D$ projection matrix $\mathbf{S}_D$ to project the target-free subspace data vector $\mathbf{r}_0(i)$ to a reduced-rank subspace vector $\bar{\mathbf{r}}_0(i) = \mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}_0(i)$. In our design, the project matrix $\mathbf{S}_D(i)$ and the reduced-rank weight vector $\bar{\omega}_{\mathrm{gsc}}(i)$ are jointly optimized by minimizing the squared error between the mainbeam output $d_0(i)$ and sidelobe output $y_0(i)$ as

$$[\bar{\omega}_{\mathrm{gsc,opt}}, \mathbf{S}_{D,\mathrm{opt}}] = \underset{\bar{\omega}_{\mathrm{gsc}}(i),\mathbf{S}_D(i)}{\arg\min}\ \mathbb{E}\{|d_0(i) - \bar{\omega}_{\mathrm{gsc}}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}_0(i)|^2\} \tag{24}$$

where $\bar{\omega}_{\mathrm{gsc}}(i) \in \mathcal{C}^D, \mathbf{S}_D(i) \in \mathcal{C}^{(M-1)\times D}$. The cost function of this optimization problem can be written as

$$\mathcal{L} = \mathbb{E}[|d_0(i) - \bar{\omega}_{\mathrm{gsc}}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}_0(i)|^2]. \tag{25}$$

By fixing $\bar{\omega}_{\mathrm{gsc}}(i)$, minimizing (25) with respect to $\mathbf{S}_D(i)$, we get the optimum projection matrix for the GSC structure

$$\mathbf{S}_D(i) = \mathbf{R}_0^{-1}\mathbf{p}_0\bar{\omega}_{\mathrm{gsc}}^{\mathrm{H}}(i)\mathbf{R}_{\bar{\omega}_{\mathrm{gsc}}}^{-1} \tag{26}$$

where $\mathbf{R}_0 = \mathbb{E}[\mathbf{r}_0(i)\mathbf{r}_0(i)^{\mathrm{H}}]$, $\mathbf{p}_0 = \mathbb{E}[\mathbf{r}_0(i)d_0^*(i)]$ and $\mathbf{R}_{\bar{\omega}_{\mathrm{gsc}}} = \mathbb{E}[\bar{\omega}_{\mathrm{gsc}}(i)\bar{\omega}_{\mathrm{gsc}}^{\mathrm{H}}(i)]$. By fixing $\mathbf{S}_D(i)$, minimizing (25) with respect to $\bar{\omega}_{\mathrm{gsc}}(i)$, we may obtain

$$\bar{\omega}_{\mathrm{gsc}}(i) = \bar{\mathbf{R}}_0^{-1}\bar{\mathbf{p}}_0 \tag{27}$$

where $\bar{\mathbf{R}}_0 = \mathbf{S}_D^{\mathrm{H}}(i)\mathbf{R}_0\mathbf{S}_D(i)$ and $\bar{\mathbf{p}}_0 = \mathbf{S}_D^{\mathrm{H}}(i)\mathbf{p}_0$.

Note that the filter expressions in (22) and (23) for the DFP structure as well as (26) and (27) for the GSC structure are not closed-form solutions for $\mathbf{S}_D(i)$ and $\bar{\omega}(i)$ since (22) and (26) are functions of $\bar{\omega}(i)$ and $\bar{\omega}_{\mathrm{gsc}}(i)$, respectively, and (23) and (27) depend on $\mathbf{S}_D(i)$. Thus it is necessary to iterate (22) and (23), (26), and (27) with an initial guess to obtain a solution. Unlike the MSWF [8] and the AVF [13] approaches which project the data onto the Krylov subspace, the proposed scheme provides an iterative exchange of information between the reduced-rank filter and the projection matrix and leads to a much simpler adaptive implementation than the MSWF and the AVF algorithms. The projection matrix reduces the dimension of the input data, whereas the reduced-rank filter attempts to estimate the desired signal. The key strategy lies in the joint iterative optimization of the filters. In the next section, we seek iterative solutions via adaptive algorithms.

B. Adaptive Implementations

Adaptive implementations of the LCMV beamformer were subsequently reported with the SG [31] and the RLS algorithms [32]. Here we describe the SG, the RLS, and a hybrid of SG and RLS algorithms [33] that adjust the parameters of the filter bank and the reduced-rank filter for the DFP structure based on the constrained minimization of the minimum value (MV) cost function, and for the GSC structure based on the minimum mean squared error (MMSE) criterion, respectively.

1) *Stochastic Gradient Algorithm*: By computing the gradient terms of (21) with respect to $\mathbf{S}_D^*(i)$ and $\bar{\omega}^*(i)$ for the DFP structure, we get

$$\begin{aligned}\nabla\mathcal{L}_{\mathrm{MV}_{\mathbf{S}_D^*(i)}}(i) &= y^*(i)\mathbf{r}(i)\bar{\omega}^{\mathrm{H}}(i) + 2\lambda^*\mathbf{s}\bar{\omega}^{\mathrm{H}}(i)\\ \nabla\mathcal{L}_{\mathrm{MV}_{\bar{\omega}^*(i)}}(i) &= y^*(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}(i) + 2\lambda^*\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{s}.\end{aligned} \tag{28}$$

By solving the above equations and introducing the convergence factors μ_ω and μ_s, the proposed jointly optimized and iterative SG algorithms for parameter estimation become

$$\mathbf{S}_D(i+1) = \mathbf{S}_D(i) - \mu_s y^*(i)[\mathbf{I} - \mathbf{s}\mathbf{s}^{\mathrm{H}}]\mathbf{r}(i)\bar{\omega}^{\mathrm{H}}(i) \tag{29}$$

$$\bar{\omega}(i+1) = \bar{\omega}(i) - \mu_\omega y^*(i)\left[\mathbf{I} - \frac{\bar{\mathbf{s}}\bar{\mathbf{s}}^{\mathrm{H}}}{\|\bar{\mathbf{s}}\|^2}\right]\bar{\mathbf{r}}(i) \tag{30}$$

where $y(i) = \bar{\omega}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}(i)$.

Subsequently, let us consider the GSC structure. The instantaneous gradient terms of (25) with respect to $\mathbf{S}_D^*(i)$ and $\bar{\omega}_{\mathrm{gsc}}^*(i)$ are given by

$$\begin{aligned}\nabla\mathcal{L}_{\mathbf{S}_D^*(i)}(i) &= (d_0(i) - \bar{\omega}_{\mathrm{gsc}}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}_0(i))^*\mathbf{r}_0(i)\bar{\omega}_{\mathrm{gsc}}^{\mathrm{H}}(i)\\ \nabla\mathcal{L}_{\bar{\omega}_{\mathrm{gsc}}^*(i)}(i) &= (d_0(i) - \bar{\omega}_{\mathrm{gsc}}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}_0(i))^*\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}_0(i).\end{aligned} \tag{31}$$

Thus, the projection matrix $\mathbf{S}_D(i)$ and the reduced-rank GSC weight vector $\bar{\omega}_{\mathrm{gsc}}(i)$ can be jointly and iteratively optimized as

$$\mathbf{S}_D(i+1) = \mathbf{S}_D(i) + \mu_s y^*(i)\mathbf{r}_0(i)\bar{\omega}_{\mathrm{gsc}}^{\mathrm{H}}(i) \tag{32}$$

$$\bar{\omega}_{\mathrm{gsc}}(i+1) = \bar{\omega}_{\mathrm{gsc}}(i) + \mu_\omega y^*(i)\bar{\mathbf{r}}_0(i) \tag{33}$$

where $y(i) = d_0(i) - \bar{\omega}_{\mathrm{gsc}}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}_0(i)$. The proposed scheme trades-off a full-rank filter against one projection matrix $\mathbf{S}_D(i)$ and one reduced-rank adaptive filter $\bar{\omega}(i)$ or $\bar{\omega}_{\mathrm{gsc}}(i)$ operating simultaneously and exchanging information. The JOINT schemes using the SG algorithm for the DFP and the GSC structures are summarized in Tables I and II, respectively.

2) *Recursive Least Squares Algorithm*: Here we derive the RLS reduced-rank adaptive algorithms for the efficient implementation of the proposed scheme using both the DFP and the GSC strucutures. To this end, let first consider that for the DFP structure whose

TABLE I
The JOINT Scheme using SG Algorithm for the DFP

Initialization	$\mathbf{S}_D(1) = [\mathbf{I}_{D\times D} \quad \mathbf{0}_{D\times(M-D)}]^T$
	$\bar{\mathbf{s}} = \mathbf{S}_D^{\mathrm{H}}(1)\mathbf{s},\ \bar{\boldsymbol{\omega}}(1) = \bar{\mathbf{s}}/\|\bar{\mathbf{s}}\|$ and $i = 1$
Recursion	$y(i) = \bar{\boldsymbol{\omega}}(i)^{\mathrm{H}}\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}(i)$
	$\mathbf{S}_D(i+1) = \mathbf{S}_D(i) - \mu_s y^*(i)[\mathbf{I} - \mathbf{s}\mathbf{s}^{\mathrm{H}}]\mathbf{r}(i)\bar{\boldsymbol{\omega}}^{\mathrm{H}}(i)$
	$y(i) = \bar{\boldsymbol{\omega}}(i)^{\mathrm{H}}\mathbf{S}_D^{\mathrm{H}}(i+1)\mathbf{r}(i)$
	$\bar{\mathbf{s}} = \mathbf{S}_D^{\mathrm{H}}(i+1)\mathbf{s},\ \bar{\mathbf{r}}(i) = \mathbf{S}_D^{\mathrm{H}}(i+1)\mathbf{r}(i)$
	$\bar{\boldsymbol{\omega}}(i+1) = \bar{\boldsymbol{\omega}}(i) - \mu_\omega y^*(i)\left[\mathbf{I} - \dfrac{\bar{\mathbf{s}}\bar{\mathbf{s}}^{\mathrm{H}}}{\|\bar{\mathbf{s}}\|^2}\right]\bar{\mathbf{r}}(i)$

TABLE II
The JOINT Scheme using SG Algorithm for the GSC

Initialization	$\mathbf{S}_D(1) = [\mathbf{I}_{D\times D} \quad \mathbf{0}_{D\times(M-1-D)}]^T$
	$\bar{\boldsymbol{\omega}}(1) = \mathbf{0}$ and $i = 1$
Recursion	$d_0(i) = \mathbf{s}^{\mathrm{H}}\mathbf{r}(i),\ \mathbf{r}_0(i) = \mathbf{B}\mathbf{r}(i)$
	$y(i) = d_0(i) - \bar{\boldsymbol{\omega}}_{\mathrm{gsc}}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}_0(i)$
	$\mathbf{S}_D(i+1) = \mathbf{S}_D(i) + \mu_s y^*(i)\mathbf{r}_0(i)\bar{\boldsymbol{\omega}}_{\mathrm{gsc}}^{\mathrm{H}}(i)$
	$y(i) = d_0(i) - \bar{\boldsymbol{\omega}}_{\mathrm{gsc}}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i+1)\mathbf{r}_0(i)$
	$\bar{\mathbf{r}}_0(i) = \mathbf{S}_D^{\mathrm{H}}(i+1)\mathbf{r}_0(i)$
	$\bar{\boldsymbol{\omega}}_{\mathrm{gsc}}(i+1) = \bar{\boldsymbol{\omega}}_{\mathrm{gsc}}(i) + \mu_\omega y^*(i)\bar{\mathbf{r}}_0(i)$

least squares (LS) unconstrained cost function is given by

$$\mathcal{L}_{\mathrm{LS}} = \sum_{n=1}^{i}\alpha^{i-n}|\bar{\boldsymbol{\omega}}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}(n))|^2 + 2\Re[\lambda^*(\bar{\boldsymbol{\omega}}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{s} - 1)] \tag{34}$$

where α is the forgetting factor chosen as a positive constant close to, but less than 1. Fixing $\bar{\boldsymbol{\omega}}(i)$ and computing the gradient of (34) with respect to $\mathbf{S}_D(i)$ equal to a zero matrix and solving for λ, we obtain

$$\mathbf{S}_D(i) = \frac{\hat{\mathbf{R}}^{-1}(i)\mathbf{s}\bar{\boldsymbol{\omega}}^{\mathrm{H}}(i)\mathbf{R}_{\bar{\omega}}^{-1}(i)}{\bar{\boldsymbol{\omega}}^{\mathrm{H}}(i)\mathbf{R}_{\bar{\omega}}^{-1}(i)\bar{\boldsymbol{\omega}}(i)\mathbf{s}^{\mathrm{H}}\hat{\mathbf{R}}^{-1}(i)\mathbf{s}} \tag{35}$$

where $\hat{\mathbf{R}}(i) = \sum_{n=1}^{i}\alpha^{i-n}\mathbf{r}(n)\mathbf{r}^{\mathrm{H}}(n)$ is the input covariance matrix and $\mathbf{R}_{\bar{\omega}}(i) = \bar{\boldsymbol{\omega}}(i)\bar{\boldsymbol{\omega}}^{\mathrm{H}}(i)$ is the reduced-rank weight matrix at time instant i. The computation of (35) includes the inversion of $\hat{\mathbf{R}}(i)$ and $\mathbf{R}_{\bar{\omega}}(i)$, which may increase significantly the complexity and create numerical problems. However, the expression in (35) can be further simplified using the constraint $\bar{\boldsymbol{\omega}}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{s} = 1$. The details of the simplification are given in the Appendix. The simplified expression for $\mathbf{S}_D(i)$ is given by

$$\mathbf{S}_D(i) = \frac{\hat{\mathbf{R}}^{-1}(i)\mathbf{s}\bar{\mathbf{s}}^{\mathrm{H}}}{\mathbf{s}^{\mathrm{H}}\hat{\mathbf{R}}^{-1}(i)\mathbf{s}}. \tag{36}$$

To avoid using the matrix inversion, we use the matrix inversion lemma [33] to estimate $\hat{\mathbf{R}}^{-1}(i)$. We define the inverse covariance matrix $\hat{\boldsymbol{\Phi}}(i) = \hat{\mathbf{R}}^{-1}(i)$ for convenience of presentation. The recursive estimation of $\hat{\boldsymbol{\Phi}}(i)$ is given by

$$\mathbf{K}(i) = \frac{\alpha^{-1}\hat{\boldsymbol{\Phi}}(i)\mathbf{r}(i)}{1 + \alpha^{-1}\mathbf{r}^{\mathrm{H}}(i)\hat{\boldsymbol{\Phi}}(i)\mathbf{r}(i)}$$
$$\hat{\boldsymbol{\Phi}}(i+1) = \alpha^{-1}\hat{\boldsymbol{\Phi}}(i) - \alpha^{-1}\mathbf{K}(i)\mathbf{r}^{\mathrm{H}}(i)\hat{\boldsymbol{\Phi}}(i) \tag{37}$$

where $\mathbf{K}(i)$ is the $M \times 1$ Kalman gain vector. We set $\hat{\boldsymbol{\Phi}}(1) = \delta^{-1}\mathbf{I}_M$ to start the recursion of (37), where δ is a positive small constant and $\mathbf{I}_M$ is an $M \times M$ identity matrix.

Assuming $\mathbf{S}_D(i)$ is known and taking the gradient of (34) with respect to $\bar{\boldsymbol{\omega}}(i)$, equating the terms to a null vector and solving for λ, we obtain the $D \times 1$ reduced-rank filter

$$\bar{\boldsymbol{\omega}}(i) = \frac{\bar{\boldsymbol{\Phi}}(i)\bar{\mathbf{s}}}{\mathbf{s}^{\mathrm{H}}\bar{\boldsymbol{\Phi}}(i)\bar{\mathbf{s}}} \tag{38}$$

where $\bar{\boldsymbol{\Phi}}(i) = \hat{\bar{\mathbf{R}}}^{-1}(i)$ and $\hat{\bar{\mathbf{R}}}(i) = \sum_{n=1}^{i}\alpha^{i-n}\bar{\mathbf{r}}(n)\bar{\mathbf{r}}^{\mathrm{H}}(n)$ is the estimate of reduced-rank input covariance matrix. In order to estimate $\bar{\boldsymbol{\Phi}}(i)$, we use the matrix inversion lemma as follows

$$\bar{\mathbf{K}}(i) = \frac{\alpha^{-1}\bar{\boldsymbol{\Phi}}(i)\bar{\mathbf{r}}(i)}{1 + \alpha^{-1}\bar{\mathbf{r}}^{\mathrm{H}}(i)\bar{\boldsymbol{\Phi}}(i)\bar{\mathbf{r}}(i)}$$
$$\bar{\boldsymbol{\Phi}}(i+1) = \alpha^{-1}\bar{\boldsymbol{\Phi}}(i) - \alpha^{-1}\bar{\mathbf{K}}(i)\bar{\mathbf{r}}^{\mathrm{H}}(i)\bar{\boldsymbol{\Phi}}(i) \tag{39}$$

where $\bar{\mathbf{K}}(i)$ is the $D \times 1$ reduced-rank gain vector and the recursion of (39) is initialized by choosing $\bar{\boldsymbol{\Phi}}(1) = \bar{\delta}^{-1}\mathbf{I}_D$, where $\bar{\delta}$ is a positive small constant.

The RLS reduced-rank algorithm for the GSC structure takes the similar procedure to minimize the LS cost function given by

$$\mathcal{L}_{\mathrm{LS}} = \sum_{n=1}^{i}\alpha^{i-n}|d_0(n) - \bar{\boldsymbol{\omega}}_{\mathrm{gsc}}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}_0(n)|^2. \tag{40}$$

Assuming that the optimum reduced-rank filter $\bar{\boldsymbol{\omega}}_{\mathrm{gsc}}$ is given. The LS design of the projection matrix $\mathbf{S}_D(i)$ can be obtained by solving the gradient of (40) with respect to $\mathbf{S}_D(i)$ equal to zero, as follows

$$\mathbf{S}_D(i) = \hat{\mathbf{R}}_0^{-1}(i)\hat{\mathbf{P}}_0(i)\mathbf{R}_{\bar{\omega}_{\mathrm{gsc}}}^{-1} \tag{41}$$

where $\mathbf{R}_{\bar{\omega}_{\mathrm{gsc}}} = \bar{\boldsymbol{\omega}}_{\mathrm{gsc}}\bar{\boldsymbol{\omega}}_{\mathrm{gsc}}^{\mathrm{H}}$, $\hat{\mathbf{R}}_0(i) = \sum_{n=1}^{i}\alpha^{i-n}\mathbf{r}_0(n)\mathbf{r}_0(n)^{\mathrm{H}}$ is the LS sidelobe covariance matrix and $\hat{\mathbf{P}}_0(i) = \sum_{n=1}^{i}\alpha^{i-n}d_0^*(n)\mathbf{r}_0(n)\bar{\boldsymbol{\omega}}_{\mathrm{gsc}}^{\mathrm{H}}$ is the LS sidelobe cross-correlation matrix. In this work, we use the time averaging of $\bar{\boldsymbol{\omega}}_{\mathrm{gsc}}(i)\bar{\boldsymbol{\omega}}_{\mathrm{gsc}}^{\mathrm{H}}(i)$ to approach the $\mathbf{R}_{\bar{\omega}_{\mathrm{gsc}}}$ and regularize the $\hat{\mathbf{R}}_{\bar{\omega}_{\mathrm{gsc}}}$ by adding a small diagonal load $\varsigma\mathbf{I}_D$, where the small positive constant ς is known as regularization parameter [33]. The regulized matrix can be written as

$$\hat{\mathbf{R}}_{\bar{\omega}_{\mathrm{gsc}}} = \sum_{n=1}^{i}\alpha^{i-n}\bar{\boldsymbol{\omega}}_{\mathrm{gsc}}(i)\bar{\boldsymbol{\omega}}_{\mathrm{gsc}}^{\mathrm{H}}(i) + \varsigma\mathbf{I}_D. \tag{42}$$

Let us fix $\mathbf{S}_D(i)$, take the gradient of (40) with the respect to $\bar{\boldsymbol{\omega}}_{\mathrm{gsc}}(i)$ and solve the equation equal to zero vector, we obtain the optimum LS reduced-rank filter

TABLE III
The JOINT Scheme using RLS Algorithm for the DFP

Initialization	$\mathbf{S}_D(1) = [\mathbf{I}_{D\times D} \quad \mathbf{0}_{D\times(M-D)}]^T$
	$\bar{\mathbf{s}} = \mathbf{S}_D^{\mathrm{H}}(1)\mathbf{s},\ \bar{\omega}(1) = \bar{\mathbf{s}}/\|\bar{\mathbf{s}}\|$
	$\hat{\bar{\mathbf{\Phi}}}(1) = \bar{\delta}^{-1}\mathbf{I}_D,\ \hat{\mathbf{\Phi}}(1) = \delta^{-1}\mathbf{I}_M$ and $i = 1$
Recursion	$y(i) = \bar{\omega}(i)^{\mathrm{H}}\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}(i)$
	$\mathbf{K}(i) = \hat{\mathbf{\Phi}}(i)\mathbf{r}(i)/(\alpha + \mathbf{r}^{\mathrm{H}}(i)\hat{\mathbf{\Phi}}(i)\mathbf{r}(i))$
	$\hat{\mathbf{\Phi}}(i+1) = \alpha^{-1}\hat{\mathbf{\Phi}}(i) - \alpha^{-1}\mathbf{K}(i)\mathbf{r}^{\mathrm{H}}(i)\hat{\mathbf{\Phi}}(i)$
	$\mathbf{S}_D(i) = \dfrac{\hat{\mathbf{R}}^{-1}(i)\mathbf{s}\bar{\mathbf{s}}^{\mathrm{H}}}{\mathbf{s}^{\mathrm{H}}\hat{\mathbf{R}}^{-1}(i)\mathbf{s}}$
	$\bar{\mathbf{s}} = \mathbf{S}_D^{\mathrm{H}}(i)\mathbf{s},\ \bar{\mathbf{r}}(i) = \mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}(i)$
	$\bar{\mathbf{K}}(i) = \hat{\bar{\mathbf{\Phi}}}(i)\bar{\mathbf{r}}(i)/(\alpha + \bar{\mathbf{r}}^{\mathrm{H}}(i)\hat{\bar{\mathbf{\Phi}}}(i)\bar{\mathbf{r}}(i))$
	$\hat{\bar{\mathbf{\Phi}}}(i+1) = \alpha^{-1}\hat{\bar{\mathbf{\Phi}}}(i) - \alpha^{-1}\bar{\mathbf{K}}(i)\bar{\mathbf{r}}^{\mathrm{H}}(i)\hat{\bar{\mathbf{\Phi}}}(i)$
	$\bar{\omega}(i) = \dfrac{\hat{\bar{\mathbf{\Phi}}}(i)\bar{\mathbf{s}}}{\mathbf{s}^{\mathrm{H}}\hat{\bar{\mathbf{\Phi}}}(i)\bar{\mathbf{s}}}$

TABLE IV
The JOINT Scheme using RLS Algorithm for the GSC

Initialization	$\mathbf{S}_D(1) = [\mathbf{I}_{D\times D} \quad \mathbf{0}_{D\times(M-D)}]^T$
	$\bar{\mathbf{s}} = \mathbf{S}_D^{\mathrm{H}}(1)\mathbf{s},\ \bar{\omega}_{\mathrm{gsc}}(1) = \mathbf{0}_D$
	$\hat{\bar{\mathbf{\Phi}}}_0(1) = \bar{\delta}^{-1}\mathbf{I}_D,\ \hat{\mathbf{\Phi}}_0(1) = \delta^{-1}\mathbf{I}_{M-1}$
	$\hat{\mathbf{R}}_{\bar{\omega}_{\mathrm{gsc}}}(1) = \varsigma\mathbf{I}_D$ and $i = 1$
Recursion	$d_0(i) = \mathbf{s}^{\mathrm{H}}\mathbf{r}(i),\ \mathbf{r}_0(i) = \mathbf{B}\mathbf{r}(i)$
	$y(i) = d_0(i) - \bar{\omega}_{\mathrm{gsc}}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}_0(i)$
	$\mathbf{K}_0(i) = \hat{\mathbf{\Phi}}_0(i)\mathbf{r}_0(i)/(\alpha + \mathbf{r}_0^{\mathrm{H}}(i)\hat{\mathbf{\Phi}}_0(i)\mathbf{r}_0(i))$
	$\hat{\mathbf{\Phi}}_0(i+1) = \alpha^{-1}\hat{\mathbf{\Phi}}_0(i) - \alpha^{-1}\mathbf{K}_0(i)\mathbf{r}_0^{\mathrm{H}}(i)\hat{\mathbf{\Phi}}_0(i)$
	$\mathbf{e}_s(i) = \bar{\omega}_{\mathrm{gsc}}^{\mathrm{H}}(i)\hat{\mathbf{R}}_{\bar{\omega}_{\mathrm{gsc}}}^{-1}(i)d_0^*(i) - \mathbf{r}_0^{\mathrm{H}}(i)\mathbf{S}_D(i)$
	$\mathbf{S}_D(i+1) = \mathbf{S}_D(i) + \mathbf{K}_0(i)\mathbf{e}_s(i)$
	$\bar{\mathbf{r}}_0(i) = \mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}_0(i)$
	$\bar{\mathbf{K}}_0(i) = \hat{\bar{\mathbf{\Phi}}}_0(i)\bar{\mathbf{r}}_0(i)/(\alpha + \bar{\mathbf{r}}_0^{\mathrm{H}}(i)\hat{\bar{\mathbf{\Phi}}}_0(i)\bar{\mathbf{r}}_0(i))$
	$\hat{\bar{\mathbf{\Phi}}}_0(i+1) = \alpha^{-1}\hat{\bar{\mathbf{\Phi}}}_0(i) - \alpha^{-1}\bar{\mathbf{K}}_0(i)\bar{\mathbf{r}}_0^{\mathrm{H}}(i)\hat{\bar{\mathbf{\Phi}}}_0(i)$
	$\bar{\omega}_{\mathrm{gsc}}(i+1) = \bar{\omega}_{\mathrm{gsc}}(i) + \bar{\mathbf{K}}_0(i)y^*(i)$
	$\hat{\mathbf{R}}_{\bar{\omega}_{\mathrm{gsc}}}(i+1) = \alpha\hat{\mathbf{R}}_{\bar{\omega}_{\mathrm{gsc}}}(i) + \bar{\omega}_{\mathrm{gsc}}(i)\bar{\omega}_{\mathrm{gsc}}^{\mathrm{H}}(i)$

as

$$\bar{\omega}_{\mathrm{gsc}}(i) = \hat{\bar{\mathbf{R}}}_0^{-1}(i)\hat{\bar{\mathbf{p}}}_0(i) \tag{43}$$

where $\hat{\bar{\mathbf{p}}}_0(i) = \sum_{n=1}^{i}\alpha^{i-n}\bar{\mathbf{r}}_0(i)d_0^*(i)$ is the LS cross-correlation vector. Thus, by applying the matrix inversion lemma to $\hat{\bar{\mathbf{\Phi}}}_0(i)$ and $\hat{\mathbf{\Phi}}_0(i)$ which denote $\hat{\bar{\mathbf{R}}}_0^{-1}(i)$ and $\hat{\mathbf{R}}_0^{-1}(i)$ respectively, we obtain them in a recursive way as

$$\hat{\bar{\mathbf{\Phi}}}_0(i+1) = \alpha^{-1}\hat{\bar{\mathbf{\Phi}}}_0(i) - \alpha^{-1}\bar{\mathbf{K}}_0(i)\bar{\mathbf{r}}_0^{\mathrm{H}}(i)\hat{\bar{\mathbf{\Phi}}}_0(i) \tag{44}$$

$$\hat{\mathbf{\Phi}}_0(i+1) = \alpha^{-1}\hat{\mathbf{\Phi}}_0(i) - \alpha^{-1}\mathbf{K}_0(i)\mathbf{r}_0^{\mathrm{H}}(i)\hat{\mathbf{\Phi}}_0(i) \tag{45}$$

where

$$\begin{aligned}\bar{\mathbf{K}}_0(i) &= \frac{\hat{\bar{\mathbf{\Phi}}}_0(i)\bar{\mathbf{r}}_0(i)}{\alpha + \bar{\mathbf{r}}_0^{\mathrm{H}}(i)\hat{\bar{\mathbf{\Phi}}}_0(i)\bar{\mathbf{r}}_0(i)}\\ \mathbf{K}_0(i) &= \frac{\hat{\mathbf{\Phi}}_0(i)\mathbf{r}_0(i)}{\alpha + \mathbf{r}_0^{\mathrm{H}}(i)\hat{\mathbf{\Phi}}_0(i)\mathbf{r}_0(i)}.\end{aligned} \tag{46}$$

Since the time average estimates of $(M-1)\times D$ cross-correlation matrix and and $D\times 1$ reduced-rank cross-correlation vector can be written in a recursive way as

$$\hat{\bar{\mathbf{p}}}_0(i) = \alpha\hat{\bar{\mathbf{p}}}_0(i) + \bar{\mathbf{r}}_0(i)d_0^*(i) \tag{47}$$

$$\hat{\mathbf{P}}_0(i+1) = \alpha\hat{\mathbf{P}}_0(i) + \mathbf{r}_0(i)\bar{\omega}_{\mathrm{gsc}}^{\mathrm{H}}(i)d_0^*(i). \tag{48}$$

By substituting (44) and (47) into (43), and (45) and (48) into (41), we obtain recursively adaptive equations for both the reduced-rank filter and the projection matrix as follows

$$\bar{\omega}_{\mathrm{gsc}}(i+1) = \bar{\omega}_{\mathrm{gsc}}(i) + \bar{\mathbf{K}}_0(i)y^*(i) \tag{49}$$

$$\mathbf{S}_D(i+1) = \mathbf{S}_D(i) + \mathbf{K}_0(i)\mathbf{e}_s(i) \tag{50}$$

where $y(i) = d_0(i) - \bar{\omega}_{\mathrm{gsc}}(i)\bar{\mathbf{r}}_0(i)$ is the output of the processor, and $\mathbf{e}_s(i) = \bar{\omega}_{\mathrm{gsc}}^{\mathrm{H}}(i)\hat{\mathbf{R}}_{\bar{\omega}_{\mathrm{gsc}}}^{-1}d_0^*(i) - \mathbf{r}_0^{\mathrm{H}}(i)\mathbf{S}_D(i-1)$ denotes a priori estimation error vector. The JOINT schemes using RLS algorithm for the DFP and the GSC structures are summarized in Tables III and IV, respectively.

3) *Hybrid Algorithm of SG and RLS*: Here we develop an algorithm for the JOINT STAP structure, so-called hybrid algorithm, to update the reduced-rank receiver $\bar{\omega}(i)$ by using the RLS algorithm and the projection matrix $\mathbf{S}_D(i)$ by using the SG algorithm, respectively, which leads to a computational saving compared with the RLS and meanwhile has better performance than the SG algorithm. Equations (29), (38), and (39) are combined to yield the hybrid algorithm of SG and RLS for the DFP. Similarly, the hybrid algorithm for the GSC can be obtained by combining (32), (44), and (49). The JOINT schemes using hybrid algorithms for the DFP and the GSC structures are summarized in Tables V and VI, respectively.

C. Complexity Analysis

We detail the computational complexity in terms of additions and multiplications per snapshot of the proposed schemes with SG and RLS and other existing algorithms, namely the full-rank SG and RLS algorithms, the MSWF with the SG and the RLS algorithms, the JDL-RLS algorithm and the AVF algorithm, for both the DFP and the GSC structures as shown in Table VII and Table VIII, respectively. Note that since we did not consider the GSC structure using the JDL-RLS and the AVF algorithms, we do not put their complexity with the GSC in Table VIII. The proposed schemes with the SG, the RLS, and the hybrid algorithms are much simpler than the full-rank RLS filter, the MSWF, and the AVF algorithms and slightly more complex than the full-rank SG algorithm (for $D \ll M$). Note that the proposed JOINT-SG

TABLE V
The JOINT Scheme using Hybrid Algorithm for the DFP

Initialization	$\mathbf{S}_D(1) = [\mathbf{I}_{D\times D} \quad \mathbf{0}_{D\times(M-D)}]^T$
	$\bar{\mathbf{s}} = \mathbf{S}_D^{\mathrm{H}}(1)\mathbf{s},\ \bar{\boldsymbol{\omega}}(1) = \bar{\mathbf{s}}/\|\bar{\mathbf{s}}\|$
	$\hat{\bar{\boldsymbol{\Phi}}}(1) = \bar{\delta}^{-1}\mathbf{I}_D$ and $i = 1$
Recursion	$y(i) = \bar{\boldsymbol{\omega}}(i)^{\mathrm{H}}\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}(i)$
	$\mathbf{S}_D(i+1) = \mathbf{S}_D(i) - \mu_s y^*(i)[\mathbf{I} - \mathbf{s}\mathbf{s}^{\mathrm{H}}]\mathbf{r}(i)\bar{\boldsymbol{\omega}}^{\mathrm{H}}(i)$
	$\bar{\mathbf{s}} = \mathbf{S}_D^{\mathrm{H}}(i+1)\mathbf{s},\ \bar{\mathbf{r}}(i) = \mathbf{S}_D^{\mathrm{H}}(i+1)\mathbf{r}(i)$
	$\bar{\mathbf{K}}(i) = \hat{\bar{\boldsymbol{\Phi}}}(i)\bar{\mathbf{r}}(i)/(\alpha + \bar{\mathbf{r}}^{\mathrm{H}}(i)\hat{\bar{\boldsymbol{\Phi}}}(i)\bar{\mathbf{r}}(i))$
	$\hat{\bar{\boldsymbol{\Phi}}}(i+1) = \alpha^{-1}\hat{\bar{\boldsymbol{\Phi}}}(i) - \alpha^{-1}\bar{\mathbf{K}}(i)\bar{\mathbf{r}}^{\mathrm{H}}(i)\hat{\bar{\boldsymbol{\Phi}}}(i)$
	$\bar{\boldsymbol{\omega}}(i) = \dfrac{\hat{\bar{\boldsymbol{\Phi}}}(i)\bar{\mathbf{s}}}{\mathbf{s}^{\mathrm{H}}\hat{\bar{\boldsymbol{\Phi}}}(i)\bar{\mathbf{s}}}$

TABLE VI
The JOINT Scheme using Hybrid Algorithm for the GSC

Initialization	$\mathbf{S}_D(1) = [\mathbf{I}_{D\times D} \quad \mathbf{0}_{D\times(M-D)}]^T$
	$\bar{\mathbf{s}} = \mathbf{S}_D^{\mathrm{H}}(1)\mathbf{s},\ \bar{\boldsymbol{\omega}}_{\mathrm{gsc}}(1) = \mathbf{0}_D$
	$\hat{\bar{\boldsymbol{\Phi}}}_0(1) = \bar{\delta}^{-1}\mathbf{I}_D$
	$\hat{\mathbf{R}}_{\bar{\omega}_{\mathrm{gsc}}}(1) = \varsigma\mathbf{I}_D$ and $i = 1$
Recursion	$d_0(i) = \mathbf{s}^{\mathrm{H}}\mathbf{r}(i),\ \mathbf{r}_0(i) = \mathbf{B}\mathbf{r}(i)$
	$y(i) = d_0(i) - \bar{\boldsymbol{\omega}}_{\mathrm{gsc}}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{r}_0(i)$
	$\mathbf{S}_D(i+1) = \mathbf{S}_D(i) - \mu_s y^*(i)[\mathbf{I} - \mathbf{s}\mathbf{s}^{\mathrm{H}}]\mathbf{r}(i)\bar{\boldsymbol{\omega}}^{\mathrm{H}}(i)$
	$y(i) = d_0(i) - \bar{\boldsymbol{\omega}}_{\mathrm{gsc}}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i+1)\mathbf{r}_0(i)$
	$\bar{\mathbf{r}}_0(i) = \mathbf{S}_D^{\mathrm{H}}(i+1)\mathbf{r}_0(i)$
	$\bar{\mathbf{K}}_0(i) = \hat{\bar{\boldsymbol{\Phi}}}_0(i)\bar{\mathbf{r}}_0(i)/(\alpha + \bar{\mathbf{r}}_0^{\mathrm{H}}(i)\hat{\bar{\boldsymbol{\Phi}}}_0(i)\bar{\mathbf{r}}_0(i))$
	$\hat{\bar{\boldsymbol{\Phi}}}_0(i+1) = \alpha^{-1}\hat{\bar{\boldsymbol{\Phi}}}_0(i) - \alpha^{-1}\bar{\mathbf{K}}_0(i)\bar{\mathbf{r}}_0^{\mathrm{H}}(i)\hat{\bar{\boldsymbol{\Phi}}}_0(i)$
	$\bar{\boldsymbol{\omega}}_{\mathrm{gsc}}(i+1) = \bar{\boldsymbol{\omega}}_{\mathrm{gsc}}(i) + \bar{\mathbf{K}}_0(i)y^*(i)$

and JOINT-hybrid algorithms have comparable computational complexity with the JDL algorithm, but outperform the JDL algorithm in terms of steady-state performance as shown in the subsequent section.

In order to illustrate the main trends in what concerns the complexity of the proposed and analyzed algorithms, we show in Figs. 5 and 6 the complexity of both the DFP and the GSC structures in terms of additions and multiplications versus the length of the filter M. The curves indicate that the proposed JOINT RLS algorithm has a complexity lower than the MSWF-RLS and the AVF algorithms for both the DFP and the GSC structures. The proposed JOINT-SG and JOINT-hybrid algorithms have a complexity that is situated between the full-rank RLS and the full-rank SG algorithms. The complexity of the GSC structure has extra $\mathcal{O}(M^2)$ more than that of the DFP structure in terms of multiplications and additions due to the block matrix in the sidelobe canceller. Obviously, there is no significant difference in complexity between the DFP and the GSC structures caused by extra operations for the RLS algorithms. However, for the SG algorithms, the GSC structure is more complex than the DFP structure.

TABLE VII
Computational Complexity of Algorithms for the DFP

Algorithm	Number of Operations per Snapshot	
	Additions	Multiplications
Full-Rank-SG	$3M + 1$	$3M + 2$
Full-Rank-RLS	$6M^2 - 8M + 3$	$6M^2 + 2M + 2$
MSWF-SG	$DM^2 - M^2$ $+3D^2 - 2$	$(D-1)M^2 + 2DM$ $+4D + 1$
MSWF-RLS	$(D+1)M^2 + 6D^2$ $-8D + 2$	$(D+1)M^2 + 2DM$ $+3D + 2$
AVF	$D(M^2 + 3(M-1)^2) - 1$ $+D(5(M-1)+1) + 2M$	$D(4M^2 + 4M + 1)$ $+4M + 2$
JDL-RLS	$DM + 4D^2 - D - 2$	$DM + 5D^2 + 5D$
JOINT-SG	$3DM + 2M$ $+2D - 2$	$3DM + M$ $+5D + 2$
JOINT-RLS	$6M^2 + 8M$ $+6D^2 - 8D + 3$	$7M^2 + 2M$ $+7D^2 + 9D$
JOINT-Hybrid	$6DM + M$ $+6D^2 + 9D + 12$	$4DM + 5D^2 - 4$

TABLE VIII
Computational Complexity of Algorithms for the GSC

Algorithm	Number of Operations per Snapshot	
	Additions	Multiplications
Full-Rank-SG	$M^2 + M - 1$	$M^2 + 2M - 2$
Full-Rank-RLS	$5M^2 - 5M$	$3M^2 - 3M + 1$
MSWF-SG	$DM^2 + (D+3)M - 2$	$(D+1)M^2$ $+(2+D)M$
MSWF-RLS	$DM^2 + (D+3)M - 3$ $+2D + 4D^2$	$(D+1)M^2 + 2M$ $+DM + 3D^2 - D$
JOINT-SG	$DM^2 + (3+D)M$ $-3 + 2D + 4D^2$	$M^2 + (1+3D)M$ $+2D + 1$
JOINT-RLS	$7M^2 - 21M + 20$ $+6D^2 - 8D$	$8M^2 - 12M + 8$ $+7D^2 + 9D$
JOINT-Hybrid	$M^2 + (1+3D)M$ $+6D^2 + 10D + 10$	$M^2 + (1+3D)M$ $+6D - 1 + 7D^2$

D. Rank Selection

Normally, the selection of the rank D is essential to most reduced-rank algorithms. Fortunately, based on our investigation, the sensitivity of the performance of our proposed algorithms to the change of rank (when $D > 3$) is less than that of the eigen-decomposition family and the Krylov-subspace family of reduced-rank algorithms. Considering the trade-off between the performance and complexity, we set the rank of our proposed algorithms to 4. Nevertheless, we consider a rank adaptation algorithm, in order to automatically determine the rank. We present a method for automatically selecting the ranks of the algorithms based on the exponentially weighted a posteriori LS type cost function described by

$$\mathcal{C}(\mathbf{S}_D(i), \bar{\boldsymbol{\omega}}_D(i)) = \sum_{l=1}^{i} \alpha^i |\bar{\boldsymbol{\omega}}_D^{\mathrm{H}}(i)\mathbf{S}_D(i)\mathbf{r}(l)|^2 \qquad (51)$$

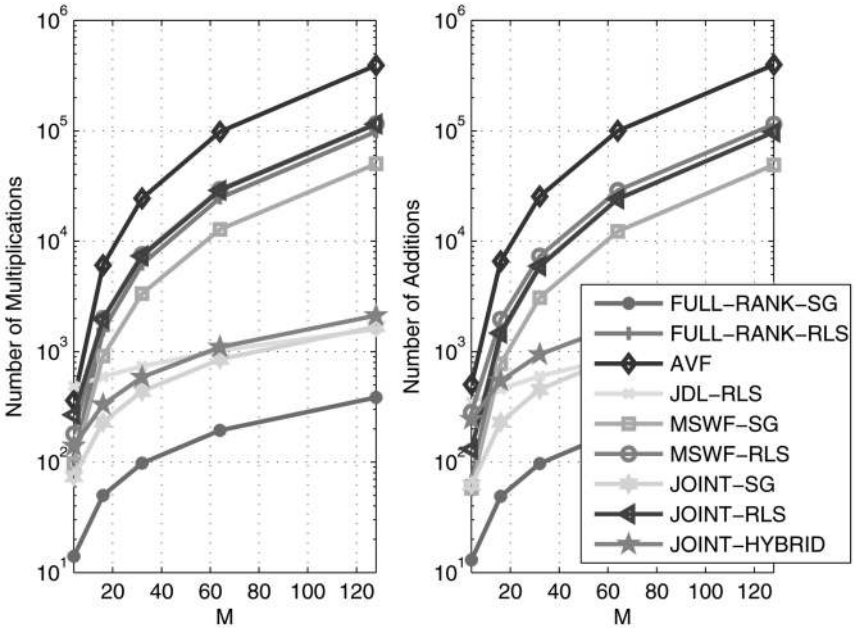

Fig. 5. Complexity for DFP in terms of additions and multiplications versus length of filter M.

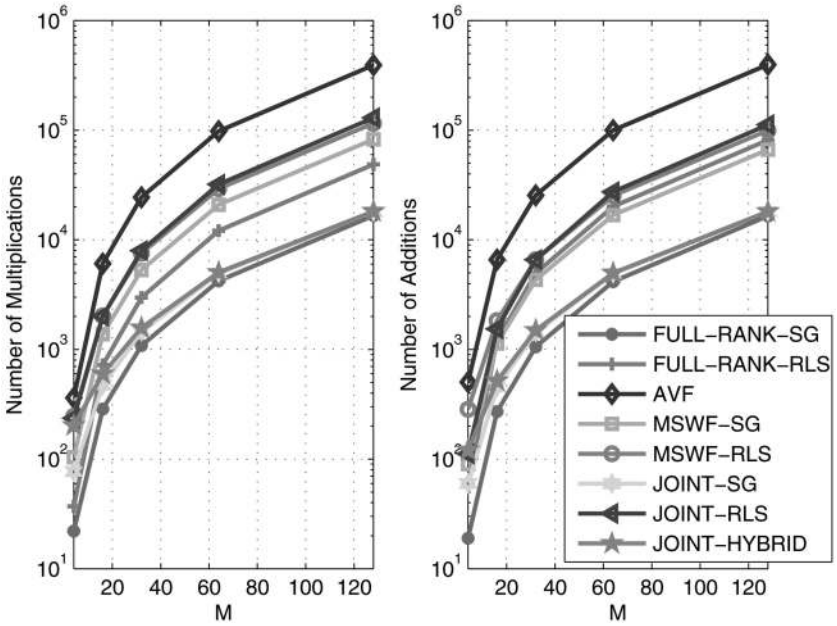

Fig. 6. Complexity for GSC in terms of additions and multiplications versus length of filter M.

where α is the forgetting factor and $\bar{\omega}_D(i)$ is the reduced-rank filter with rank D. For each time interval i, we can select the rank D_{opt} which minimizes $\mathcal{C}(\mathbf{S}_D(i),\bar{\omega}_D(i))$ and the exponential weighting factor α is required as the optimal rank varies as a function of the data record. The key quantities to be updated are the projection matrix $\mathbf{S}_D(i)$, the reduced-rank filter $\bar{\omega}_D(i)$, the associated reduced-rank steering vector $\bar{\mathbf{s}}$, and the inverse of the reduced-rank covariance matrix $\bar{\mathbf{\Phi}}(i)$ (for the proposed RLS algorithm). To this end, we define the following extended projection matrix $\mathbf{S}_D(i)$ and the extended reduced-rank filter weight vector $\bar{\omega}_D(i)$ as follows:

$$\mathbf{S}_D(i) = \begin{bmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,D_{\min}} & \cdots & s_{1,D_{\max}} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{M,1} & s_{M,2} & \cdots & s_{M,D_{\min}} & \cdots & s_{M,D_{\max}} \end{bmatrix} \tag{52}$$

$$\bar{\omega}_D(i) = [w_1 \; w_2 \cdots w_{D_{\min}} \cdots w_{D_{\max}}]^T.$$

The extended projection matrix $\mathbf{S}_D(i)$ and the extended reduced-rank filter weight vector $\bar{\omega}_D(i)$ are updated along with the associated quantities $\bar{\mathbf{s}}$ and $\bar{\mathbf{\Phi}}(i)$ (only for the RLS) for the maximum allowed rank $D_{\max}$ and then the proposed rank adaptation algorithm

determines the rank that is best for each time instant i using the cost function in (51). The proposed rank adaptation algorithm is then given by

$$D_{\text{opt}} = \arg \min_{D_{\min} \le d \le D_{\max}} \mathcal{C}(\mathbf{S}_D(i), \bar{\mathbf{w}}_D(i)) \tag{53}$$

where d is an integer, $D_{\min}$ and $D_{\max}$ are the minimum and maximum ranks allowed for the reduced-rank filter, respectively. Note that a smaller rank may provide faster adaptation during the initial stages of the estimation procedure and a greater rank usually yields a better steady-state performance. Our studies reveal that the range for which the rank D of the proposed algorithms have a positive impact on the performance of the algorithms is limited, being from $D_{\min} = 3$ to $D_{\max} = 8$ for the reduced-rank filter recursions. A drawback of this method is that it may increase the length of the filters, resulting in higher complexity. Another issue of the rank selection is that some prior knowledge about the clutter should be used in order to set $D_{\min}$ and $D_{\max}$. The development of cost-effective methods for rank selection remains an interesting area for investigation.

IV. CONVEXITY ANALYSIS OF THE PROPOSED SCHEME

Since the GSC structure is an alternative implementation of the LCMV DFP beamforming algorithm [19], in this section, only the convexity analysis of the proposed reduced-rank method employing the DFP structure is carried out for the sake of simplification. Our approach is based on expressing the output of the proposed scheme and the proposed constraint in a convenient form that renders itself to analysis. Let us rewrite the proposed constrained optimization method in (20) using the method of Lagrange multipliers and express it by the Lagrangian

$$\begin{aligned}\mathcal{L} &= \mathbb{E}[|\bar{\boldsymbol{\omega}}^{\text{H}}(i)\mathbf{S}_D^{\text{H}}(i)\mathbf{r}(i)|^2] + 2\Re[\lambda^*(\bar{\boldsymbol{\omega}}^{\text{H}}(i)\mathbf{S}_D^{\text{H}}(i)\mathbf{s} - 1)] \\ &= \mathbb{E}[|y(i)|^2] + 2\Re[\lambda^*(\bar{\boldsymbol{\omega}}^{\text{H}}(i)\mathbf{S}_D^{\text{H}}(i)\mathbf{s} - 1)].\end{aligned} \tag{54}$$

In order to proceed, let us express $y(i)$ in an alternative and more convenient form as

$$\begin{aligned}y(i) &= \bar{\boldsymbol{\omega}}^{\text{H}}(i)\mathbf{S}_D^{\text{H}}(i)\mathbf{r}(i) \\ &= \bar{\boldsymbol{\omega}}^{\text{H}}(i)\begin{bmatrix}\mathbf{r}(i) & 0 & \cdots & 0 \\ 0 & \mathbf{r}(i) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{r}(i)\end{bmatrix}^T \begin{bmatrix}\mathbf{s}_1^*(i) \\ \mathbf{s}_2^*(i) \\ \vdots \\ \mathbf{s}_D^*(i)\end{bmatrix} \\ &= \bar{\boldsymbol{\omega}}^{\text{H}}(i)\mathbb{R}^T(i)\mathbf{s}_v^*(i)\end{aligned} \tag{55}$$

where $\mathbb{R}(i)$ is a $DM \times D$ block diagonal matrix with the input data vector $\mathbf{r}(i)$ and $\mathbf{s}_v^*(i)$ is a $DM \times 1$ vector with the columns of $\mathbf{S}_D(i)$ stacked on top of each other.

In order to analyze the proposed joint optimization procedure, we can rearrange the terms in $y(i)$ and define a single $D(M+1) \times 1$ parameter vector $\boldsymbol{f}(i) = [\bar{\boldsymbol{\omega}}^T(i) \;\; \mathbf{s}_v^*(i)]^{\text{H}}$. We can therefore further express $y(i)$ as

$$\begin{aligned}y(i) &= \boldsymbol{f}^{\text{H}}(i)\begin{bmatrix}\mathbf{0}_{D\times D} & \mathbf{0}_{D\times DM} \\ \mathbb{R}(i) & \mathbf{0}_{DM\times DM}\end{bmatrix}\boldsymbol{f}(i) \\ &= \boldsymbol{f}^{\text{H}}(i)\mathbf{G}(i)\boldsymbol{f}(i)\end{aligned} \tag{56}$$

where $\mathbf{G}(i)$ is a $D(M+1) \times D(M+1)$ matrix which contains $\mathbb{R}(i)$. Now let us perform a similar linear algebra transformation with the proposed constraint $\bar{\boldsymbol{\omega}}^{\text{H}}(i)\mathbf{S}_D^{\text{H}}(i)\mathbf{s} = 1$ and express it as

$$\bar{\boldsymbol{\omega}}^{\text{H}}(i)\mathbf{S}_D^{\text{H}}(i)\mathbf{s} = \boldsymbol{f}^{\text{H}}(i)\mathbf{A}_s\boldsymbol{f}(i) \tag{57}$$

where $\mathbf{A}_s$ is a $D(M+1) \times D(M+1)$ matrix structured as

$$\mathbf{A}_s = \begin{bmatrix}\mathbf{0}_{D\times D} & \mathbf{0}_{D\times DM} \\ \mathbb{R}_s & \mathbf{0}_{DM\times DM}\end{bmatrix} \tag{58}$$

and the $DM \times D$ block diagonal matrix $\mathbb{R}_s$ with the steering vector $\mathbf{s}$ constructed as

$$\mathbb{R}_s = \begin{bmatrix}\mathbf{s} & 0 & \cdots & 0 \\ 0 & \mathbf{s} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{s}\end{bmatrix}. \tag{59}$$

At this point, we can alternatively express the Lagrangian in (54) as

$$\mathcal{L} = \mathbb{E}[|\boldsymbol{f}^{\text{H}}(i)\mathbf{G}(i)\boldsymbol{f}(i)|^2] + 2\Re[\lambda^*(\boldsymbol{f}^{\text{H}}(i)\mathbf{A}_s\boldsymbol{f}(i) - 1)]. \tag{60}$$

We can examine the convexity of the above Lagrangian by computing the Hessian ($\boldsymbol{H}_e$) with respect to $\boldsymbol{f}(i)$ using the expression [34]

$$\boldsymbol{H}_e = \frac{\partial}{\partial \boldsymbol{f}^{\text{H}}(i)}\frac{\partial \mathcal{L}}{\partial \boldsymbol{f}(i)} \tag{61}$$

and testing if the terms are positive semi-definite. Specifically, $\boldsymbol{H}_e$ is positive semi-definite if $\boldsymbol{v}^{\text{H}}\boldsymbol{H}_e\boldsymbol{v} \ge 0$ for all non-zero, $\boldsymbol{v} \in \mathcal{C}^{D(M+1)\times D(M+1)}$ [35]. Therefore, the optimization problem is convex if the Hessian $\boldsymbol{H}_e$ is positive semi-definite.

Evaluating the partial differentiation in the expression given in (61) yields

$$\begin{aligned}\boldsymbol{H}_e &= 2\mathbb{E}[\boldsymbol{f}^{\text{H}}(i)\mathbf{G}(i)\boldsymbol{f}(i)\mathbf{G}^{\text{H}}(i) + \mathbf{G}(i)\boldsymbol{f}(i)\boldsymbol{f}^{\text{H}}(i)\mathbf{G}^{\text{H}}(i)] \\ &\quad + 2\lambda\mathbf{A}_s.\end{aligned} \tag{62}$$

By examining $\boldsymbol{H}_e$, we verify that the second term is positive semi-definite, whereas the first terms is indefinite. The third term depends on the constraint, which is typically positive in the proposed scheme as verified in our studies, yielding a positive semi-definite matrix. Therefore, the optimization problem cannot be classified as convex. It is however important to remark that our studies indicate that there

are no local minima and there exists multiple solutions (which are possibly identical).

In order to support this claim, we have checked the impact on the proposed algorithms of different initializations. This study confirmed that the algorithms are not subject to performance degradation due to the initialization although we have to bear in mind that the initialization $\mathbf{S}_D(1) = \mathbf{0}_{M\times D}$ annihilates the signal and must be avoided. We have also studied a particular case of the proposed scheme when $M = 1$ and $D = 1$, which yields the Lagrangian

$$\mathcal{L}(\bar{\omega}(i), S_D(i)) = \mathbb{E}[|\bar{\omega}^{\mathrm{H}}(i)S_D^{\mathrm{H}}(i)r(i)|^2] + 2\Re[\lambda^*(\bar{\omega}^{\mathrm{H}}(i)S_D^{\mathrm{H}}(i)s - 1)]. \quad (63)$$

By choosing $\mathbf{S}_D = 1$, it is evident that the resulting function

$$\mathcal{L}(\bar{\omega}(i), r(i)) = \mathbb{E}[|\bar{\omega}^{\mathrm{H}}(i)r(i)|^2] + 2\Re[\lambda^*(\bar{\omega}^{\mathrm{H}}(i)s - 1)] \quad (64)$$

is a convex one. In contrast to that, for a time-varying projection $\mathbf{S}_D$ the plots of the function indicate that the function is no longer convex but it also does not exhibit local minima. The problem at hand can be generalized to the vector case, however, we can no longer verify the existence of local minima due to the multi-dimensional surface. This remains as an interesting open problem to be studied.

V. PERFORMANCE ASSESSMENT

In this section we evaluate the performance of the proposed JOINT STAP algorithm in an airborne radar application using the simulated collected radar data. The parameters of the simulated radar platform are shown in the Table IX. For all simulations, we assume the presence of a mixture of two broadband jammers at $-45°$ and $60°$ with jammer-to-noise-ratio (JNR) equal to 30 dB. The clutter-to-noise-ratio (CNR) is fixed at 40 dB. The signal-to-noise-ratio (SNR) is set at 10 dB. In particular, the performance of the proposed scheme with the SG, the RLS, and the hybrid algorithms is compared with existing techniques, namely the full-rank SG and RLS algorithms, the reduced-rank algorithms with $\mathbf{S}_D(i)$ designed according to the MSWF and the AVF algorithms, the reduced-dimension algorithm JDL with specified localized processing region (LPR), and the optimal linear filter that assumes the knowledge of the covariance matrix for both the DFP and the GSC structures. To this end, the algorithms are compared in terms of the SINR, which is defined for the reduced-rank scheme as

$$\mathrm{SINR}(i) = \frac{|\bar{\omega}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{s}|^2}{\bar{\omega}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{R}\mathbf{S}_D(i)\bar{\omega}(i)} \quad (65)$$

and the probability of detection with specific false alarm rate. In all the experiments, we set $D = 6$ for the AVF and the MSWF-RLS algorithms, $D = 4$ for the MSWF-SG algorithm, our proposed JOINT scheme with the SG, the RLS, and the hybrid algorithms. The performance of the JDL algorithm with $D = 3 \times 3$ LPR is investigated. Note that the Ds are selected via a large number of experiments to optimize the algorithms in terms of SINR.

TABLE IX
Radar System Parameters

Parameter	Value
Carrier frequency (f_c)	450 MHz
Transmit pattern	Uniform
PRF (f_r)	300 Hz
Platform velocity (v)	50 m/s
Platform height (h)	9000 m
Clutter-to-noise ratio (CNR)	40 dB
Jammer-to-noise ratio (JNR)	30 dB
Elements of sensors (N)	8
Number of Pulses (K)	8

A. SINR Performance

In the first two experiments, we evaluate the SINR performance of the considered algorithms against the number of snapshots for both the DFP and the GSC structures, and their curves are shown in Figs. 7 and 8, respectively. $L = 500$ snapshots are simulated and all presented results are averages over 100 independent Monte-Carlo runs. The parameters used to obtain these curves are also shown and the rank D shown in the figures are selected to optimize the algorithms and we suppose the target injected in the boresight ($0°$) with Doppler frequency 100 Hz. The SG-version MSWF is known to have problems in these situations since it does not tridiagonalize its covariance matrix [9], being unable to approach the optimal performance. The curves show an excellent performance for the proposed scheme with the RLS and the hybrid algorithms which converge much faster than the full-rank SG and RLS algorithms, and are also better than the more complex MSWF-RLS and AVF schemes. The low complexity JOINT-SG algorithm performs better than the MSWF-RLS algorithm but slightly worse than the AVF algorithm. Although the JDL algorithm converges a little faster than our proposed algorithms, its steady-state performance is inferior to the reduced-rank algorithms and the full-rank RLS algorithm. Note that the performance for the DFP and GSC structures is similar and the proposed scheme using RLS and hybrid algorithms is more promising than others for the designer because of their good convergence performance and reduced complexity.

In the next two experiments with the same scenario as the previous ones, we evaluate the SINR performance against the target Doppler frequency at the mainbeam look angle for both the DFP and the GSC structures which are illustrated in Figs. 9 and 10,

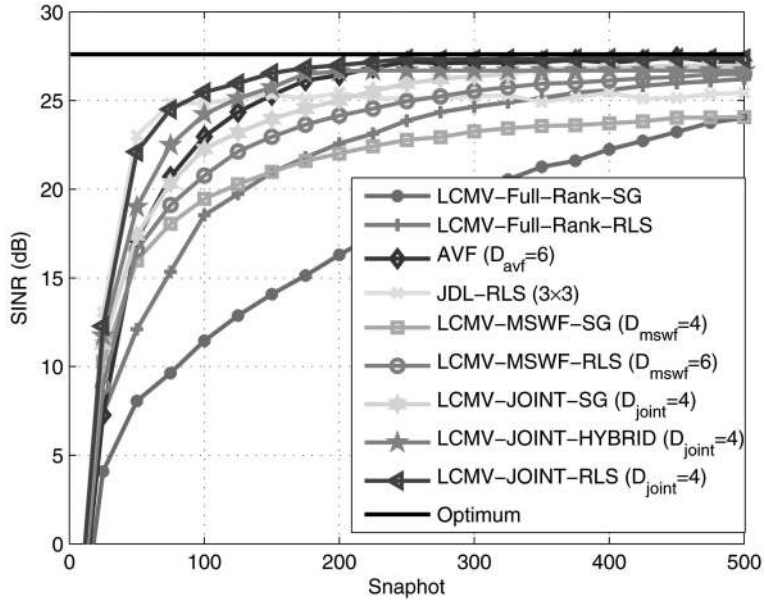

Fig. 7. SINR performance against snapshot for FDP structure with $M = 64$, SNR = 10 dB, $\alpha = 0.9998$, $\mu_s = 0.0003$, $\mu_\omega = 0.001$.

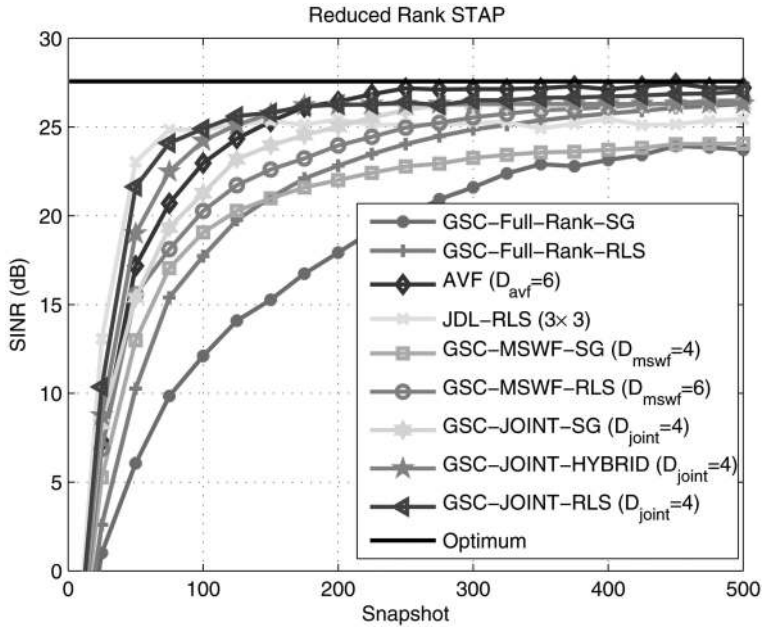

Fig. 8. SINR performance against snapshot for GSC structure with $M = 64$, SNR = 10 dB, $\alpha = 0.9998$, $\mu_s = 0.0003$, $\mu_\omega = 0.001$.

respectively. The potential Doppler frequency space from -100 to 100 Hz is examined. To illustrate that our proposed scheme supports short data record, 100 snapshots are simulated and all results are averages over 100 independent Monte-Carlo runs. The plots show that our algorithms converge and approach the optimum in a very short time, and form a deep null to cancel the mainbeam clutter. It is worth to note that the proposed JOINT-RLS and JOINT-hybrid algorithms in the GSC structure have much deeper notch in the low value Doppler space but less SINR gain in the high value Doppler space than those in the DFP structure, which means that they have better performance to cancel the mainbeam clutter but less response to the moving target with relative high velocity.

B. Probability of Detection

In this part, the probability of detection (P_D) against SNR is evaluated for all algorithms with limited training, say $L = 100$ snapshots, and the results for the DFP and the GSC structures are shown in Figs. 11 and 12, respectively. The false alarm probability P_{FA} is set to 1e-6 and we suppose the target injected in the boresight (0°) with Doppler frequency 100 Hz. The figures illustrate that the best

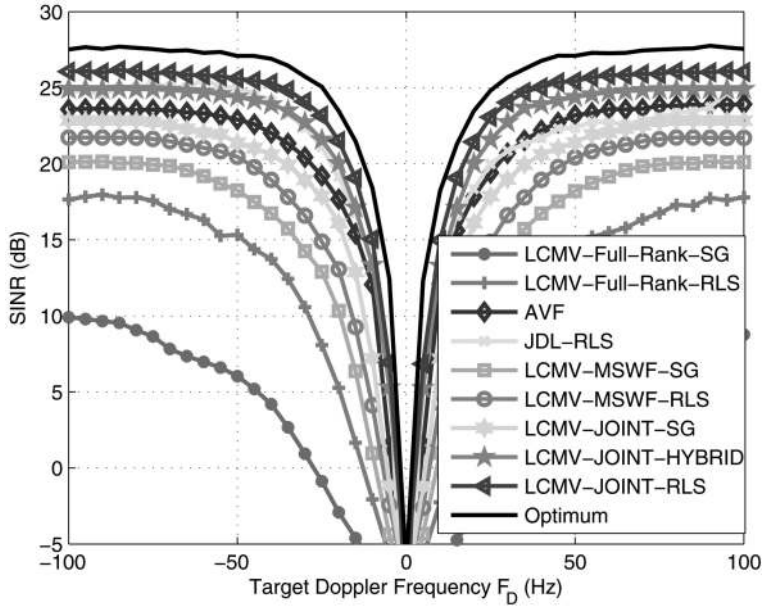

Fig. 9. SINR performance against target Doppler frequency (F_D) for FDP structure with 100 snapshots, $M = 64$, SNR = 10 dB, $\alpha = 0.9998$, $\mu_s = 0.0003$, $\mu_\omega = 0.001$.

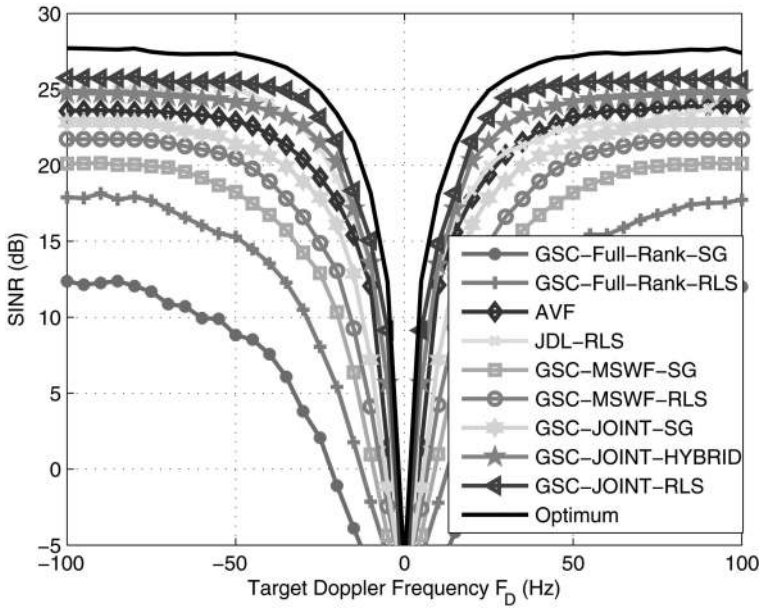

Fig. 10. SINR performance against target Doppler frequency (F_D) for GSC structure with 100 snapshots, $M = 64$, SNR = 10 dB, $\alpha = 0.9998$, $\mu_s = 0.0003$, $\mu_\omega = 0.001$.

detection performance is provided by the JOINT-RLS, the JDL-RLS, and the JOINT-hybrid algorithms, followed by the AVF and the proposed JOINT-SG algorithms. As mentioned before, the proposed algorithms in the GSC structure have less response to the moving target with relative high velocity than those in the DFP structure. Thus, in this experiment, the P_D performance of the proposed algorithms in the DFP structure is better. We also note that with limited training, the JDL algorithm has similar detection performance to our proposed algorithm.

VI. CONCLUSIONS

We developed a novel reduced-rank scheme based on joint iterative optimization of adaptive filters and a reduced-rank STAP technique based on the proposed JOINT algorithm with a low-complexity implementation using the the SG, the RLS, and the hybrid adaptive algorithms for airborne radar applications. In the proposed scheme, a bank of full-rank adaptive filters forms the projection matrix and a small reduced-rank filter is responsible for estimating the desired signal. We developed

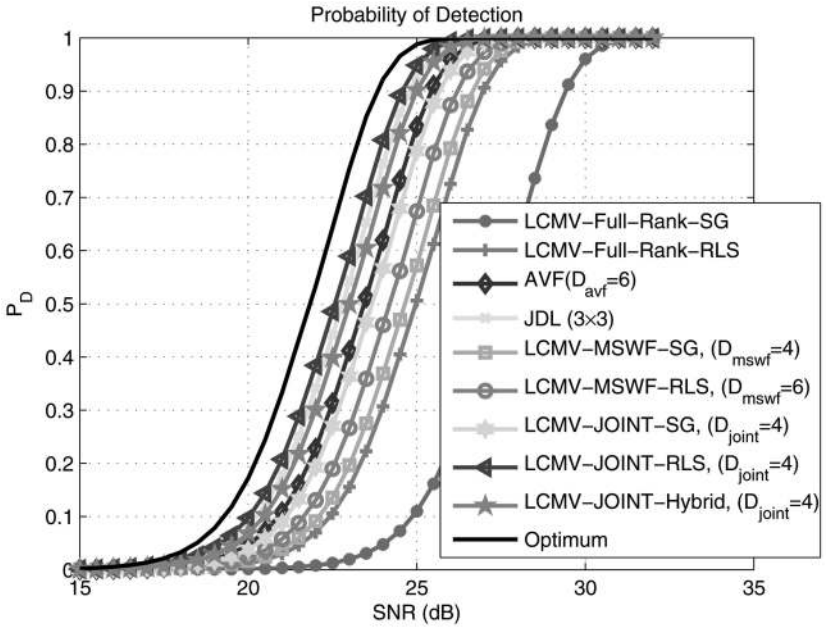

Fig. 11. Probability of detection performance against normalized SINR for FDP structure with 100 snapshots, P_{FA} = 1e-6, M = 64, $\alpha = 0.9998$, $\mu_s = 0.0003$, $\mu_\omega = 0.001$.

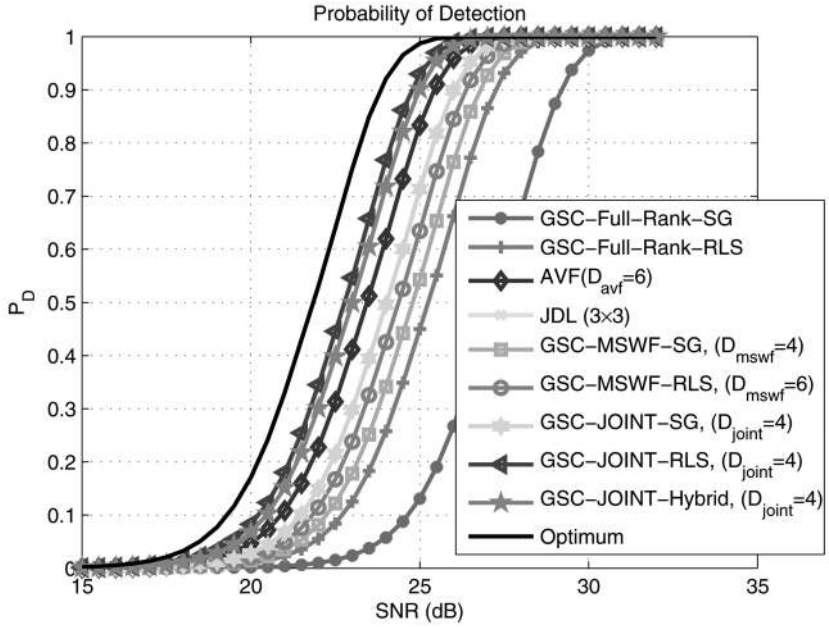

Fig. 12. Probability of detection performance against normalized SINR for GSC structure with 100 snapshots, P_{FA} = 1e-6, M = 64, $\alpha = 0.9998$, $\mu_s = 0.0003$, $\mu_\omega = 0.001$.

cost-efficient adaptive algorithms for the proposed method for both the DFP and the GSC structures and the complexity analysis was shown in terms of the number of multiplications and additions per snapshot. Furthermore, the convexity analysis of the proposed method was carried out. The results for radar clutter and jamming suppression show a performance significantly better and lower complexity than existing reduced-rank schemes.

APPENDIX. SIMPLIFICATION OF $S_D(i)$

In this Appendix, we show the details of the simplification of the filter $\mathbf{S}_D(i)$ in (35) for reducing the computational complexity. Observing the terms $\bar{\omega}^{\mathrm{H}}(i)\mathbf{R}_{\bar{\omega}}^{-1}(i)\bar{\omega}(i)$ in the denominator of (35) and employing the constrained condition $\bar{\omega}^{\mathrm{H}}(i)\mathbf{S}_D^{\mathrm{H}}(i)\mathbf{s} = 1$, we have

$$\begin{aligned}\bar{\omega}^{\mathrm{H}}(i)\mathbf{R}_{\bar{\omega}}^{-1}(i)\bar{\omega}(i) &= \mathbf{s}^{\mathrm{H}}\mathbf{S}_D^{\mathrm{H}}(i)\bar{\omega}(i)\bar{\omega}^{\mathrm{H}}(i)\mathbf{R}_{\bar{\omega}}^{-1}(i)\bar{\omega}(i)\\ &= \mathbf{s}^{\mathrm{H}}\mathbf{S}_D^{\mathrm{H}}(i)\bar{\omega}(i) = 1. \end{aligned} \quad (66)$$

Besides, we pay attention to the reduced-rank weight matrix $\mathbf{R}_{\bar{\omega}}(i) = \bar{\omega}(i)\bar{\omega}^{\mathrm{H}}(i)$. We can operate the transformations as follow:

$$\bar{\mathbf{s}}^{\mathrm{H}}\mathbf{R}_{\bar{\omega}}(i)\bar{\mathbf{s}} = \bar{\mathbf{s}}^{\mathrm{H}}\bar{\omega}(i)\bar{\omega}^{\mathrm{H}}(i)\bar{\mathbf{s}} = 1. \tag{67}$$

Thus, performance of the same operation in (eq:step1) again, we have

$$\bar{\mathbf{s}}\bar{\mathbf{s}}^{\mathrm{H}}\mathbf{R}_{\bar{\omega}}(i)\bar{\mathbf{s}}\bar{\mathbf{s}}^{\mathrm{H}} = \bar{\mathbf{s}}\bar{\mathbf{s}}^{\mathrm{H}}. \tag{68}$$

By multiplying the inversion $(\bar{\mathbf{s}}\bar{\mathbf{s}}^{\mathrm{H}})^{-1}$ on the both sides of (68) and combining the indentity matrix, the result is

$$\mathbf{R}_{\bar{\omega}}^{-1}(i) = \bar{\mathbf{s}}\bar{\mathbf{s}}^{\mathrm{H}}. \tag{69}$$

Substituting (66) and (69) into (35) and using the constrained condition again, we can get a simpler expression for the projection matrix as (36)

$$\mathbf{S}_D(i) = \frac{\hat{\mathbf{R}}^{-1}(i)\mathbf{s}\bar{\omega}(i)\bar{\mathbf{s}}\bar{\mathbf{s}}^{\mathrm{H}}}{\mathbf{s}^{\mathrm{H}}\hat{\mathbf{R}}^{-1}(i)\mathbf{s}} = \frac{\hat{\mathbf{R}}^{-1}(i)\mathbf{s}\bar{\mathbf{s}}^{\mathrm{H}}}{\mathbf{s}^{\mathrm{H}}\hat{\mathbf{R}}^{-1}(i)\mathbf{s}}. \tag{70}$$

REFERENCES

[1] Brennan, L. E. and Reed, I. S.
Theory of adaptive radar.
IEEE Transactions on Aerospace and Electronic Systems, **AES-9**, 2 (1973), 237–252.

[2] Ward, J.
Space-time adaptive processing for airborne radar.
MIT Lincoln Laboratory, Lexington, MA, Technical Report 1015, Dec. 1994.

[3] Klemm, R.
Principle of Space-Time Adaptive Processing.
Bodmin, UK: IEE Press, 2002.

[4] Haimovich, A. M. and Bar-Ness, Y.
An eigenanalysis interference canceler.
IEEE Transactions on Signal Processing, **39**, 1 (1991), 76–84.

[5] Haimovich, A.
The eigencanceler: Adaptive radar by eigenanalysis methods.
IEEE Transactions on Aerospace and Electronic Systems, **32**, 2 (1996), 532–542.

[6] Goldstein, J. S. and Reed, I. S.
Reduced-rank adaptive filtering.
IEEE Transactions Signal Processing, **45**, 2 (1997), 492–496.

[7] Goldstein, J. S. and Reed, I. S.
Subspace selection for partially adaptive sensor array processing.
IEEE Transactions on Aerospace and Electronic Systems, **33**, 2 (1997), 539–544.

[8] Goldstein, J. S., Reed, I. S., and Scharf, L. L.
A multistage representation of the Wiener filter based on orthogonal projections.
IEEE Transactions on Information Theory, **44**, 7 (1998), 2943–2959.

[9] Honig, M. L. and Goldstein, J. S.
Adaptive reduced-rank interference suppression based on the multistage Wiener filter.
IEEE Transactions on Communications, **50**, 6 (2002), 986–994.

[10] Scharf, L. L., et al.
Subspace expansion and the equivalence of conjugate direction and multistage Wiener filters.
IEEE Transactions on Signal Processing, **56**, 10 (2008), 5013–5019.

[11] Pados, D. A. and Batalama, S. N.
Joint space-time auxiliary-vector filtering for DS-CDMA systems with antenna arrays.
IEEE Transactions on Communications, **47**, 9 (1999), 1406–1415.

[12] Pados, D. A. and Karystinos, G. N.
An iterative algorithm for the computation of the MVDR filter.
IEEE Transactions on Signal Processing, **49**, 2 (Feb. 2001), 290–300.

[13] Pados, D. A., et al.
Short-data-record adaptive detection.
In *Proceedings of the 2007 IEEE Radar Conference*, Boston, MA, Apr. 17–20, 2007, 357–361.

[14] de Lamare, R. C. and Sampaio-Neto, R.
Reduced-rank adaptive filtering based on joint iterative optimization of adaptive filters.
IEEE Signal Processing Letters, **14**, 12 (2007), 980–983.

[15] de Lamare, R. C. and Sampaio-Neto, R.
Adaptive reduced-rank MMSE filtering with interpolated FIR filters and adaptive interpolators.
IEEE Signal Processing Letters, **12**, 3 (Mar. 2005).

[16] de Lamare, R. C. and Sampaio-Neto, R.
Reduced-rank interference suppression for DS-CDMA based on interpolated FIR filters.
IEEE Communication Letters, **9**, 3 (Mar. 2005).

[17] de Lamare, R. C. and Sampaio-Neto, R.
Adaptive interference suppression for CDMA based on interpolated FIR filters in multipath channels.
IEEE Transactions on Vehicular Technology (Sept. 2007), 2457–2474.

[18] Applebaum, S. and Chapman, D.
Adaptive arrays with main beam constraints.
IEEE Transactions on Antennas and Propagation, **AP-24**, 5 (1976), 650–662.

[19] Griffiths, L. and Jim, C.
An alternative approach to linearly constrained adaptive beamforming.
IEEE Transactions on Antennas and Propagation, **AP-30**, 1 (1982), 27–34.

[20] Goldstein, J. S. and Reed, I. S.
Theory of partially adaptive radar.
IEEE Transactions on Aerospace and Electronic Systems, **33**, 4 (1997), 1309–1325.

[21] Reed, I. S., Mallett, J. D., and Brennan, L. E.
Rapid convergence rate in adaptive arrays.
IEEE Transactions on Aerospace and Electronics Systems, **AES-10**, 6 (1974), 853–863.

[22] Goldstein, J. S., Zulch, P. A., and Reed, I. S.
Reduced rank space-time adaptive radar processing.
In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP 1996), vol. 2, Atlanta, GA, 1173–1176.

[23] Guerci, J. R., Goldstein, J. S., and Reed, I. S.
Optimal and adaptive reduced-rank STAP.
IEEE Transactions on Aerospace and Electronic Systems, **36**, 2 (2000), 647–663.

[24] Wang, H. and Cai, L.
On adaptive spatial-temporal processing for airborne surveillance radar systems.
IEEE Transactions on Aerospace and Electronic Systems, **30**, 3 (1994), 66–70.

[25] Adve, R. S., Hale, T. B., and Wicks, M. C.
Practical joint domain localised adaptive processing in homogeneous and nonhomogeneous environments, Part 1: Homogeneous environments.
IEE Proceedings—Radar, Sonar and Navigation, **147**, 2 (2000), 575.

[26] Adve, R. S., Hale, T. B., and Wicks, M. C.
Practical joint domain localised adaptive processing in homogeneous and nonhomogeneous environments, Part 2: Nonhomogeneous environments.
IEE Proceedings—Radar, Sonar and Navigation, **147**, 2 (2000), 664.

[27] Wong, R. Y. M. and Adve, R. S.
Reduced-rank adaptive filtering using localized processing for CDMA systems.
IEEE Transactions on Vehicular Technology, **56**, 6 (Nov. 2007), 3846–3856.

[28] Fa, R., de Lamare, R. C., and Zanatta-Filho, D.
Reduced-rank STAP algorithm for adaptive radar based on joint iterative optimization of adaptive filters.
Presented at the 42nd Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, 2008.

[29] Burykh, S. and Abed-Meraim, K.
Reduced-rank adaptive filtering using Krylov subspace.
EURASIP Journal on Applied Signal Processing, **12** (2002), 1387–1400.

[30] Chen, W. S., Mitra, U., and Schniter, P.
On the equivalence of three reduced rank linear estimators with applications to DS-CDMA.
IEEE Transactions Information Theory, **48**, 9 (2002), 2609–2614.

[31] Frost, O. L.
An algorithm for linearly constrained adaptive array processing.
Proceedings of the IEEE, **60**, 8 (1972), 926–935.

[32] Resende, L. S., Romano, J. M. T., and Bellanger, M. G.
A fast least-squares algorithm for linearly constrained adaptive filtering.
IEEE Transactions on Signal Processing, **44**, 5 (1996), 1168–1174.

[33] Haykin, S.
Adaptive Filter Theory (4th ed.).
Upper Saddle River, NJ: Prentice-Hall, 2002.

[34] Bertsekas, D. P.
Nonlinear Programming (2nd ed.).
Belmont, MA: Athena Scientific, 1999.

[35] Golub, G. H. and van Loan, C. F.
Matrix Computations (3rd ed.).
Baltimore, MD: The Johns Hopkins University Press, 1996.

[36] Van Trees, H. L.
Optimum Array Processing.
Hoboken, NJ: Wiley, 2002.

Rui Fa received his Bachelor and Master degrees in electronic and electrical engineering from Nanjing University of Science and Technology (NUST), China, in 2000 and 2003, respectively, and his Ph.D. degree in electrical engineering from University of Newcastle, UK, in 2007.

Since January 2008, he has been with the Communications Research Group, Department of Electronics, University of York, where he is currently a research associate. His current research interests include radar signal processing, low-complexity interference cancellation, iterative decoding, and detection for wireless communications.

Rodrigo C. de Lamare received his electronic engineering degree from the Federal University of Rio de Janeiro (UFRJ), Brazil, in 1998 and the M.Sc. and Ph.D. degrees in electrical engineering from the Pontifical Catholic University of Rio de Janeiro (PUC-RIO), Brazil, in 2001 and 2004, respectively.

Since January 2006, he has been with the Communications Research Group, Department of Electronics, University of York, where he is currently a lecturer in communications engineering. His research interests lie in communications and signal processing.

In the areas of his research interests, Dr. de Lamare has published about 170 papers in refereed journals and conferences. Dr. de Lamare serves as associate editor for the *EURASIP Journal on Wireless Communications and Networking*. He has also been the General Chair of the 7th IEEE International Symposium on Wireless Communications Systems, held in York, UK in September 2010.