# DATA-ADAPTIVE REDUCED-DIMENSION ROBUST CAPON BEAMFORMING

Samuel D. Somasundaram<sup>\*</sup>, Nigel H. Parsons

General Sonar Studies Thales Underwater Systems Cheadle Heath, Stockport, Cheshire, U.K.

## ABSTRACT

We present low complexity, quickly converging robust adaptive beamformers that combine robust Capon beamformer (RCB) methods and data-adaptive Krylov subspace dimensionality reduction techniques. We extend a recently proposed reduced-dimension RCB framework, which ensures proper combination of RCBs with any form of dimensionality reduction that can be expressed using a fullrank dimension reducing transform, providing new results useful for data-adaptive dimensionality reduction. We consider Krylov subspace methods computed with the Powers-of-R (PoR) and Conjugate Gradient (CG) techniques, illustrating how a fast CG-based algorithm can be formed by beneficially exploiting that the CGalgorithm yields a diagonal reduced-dimension covariance matrix. Our simulations show the benefits of the proposed approaches.

*Index Terms*— Robust adaptive beamforming, dimensionality reduction, Krylov subspace methods.

### 1. INTRODUCTION AND PRELIMINARIES

When implementing adaptive beamforming on arrays with large apertures and many elements that operate in dynamic environments, reduced-dimension techniques are often needed to speed-up the convergence of the algorithms and reduce their computational complexity [1]. This is of fundamental importance in applications found in passive sonar and radar systems. Furthermore, robust adaptive techniques are often required to alleviate the deleterious effects of array steering vector (ASV) mismatch, e.g., caused by calibration and pointing errors. A popular class of these are the robust Capon beamformers (RCBs) that exploit ellipsoidal, including spherical, uncertainty sets of the ASV [2-6]. In [1, 7], a framework for combining reduced-dimension and RCB techniques was derived, allowing rapidly converging, low complexity robust adaptive reduced-dimension robust Capon beamformers (RDRCBs) to be formed. A key contribution of that work was the derivation of a complex propagation theorem that allows a reduced-dimension ellipsoid to be derived from an element-space ellipsoid and any full-rank dimension reducing transform. The reduced-dimension ellipsoid may then be exploited by using RCB techniques in the reduced-dimension space. In [1, 7], only data-independent dimensionality reduction was considered. Here, we extend the framework developed in [1,7] to data-adaptive dimensionality reduction, providing new results useful for exploiting in a variety of scenarios that occur in practical applications of robust beamforming algorithms.

The problem under consideration is the design of RDRCBs that are suitable for large arrays. We consider Krylov subspace

Peng Li, Rodrigo C. de Lamare

Communications Research Group Department of Electronics University of York, U.K.

techniques for data-adaptive dimensionality reduction and use the Powers-of-R (PoR) [8–10] and Conjugate-Gradient (CG) [11, 12] algorithms to compute dimension reducing transforms, which are used to form novel CG-based and PoR-based RDRCBs suitable for beamforming large arrays. We exploit the fact that the CG algorithm results in a diagonal reduced-dimension sample covariance matrix to give particularly low-complexity data-adaptive beamforming algorithms. Scenarios with large planar arrays are investigated along with both non-degenerate ellipsoidal and spherical uncertainty sets.

In the following,  $E\{\cdot\}, (\cdot)^T, (\cdot)^H, (\cdot)^{-1}$  and  $(\cdot)^{\dagger}$  denote the expectation, transpose, Hermitian transpose, inverse and Moore-Penrose pseudo-inverse operators, respectively. Furthermore,  $\|\cdot\|_2$ ,  $\mathbf{N}_{\mathbf{X}}^l$ ,  $\mathbf{\Pi}_{\mathbf{X}}$  and  $\mathbf{\Pi}_{\mathbf{X}}^{\perp}$  denote the two-norm, a basis for the left null-space of  $\mathbf{X}$ , the orthogonal projector onto the range space of  $\mathbf{X}$  and the orthogonal projector onto the space perpendicular to the range space of  $\mathbf{X}$ , respectively. Moreover,  $\mathbf{X} \ge 0$  or  $\mathbf{X} > 0$  mean that the Hermitian matrix  $\mathbf{X}$  is +ve semi-definite or +ve definite.

## 1.1. Robust Capon Beamforming

We model the *k*th element-space array snapshot  $\mathbf{x}_k \in \mathbb{C}^M$  as

$$\mathbf{x}_k = \mathbf{a}_0 s_{0,k} + \mathbf{n}_k,\tag{1}$$

where  $\mathbf{a}_0, s_{0,k}$  and  $\mathbf{n}_k$  denote the true signal-of-interest (SOI) ASV, the scalar SOI complex amplitude and an additive zero-mean complex Gaussian vector that incorporates the noise and the interference. Assuming  $s_{0,k}$  is zero mean and uncorrelated with  $\mathbf{n}_k$ , the array covariance matrix can be written as  $\mathbf{R}_{\mathbf{x}} = E \{\mathbf{x}_k \mathbf{x}_k^H\} = \sigma_0^2 \mathbf{a}_0 \mathbf{a}_0^H + \mathbf{Q}_{\mathbf{x}}$ , where  $\mathbf{R}_{\mathbf{x}} > 0, \sigma_0^2 = E \{|s_{0,k}|^2\}$  is the SOI power and  $\mathbf{Q}_{\mathbf{x}} = E \{\mathbf{n}_k \mathbf{n}_k^H\}$  is the noise plus interference covariance matrix. In practice,  $\mathbf{R}_{\mathbf{x}}$  is replaced by the sample covariance matrix

$$\hat{\mathbf{R}}_{\mathbf{x}} = \frac{1}{K} \sum_{k=1}^{K} \mathbf{x}_k \mathbf{x}_k^H, \qquad (2)$$

formed from K snapshots. In [3] (see, also [6]), RCBs were derived by solving  $\max_{\sigma^2, \mathbf{a}} \sigma^2$  s.t.  $\mathbf{R}_{\mathbf{x}} - \sigma^2 \mathbf{a} \mathbf{a}^H \ge 0$ ,  $\mathbf{a} \in \mathcal{E}_M(\bar{\mathbf{a}}, \mathbf{E})$ , which can be reduced to [3]

$$\min_{\mathbf{a}} \mathbf{a}^{H} \mathbf{R}_{\mathbf{x}}^{-1} \mathbf{a} \quad \text{s.t.} \quad \mathbf{a} \in \mathcal{E}_{M}(\bar{\mathbf{a}}, \mathbf{E}).$$
(3)

The *M*-dimensional element-space ellipsoid  $\mathcal{E}_M(\bar{\mathbf{a}}, \mathbf{E})$  is parameterized by  $\bar{\mathbf{a}}$ , which often represents the assumed ASV, and  $\mathbf{E} \ge 0 \in \mathbb{C}^{M \times M}$ , and can be written as

$$\mathcal{E}_{M}(\bar{\mathbf{a}}, \mathbf{E}) = \left\{ \mathbf{a} \in \mathbb{C}^{M} \mid [\mathbf{a} - \bar{\mathbf{a}}]^{H} \mathbf{E}[\mathbf{a} - \bar{\mathbf{a}}] \leq 1 \right\}.$$
(4)

For non-degenerate sets,  $\mathbf{E} > 0$ . To solve (3), we assume that

$$\bar{\mathbf{a}}^H \mathbf{E} \bar{\mathbf{a}} > 1. \tag{5}$$

<sup>\*</sup>This work was supported by MOD under contract from the Centre for Defence Enterprise.

When  $\mathbf{E} = (1/\epsilon)\mathbf{I}$ , (4) reduces to a spherical uncertainty set,  $\|\mathbf{a} - \bar{\mathbf{a}}\|_2^2 \leq \epsilon$ , with radius  $\sqrt{\epsilon}$  and (5) becomes  $\|\bar{\mathbf{a}}\|_2^2 > \epsilon$ . For non-degenerate ellipsoids, we can factor  $\mathbf{E} = \mathbf{E}^{\frac{H}{2}} \mathbf{E}^{\frac{1}{2}}$  and form  $\breve{\mathbf{a}} = \mathbf{E}^{\frac{1}{2}} \mathbf{a}, \ \breve{\mathbf{a}} = \mathbf{E}^{\frac{1}{2}} \bar{\mathbf{a}} \text{ and } \ \breve{\mathbf{R}} = \mathbf{E}^{\frac{1}{2}} \mathbf{R} \mathbf{E}^{\frac{H}{2}}.$  Then, (3) can be re-written using the following spherical constraint [4]

$$\min_{\mathbf{x}} \mathbf{\breve{a}}^H \mathbf{\breve{R}}^{-1} \mathbf{\breve{a}} \text{ s.t. } \|\mathbf{\breve{a}} - \mathbf{\breve{a}}\|_2^2 \le 1.$$
 (6)

As shown in [4], (6) can be solved via the eigenvalue decomposition (EVD) of  $\check{\mathbf{R}}$ , where computing the EVD is the most computationally expensive step. Denoting  $\ddot{a}$  as the solution to (6), the solution to (3) is formed as  $\hat{\mathbf{a}}_{0,\text{RCB}} = \mathbf{E}^{-\frac{1}{2}} \hat{\mathbf{a}}$ . The RCB power estimate is formed as  $\hat{\sigma}_{0,\text{RCB}}^2 = \frac{\|\hat{\mathbf{a}}_{0,\text{RCB}}\|_2^2/M}{\hat{\mathbf{a}}_{0,\text{RCB}}^H \mathbf{R}_x^{-1} \hat{\mathbf{a}}_{0,\text{RCB}}}$  and the weight vector as  $\hat{\mathbf{w}}_{\text{RCB}} = \frac{\mathbf{R}_x^{-1} \hat{\mathbf{a}}_{0,\text{RCB}}}{\hat{\mathbf{a}}_{0,\text{RCB}}^H \mathbf{R}_x^{-1} \hat{\mathbf{a}}_{0,\text{RCB}}}$ .

#### 2. A DATA-ADAPTIVE REDUCED-DIMENSION ROBUST CAPON BEAMFORMING FRAMEWORK

In reduced-dimension methods, the kth element-space snapshot,  $\mathbf{x}_k \in \mathbb{C}^M$ , is projected onto an N-dimensional subspace (with N < M) using a dimension reducing transform  $\mathbf{D} \in \mathbb{C}^{M \times N}$ , yielding the reduced-dimension reducing transform  $\mathbf{D} \in \mathbb{C}^{H}$ ,  $\mathbf{y}_{k} \in \mathbb{C}^{N}$ . As shown in [1,7], this leads to the following RDRCB problem  $\max_{\sigma^{2},\mathbf{b}} \sigma^{2}$  s.t.  $\mathbf{R}_{\mathbf{y}} - \sigma^{2}\mathbf{b}\mathbf{b}^{H} \geq 0$ ,  $\mathbf{b} \in \mathcal{E}_{N}(\bar{\mathbf{b}}, \mathbf{F})$ , where  $\mathbf{b} = \mathbf{D}^H \mathbf{a}, \mathbf{R}_y = \mathbf{D}^H \mathbf{R}_x \mathbf{D}$  and  $\mathcal{E}_N(\bar{\mathbf{b}}, \mathbf{F})$  denote the reduced-dimension array steering vector, covariance and uncertainty ellipsoid, respectively, which can be reduced to

$$\min_{\mathbf{b}} \mathbf{b}^{H} \mathbf{R}_{\mathbf{y}}^{-1} \mathbf{b} \quad \text{s.t.} \quad \mathbf{b} \in \mathcal{E}_{N}(\bar{\mathbf{b}}, \mathbf{F}).$$
(7)

The following theorem is used to derive  $\mathcal{E}_N(\bar{\mathbf{b}}, \mathbf{F})$ .

Propagation Theorem: [1, 7] The propagation of the element-space ellipsoid (4), with  $\mathbf{E} \ge 0 \in \mathbb{C}^{M \times M}$ , through the mapping  $\mathbf{D}^H \mathbf{a} - \mathbf{I}_N \mathbf{b} = \mathbf{0}$ , where  $\mathbf{D} \in \mathbb{C}^{M \times N}$  has full column rank, yields the ellipsoid  $\mathcal{E}_N(\bar{\mathbf{b}}, \mathbf{F})$  [see (4)] with

$$\bar{\mathbf{b}} = \mathbf{D}^{H} \bar{\mathbf{a}}$$

$$\mathbf{F} = \mathbf{D}^{\dagger} (\mathbf{E} - \mathbf{E} \mathbf{N}_{\mathbf{D}}^{l} [(\mathbf{N}_{\mathbf{D}}^{l})^{H} \mathbf{E} \mathbf{N}_{\mathbf{D}}^{l}]^{\dagger} (\mathbf{N}_{\mathbf{D}}^{l})^{H} \mathbf{E}) (\mathbf{D}^{\dagger})^{H}.$$
(8)

For data-adaptive dimensionality reduction,  $\bar{\mathbf{b}}$  and  $\mathbf{F}$  need updating each time D is updated. If we use (9) for updating, then we observe that  $\mathbf{N}_{\mathbf{D}}^{l}$ ,  $[(\mathbf{N}_{\mathbf{D}}^{l})^{H}\mathbf{E}\mathbf{N}_{\mathbf{D}}^{l}]^{\dagger}$  and  $\mathbf{D}^{\dagger}$  are needed, which are expensive to compute. Fortunately, if the original element-space ellipsoid is non-degenerate, such that  $\mathbf{E} > 0$ , we can simplify (9). Then,  $[(\mathbf{N}_{\mathbf{D}}^{l})^{H}\mathbf{E}\mathbf{N}_{\mathbf{D}}^{l}]^{\dagger} = [(\mathbf{N}_{\mathbf{D}}^{l})^{H}\mathbf{E}\mathbf{N}_{\mathbf{D}}^{l}]^{-1}$  and we can write

$$\mathbf{F} = \mathbf{D}^{\dagger} \mathbf{E}^{\frac{1}{2}} \mathbf{\Pi}_{\mathbf{E}^{\frac{1}{2}} \mathbf{N}_{\mathbf{D}}^{L}}^{\perp} \mathbf{E}^{\frac{1}{2}} (\mathbf{D}^{\dagger})^{H}$$
$$= \mathbf{D}^{\dagger} \mathbf{E}^{\frac{1}{2}} \mathbf{\Pi}_{\mathbf{E}^{-\frac{1}{2}} \mathbf{D}}^{\frac{1}{2}} \mathbf{E}^{\frac{1}{2}} (\mathbf{D}^{\dagger})^{H}$$
$$= \left[ \mathbf{D}^{H} \mathbf{E}^{-1} \mathbf{D} \right]^{-1}, \qquad (10)$$

where  $\Pi_{\mathbf{E}^{\frac{1}{2}}\mathbf{N}_{\mathbf{D}}^{l}}^{\perp} = \mathbf{I} - \mathbf{E}^{\frac{1}{2}}\mathbf{N}_{\mathbf{D}}^{l}[(\mathbf{N}_{\mathbf{D}}^{l})^{H}\mathbf{E}\mathbf{N}_{\mathbf{D}}^{l}]^{-1}(\mathbf{N}_{\mathbf{D}}^{l})^{H}\mathbf{E}^{\frac{1}{2}}$ . The  $M \times M$  inverse  $\mathbf{E}^{-1}$  can be computed offline and therefore, the online computation of  $\mathbf{F}$  reduces to the computation of an  $N \times N$ inverse. Note that, in general, we will need to compute  $\mathbf{F}^{\frac{1}{2}}, \mathbf{F}^{\frac{H}{2}}$  and  $\mathbf{F}^{-\frac{1}{2}}$  [see Section 1.1], which can all be obtained from the EVD of

 $[\mathbf{D}^{H}\mathbf{E}^{-1}\mathbf{D}]$ . Note also that, in general, we will require the EVD of  $\mathbf{\ddot{R}_y} = \mathbf{F}^{\frac{1}{2}} \mathbf{R_y} \mathbf{F}^{\frac{H}{2}} = \mathbf{F}^{\frac{1}{2}} \mathbf{D}^H \mathbf{R_x} \mathbf{D} \mathbf{F}^{\frac{H}{2}}$ . Thus, in general, two Ndimensional EVDs will be required, one decomposing  $\mathbf{\breve{R}_v}$  and one decomposing  $[\mathbf{D}^{H}\mathbf{E}^{-1}\mathbf{D}]$ . When the element-space uncertainty set is a sphere, so that in (4),  $\mathbf{E} = \frac{1}{\epsilon} \mathbf{I}$ , then

$$\mathbf{F} = \frac{1}{\epsilon} (\mathbf{D}^H \mathbf{D})^{-1}.$$
 (11)

In this case, if **D** is orthogonal, **F** in (11) reduces to  $\mathbf{F} = \frac{1}{\epsilon} \mathbf{I}_N$ . Thus, if the element-space set is a sphere and the dimension reducing transform is orthogonal, then F can be written analytically and only one EVD is required. Denoting  $\hat{\mathbf{b}}_0$  as the solution to (7), we form the RDRCB weight vector as

$$\hat{\mathbf{w}}_{\text{RDRCB}} = \frac{\mathbf{R}_{\mathbf{y}}^{-1} \hat{\mathbf{b}}_0}{\hat{\mathbf{b}}_0^H \mathbf{R}_{\mathbf{y}}^{-1} \hat{\mathbf{b}}_0}.$$
 (12)

The weight vector (12) operates on the reduced-dimension data, whilst  $\hat{\mathbf{w}}_{\text{RDRCB,ES}} = \mathbf{D}\hat{\mathbf{w}}_{\text{RDRCB}}$  operates on the original elementspace data. An estimate of  $\mathbf{a}_0$  can be formed as  $\hat{\mathbf{a}}_0 = (\mathbf{D}^H)^{\dagger} \hat{\mathbf{b}}_0 =$  $\mathbf{D}(\mathbf{D}^H\mathbf{D})^{-1}\hat{\mathbf{b}}_0$ , which indicates that  $\hat{\mathbf{a}}_0$  belongs to the column space of **D**. Given  $\hat{\mathbf{a}}_0$ , we form the RDRCB SOI power estimate as

$$\hat{\sigma}_{0,\text{RDRCB}}^{2} = \frac{\left( \|\hat{\mathbf{a}}_{0}\|_{2}^{2} / M \right)}{\hat{\mathbf{b}}_{0}^{H} \mathbf{R}_{\mathbf{y}}^{-1} \hat{\mathbf{b}}_{0}} = \frac{\hat{\mathbf{b}}_{0}^{H} (\mathbf{D}^{H} \mathbf{D})^{-1} \hat{\mathbf{b}}_{0}}{M \hat{\mathbf{b}}_{0}^{H} \mathbf{R}_{\mathbf{y}}^{-1} \hat{\mathbf{b}}_{0}}.$$
 (13)

## 3. DATA-ADAPTIVE DIMENSIONALITY REDUCTION

Here, we consider Krylov methods that use the PoR and CG algorithms to compute dimension reducing transforms. The associated RDRCBs can be formed by exploiting these transforms in the framework presented in Section 2.

### 3.1. Non-Orthogonal PoR Krylov Basis

The standard PoR-based dimension reducing transform is given by

$$\mathbf{D} = \begin{bmatrix} \frac{\bar{\mathbf{a}}}{\|\bar{\mathbf{a}}\|_2} & \frac{\hat{\mathbf{R}}_{\mathbf{x}}\bar{\mathbf{a}}}{\|\bar{\mathbf{R}}_{\mathbf{x}}\bar{\mathbf{a}}\|_2} & \cdots & \frac{\hat{\mathbf{R}}_{\mathbf{x}}^{N-1}\bar{\mathbf{a}}}{\|\bar{\mathbf{R}}_{\mathbf{x}}^{N-1}\bar{\mathbf{a}}\|_2} \end{bmatrix}, \quad (14)$$

which we term the non-orthogonal PoR (NO-PoR) dimension reducing transform, and can be formed iteratively. That is, starting with  $\mathbf{\kappa}_1 = \bar{\mathbf{a}}$ , and  $\mathbf{D}_1 = \frac{\bar{\mathbf{a}}}{\|\bar{\mathbf{a}}\|_2}$ , for  $i = 2, \dots, N$ , calculate

$$\boldsymbol{\kappa}_i = \hat{\mathbf{R}}_{\mathbf{x}} \boldsymbol{\kappa}_{i-1},\tag{15}$$

$$\mathbf{d}_i = \frac{\boldsymbol{\kappa}_i}{\|\boldsymbol{\kappa}_i\|_2} \tag{16}$$

and

$$\mathbf{D}_{i} = \begin{bmatrix} \mathbf{D}_{i-1} & \mathbf{d}_{i} \end{bmatrix}.$$
(17)

The cost of calculating  $\kappa_i$  from  $\kappa_{i-1}$  is  $\mathcal{O}(M^2)$  and calculating  $\mathbf{d}_i$ is  $\mathcal{O}(M)$ . Thus, calculating the NO-PoR dimension reducing transform costs  $\mathcal{O}(NM[M+1])$  flops. Due to the non-orthogonal nature of the transform, two N-dimensional EVDs are required to compute the NO-PoR RDRCB, even if the element-space set is spherical.

#### 3.2. Orthogonal PoR Krylov Basis

In [9], the orthogonal PoR (O-PoR) Krylov subspace technique was proposed and suggested for applications where the model order is highly variable and time-varying. To form the O-PoR dimension reducing transform, let  $\kappa_1 = \bar{\mathbf{a}}$ ,  $\mathbf{D}_1 = \bar{\mathbf{a}}/\|\bar{\mathbf{a}}\|_2$ , and for  $i = 2, \ldots, N$ , calculate

$$\boldsymbol{\kappa}_i = \boldsymbol{\Pi}_{\mathbf{D}_{i-1}}^{\perp} \mathbf{\hat{R}}_{\mathbf{x}} \boldsymbol{\kappa}_{i-1}$$
(18)

$$\mathbf{d}_i = \frac{\boldsymbol{\kappa}_i}{\|\boldsymbol{\kappa}_i\|_2} \tag{19}$$

and

$$\begin{bmatrix} \mathbf{D}_{i-1} & \mathbf{d}_i \end{bmatrix}, \tag{20}$$

where  $\Pi_{\mathbf{D}_{i-1}}^{\perp} = \mathbf{I} - \sum_{k=1}^{i-1} \mathbf{d}_k \mathbf{d}_k^H$  can be updated efficiently in  $\mathcal{O}(M^2)$  operations using  $\Pi_{\mathbf{D}_i}^{\perp} = \Pi_{\mathbf{D}_{i-1}}^{\perp} - \mathbf{d}_i \mathbf{d}_i^H$ . Given  $\Pi_{\mathbf{D}_{i-1}}^{\perp}$ , updating  $\kappa_i$  and  $\mathbf{d}_i$  costs  $\mathcal{O}(2M^2)$  and  $\mathcal{O}(M)$ . Thus, the calculation of one new column of  $\mathbf{D}$  costs  $\mathcal{O}(3M^2 + M)$ , so that calculation of the O-PoR dimension reducing transform costs  $\mathcal{O}(NM[3M + 1])$ , which is roughly three times more expensive than calculating the NO-PoR transform. Since the resulting dimension reducing transform is orthogonal, as discussed earlier, for spherical uncertainty sets only one EVD is required to compute the O-PoR RDRCB.

 $\mathbf{D}_i =$ 

#### 3.3. Conjugate Gradient Method

Using the approach outlined in [11], the CG dimension reducing transform  $\mathbf{D} = [\mathbf{d}_1 \dots \mathbf{d}_N]$  can be formed by setting,  $\mathbf{d}_1 = \bar{\mathbf{a}}$ ,  $\mathbf{r}_1 = -\bar{\mathbf{a}}$ , and then for  $i = 2, \dots, N$ , update using

$$\alpha_i = -\frac{\mathbf{d}_i^H \mathbf{r}_i}{\mathbf{d}_i^H \hat{\mathbf{R}}_{\mathbf{x}} \mathbf{d}_i},\tag{21}$$

$$\mathbf{r}_{i+1} = \mathbf{r}_i + \alpha_i \hat{\mathbf{R}}_{\mathbf{x}} \mathbf{d}_i, \tag{22}$$

$$\beta_i = \frac{\mathbf{d}_i^H \hat{\mathbf{R}}_{\mathbf{x}} \mathbf{r}_{i+1}}{\mathbf{d}_i^H \hat{\mathbf{R}}_{\mathbf{x}} \mathbf{d}_i}$$
(23)

and

$$\mathbf{d}_{i+1} = -\mathbf{r}_{i+1} + \beta_i \mathbf{d}_i. \tag{24}$$

The cost of computing  $\hat{\mathbf{R}}_{\mathbf{x}} \mathbf{d}_i$  is  $\mathcal{O}(M^2)$ . Given  $\hat{\mathbf{R}}_{\mathbf{x}} \mathbf{d}_i$ , the cost of computing  $\alpha_i$  is  $\mathcal{O}(2M)$ . Updating  $\mathbf{r}_{i+1}$  is  $\mathcal{O}(M)$ . The cost of computing  $\beta_i$ , given  $\hat{\mathbf{R}}_{\mathbf{x}} \mathbf{d}_i$  and the denominator of  $\alpha_i$  is  $\mathcal{O}(M)$ . Then, updating  $\mathbf{d}_{i+1}$  is  $\mathcal{O}(M)$ . Thus, the total cost to compute a new column of the CG dimension reducing transform is  $\mathcal{O}(M^2 + 5M)$ . The total cost to calculate the CG transform is  $\mathcal{O}(NM[M + 5])$ , which is almost the same as calculating the NO-POR transform. Since the CG transform is non-orthogonal, we would expect that we would need two EVDs to compute the CG-RDRCB. However, in the next section, we illustrate how a fast CG-based RDRCB can be obtained by exploiting that the CG transform diagonalizes the sample covariance matrix so that

$$\hat{\mathbf{R}}_{\mathbf{y}} = \mathbf{D}^H \hat{\mathbf{R}}_{\mathbf{x}} \mathbf{D} = \mathbf{\Lambda}_{\mathrm{CG}},$$
(25)

where  $\Lambda_{\text{CG}}$  is a diagonal matrix.

## 4. FAST CONJUGATE-GRADIENT RDRCB

Here, we illustrate how only one N-dimensional EVD is required to solve the CG-RDRCB under either spherical or non-degenerate uncertainty. In general, we will be solving

$$\min_{\mathbf{b}} \mathbf{b}^{H} \mathbf{R}_{\mathbf{y}}^{-1} \mathbf{b} \text{ s.t. } \left[ \mathbf{b} - \bar{\mathbf{b}} \right]^{H} \mathbf{F} \left[ \mathbf{b} - \bar{\mathbf{b}} \right] \le 1.$$
 (26)

Usually, at this stage one would transform with  $\mathbf{F}^{\frac{1}{2}}$  to give a spherical uncertainty set. However, from (25), we observe that  $\mathbf{R}_{\mathbf{y}}^{-1} = \mathbf{\Lambda}_{CG}^{-1}$ , so that (26) can be written as

$$\min_{\mathbf{b}} \mathbf{b}^{H} \mathbf{\Lambda}_{CG}^{-1} \mathbf{b} \text{ s.t. } \left[\mathbf{b} - \bar{\mathbf{b}}\right]^{H} \mathbf{F} \left[\mathbf{b} - \bar{\mathbf{b}}\right] \leq 1.$$
(27)

Noting (10), we let  $\mathbf{M} = \mathbf{\Lambda}_{CG}^{-\frac{1}{2}} \mathbf{D}^{H} \mathbf{E}^{-1} \mathbf{D} \mathbf{\Lambda}_{CG}^{-\frac{H}{2}}$ ,  $\check{\mathbf{b}} = \mathbf{\Lambda}_{CG}^{-\frac{1}{2}} \mathbf{b}$  and  $\check{\mathbf{b}} = \mathbf{\Lambda}_{CG}^{-\frac{1}{2}} \bar{\mathbf{b}}$ , and rewrite (27) as

$$\min_{\check{\mathbf{b}}}\check{\mathbf{b}}^{H}\check{\mathbf{b}} \text{ s.t. } \left[\check{\mathbf{b}}-\check{\bar{\mathbf{b}}}\right]^{H}\mathbf{M}^{-1}\left[\check{\mathbf{b}}-\check{\bar{\mathbf{b}}}\right] \leq 1.$$
(28)

We form the Lagrangian using the real Lagrange multiplier  $\mu$ 

$$L(\check{\mathbf{b}},\mu) = \check{\mathbf{b}}^{H}\check{\mathbf{b}} + \mu\left(\left[\check{\mathbf{b}} - \check{\bar{\mathbf{b}}}\right]^{H}\mathbf{M}^{-1}\left[\check{\mathbf{b}} - \check{\bar{\mathbf{b}}}\right] - 1\right).$$
 (29)

Setting  $\frac{\partial L(\check{\mathbf{b}},\mu)}{\partial \check{\mathbf{b}}^H} = \mathbf{0}$  yields

$$\hat{\mathbf{b}} = \left(\frac{\mathbf{M}}{\mu} + \mathbf{I}\right)^{-1} \check{\mathbf{b}} = \check{\mathbf{b}} - \left[\mu \mathbf{M}^{-1} + \mathbf{I}\right]^{-1} \check{\mathbf{b}}, \qquad (30)$$

where we have used the matrix inversion lemma [13] to obtain the term after the second equality. Using (30) in the constraint equation in (28) yields

$$h(\hat{\mathbf{b}},\mu) = \check{\mathbf{b}}^H \left[\mu \mathbf{M}^{-1} + \mathbf{I}\right]^{-1} \mathbf{M}^{-1} \left[\mu \mathbf{M}^{-1} + \mathbf{I}\right]^{-1} \check{\mathbf{b}}.$$
 (31)

Letting  $\mathbf{M} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$  denote the EVD of  $\mathbf{M}$ , where  $\mathbf{\Lambda} = \text{diag}\left\{ \begin{bmatrix} \lambda_1 & \dots & \lambda_N \end{bmatrix} \right\}$  is a diagonal matrix containing the eigenvalues in non-increasing order on its main diagonal and  $\mathbf{U}$  contains the associated eigenvectors, we can write (31) as

$$h(\hat{\mathbf{b}},\mu) = \sum_{n=1}^{N} \frac{\lambda_n |c_n|^2}{(\mu + \lambda_n)^2},$$
(32)

where  $c_n$  is the *n*th element of  $\mathbf{c} = \mathbf{U}^H \check{\mathbf{b}}$ . Since we can write  $\mathbf{M} = \mathbf{M}^{\frac{1}{2}} \mathbf{M}^{\frac{H}{2}}$ , where  $\mathbf{M}^{\frac{1}{2}} = \mathbf{\Lambda}_{CG}^{-\frac{1}{2}} \mathbf{D}^H \mathbf{E}^{-\frac{1}{2}}$ , we know that  $\mathbf{M}$  is non-negative definite [13, 14] and therefore, it has non-negative eigenvalues. Thus,  $h(\hat{\mathbf{b}}, \mu)$  is a monotonically decreasing function of  $\mu > 0$ . For  $\mu = 0$ , we obtain

$$h(\hat{\mathbf{b}},0) = \check{\mathbf{b}}^H \mathbf{M}^{-1} \check{\mathbf{b}} = \bar{\mathbf{b}}^H \left[ \mathbf{D}^H \mathbf{E}^{-1} \mathbf{D} \right]^{-1} \bar{\mathbf{b}} = \bar{\mathbf{b}}^H \mathbf{F} \bar{\mathbf{b}}.$$
 (33)

Note that, to exclude a non-trivial solution, we require that  $\mathbf{\bar{b}}^H \mathbf{F} \mathbf{\bar{b}} > 1$ . Since we require  $h(\hat{\mathbf{b}}, \mu) = 1$ , it is clear that  $\mu \neq 0$ . Further, it is clear that  $\lim_{\mu \to \infty} h(\hat{\mathbf{b}}, \mu) = 0$ , therefore, there is a unique solution  $\mu > 0$  to  $h(\hat{\mathbf{b}}, \mu) = 1$ , which can be found, e.g., by Newton search. Once  $\mu$  has been found,  $\hat{\mathbf{b}}$  is found using (30) and the solution to (27) is formed as  $\hat{\mathbf{b}}_0 = \mathbf{A}_{CG}^{\frac{1}{2}}\hat{\mathbf{b}}$ . We can use  $\hat{\mathbf{b}}_0$  and  $\mathbf{R}_{\mathbf{y}}^{-1} = \mathbf{A}_{CG}^{-1}$  in (12) to form the adaptive weights. To form the power estimate using (13), we need to evaluate  $[\mathbf{D}^H \mathbf{D}]^{-1}$ . If the uncertainty set is spherical, then we can evaluate this quantity from the EVD of M and  $\mathbf{A}_{CG}$ , which are already available. For a general, non-degenerate ellipsoid this quantity will need computing.

Fig. 1 shows the relative complexities as N is increased from 1 to M, for M = 320, where for the NO-PoR and CG-based RDRCBs under non-degenerate (ND) uncertainty, we distinguish between applications that only require weight estimation and those which also need power estimation. The results illustrate that the CG-based algorithms are significantly cheaper than the other methods.



Fig. 1. Relative complexities of different data-dependent RDRCBs.

#### 5. NUMERICAL EXAMPLES

In this section, we assess the performance of the proposed algorithms through numerical examples. For an M = 320, halfwavelength spaced planar array with  $M_h = 40$  elements in a row and  $M_v = 8$  rows, we simulated data with covariance matrix  $\mathbf{R}_{\mathbf{x}} = \sigma_0^2 \mathbf{a}_0 \mathbf{a}_0^H + \mathbf{Q}_{\mathbf{x}}$ , with  $\mathbf{Q}_{\mathbf{x}} = \sum_{i=1}^d \sigma_i^2 \mathbf{a}_i \mathbf{a}_i^H + \sigma_s^2 \mathbf{I} + \sigma_{iso}^2 \mathbf{Q}_{iso}$ , where  $\mathbf{Q}_{\mathbf{x}}$  consists of terms due to *d* zero-mean uncorrelated interferences, where for the *i*th interferer  $\sigma_i^2$  and  $\mathbf{a}_i$  denote the source power and ASV, a term modeling sensor noise  $\sigma_s^2 \mathbf{I}$ , with sensor noise power  $\sigma_s^2$ , and a term modeling an isotropic ambient noise  $\sigma_{iso}^2 \mathbf{Q}_{iso}$ , with power  $\sigma_{iso}^2$ . The isotropic noise covariance is given by  $[\mathbf{Q}_{iso}]_{m,n} = \operatorname{sinc}[\pi g_{mn}]$ , where  $g_{mn}$  is the distance between the mth and nth sensors in units of wavelength. The *i*th source (SOI or interference) ASV is simulated according to  $\mathbf{a}_i = \mathbf{a}(\bar{\boldsymbol{\theta}}_i + \boldsymbol{\delta}_i) + \sigma_{e,i}\mathbf{e}_i$ , where  $\bar{\boldsymbol{\theta}}_i = [\bar{\theta}_i, \bar{\phi}_i]$ denotes the assumed AOA, comprising the assumed azimuth angle  $\bar{\theta}_i$  and the assumed elevation angle  $\bar{\phi}_i$ ,  $\boldsymbol{\delta}_i = [\delta_{\theta_i}, \delta_{\phi_i}]^{\dagger}$ denotes the mismatch in the assumed AOA, comprising the mismatch in azimuth angle  $\delta_{\theta_i}$  and the mismatch in elevation angle  $\delta_{\phi_i}$ ,  $\sigma_{e,i}$  denotes the length of the arbitrary error vector and  $\mathbf{e}_i$  is a zero-mean complex circularly symmetric random vector with unit norm. Unless stated otherwise, in the following, we have assumed d = 3,  $\bar{\theta}_0 = 89.9^\circ$ ,  $\bar{\phi}_0 = 94.8^\circ$ ,  $\sigma_0^2 = 40$  dB,  $\bar{\theta}_1 = 70^\circ$ ,  $\bar{\phi}_1 = 94.8^\circ$ ,  $\sigma_1^2 = 30$  dB,  $\bar{\theta}_2 = 85^\circ$ ,  $\bar{\phi}_2 = 94.8^\circ$ ,  $\sigma_2^2 = 20$  dB,  $\bar{\theta}_3 = 130^\circ$ ,  $\bar{\phi}_3 = 94.8^\circ$  and  $\sigma_3^2 = 15$  dB. We assume that  $\delta_{\theta_0}$  is a random variable uniformly distributed over the interval  $[-0.72, 0.72]^{\circ}$  and  $\delta_{\phi_0}$  is uniformly distributed over  $[-3.6, 3.6]^{\circ}$ , whilst for i = 1 to d,  $\delta_i = 0$ . For i = 0 to d,  $\sigma_{e,i} = 1$ . Therefore, the SOI ASV is subjected to both AOA and arbitrary errors. We assume azimuth and elevation beams spaced at  $1/M_h$  and  $1/M_v$  in cosine space and, using the methods described in [15], design nondegenerate minimum volume ellipsoidal (NDMVE) sets, whose error sphere radii are set to  $\check{\epsilon} = 10$ , and tight-spherical uncertainty sets, based on the expected AOA errors given the spacing of the beams. We refer the reader to [15] for further details. Thus, we examine the following RDRCBs: NO-PoR-Spherical, NO-PoR-NDMVE, O-PoR-Spherical, O-PoR-NDMVE, CG-Spherical and CG-NDMVE, where we append Spherical or NDMVE to denote that we are exploiting a tight spherical uncertainty set or a non-degenerate NDMVE uncertainty set. We also examine the standard non-robust



Fig. 2. Azimuth spectra, for the horizontal elevation at  $94.8^{\circ}$ , where the true AOAs and powers are shown by the black circles.



Fig. 3. SINR versus SOI power.

reduced-dimension MVDR (RDMVDR) versions: NO-PoR-MVDR, O-PoR-MVDR and CG-MVDR. RDMVDR weights can be formed using (12), but with  $\bar{\mathbf{b}}$  in place of  $\hat{\mathbf{b}}_0$ . We assume that N = 5and that only K = 80 snapshots are available to estimate the sample covariance matrix (2), representing a highly snapshot deficient scenario, not uncommon in, e.g., passive sonar. Fig. 2 shows the azimuth spectra, illustrating that the RDRCB variants are able to correctly estimate the power, whereas the non-robust RDMVDRbased variants suffer severe SOI cancellation. Fig. 3 shows the output SINR versus the SOI power, obtained using MC = 200 Monte Carlo simulations. The results for the CG and O-PoR methods are the same, whilst the NO-PoR results diverge for very high SOI powers. This divergence is a result of numerical instability in the NO-PoR algorithm. It is clear that the robust RDRCB versions, exploiting spherical or non-degenerate NDMVE sets, provide much better robustness at higher SOI powers compared to the standard RDMVDR-based implementations.

#### 6. REFERENCES

- S. D. Somasundaram, "Reduced Dimension Robust Capon Beamforming for Large Aperture Passive Sonar Arrays," *IET Radar, Sonar Navig.*, vol. 5, no. 7, pp. 707–715, Aug. 2011.
- [2] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust Adaptive Beamforming Using Worst-Case Performance Optimization: A Solution to the Signal Mismatch Problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.
- [3] P. Stoica, Z. Wang, and J. Li, "Robust Capon Beamforming," *IEEE Sig. Process. Lett.*, vol. 10, no. 6, pp. 172–175, Jun. 2003.
- [4] J. Li, P. Stoica, and Z. Wang, "On Robust Capon Beamforming and Diagonal Loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.
- [5] R. G. Lorenz and S. P. Boyd, "Robust Minimum Variance Beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684–1696, May 2005.
- [6] J. Li and P. Stoica, *Robust Adaptive Beamforming*, Wiley, New York, 2005.
- [7] S. D. Somasundaram, "A Framework for Reduced Dimension Robust Capon Beamforming," in *Proc. IEEE Workshop on Statistical Signal Processing*, Nice, France, Jun. 28-30 2011, pp. 425–428.
- [8] M. L. Honig and J. S. Goldstein, "Adaptive Reduced-Rank Interference Suppression Based on the Multistage Wiener Filter," *IEEE Transactions on Communications*, vol. 50, no. 6, pp. 986–994, Jun. 2002.
- [9] H. Ge, I. P. Kirsteins, and L. L. Scharf, "Data Dimension Reduction Using Krylov Subspaces: Making Adaptive Beamformers Robust to Model Order-Determination," in 31<sup>st</sup> IEEE International Conference on Acoustics, Speech and Signal Processing, Toulouse, Fr, May. 2006, vol. 4, pp. 1001–1004.
- [10] 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.
- [11] G. Dietl, "Conjugate Gradient Implementation of Multi-Stage Nested Wiener Filter for Reduced-Dimension Processing," in *MSc. Dissertation, Munich University of Technology*, Munich, Germany, May 2001.
- [12] L. Wang and R. C. de Lamare, "Constrained adaptive filtering algorithms based on conjugate gradient techniques for beamforming," *IET Signal Processing*, vol. 4, pp. 686–697, 2010.
- [13] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*, Prentice Hall, 2005.
- [14] G. Strang, *Linear Algebra and its Applications (3rd Edition)*, Harcourt Brace Jovanovich, 1988.
- [15] S. D. Somasundaram, A. Jakobsson, and N. H. Parsons, "Robust and Automatic Data-Adaptive Beamforming for Multi-Dimensional Arrays," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4642–4656, Nov. 2012.