Low-Complexity Uncertainty-Set-Based Robust Adaptive Beamforming for Passive Sonar

Somasundaram, Samuel; Butt, Naveed; Jakobsson, Andreas; Hart, Les

Published in:
IEEE Journal of Oceanic Engineering

DOI:
10.1109/JOE.2015.2474495

2015

LUND UNIVERSITY

Low-Complexity Uncertainty Set Based Robust Adaptive Beamforming for Passive Sonar

SAMUEL D. SOMASUNDARAM, NAVEED R. BUTT,
ANDREAS JAKOBSSON, AND LES HART

Published in: IEEE Journal of Oceanic Engineering
doi:10.1109/JOE.2015.2474495

Lund 2015

Mathematical Statistics
Centre for Mathematical Sciences
Lund University

Low-Complexity Uncertainty Set Based Robust Adaptive Beamforming for Passive Sonar

Samuel D. Somasundaram, *Member, IEEE*, Naveed R. Butt, *Member, IEEE*, Andreas Jakobsson, *Senior Member, IEEE*, and Les Hart

Abstract

Recent work has highlighted the potential benefits of exploiting ellipsoidal uncertainty set based robust Capon beamformer (RCB) techniques in passive sonar. Regrettably, the computational complexity of forming RCB weights is cubic in the number of adaptive degrees of freedom, which is often prohibitive in practice. For this reason, several low-complexity techniques for computing RCB weights, or equivalent worst-case robust adaptive beamformer (WC-RAB) weights, have recently been developed, whose complexities are only quadratic in the number of adaptive degrees of freedom. In this work, we review several such techniques for passive sonar, evaluating them initially on simulated data. The best performing methods are then evaluated on in-water recorded passive sonar data.

Index Terms

Efficient robust adaptive beamforming, robust Capon beamforming, worst-case robust adaptive beamforming, robust adaptive beamforming, underwater acoustics.

I. Introduction

In this work, we are interested in examining techniques for efficiently computing ellipsoidal uncertainty set based robust Capon beamformer (RCB) weights, or equivalent worst-case robust adaptive beamformer (WC-RAB) weights, for the purposes of implementing recent RCB-based passive sonar beamformers with low computational complexity.

As is well known, an adaptive beamformer can outperform a conventional delay-and-sum beamformer providing that it a) is robust to array steering vector errors, b) can converge sufficiently on the, often limited, data available, and c) is low enough complexity to be implementable on the, typically limited, computational resources that are available [1], [2]. The model of the signal-of-interest spatial signature, or array steering vector, that is used by a beamformer is usually derived under plane-wave assumptions, as a set of complex phase delays, and is subject to several sources of error. These include calibration errors, due to sensor gain/phase errors, mutual coupling effects and sensor position errors, pointing errors, caused by differences in the assumed angle-of-arrival and the true one, and deviations from plane wave assumptions, e.g., caused by inhomogeneity in the ocean and/or multipath propagation. If not properly dealt with, these errors lead to cancellation of the desired signal and a consequent loss of array output signal to noise ratio [3]. A wealth of robust adaptive techniques have been proposed to deal with steering vector errors, including using extra linear constraints [4]–[8], diagonal loading [9], white noise gain or weight norm constraints [10], and more recently ellipsoidal uncertainty set based robust Capon beamforming [2], [11], [12] or, equivalently, worst-case robust adaptive beamforming (WC-RAB) [13], [14]. The main practical drawback with many robust adaptive algorithms is that the user is a required to select a parameter that is not directly related to the steering vector uncertainty, making the parameter choice difficult and often ad hoc. For instance, it is unclear how to choose a white noise gain constraint, which is a constraint on the norm of the weights, based on the array steering vector uncertainty [15]. The main benefit of the

This work was supported in part by Thales U.K., the Swedish Research Council, and Carl Trygger's foundation.

S. D. Somasundaram and L. Hart are with the General Sonar Studies Group, Thales U.K., Dolphin House, Ashurst Drive, Bird Hall Lane, Cheadle Heath, Stockport, Cheshire, SK3 0XB, U.K. (emails: sdsomasundaram@hotmail.com; les.hart@uk.thalesgroup.com).

N. R. Butt and A. Jakobsson are with the Department of Mathematical Statistics, Lund University, 221 00 Lund, Sweden (emails: naveed@maths.lth.se; aj@maths.lth.se).

RCB-type techniques, which are the focus of this work, is that their constraint is the ellipsoidal uncertainty set of the array steering vector, which can be defined in a systematic manner to allow for mismatch in passive sonar applications [16]–[19]. We remark that the benefits of using a narrowband robust Capon beamformer (NBRCB) over the commonly used white noise gain constrained MVDR beamformer, with or without extra linear constraints, were shown in [16] for passive sonar.

To provide sufficient sensitivity, passive sonar arrays are typically large and contain many sensors and therefore, if independent adaptive weights were calculated for each sensor, would require a large number of independent snapshots for good covariance matrix estimation. However, passive sonars often operate in highly dynamic environments [20], which limits the amount of stationary data available for covariance matrix estimation, leading to a snapshot deficient situation. Since covariance matrix estimation error can be viewed as a type of array steering vector error [21], many of the robust adaptive techniques that are used to deal with array steering vector errors, including robust Capon beamforming, can also, up to a point, alleviate the effects of snapshot deficiency. Other notable approaches, which can be combined with robust Capon beamforming, when needed, include using reduced-dimension techniques, such as beamspace [22] or sub-array [23]–[26] preprocessing, eigenvalue [27] and Krylov methods [28], [29], dominant mode rejection [27], [30], [31], null broadening [32], and wideband integration techniques [19]. For instance in [18], [33], a framework for combining any reduced-dimension technique with robust Capon beamforming was derived for producing reduced-dimension RCBs (RDRCBs), whereas in [19], the wideband nature of sonar signals was exploited to speed-up algorithm convergence and reduce computational complexity, by using the steered covariance matrix method with the RCB to form wideband RCBs (WBRCBs).

At the core of the NBRCB, WBRCB, and RDRCB algorithms is an RCB algorithm that computes RCB weights by solving an RCB optimization problem using the Lagrange multiplier method outlined in [2], [11], [12] via eigenvalue decomposition (EVD) of the sample covariance matrix, which we term RCB-EVD. Due to the required eigenvalue decomposition, RCB-EVD is $\mathcal{O}(M^3)$ complexity, where M denotes the number of adaptive degrees of freedom, which is often computationally prohibitive in practice. To alleviate this problem, the aim of this work is to examine recent low complexity techniques for computing RCB weights efficiently[1]. As shown in [34], the WC-RAB and RCB weights are equivalent[2], even though their optimization problems are formulated quite differently. Therefore, we examine recent efficient implementations, whose complexities are only quadratic in M, that start from either the WC-RAB or the RCB formulation. Those that start from a WC-RAB formulation are the (second-order) constrained Kalman filter implementation of the WC-RAB [36], which we term WC-KF and re-writes the WC-RAB optimization as a minimum mean-squared error problem that can be solved using Kalman filtering techniques, and a modified version of the iterative gradient-based implementation of the WC-RAB [37], which we term WC-IG and solves the WC-RAB optimization problem iteratively using gradient-based methods. We examine three methods that solve the RCB optimization. Two of these use steepest descent methods, with one exploiting variable diagonal loading [38], which we term RCB-VDL-SD, and the other exploiting scaled projections [39], which we term RCB-SP-SD. The third, which we term the Krylov-RDRCB, solves the RCB problem by first projecting the data onto a reduced-rank Krylov-subspace and then solving a reduced-dimension RCB problem [29]. For quick reference, we refer the reader to Table II, which summarizes the complexity of the five efficient algorithms and the eigenvalue decomposition based

[1]In [34], it was suggested that a recursive eigenvalue decomposition (EVD) approach could be used to implement the RCB. Although such schemes may be acceptable when only adding snapshots to the sample covariance matrix, they often become unstable and more expensive when simultaneous addition and deletion of snapshots are required, as is the case for passive sonar. Further, due to the overheads involved, these techniques are only beneficial when M is sufficiently large, which is typically not the case in passive sonar applications.

[2]Due to the equivalence of the RCB and WC-RAB weights, one could instead consider implementations based on the WC-RAB [13], [14], [35], which can be formulated as a convex second-order cone program and solved using the interior-point method in $\rho\mathcal{O}(M^3)$ complexity [13], where ρ is the number of required iterations (typically less than 10), or solved using the Lagrange multiplier method in $\mathcal{O}(M^3)$ complexity [14], [35]. However, as was shown in [12], the RCB weights are simpler to compute and the RCB method provides a means for addressing the scaling ambiguity issue occurring in WC-RABs. For this reason, several of our previous works have focused on RCB based beamforming [16]–[19].

RCB-EVD. Initially, we compare and contrast these efficient algorithms using simulated passive sonar data and then select the most promising algorithms to implement the RCB-based NBRCB algorithm used in [16], [19]. These efficient implementations are evaluated on in-water recorded experimental data and compared with the original RCB-EVD based implementation. We note that these techniques could equally well be used to implement other RCB-based algorithms such as the WBRCBs or RDRCBs.

The remainder of this paper is organized as follows: In Section II, we briefly review the data model, the NBRCB approach, the uncertainty set design, and the RCB and WC-RAB formulations. In Sections III–VII, we summarize the efficient algorithms, before examining their complexities in Section VIII. In Sections IX and X, we evaluate the efficient algorithms on simulated and experimental in water passive sonar data, respectively. 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, $\|\cdot\|_2$ denotes the two-norm. Moreover, $\mathbf{X} \geq 0$ ($\mathbf{X} > 0$) implies that the Hermitian matrix $\mathbf{X}$ is positive semi-definite (positive definite). Part of this work was presented at UDT 2014 [40].

II. Data Model and Uncertainty Set Based Robust Adaptive Beamforming

To apply narrowband beamforming methods to wideband sonar signals, the time-series sensor data are Fourier transformed, using the fast Fourier transform (FFT), to a frequency resolution, B_w, satisfying

$$B_w < \frac{1}{10\, T_{\text{transit}}} \tag{1}$$

where T_{transit} denotes the transit time across the array, so that the signals in each FFT bin can be assumed to be narrowband. The kth snapshot from FFT bin with center frequency f can be modeled as (see, e.g., [19])

$$\begin{aligned} \mathbf{x}_{k,f} &\triangleq \left[\begin{array}{ccc} x_{1,k}(f) & \ldots & x_{M,k}(f) \end{array}\right]^T \\ &= \mathbf{a}_0(f,\boldsymbol{\theta}_0)s_{0,k}(f) + \mathbf{n}_{k,f} \end{aligned} \tag{2}$$

where $x_{m,k}(f)$ and $s_{0,k}(f)$ are the kth Fourier coefficients associated with the mth sensor and signal-of-interest (SOI) waveform, respectively, $\mathbf{n}_{k,f}$ is a zero-mean complex Gaussian noise vector, uncorrelated with the SOI and defined similarly to $\mathbf{x}_{k,f}$, and $\mathbf{a}_0(f,\boldsymbol{\theta}_0)$ is the true SOI array steering vector (ASV) at center frequency f. For a source at $\boldsymbol{\theta}$, the ASV can be modeled as

$$\mathbf{a}(f,\boldsymbol{\theta}) \triangleq \left[\begin{array}{ccc} e^{-i2\pi f\tau_1(\boldsymbol{\theta})} & \ldots & e^{-i2\pi f\tau_M(\boldsymbol{\theta})} \end{array}\right]^T, \tag{3}$$

where $\tau_m(\boldsymbol{\theta})$ is the propagation delay to the mth sensor, relative to some reference point, for a signal impinging from a location described by $\boldsymbol{\theta}$. The cross-spectral density matrix (CSDM) at frequency f is given by

$$\mathbf{R}_f \triangleq E\left\{\mathbf{x}_{k,f}\mathbf{x}_{k,f}^H\right\} = \sigma_{0,f}^2\mathbf{a}_0(f,\boldsymbol{\theta}_0)\mathbf{a}_0^H(f,\boldsymbol{\theta}_0) + \mathbf{Q}_f \tag{4}$$

where $\sigma_{0,f}^2 = E\left\{|s_{0,k}(f)|^2\right\}$ is the desired signal power at frequency f and $\mathbf{Q}_f = E\left\{\mathbf{n}_{k,f}\mathbf{n}_{k,f}^H\right\}$ is the noise-plus-interference covariance. In practice, the true CSDM is not available, and $\mathbf{R}_f$ is replaced by the sample CSDM estimate

$$\hat{\mathbf{R}}_f = \frac{1}{K}\sum_{k=1}^{K}\mathbf{x}_{k,f}\mathbf{x}_{k,f}^H \tag{5}$$

In the following, we will specify the different algorithms assuming $\mathbf{R}_f$, but note that in practice they will be implemented using $\hat{\mathbf{R}}_f$.

A. The Narrowband Robust Capon Beamformer (NBRCB)

For the bth beam, with an assumed steer direction of $\bar{\boldsymbol{\theta}}_b$, and center frequency f the NBRCB is obtained from $\max_{\sigma^2,\mathbf{a}}\ \sigma^2$ s.t. $\mathbf{R}_f - \sigma^2\mathbf{a}\mathbf{a}^H \geq 0,\ \mathbf{a} \in \mathcal{E}_M(\bar{\mathbf{a}}_{\bar{\boldsymbol{\theta}}_b,f}, \mathbf{E}_{\bar{\boldsymbol{\theta}}_b,f})$, which can be reduced to [11], [12]

$$\min_{\mathbf{a}} \mathbf{a}^H\mathbf{R}_f^{-1}\mathbf{a} \quad \text{s.t.} \quad \mathbf{a} \in \mathcal{E}_M(\bar{\mathbf{a}}_{\bar{\boldsymbol{\theta}}_b,f}, \mathbf{E}_{\bar{\boldsymbol{\theta}}_b,f}) \tag{6}$$

where $\mathcal{E}_M(\bar{\mathbf{a}}_{\bar{\boldsymbol{\theta}}_b,f}, \mathbf{E}_{\bar{\boldsymbol{\theta}}_b,f})$ denotes the uncertainty ellipsoid for the bth beam and the frequency bin with center frequency f. The ellipsoid is parameterized by its center $\bar{\mathbf{a}}_{\bar{\boldsymbol{\theta}}_b,f}$, which often represents the assumed ASV, and by its principal semi-axes, described by the unit-norm left singular vectors of $\mathbf{E}_{\bar{\boldsymbol{\theta}}_b,f}^{-1/2}$ scaled by the corresponding singular values, where $\mathbf{E}_{\bar{\boldsymbol{\theta}}_b,f} \geq 0 \in \mathbb{C}^{M\times M}$, and can be written as $\mathcal{E}_M(\bar{\mathbf{a}}_{\bar{\boldsymbol{\theta}}_b,f}, \mathbf{E}_{\bar{\boldsymbol{\theta}}_b,f}) = \left\{\mathbf{a} \;\middle|\; [\mathbf{a}-\bar{\mathbf{a}}_{\bar{\boldsymbol{\theta}}_b,f}]^H\mathbf{E}_{\bar{\boldsymbol{\theta}}_b,f}[\mathbf{a}-\bar{\mathbf{a}}_{\bar{\boldsymbol{\theta}}_b,f}] \leq 1\right\}$. If $\mathbf{E}_{\bar{\boldsymbol{\theta}}_b,f} = \epsilon_{\bar{\boldsymbol{\theta}}_b,f}^{-1}\mathbf{I}$, the ellipsoid reduces to a spherical uncertainty set $\left\|\mathbf{a}-\bar{\mathbf{a}}_{\bar{\boldsymbol{\theta}}_b,f}\right\|_2^2 \leq \epsilon_{\bar{\boldsymbol{\theta}}_b,f}$ with radius $\sqrt{\epsilon_{\bar{\boldsymbol{\theta}}_b,f}}$. The uncertainty sets must be specified for each beam and FFT bin (see, e.g., [2], [14], [16], [17], [19], [41] for ellipsoidal uncertainty set design techniques). Denoting the solution to (6) $\hat{\mathbf{a}}_{0,\bar{\boldsymbol{\theta}}_b,f}$ (see also [2], [11], [12]), to overcome the scaling ambiguity, the ASV is scaled so that its Euclidean length squared equals M, forming

$$\hat{\hat{\mathbf{a}}}_{0,\bar{\boldsymbol{\theta}}_b,f} = \sqrt{M}\hat{\mathbf{a}}_{0,\bar{\boldsymbol{\theta}}_b,f} / \left\|\hat{\mathbf{a}}_{0,\bar{\boldsymbol{\theta}}_b,f}\right\|_2 \tag{7}$$

Then, the NBRCB power and weight estimates are given by $\hat{\sigma}^2_{0,\bar{\boldsymbol{\theta}}_b,f} = \frac{1}{\hat{\mathbf{a}}^H_{0,\bar{\boldsymbol{\theta}}_b,f}\mathbf{R}_f^{-1}\hat{\mathbf{a}}_{0,\bar{\boldsymbol{\theta}}_b,f}}$ and $\mathbf{w}_{\bar{\boldsymbol{\theta}}_b,f,\text{NBRCB}} = \frac{\mathbf{R}_f^{-1}\hat{\hat{\mathbf{a}}}_{0,\bar{\boldsymbol{\theta}}_b,f}}{\hat{\mathbf{a}}^H_{0,\bar{\boldsymbol{\theta}}_b,f}\mathbf{R}_f^{-1}\hat{\hat{\mathbf{a}}}_{0,\bar{\boldsymbol{\theta}}_b,f}}$, where the weights can be used to filter the data in the frequency bin with center frequency f, $\mathbf{x}_{k,f}$.

B. Uncertainty Set Design

We now describe the approach outlined in [16], [17] for designing a spherical ASV uncertainty set for the bth beam and for the frequency bin with center frequency f, assuming azimuth and elevation angle steering, so that $\boldsymbol{\theta} = [\theta, \phi]^T$, where θ and ϕ denote azimuth and elevation. For the bth beam, we assume that the possible SOI angle-of-arrival (AOA) angles belong to a rectangle in azimuth/elevation space that is described by azimuth and elevation angles in the ranges $[\bar{\theta}_b - 0.5\Delta_{\text{Az}}, \bar{\theta}_b + 0.5\Delta_{\text{Az}}]$ and $[\bar{\phi}_b - 0.5\Delta_{\text{El}}, \bar{\phi}_b + 0.5\Delta_{\text{El}}]$. The uncertainty set is derived from $N_{\text{Az}}N_{\text{El}}$ equally spaced samples of the array response over this rectangle. It is assumed that N_{Az} and N_{El} are large enough to give a sufficiently dense sampling of the rectangle in azimuth angle and elevation angle, respectively. We find the smallest possible spherical uncertainty set [15]–[17] so that

$$\epsilon_{\bar{\boldsymbol{\theta}}_b,f} = \sup_{k,j} \left\| \mathbf{a}(f,\theta_k,\phi_j) \left(\frac{\mathbf{a}(f,\theta_k,\phi_j)^H\bar{\mathbf{a}}_{\bar{\boldsymbol{\theta}}_b,f}}{\bar{\mathbf{a}}^H_{\bar{\boldsymbol{\theta}}_b,f}\mathbf{a}(f,\theta_k,\phi_j)} \right)^{1/2} - \bar{\mathbf{a}}_{\bar{\boldsymbol{\theta}}_b,f} \right\|_2^2, \tag{8}$$

for $k = 1,\ldots,N_{\text{Az}}$ and $j = 1,\ldots,N_{\text{El}}$, where $\mathbf{a}(f,\theta_k,\phi_j)$ [c.f. (3)] is calculated from the delays of a plane-wave signal with frequency f impinging on the array from azimuth θ_k and elevation ϕ_j, with $\theta_k = \bar{\theta}_b + \left(\frac{k-1}{N_{\text{Az}}-1} - \frac{1}{2}\right)\Delta_{\text{Az}}$ and $\phi_j = \bar{\phi}_b + \left(\frac{j-1}{N_{\text{El}}-1} - \frac{1}{2}\right)\Delta_{\text{El}}$. The sphere center is given by $\bar{\mathbf{a}}_{\bar{\boldsymbol{\theta}}_b,f} = \mathbf{a}(f,\bar{\theta}_b,\bar{\phi}_b)$.

So far, these sets have been designed assuming only AOA errors. As shown in [16], the uncertainty sets need to be large enough to allow for the effects of covariance matrix estimation errors, which can be interpreted as additional ASV errors (see, e.g. [21]), and for calibration errors. To allow for these other errors, we impose a minimum value on the size of the uncertainty sets. As shown in [17], and as was also done in [16], we can use ASV estimates obtained using the RCB itself on strong sources of opportunity to determine the calibration errors and set a minimum value for allowing for these types of error. As was shown in [16], one can use simulations with the sample covariance matrices to determine the minimum sphere radius to cope with covariance matrix estimation errors. Note how one can systematically look at each of the errors and include this information in the uncertainty sets in a non ad hoc manner.

C. Robust Capon Beamformer (RCB) Optimization

At the core of the NBRCB (and WBRCB and RDRCB) is an RCB optimization (6) . Without loss of generality, we consider spherical uncertainty, which gives [11], [12]

$$\min_{\mathbf{a}} \mathbf{a}^H \mathbf{R}^{-1} \mathbf{a} \quad \text{s.t.} \quad \|\mathbf{a} - \bar{\mathbf{a}}\|_2^2 = \epsilon, \tag{9}$$

where, for notational simplicity, we have dropped the dependence on frequency and beam. The associated Lagrangian function is given by

$$L(\lambda, \mathbf{a}) = \mathbf{a}^H \mathbf{R}^{-1} \mathbf{a} + \lambda \left(\|\mathbf{a} - \bar{\mathbf{a}}\|_2^2 - \epsilon \right), \tag{10}$$

where λ denotes a real-valued Lagrange multiplier. As shown in [2], [11], [12], (10) can be solved exactly for $\mathbf{a}$ and λ via eigenvalue decomposition. We term the resulting $\mathcal{O}(M^3)$ algorithm the RCB eigenvalue decomposition (RCB-EVD) algorithm. The RCB weight vector is given by

$$\mathbf{w}_{\text{RCB}} = \frac{\mathbf{R}^{-1} \hat{\mathbf{a}}}{\hat{\mathbf{a}}^H \mathbf{R}^{-1} \hat{\mathbf{a}}}, \tag{11}$$

where $\hat{\mathbf{a}}$ is the solution to (9).

D. WC Robust Adaptive Beamformer (WC-RAB) Optimization

In [13], the worst-case robust adaptive beamformer (WC-RAB) problem, under spherical uncertainty, is formulated as

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{R} \mathbf{w} \text{ s.t. } \left| \mathbf{w}^H \mathbf{a} \right| \geq 1 \ \forall \ \mathbf{a} \| \mathbf{a} - \bar{\mathbf{a}} \|_2^2 \leq \epsilon \tag{12}$$

where the constraints ensure that the distortionless constraint is maintained for the worst-case steering vector contained in the set, i.e., for the steering vector $\mathbf{a}$ such that $\left| \mathbf{w}^H \mathbf{a} \right|$ has the smallest value. After some manipulation, (12), which contains an infinite number of non-linear and non-convex constraints, can be re-written equivalently as [13]

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{R} \mathbf{w} \text{ s.t. } \mathbf{w}^H \bar{\mathbf{a}} = 1 + \sqrt{\epsilon} \left\| \mathbf{w} \right\|_2 \tag{13}$$

providing that $\left| \mathbf{w}^H \bar{\mathbf{a}} \right| > \sqrt{\epsilon} \left\| \mathbf{w} \right\|_2$. The weights that solve (13), here termed the WC-RAB weights, have been shown to be equivalent to the RCB weights in (11) [12]. Thus, in the following sections, we examine efficient implementations of both the RCB and the WC-RAB.

III. Worst-Case Robust Adaptive Beamformer Using Iterative Gradient Minimization (WC-IG)

Here, we present a modified version of the iterative gradient (IG) minimization based approach proposed in [37], fixing an issue with the Lagrange multiplier evaluation in the original algorithm and also providing an ad hoc fix to the weight normalization problem. The Lagrange function for (13) can be written as

$$J(\mathbf{w}, \lambda) = \mathbf{w}^H \mathbf{R} \mathbf{w} - \lambda \Big(\mathbf{w}^H \bar{\mathbf{a}} - 1 - \sqrt{\epsilon} \left\| \mathbf{w} \right\|_2 \Big) \tag{14}$$

where λ denotes the Lagrange multiplier. Instead of solving (14) analytically for $\mathbf{w}$ and λ, which would require at least an $\mathcal{O}(M^3)$ complexity for eigenvalue decomposition [2], [14], [37], a numerical iterative gradient (IG) minimization scheme is used to update the weight vector as

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \mu_k \boldsymbol{\delta}_k \tag{15}$$

where, for the kth snapshot, μ_k and $\boldsymbol{\delta}_k$ denote the step-size parameter and the gradient vector of (14), respectively, with

$$\boldsymbol{\delta}_k = \mathbf{R}_k \mathbf{w}_k - \lambda \left(\bar{\mathbf{a}} - \sqrt{\epsilon} \frac{\mathbf{w}_k}{\left\| \mathbf{w}_k \right\|_2} \right) \tag{16}$$

Substituting (16) into (15) yields

$$\begin{aligned}\mathbf{w}_{k+1} &= \mathbf{w}_k - \mu_k \mathbf{R}_k \mathbf{w}_k + \mu_k \lambda \left(\bar{\mathbf{a}} - \sqrt{\epsilon} \frac{\mathbf{w}_k}{\|\mathbf{w}_k\|_2} \right) \\ &\triangleq \hat{\mathbf{w}}_{k+1} + \mu_k \lambda \mathbf{p}_k \end{aligned} \tag{17}$$

with

$$\hat{\mathbf{w}}_{k+1} = \mathbf{w}_k - \mu_k \mathbf{R}_k \mathbf{w}_k \tag{18}$$

$$\mathbf{p}_k = \bar{\mathbf{a}} - \sqrt{\epsilon} \frac{\mathbf{w}_k}{\|\mathbf{w}_k\|_2} \tag{19}$$

where $\hat{\mathbf{w}}_{k+1}$ solves the unconstrained minimization. The optimum step-size is given by [37]

$$\mu_{\text{opt},k} = \alpha \frac{\mathbf{w}_k^H \mathbf{R}_k^2 \mathbf{w}_k}{\mathbf{w}_k^H \mathbf{R}_k^3 \mathbf{w}_k} \tag{20}$$

where $0 < \alpha < 1$ is added to improve the numerical stability of the algorithm. If $\hat{\mathbf{w}}_{k+1}$ in (18) satisfies the constraint in (13), the weights are accepted and one sets $\mathbf{w}_{k+1} = \hat{\mathbf{w}}_{k+1}$; otherwise, (17) is substituted into the following inequality constraint to estimate the Lagrange multiplier yielding

$$\text{Re}\left\{ \left[\hat{\mathbf{w}}_{k+1} + \mu_k \lambda \mathbf{p}_k \right]^H \bar{\mathbf{a}} \right\} \geq \sqrt{\epsilon} \left\| \hat{\mathbf{w}}_{k+1} + \mu_k \lambda \mathbf{p}_k \right\|_2 + 1 \tag{21}$$

Rearranging (21), squaring both sides and noting that the inequality can be replaced by equality, yields

$$\left(\left[\text{Re}\left\{ \hat{\mathbf{w}}_{k+1}^H \bar{\mathbf{a}} \right\} - 1 \right] + \lambda \mu_k \text{Re}\left\{ \mathbf{p}_k^H \bar{\mathbf{a}} \right\} \right)^2 = \epsilon \left\| \hat{\mathbf{w}}_{k+1} + \lambda \mu_k \mathbf{p}_k \right\|_2^2$$

which can be expressed as

$$a\lambda^2 + b\lambda + c = 0, \tag{22}$$

where

$$a = \mu_k^2 \left[\left(\text{Re}\left\{ \mathbf{p}_k^H \bar{\mathbf{a}} \right\} \right)^2 - \epsilon \left\| \mathbf{p}_k \right\|_2^2 \right] \tag{23}$$

$$b = 2\mu_k \left[\mathcal{X} \text{Re}\left\{ \mathbf{p}_k^H \bar{\mathbf{a}} \right\} - \epsilon \text{Re}\left\{ \hat{\mathbf{w}}_{k+1}^H \mathbf{p}_k \right\} \right] \tag{24}$$

$$c = \mathcal{X}^2 - \epsilon \left\| \hat{\mathbf{w}}_{k+1} \right\|_2^2 \tag{25}$$

with $\mathcal{X} = \text{Re}\left\{ \hat{\mathbf{w}}_{k+1}^H \bar{\mathbf{a}} \right\} - 1$. Using the quadratic formula, the Lagrange multiplier is given by

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{26}$$

In the original derivation in [37], it was assumed that $c < 0$ on the basis that the Lagrange multiplier is only estimated if $\mathcal{X} < \sqrt{\epsilon} \left\| \hat{\mathbf{w}}_{k+1} \right\|_2$; however, this is only true if $\mathcal{X}$ is non-negative which it may not be. Thus, we include additional solutions for the cases of $c = 0$ and $c > 0$. Also, we disagree with the solutions given in [37] for the $c < 0$ case and therefore, give alternative solutions for this case. Further, we remark that λ is a real constant, and therefore, $b^2 - 4ac \geq 0$. We treat the cases of $b^2 - 4ac > 0$ and $b^2 - 4ac = 0$ separately. Table I summarizes the signs of the roots under the various different conditions on a, b and c. When $b^2 > 4ac$, there are two distinct roots. If there is a negative and a positive root, the positive root is selected to ensure positive definiteness of the loaded covariance matrix. When there are two positive roots, the smallest root is selected to guarantee algorithm stability. Since λ is a Lagrange multiplier, it should satisfy $\lambda \geq 0$, however, we do sometimes find that we get two negative roots due to numerical issues. In such cases, we select the largest (or least negative) root and either reduce the uncertainty sphere radius or α in (20) until we stop getting two negative roots; we stress that we do not force the negative roots to be zero. When $c = 0$, there will always be one zero root. If there is zero root

Algorithm 1 The WC-IG algorithm

1: Update the sample covariance matrix $\hat{\mathbf{R}}$.
2: Compute $\mu_{\text{opt},k}$ using (20).
3: Update the unconstrained MV weight vector (18).
4: **if** $\mathcal{X} < \sqrt{\epsilon}\,\|\hat{\mathbf{w}}_{k+1}\|_2$ **then**
5: Compute λ using (26) and update weights as $\mathbf{w}_{k+1} = \hat{\mathbf{w}}_{k+1} + \mu_k \lambda \mathbf{p}_k$
6: **else**
7: Set $\mathbf{w}_{k+1} = \hat{\mathbf{w}}_{k+1}$.
8: **end if**
9: Rescale $\mathbf{w}_{k+1}$ using (29) prior to filtering the data.

and a negative root, we select the zero root. If there is a zero root and a positive root, we select the positive root. We remark that when $b^2 = 4ac$, the two roots are identical. The solution λ is inserted into (17) to form $\mathbf{w}_{k+1}$.

An important consideration in passive sonar is appropriate normalization of the weights, not only for accurate power estimation, which is well understood, but also for, e.g., time-series analysis of beam data when several weight updates have been applied over the duration of the time-series to be analyzed. If weight normalization is not consistent, so that the response of the weights to the signal ASV varies significantly at each weight update, the segments of the time series will have varying magnitudes or power due to this, which would then affect, e.g., spectral analysis. In the RCB-based algorithms, this issue can be addressed as one can access and re-scale the estimated ASV in a consistent manner such as in (7). This was not considered in the original WC-based algorithms, however, in [12], it was suggested that given the optimal worst-case weights $\mathbf{w}_0$, the scaling ambiguity could be resolved by obtaining an estimate of the ASV using

$$\mathbf{a}_0 = \frac{\mathbf{R}\mathbf{w}_0}{\mathbf{w}_0^H \mathbf{R} \mathbf{w}_0} \tag{27}$$

Replacing the above quantities with their estimates, the worst-case weights could then be scaled as

$$\hat{\hat{\mathbf{w}}}_0 = \hat{\mathbf{w}}_0 \|\hat{\mathbf{a}}_0\|_2 / \sqrt{M}. \tag{28}$$

Later, in Section IX, we show why this in not necessarily a good fix in practice. Here, we propose an alternative ad hoc scaling of the weights. Note that $\bar{\mathbf{a}}$ is contained within the uncertainty ellipsoid and so $\mathbf{w}^H \bar{\mathbf{a}} \geq 1$, that is, the worst-case beamformer will not steer a null towards $\bar{\mathbf{a}}$ and therefore, we expect that the beampattern will not vary drastically between the true ASV and the assumed ASV. Thus, we constrain the weights so that the response to $\bar{\mathbf{a}}$ is unity, forming

$$\hat{\hat{\mathbf{w}}} = \hat{\mathbf{w}} \frac{e^{i\angle \hat{\mathbf{w}}^H \bar{\mathbf{a}}}}{|\hat{\mathbf{w}}^H \bar{\mathbf{a}}|}. \tag{29}$$

After initializing with $\mathbf{R}_0 = \mathbf{I}$, $\mathbf{w}_0 = \bar{\mathbf{a}}$, and $\alpha = 1$, the WC-IG algorithm iterates the steps given in Algorithm 1. The first step requires $2M^2$ operations, the second $2M^2 + 2M$, and the third M. Finally, the remaining steps require $7M$ operations. We term the resulting algorithm the gradient minimization based WC-RAB (WC-IG), which has $\mathcal{O}(4M^2 + 10M)$ complexity.

IV. Worst-Case Robust Adaptive Beamformer Using Kalman Filtering (WC-KF)

The WC-RAB problem (13) can be written as

$$\min_{\mathbf{w}} \text{MSE} \text{ s.t. } h_2(\mathbf{w}) = 1 \tag{30}$$

TABLE I
SIGNS OF ROOTS FOR DIFFERENT CONDITIONS ON a, b AND c.

	$a>0, b>0$	$a<0, b>0$	$a>0, b<0$	$a<0, b<0$
$c<0,\ b^2>4ac$	$+-$	$++$	$+-$	$--$
$c>0,\ b^2>4ac$	$--$	$+-$	$++$	$+-$
$c=0,\ b^2>4ac$	$0-$	$+0$	$+0$	$0-$
$b^2=4ac$	$--$	$++$	$++$	$--$

where $\text{MSE} = E\left\{\left|0-\mathbf{w}^H\mathbf{x}\right|^2\right\} = \mathbf{w}^H\mathbf{R}\mathbf{w}$ is the mean square error (MSE) between a desired signal of 0 and the beamformer output, and where

$$h_2(\mathbf{w}) \stackrel{\triangle}{=} \epsilon\left\|\mathbf{w}\right\|_2^2 - \mathbf{w}^H\bar{\mathbf{a}}\bar{\mathbf{a}}^H\mathbf{w} + \mathbf{w}^H\bar{\mathbf{a}} + \bar{\mathbf{a}}^H\mathbf{w} = 1 \tag{31}$$

is equivalent to the constraint $\left|1-\mathbf{w}^H\bar{\mathbf{a}}\right|^2 = \left|-\sqrt{\epsilon}\left\|\mathbf{w}\right\|_2\right|^2$. Since the Kalman filter is a minimum MSE (MMSE) filter, it can be used to solve (30) as shown in [36]. We proceed to summarize the algorithm.

The system equation is given by the weight update equation $\mathbf{w}_{k+1} = \gamma\mathbf{w}_k + \mathbf{v}_{s,k}$, where γ is a fixed parameter of the model and $\mathbf{v}_{s,k}$ is the process noise, which is assumed to be a zero mean white Gaussian process with covariance $\mathbf{R}_{\mathbf{v}_s} = E\left\{\mathbf{v}_{s,k}\mathbf{v}_{s,k}^H\right\} = \sigma_{s,\text{KF}}^2\mathbf{I}$. The measurement equation is given by $\mathbf{z} = \mathbf{h}\left(\mathbf{w}_k\right) + \mathbf{v}_{m,k}$, with $\mathbf{z} = \left[\begin{array}{cc} 0 & 1 \end{array}\right]^T$, $\mathbf{h}\left(\mathbf{w}_k\right) = \left[\begin{array}{cc} \mathbf{w}_k^H\mathbf{x}_k & h_2^*(\mathbf{w}_k) \end{array}\right]^H$ and $\mathbf{v}_{m,k} = \left[\begin{array}{cc} v_{1,k} & v_{2,k} \end{array}\right]^T$, where $v_{1,k}$ and $v_{2,k}$ are the residual and constraint errors, respectively. These errors are modeled as zero-mean, independent white noise sequences with covariance matrix

$$\mathbf{R}_{\mathbf{v}_m} = E\left\{\mathbf{v}_{m,k}\mathbf{v}_{m,k}^H\right\} = \text{diag}\left\{\left[\begin{array}{cc} \sigma_{1,\text{KF}}^2 & \sigma_{2,\text{KF}}^2 \end{array}\right]\right\} \tag{32}$$

Due to the non-linearity of the measurement equation, the second-order extended Kalman filter is used to obtain the following recursion for the weight vector estimate [36]

$$\hat{\mathbf{w}}_k = \hat{\mathbf{w}}_{k-1} + \mathbf{G}_k\left[\mathbf{z} - \hat{\mathbf{z}}_{k|k-1}\right] \tag{33}$$

where $\mathbf{G}_k$ and $\hat{\mathbf{z}}_{k|k-1}$ denote the filter gain and the predicted measurement.

The user must select the following parameters: γ, $\sigma_{s,\text{KF}}^2$, $\mathbf{H}_{\mathbf{ww}}^{(2)} = \epsilon\mathbf{I} - \bar{\mathbf{a}}\bar{\mathbf{a}}^H$, and $\sigma_{1,\text{KF}}^2$ and $\sigma_{2,\text{KF}}^2$. For a stationary environment, one should set $\gamma = 1$ and $\sigma_{s,\text{KF}}^2 = 0$, whereas for a non-stationary environment, increasing these values will enable better tracking of the environment and typical values are γ set slightly larger than 1 and $\sigma_{s,\text{KF}}^2 = 10^{-4}$. The parameter $\sigma_{1,\text{KF}}^2$ should be selected to be of the same order as the optimal array output power, which is a quantity often estimated by the beamformer. It is suggested in [36] that one may approximate the power as $\left\|\mathbf{w}\right\|_2^2\left(M\sigma^2 + \sigma_n^2\right)$, where σ^2 and σ_n^2 denote the desired signal and sensor noise powers, respectively. However, we note that these quantities are not available *a priori* so, instead, we here use the DAS estimate. It is noted in [36] that the beamformer is not sensitive to the choice of $\sigma_{1,\text{KF}}^2$ because the norm of the weights is chosen by the filter so that the output power of the filter matches $\sigma_{1,\text{KF}}^2$. The value of $\sigma_{2,\text{KF}}^2$ should be chosen very small to ensure that the robustness constraint is satisfied with high accuracy. The uncertainty sphere radius $\sqrt{\epsilon}$ is chosen according to the uncertainty in the ASV.

Given the selected parameters, the algorithm iterates the steps given in Algorithm 2. Initial values $\hat{\mathbf{w}}_0$, $\hat{\mathbf{z}}_{1|0}$ and $\mathbf{P}_{1|0}$ are required. $\hat{\mathbf{w}}_0$ is set to a random vector. To obtain an initial $\mathbf{P}_{1|0}$, first form $h_2(\gamma\hat{\mathbf{w}}_0)$ using (31) and

$$\mathbf{H}_{\mathbf{w}}(1, \gamma\hat{\mathbf{w}}_0) = \left[\begin{array}{c} \mathbf{x}_1^H \\ \epsilon\gamma\hat{\mathbf{w}}_0^H - \gamma\left[\bar{\mathbf{a}}\bar{\mathbf{a}}^H\hat{\mathbf{w}}_0\right]^H + \bar{\mathbf{a}}^H \end{array}\right] \tag{34}$$

Then, set $\mathbf{P}_{1|0} = \beta\mathbf{I}$, where

$$\beta = \frac{1}{\epsilon_{v,1}}\left[\mathbf{z} - \hat{\mathbf{z}}_{1|0}\right]^H\left[\mathbf{H}_{\mathbf{w}}(1, \gamma\hat{\mathbf{w}}_0)\mathbf{H}_{\mathbf{w}}^H(1, \gamma\hat{\mathbf{w}}_0)\right]^{-1}\left[\mathbf{z} - \hat{\mathbf{z}}_{1|0}\right] \tag{35}$$

Algorithm 2 Kalman Filtering Implementation of WC-RAB

Initiate $\hat{\mathbf{z}}_{1|0}$, $\hat{\mathbf{w}}_0$ and $\mathbf{P}_{1|0}$ as indicated.
for k=1,... **do**
  $\mathbf{P}_{k|k-1} = \gamma^2 \mathbf{P}_{k-1|k-1} + \mathbf{R}_{\mathbf{v}_s}$
  $\mathbf{H}_{\mathbf{w}}(k, \gamma\hat{\mathbf{w}}_{k-1}) = \begin{bmatrix} \mathbf{x}_k^H \\ \epsilon\gamma\hat{\mathbf{w}}_{k-1}^H - \left[\bar{\mathbf{a}}\bar{\mathbf{a}}^H\gamma\hat{\mathbf{w}}_{k-1}\right]^H + \bar{\mathbf{a}}^H \end{bmatrix}$
  $\mathbf{S}_k = \mathbf{H}_{\mathbf{w}}(k, \gamma\hat{\mathbf{w}}_{k-1})\mathbf{P}_{k|k-1}\mathbf{H}_{\mathbf{w}}^H(k, \gamma\hat{\mathbf{w}}_{k-1}) + \frac{1}{2}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\left[\epsilon^2 \text{tr}\left\{\mathbf{P}_{k-1|k-1}^2\right\} - 2\epsilon\left\|\mathbf{P}_{k|k-1}\bar{\mathbf{a}}\right\|_2^2 + |\bar{\mathbf{a}}^H\mathbf{P}_{k|k-1}\bar{\mathbf{a}}|^2\right] + \mathbf{R}_{\mathbf{v}_m}$
  $\mathbf{G}_k = \mathbf{P}_{k|k-1}\mathbf{H}_{\mathbf{w}}^H(k, \gamma\hat{\mathbf{w}}_{k-1})\mathbf{S}_k^{-1}$
  $\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{G}_k\mathbf{S}_k\mathbf{G}_k^H$
  $h_2(\gamma\hat{\mathbf{w}}_{k-1}) = \epsilon\gamma^2\hat{\mathbf{w}}_{k-1}^H\hat{\mathbf{w}}_{k-1} - \gamma^2\hat{\mathbf{w}}_{k-1}^H\bar{\mathbf{a}}\bar{\mathbf{a}}^H\hat{\mathbf{w}}_{k-1} + \gamma\hat{\mathbf{w}}_{k-1}^H\bar{\mathbf{a}} + \bar{\mathbf{a}}^H\hat{\mathbf{w}}_{k-1}$
  $\hat{\mathbf{z}}_{k|k-1} = \begin{bmatrix} \gamma\mathbf{x}_k^H\hat{\mathbf{w}}_{k-1} \\ h_2(\gamma\hat{\mathbf{w}}_{k-1}) + \frac{1}{2}\text{tr}\left\{\mathbf{H}_{\mathbf{ww}}^{(2)}\mathbf{P}_{k|k-1}\right\} \end{bmatrix}$
  $\hat{\mathbf{w}}_k = \hat{\mathbf{w}}_{k-1} + \mathbf{G}_k\left[\mathbf{z} - \hat{\mathbf{z}}_{k|k-1}\right]$
  $k \leftarrow k+1$
end for

where $\epsilon_{v,1} = 3$ was used as suggested in [36], and where

$$\hat{\mathbf{z}}_{1|0} = \begin{bmatrix} \gamma\mathbf{x}_1^H\hat{\mathbf{w}}_0 \\ h_2(\gamma\hat{\mathbf{w}}_0) + \frac{1}{2}\text{tr}\left\{\mathbf{H}_{\mathbf{ww}}^{(2)}\right\} \end{bmatrix} \tag{36}$$

Earlier, it was noted that the norm of the weights is chosen so that the output power matches $\sigma_{1,\text{KF}}^2$. Since we have chosen $\sigma_{1,\text{KF}}^2$ to be equal to the DAS estimate, it means that the spatial spectrum would be similar to the DAS spectrum. We apply the ad hoc rescaling of the weights (29) noting that this will not affect the output SNR. In the following, we term this implementation the Kalman Filtering Implementation of WC-RAB (WC-KF). Ignoring $\mathcal{O}(M)$ terms, the complexity of updating the weights is approximately $\mathcal{O}(12M^2)$.

V. The Krylov-RDRCB

In [29], a family of Krylov-subspace reduced-dimension robust Capon beamformers (Krylov-RDRCBs) was derived, which project the data onto a reduced-rank Krylov subspace and solve an RCB problem in the reduced-dimension space. Here, we describe the Conjugate Gradient RDRCB, which uses the conjugate algorithm [42], [43] to expand a rank-N Krylov-subspace basis $\mathbf{D}_N = \begin{bmatrix} \mathbf{d}_1 & \dots & \mathbf{d}_N \end{bmatrix}$ by initializing with $\mathbf{d}_1 = \bar{\mathbf{a}}$, $\mathbf{r}_1 = -\bar{\mathbf{a}}$ and then for $i = 1, \dots, N-1$ calculates $\alpha_i = -\frac{\mathbf{d}_i^H\mathbf{r}_i}{\mathbf{d}_i^H\mathbf{R}\mathbf{d}_i}$, $\mathbf{r}_{i+1} = \mathbf{r}_i + \alpha_i\mathbf{R}\mathbf{d}_i$, $\beta_i = \frac{\mathbf{d}_i^H\mathbf{R}\mathbf{r}_{i+1}}{\mathbf{d}_i^H\mathbf{R}\mathbf{d}_i}$ and $\mathbf{d}_{i+1} = -\mathbf{r}_{i+1} + \beta_i\mathbf{d}_i$. The framework in [29] (see, also, [18]) shows how to produce a reduced-dimension ellipsoid, from the full-dimension (or element-space) ellipsoid and the Krylov basis $\mathbf{D}_N$, which can be used in the reduced-rank subspace. Here, without loss of generality, we only consider spherical full-dimension uncertainty, which leads to a reduced-dimension ellipsoid $\mathcal{E}_N(\bar{\mathbf{b}}, \mathbf{F})$, with $\bar{\mathbf{b}} = \mathbf{D}_N^H\bar{\mathbf{a}}$ and $\mathbf{F} = \epsilon^{-1}(\mathbf{D}_N^H\mathbf{D}_N)^{-1}$, where ϵ and $\bar{\mathbf{a}}$ denote the radius squared and center of the full-dimension spherical uncertainty set. This leads to the following rank-N reduced-dimension RCB problem

$$\min_{\mathbf{b}} \mathbf{b}^H\mathbf{\Lambda}_{\text{CG}}^{-1}\mathbf{b} \text{ s.t. } \left[\mathbf{b} - \bar{\mathbf{b}}\right]^H\mathbf{F}\left[\mathbf{b} - \bar{\mathbf{b}}\right] \leq 1, \tag{37}$$

where $\mathbf{\Lambda}_{\text{CG}} = \mathbf{D}_N^H\mathbf{R}\mathbf{D}_N$ is a diagonal reduced-dimension covariance matrix. We refer the reader to [29] on details on how to solve (37) via the Lagrange multiplier method. Denoting the solution to (37) as $\hat{\mathbf{b}}_0$,

Algorithm 3 The RCB-VDL-SD algorithm

Update inverse covariance matrix using
$\mathbf{q}_k = \eta^{-1}\mathbf{R}_{k-1}^{-1}\mathbf{x}_k\left(1+\eta^{-1}\mathbf{x}_k^H\mathbf{R}_{k-1}^{-1}\mathbf{x}_k\right)^{-1}$
$\mathbf{R}_k^{-1} = \eta^{-1}\mathbf{R}_{k-1}^{-1} - \eta^{-1}\mathbf{q}_k\mathbf{x}_k^H\mathbf{R}_{k-1}^{-1}$
Update the gradient vector using
$\bar{\mathbf{g}}_k = \mathbf{R}_k^{-1}\hat{\mathbf{a}}_{k-1}$
$\mathbf{d}_{k-1} = \hat{\mathbf{a}}_{k-1} - \bar{\mathbf{a}}$
$\mathbf{g}_k = \bar{\mathbf{g}}_k + \lambda_{k-1}\mathbf{d}_{k-1}$
Calculate step-size parameter using
$\mu_{\text{SD},k} = \alpha\mathbf{g}_k^H\mathbf{g}_k\left(\text{Re}\left\{\mathbf{g}_k^H\mathbf{R}_k^{-1}\mathbf{g}_k\right\}+\delta\right)^{-1}$
Calculate unconstrained ASV using
$\tilde{\mathbf{a}}_k = \hat{\mathbf{a}}_{k-1} - \mu_{\text{SD},k}\bar{\mathbf{g}}_k$
if $\|\tilde{\mathbf{a}}_k - \bar{\mathbf{a}}\|_2^2 \leq \epsilon$ **then**
  $\hat{\mathbf{a}}_k = \tilde{\mathbf{a}}_k$ and $\lambda_k = 0$
else
  Calculate optimal loading and constrained ASV estimate using
  $a_1 = \mu_{\text{SD},k}^2\|\hat{\mathbf{a}}_{k-1} - \bar{\mathbf{a}}\|_2^2$
  $b_1 = \mu_{\text{SD},k}\text{Re}\left\{[\tilde{\mathbf{a}}_k - \bar{\mathbf{a}}]^H\mathbf{d}_{k-1}\right\}$
  $c_1 = \|\tilde{\mathbf{a}}_k - \bar{\mathbf{a}}\|_2^2 - \epsilon$
  $\lambda_k = \frac{-b_1 - \text{Re}\left\{\sqrt{b_1^2 - a_1c_1}\right\}}{a_1}$
  $\hat{\mathbf{a}}_k = \tilde{\mathbf{a}}_k - \mu_{\text{SD},k}\lambda_k\mathbf{d}_{k-1}$
end if
Form weights as $\mathbf{w}_k = \frac{\mathbf{R}_k^{-1}\hat{\mathbf{a}}_k}{\hat{\mathbf{a}}_k^H\mathbf{R}_k^{-1}\hat{\mathbf{a}}_k}$

the weight vector is formed as $\mathbf{w} = \frac{\mathbf{\Lambda}_{\text{CG}}^{-1}\hat{\mathbf{b}}_0}{\hat{\mathbf{b}}_0^H\mathbf{\Lambda}_{\text{CG}}^{-1}\hat{\mathbf{b}}_0}\sqrt{\frac{\hat{\mathbf{b}}_0^H(\mathbf{D}_N^H\mathbf{D}_N)^{-1}\hat{\mathbf{b}}_0}{M}}$. The operation counts for the Conjugate-Gradient RDRCB under spherical uncertainty are given in Table V in [29] and are summarized here in Table II. In the following, we will refer to the algorithm simply as the Krylov-RDRCB.

VI. The RCB-VDL-SD Algorithm

We proceed to examine the steepest-descent based RCB with variable diagonal loading (RCB-VDL-SD) introduced in [38], where the ASV is updated recursively using

$$\hat{\mathbf{a}}_k = \hat{\mathbf{a}}_{k-1} - \mu_{\text{SD},k}\mathbf{g}_k \tag{38}$$

where $\mathbf{g}_k$ is the gradient, obtained by differentiating (10) w.r.t $\mathbf{a}^H$,

$$\mathbf{g}_k = \mathbf{R}_k^{-1}\hat{\mathbf{a}}_{k-1} + \lambda(\hat{\mathbf{a}}_{k-1} - \bar{\mathbf{a}}) \tag{39}$$

The optimal step size is given by

$$\mu_{\text{SD},k} = \frac{\alpha_{\text{VDL}}\|\mathbf{g}_k\|_2^2}{\mathbf{g}_k^H\mathbf{R}_k^{-1}\mathbf{g}_k + \delta} \tag{40}$$

Initializing with $\mathbf{R}_0^{-1} = \mathbf{I}$, $\hat{\mathbf{a}}_0 = \bar{\mathbf{a}}$, $\eta = 1$, $\lambda_0 = 0$, $\mathbf{g}_0 = \bar{\mathbf{a}}$, and $\alpha_{\text{VDL}} = 0.01$, the RCB-VDL-SD algorithm iterates the steps given in Algorithm 3. Updating the inverse covariance matrix, the gradient vector, the step-size parameter, and the unconstrained ASV estimate requires $\mathcal{O}(2M^2+2M)$, $\mathcal{O}(M^2+2M)$, $\mathcal{O}(M^2+2M)$, and $\mathcal{O}(M)$ operations, respectively. Testing whether the unconstrained estimate if feasible requires around $2M$ operations and then, if needed, calculating the constrained ASV estimate requires a further $O(4M)$ operations. Forming the weights requires a further $\mathcal{O}(M^2+M)$ operations. Therefore, the total cost of updating the weights is roughly $\mathcal{O}(5M^2+14M)$.

VII. The RCB-SP-SD Algorithm

In [39], the RCB optimization problem (9) is solved by using a steepest descent algorithm to recursively update the ASV using

$$\tilde{\mathbf{a}}_k = \mathbf{a}_{k-1} - \mu_{\text{SD},k}\mathbf{g}_k \tag{41}$$

where $\mathbf{a}_{k-1}$ is the steering vector estimate from the previous iteration, $\mu_{\text{SD},k} = \frac{1}{\text{tr}\{\mathbf{R}_k^{-1}\}}$ and $\mathbf{g}_k = \mathbf{R}_k^{-1}\mathbf{a}_{k-1}$ is the unconstrained minimizer of (9). In the same way as was done in the RCB-VDL-SD algorithm, the inverse covariance matrix estimate at the kth iteration $\mathbf{R}_k^{-1}$ is updated using the matrix inversion lemma as new snapshots are added to the sample covariance matrix at each new time step. To satisfy the constraint in (9), $\tilde{\mathbf{a}}_k$ is projected onto the uncertainty set constraint boundary, yielding

$$\mathbf{a}_k = \frac{\sqrt{\epsilon}\,(\tilde{\mathbf{a}}_k - \bar{\mathbf{a}})}{\|\tilde{\mathbf{a}}_k - \bar{\mathbf{a}}\|_2} + \bar{\mathbf{a}} \tag{42}$$

The weights are then formed as

$$\mathbf{w}_k = \frac{\mathbf{R}_k^{-1}\mathbf{a}_k}{\mathbf{a}_k^H\mathbf{R}_k^{-1}\mathbf{a}_k} \tag{43}$$

We remark that one can scale $\mathbf{a}_k$ to have norm $\sqrt{M}$ and use this in the equation for the weights. Updating $\mathbf{g}_k$, $\mu_{\text{SD},k}$ and $\tilde{\mathbf{a}}_k$ requires $\mathcal{O}(M^2)$, $\mathcal{O}(M)$ and $\mathcal{O}(M)$ operations, respectively. Then, updating $\mathbf{a}_k$ and the weights costs $\mathcal{O}(3M)$ and $\mathcal{O}(M^2 + M)$ operations, respectively. Thus, updating the weights at each time step requires a total of $\mathcal{O}(4M^2 + 8M)$ operations, assuming $O(2M^2 + 2M)$ operations are needed to update the inverse covariance matrix.

VIII. Comparison of Computational Complexity

Table II summarizes the approximate number of floating point operations required to update the weights, given that either the covariance matrix, in RCB-EVD, Krylov-RDRCB and WC-IG, or the inverse covariance matrix, in RCB-VDL-SD and RCB-SP-SD, has already been updated. For the RCB-EVD algorithm, we use the online operation counts for spherical uncertainty from Table I in [29], where n_{iter} is the number of iterations used in the Newton search to find the Lagrange multiplier. For the Krylov-RDRCB, we use the operation counts for the Conjugate-Gradient RDRCB under spherical uncertainty from Table V in [29], where N is the rank of the Krylov-subspace and $\tilde{n}_{\text{iter}}$ is the number of Newton iteration used to find the Lagrange multiplier. With $N = 5$, $n_{\text{iter}} = \tilde{n}_{\text{iter}} = 15$ and $M = 40$, we find that the Krylov-RDRCB, WC-IG, WC-KF, RCB-VDL-SD and RCB-SP-SD are around 20, 80, 15, 50 and 80 times less complex than RCB-EVD, respectively.

The complexity of the RCB-EVD algorithm is dominated by the $\mathcal{O}(4\frac{1}{3}M^3)$ complexity eigenvalue decomposition, which is neither easily parallelizable nor efficient on parallel hardware such as graphics processing units (GPUs). Consequently, as reported in [40], our GPU implementation of RCB-EVD runs tens of times slower than our CPU version. However, because the WC-IG algorithm is made up much simpler operations that can be parallelized on GPUs, as are the other efficient implementations we have considered, it runs more than 20 CHECK WITH ROGER times faster on GPUs than on a CPU, providing yet another advantage for implementation.

IX. Simulated Data Examples

Initially, we evaluate the above discussed algorithms on a simulated half-wavelength spaced uniform linear array with $M = 64$ elements, recreating the simulated scenario described in [18]. The data were simulated using array covariance $\mathbf{R} = \sigma_0^2\mathbf{a}_0\mathbf{a}_0^H + \mathbf{Q}$, where σ_0^2 and $\mathbf{a}_0$ denote the true SOI power and the true SOI ASV, respectively, and where $\mathbf{Q} = \sum_{i=1}^{d}\sigma_i^2\mathbf{a}_i\mathbf{a}_i^H + \sigma_s^2\mathbf{I} + \sigma_{\text{iso}}^2\mathbf{Q}_{\text{iso}}$. The simulated noise plus interference covariance, $\mathbf{Q}$, consists of terms due to d zero-mean uncorrelated interfering sources, where, for the ith interferer, σ_i^2 and $\mathbf{a}_i$ denote the source power and the ASV, respectively, as well as a term

TABLE II
COMPLEXITY REQUIRED TO UPDATE WEIGHTS, GIVEN $\hat{\mathbf{R}}$ OR $\hat{\mathbf{R}}^{-1}$ HAVE BEEN UPDATED ALREADY.

Algorithm	Complexity
RCB-EVD	$\mathcal{O}(4\frac{1}{3}M^3 + 3M^2 + [6n_{\text{iter}} + 3]M)$
WC-IG	$\mathcal{O}(2M^2 + 10M)$
WC-KF	$\mathcal{O}(10M^2 + 19M)$
Krylov-RDRCB	$\mathcal{O}([N+1]M^2 + [N^2 + 6N + 1]M + [4\frac{1}{3}N^3 + 4N^2 + (6\tilde{n}_{\text{iter}} + 9)N])$
RCB-VDL-SD	$\mathcal{O}(3M^2 + 12M)$
RCB-SP-SD	$2M^2 + 6M$

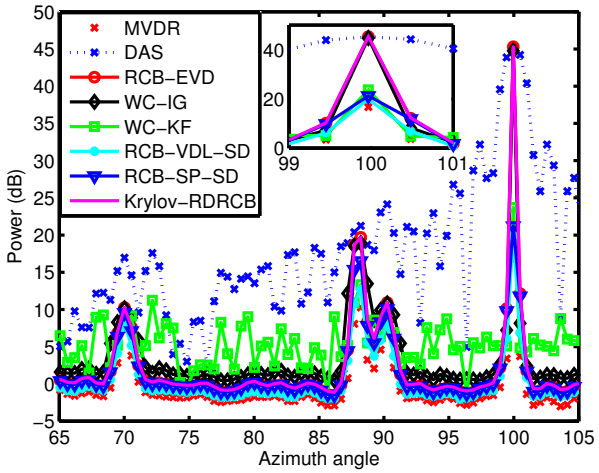

Fig. 1. Spatial spectra for simulated data, assuming $K = 128$.

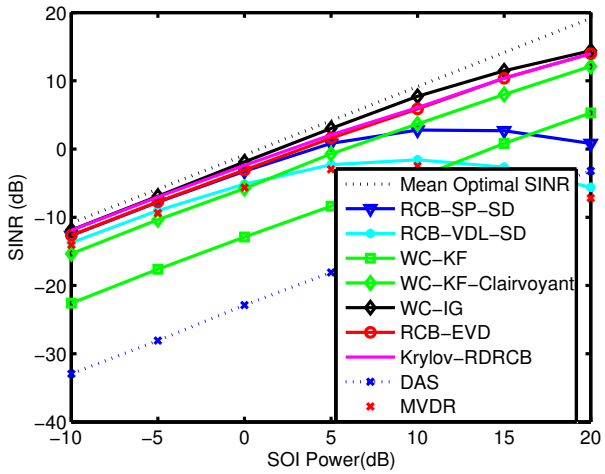

Fig. 2. SINR versus SOI power for $K = 128$.

modeling the sensor noise $\sigma_s^2\mathbf{I}$, with sensor noise power σ_s^2, and a term modeling an isotropic ambient noise, $\sigma_{\text{iso}}^2\mathbf{Q}_{\text{iso}}$, with power σ_{iso}^2. The isotropic noise covariance is given by $[\mathbf{Q}_{\text{iso}}]_{m,n} = \text{sinc}[\pi\lambda(m-n)]$. In the following, unless otherwise stated, $d = 3$, $\sigma_s^2 = 0$ dB, $\sigma_{\text{iso}}^2 = 1$ dB, $\sigma_0^2 = 10$ dB, $\sigma_1^2 = 10$ dB, $\sigma_2^2 = 20$ dB, and $\sigma_3^2 = 45$ dB. The AOAs of the discrete interferers are $\theta_1 = 70°$, $\theta_2 = 88°$, and $\theta_3 = 100°$, where the angle of arrival is measured from along the axis of the line array, i.e., from end-fire, so that $\theta \in [0, 180°]$. The azimuth space is sampled using $3M = 192$ equally cosine-spaced beams, where uncertainty sets are designed for each beam as described in Section II-B. The SOI is assumed to belong to the beam whose center is at $\bar{\theta}_0 = 90.25°$, with uncertainty sphere radius squared $\epsilon = 4.2029$. To allow for the typical case

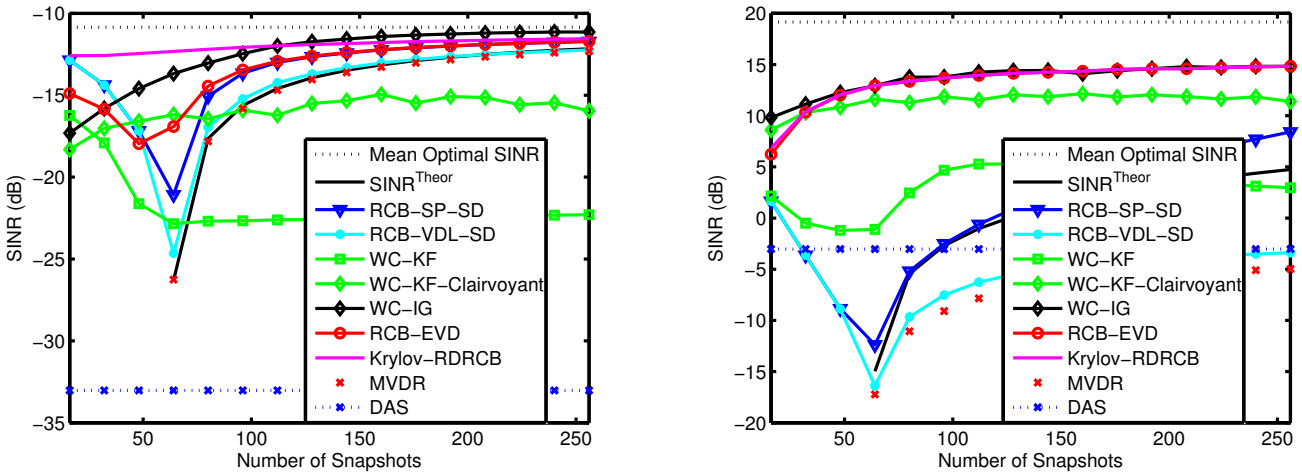

Fig. 3. SINR vs K, for a SOI with AOA errors and arbitrary ASV errors, when (a) $\sigma_0^2 = -10$ dB, (b) $\sigma_0^2 = 20$ dB.

that AOA errors exist, the SOI is simulated anywhere in the interval $[\theta_l, \theta_u]$, where θ_l (θ_u) is the angle midway between the center of the SOI beam and the center of the adjacent beam with lower (higher) angle. Furthermore[3], we assume that the SOI and interference ASVs are subject to independent arbitrary errors and, at each Monte-Carlo simulation, add to each ASV an arbitrary error vector $\mathbf{e} = \tilde{\mathbf{e}}/\|\tilde{\mathbf{e}}\|_2$, where each element of $\tilde{\mathbf{e}}$ is drawn from a zero-mean circularly symmetric distribution with unit variance. In the following, we examine the beamformer signal-to-interference-plus-noise ratio (SINR), defined as $\text{SINR} = \frac{\sigma_0^2 |\mathbf{w}^H \mathbf{a}_0|^2}{\mathbf{w}^H \mathbf{Q} \mathbf{w}}$. It is well known that the optimal SINR is given by $\text{SINR}_{\text{opt}} = \sigma_0^2 \mathbf{a}_0^H \mathbf{Q}^{-1} \mathbf{a}_0$. Since we simulate mismatch in the source ASVs, we compare the beamformer SINRs to the mean optimal SINR, obtained by averaging SINR_{opt} over the Monte-Carlo simulations. Further, we will also sometimes compare to $\text{SINR}^{\text{Theor}}$, which is obtained from the analytical expression for the expected SINR as a function of K given the true ASV (without mismatch), derived in [21]. Unless otherwise stated, we assume that $K = 2M = 128$ snapshots are available for covariance matrix estimation. In WC-KF, we set $\gamma = 1$, $\sigma_{s,\text{KF}}^2 = 0$, $\sigma_{1,\text{KF}}^2 = \hat{\sigma}_{\text{DAS}}^2 = M^{-2} \bar{\mathbf{a}}^H \mathbf{R} \bar{\mathbf{a}}$ and $\sigma_{2,\text{KF}}^2 = 10^{-12}$. We set $\alpha = 1$ in WC-RAB and $\alpha_{\text{VDL}} = 0.01$ in RCB-VDL-SD. In RCB-VDL-SD and RCB-SP-SD, we set $\eta = 1$ and, when estimating the spatial spectrum, we re-scale the estimated SOI ASV to have norm $\sqrt{M}$ before calculating the SOI power. We have used a fixed rank of $N = 5$ in the Krylov-RDRCB.

Fig. 1 illustrates the spatial spectra, and, as can be seen, particularly in the zoomed in panel, the standard RCB-EVD, Krylov-RDRCB, and WC-IG, which exploits the ad hoc weights normalization (29), provide the best power estimates, whereas the MVDR, WC-KF, RCB-SP-SD, and RCB-VDL-SD all exhibit severe SOI cancellation. We remark that without the additional weights normalization step (29), the Kalman-filter based WC-KF provides spatial power estimates similar to the DAS, as we used the DAS power estimate for $\sigma_{1,\text{KF}}^2$ in (32).

Fig. 2 shows the SINR versus the SOI power, clearly showing that WC-IG performs the best out of the efficient techniques and even better than the more complex standard RCB-EVD. The performance of the Krylov-RDRCB is indistinguishable from RCB-EVD. Figs. 3(a) and (b) illustrate the SINR as a function of the number of snapshots K for a weak SOI (low SNR) and a strong SOI (high SNR), clearly indicating that WC-IG and Krylov-RDRCB are also here showing preferable performance as compared to the other algorithms. Fig. 3(a) illustrates the faster convergence properties of the Krylov-RDRCB compared with the RCB-EVD. In Fig. 2 and Figs. 3(a) and (b), we have also added an additional curve for the WC-KF, which instead of using the DAS estimate for $\sigma_{1,\text{KF}}^2$, uses $\sigma_{1,\text{KF}}^2 = \|\mathbf{w}_{k-1}\|_2^2 (M\sigma_0^2 + \sigma_{\text{NPI}}^2)$, where $\mathbf{w}_{k-1}$ denotes

[3]We note that our simulated scenario differs slightly to that in [18], as here we have added arbitrary ASV errors to the source ASVs.

the weight vector at the previous iteration and $\sigma^2_{\text{NPI}} = \text{trace}\{\mathbf{Q}\}/M$. This clairvoyant version, which is not implementable on measured data, is termed WC-KF-Clairvoyant and though it performs much better than the one that exploits the DAS estimate, it is still poorer performing than both the Krylov-RDRCB and the WC-IG.

In summary, on simulated data, the WC-IG and Krylov-RDRCB approaches perform the best out of the efficient schemes examined.

X. Experimental Data Results

We proceed to examine the performance using experimental data. Here, we are interested in azimuth beamforming of a hull-mounted sonar array, which consists of 40 staves of sensors, where each stave is on the arc of a circle and contains 8 sensors. As we are only interested in azimuth beamforming, each stave of sensors is first beamformed conventionally, at a prescribed elevation, to give a pseudo-row of hydrophones focused at this elevation that are then processed adaptively. Due to some faulty sensors in the experimental array, the pseudo-row contains only $M = 35$ sensors, with its geometry given in Fig. 1(b) (in arbitrary units) in [19]. We examine results from two different data sets, two of which were used in [16] and one of which was used in [19]. The time-series sensor data were 50% overlapped, Hann shaded, and Fourier transformed to a resolution satisfying (1). Actual parameter values, such as the used frequencies and sensor spacings, are proprietary, but we remark that the array is not aliased. We chose to look at a subband consisting of 96 frequency cells and evaluated 192 equally spaced (in degrees) azimuth beams. The beamformed frequency-domain data is inverse Fourier transformed to produce time-series beam data, which is integrated and then magnitude squared, to produce a bearing time record (BTR). From the BTR samples, we obtain estimates of the detection index, also known as the linear deflection measure of SNR, at sample t using

$$SNR(t) = \frac{\left[\hat{\bar{s}}(t) - \hat{\bar{n}}(t)\right]^2}{\hat{\sigma}_n^2(t)} \tag{44}$$

where, at sample t, $\hat{\bar{s}}(t)$, $\hat{\bar{n}}(t)$, and $\hat{\sigma}_n^2(t)$ denote the mean signal-plus-noise level, the mean noise level and the noise variance estimates, respectively, which are given by

$$\hat{\bar{s}}(t) \;=\; \frac{1}{T}\sum_{k=t-T+1}^{t} s(k) \tag{45}$$

and

$$\hat{\bar{n}}(t) \;=\; \frac{1}{TN_n}\sum_{k=t-T+1}^{t}\sum_{i=1}^{N_n} n_i(k) \tag{46}$$

$$\hat{\sigma}_n^2(t) \;=\; \frac{1}{TN_n - 1}\sum_{k=t-T+1}^{t}\sum_{i=1}^{N_n}(n_i(k) - \hat{\bar{n}}(t))^2 \tag{47}$$

with $s(k)$ and $\{n_i(k)\}_{i=1}^{N_n}$ denoting the assumed signal-plus-noise sample, and the N_n assumed noise samples at the kth BTR time sample. To improve the quality of the estimates, they are obtained by averaging over a time window of length T, where T is selected small enough to reasonably assume stationarity over the window. The simulated data results in Section IX indicated that the WC-IG and Krylov-RDRCB algorithms performed the best out of the efficient algorithms. We thus proceed to compare results obtained from the EVD-based NBRCB used in [16], [19] and summarized in Section II-A, here denoted NBRCB-EVD, with an implementation based on the WC-IG, which we denote NBWC-IG, and another based on the Krylov-RDRCB, which we term NBKrylov-RDRCB. In the robust adaptive methods, we used the method described in Section II-B to design uncertainty sets, where the minimum sphere radius was set equal to 3. In the following, we assume $K = 96$ frequency-domain snapshots per frequency-bin are

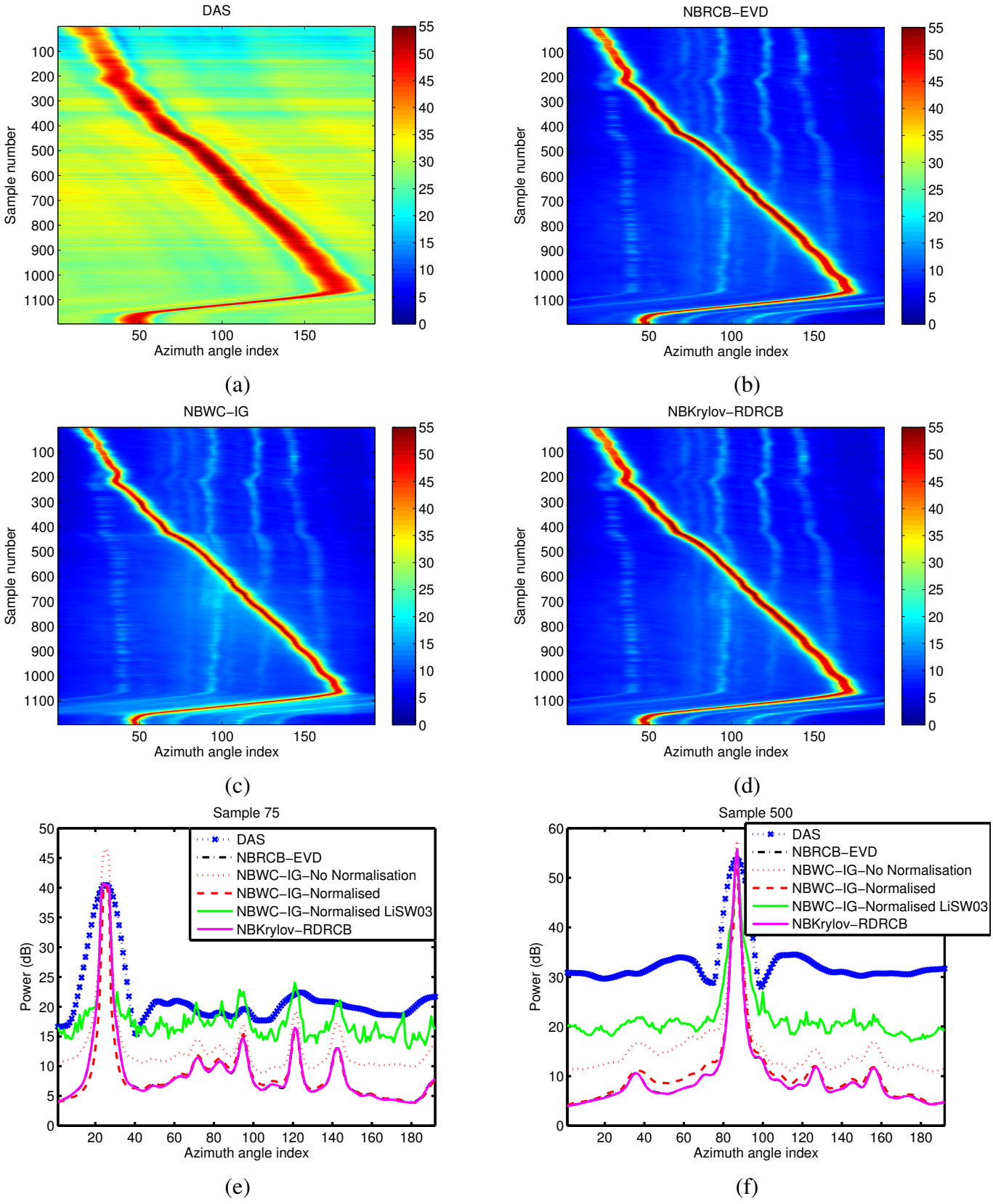

Fig. 4. For the strong controlled source run, bearing time records with $K = 96$ for (a) DAS, (b) NBRCB-EVD, (c) NBWC-IG with normalization, and (d) NBKrylov-RDRCB. Azimuth spectra for samples (e) 75 and (f) 500.

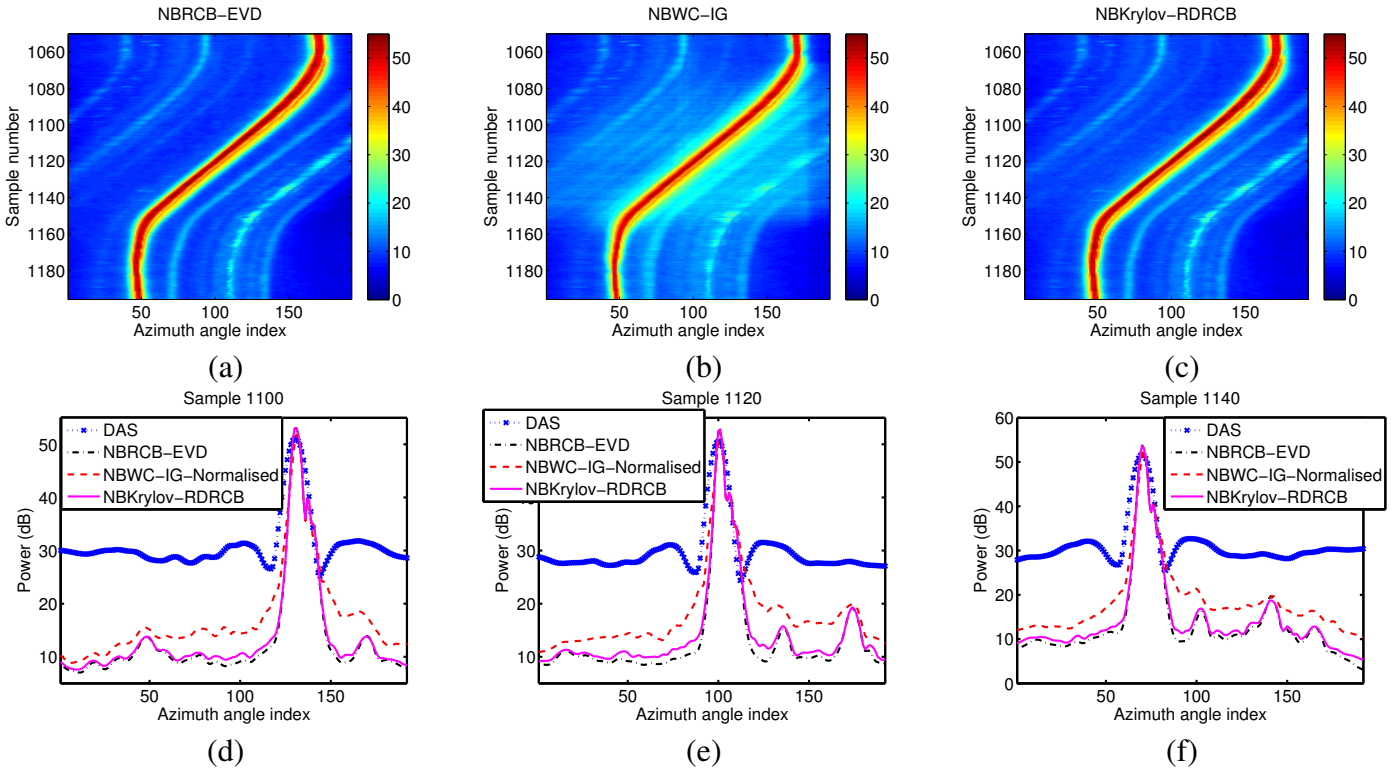

Fig. 5. For a fast section of the strong controlled source run, bearing time records with $K = 96$ for (a) NBRCB-EVD, (b) NBWC-IG with normalization, (c) NBKrylov-RDRCB. Azimuth spectra for samples (d) 1100, (e) 1120, and (f) 1140.

available for covariance estimation. Due to the Hann shading and 50% overlap, the degree of correlation between the frequency-domain snapshots is 0.1667 [44], yielding $K_{\text{eff}} = (1 - 0.1667)K = 80$ effective independent snapshots. As there are $M = 35$ adaptive degrees of freedom, since $K_{\text{eff}} \geq 2M$, we are not concerned with snapshot deficiency here. We also examine results obtained from a (Taylor) shaded DAS (SDAS) beamformer.

A. Strong Controlled Acoustic Source Run

The first data set contains signals from a known, strong, controlled acoustic source in the far-field of the array, on which we were able to check if the algorithms protected against (desired) signal cancellation, when pointing towards the source, and also their ability to null out the strong source when pointing away from it. The experiment was conducted in deep water, late afternoon in midwinter between sea state conditions 1 and 2. Fig. 4(a)–(d) shows the bearing time records (BTRs), whose axes are in arbitrary units, clearly showing that the robust adaptive methods improve the output SNR and spatial resolution compared to the SDAS. Notably, when using robust adaptive beamforming, it is possible to see four weaker sources around beams 40, 95, 120, and 140, which are masked by high sidelobes when using SDAS.

Fig. 4(e) and (f), shows azimuth spectra for samples 75 and 100, where we have also included curves when the proposed ad hoc normalization (29) is not used in the NBWC-IG beamformer and also when we instead use the normalization (28) proposed in [34], which we term the "LiSW03" normalization. For the strong controlled source, the NBRCB-EVD, NBWC-IG with the proposed normalization, and the NBKrylov-RDRCB power estimates converge to the DAS estimate. Since the strong controlled source power is at least 20 dB greater than any other source or background noise, and is well separated from any other source at sample 75, we expect that its DAS power estimate will be close to the true power [16]. When no normalization is used, the NBWC-IG power estimate does not converge to the DAS estimate, as a result of the scaling ambiguity that occurs in the worst-case based algorithms. Even though using the normalization (28) proposed in [34] results in convergence to the DAS estimate for the strong source,

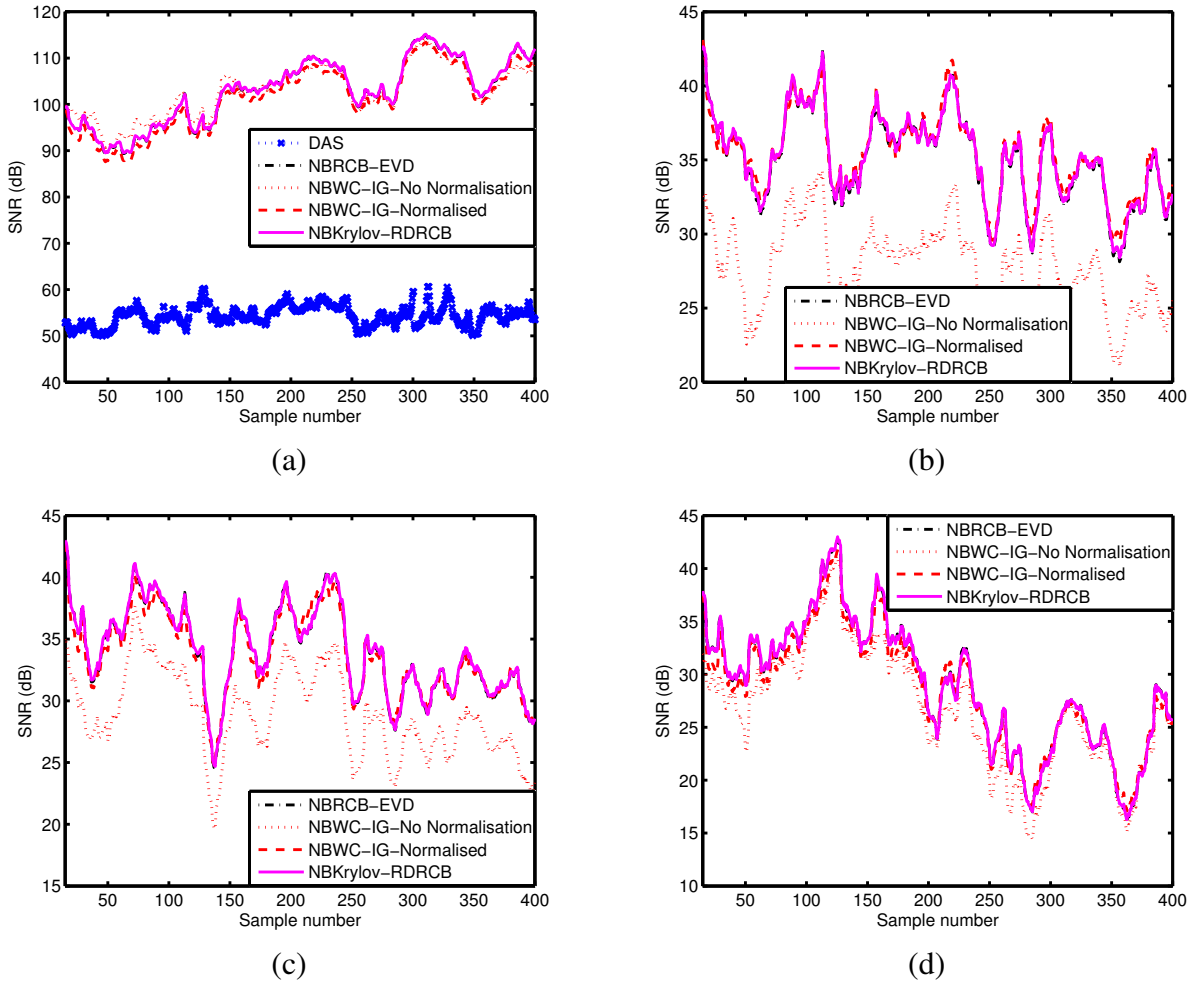

Fig. 6. For the strong controlled source run, SNR for (a) the strong controlled source and the lines at around azimuth indices (b) 95, (c) 120, and (d) 140.

one can see that a large amount of noise is added the spatial spectrum, which leads to large loss in the (deflection) SNR as compared to the unnormalized case. It is clear that our proposed ad hoc normalization (29) is preferable.

Even though the BTRs in Fig. 4(b)–(d) are similar, there are some important differences especially during the fast moving turn towards the end of the data. For samples 1050 onwards, Fig. 5(a), (b) and (c) show the BTRs for NBRCB-EVD, NBWC-IG, and NBKrylov-RDRCB, respectively, indicating that NBWC-IG sidelobes are comprised during the fast turn, due to the fact that the iterative scheme is unable to track the minimum of the optimization fast enough. This issue does not occur in the NBKrylov-RDRCB or NBRCB-EVD which effectively solve the optimization afresh for each new snapshot. Fig. 5(d), (e) and (f) show the azimuth spectra for samples 1100, 1120 and 1140, respectively, showing the issue more clearly.

From the first 400 samples of the BTRs, and using (44), we extract the linear deflection measure of SNR for the strong controlled acoustic source and the weaker sources that are around azimuth indices (beams) 94, 120, and 140. At instant k, the peak sample from each source is used as the signal-plus-noise sample $s(k)$ in (45). We assume that beams 160-180 contain only noise, from which we can measure the noise statistics in (46) and (47). Here, we use an integration window of $T = 10$ BTR samples when obtaining the estimates in (45)–(47). Figs. 6 show the SNRs as a function of sample number for the strong controlled source and the weaker sources at around azimuth indices 95, 120, and 140, respectively. Figs. 7 show the SNRs normalized with respect to the NBRCB-EVD SNR. We have only plotted the SDAS SNR

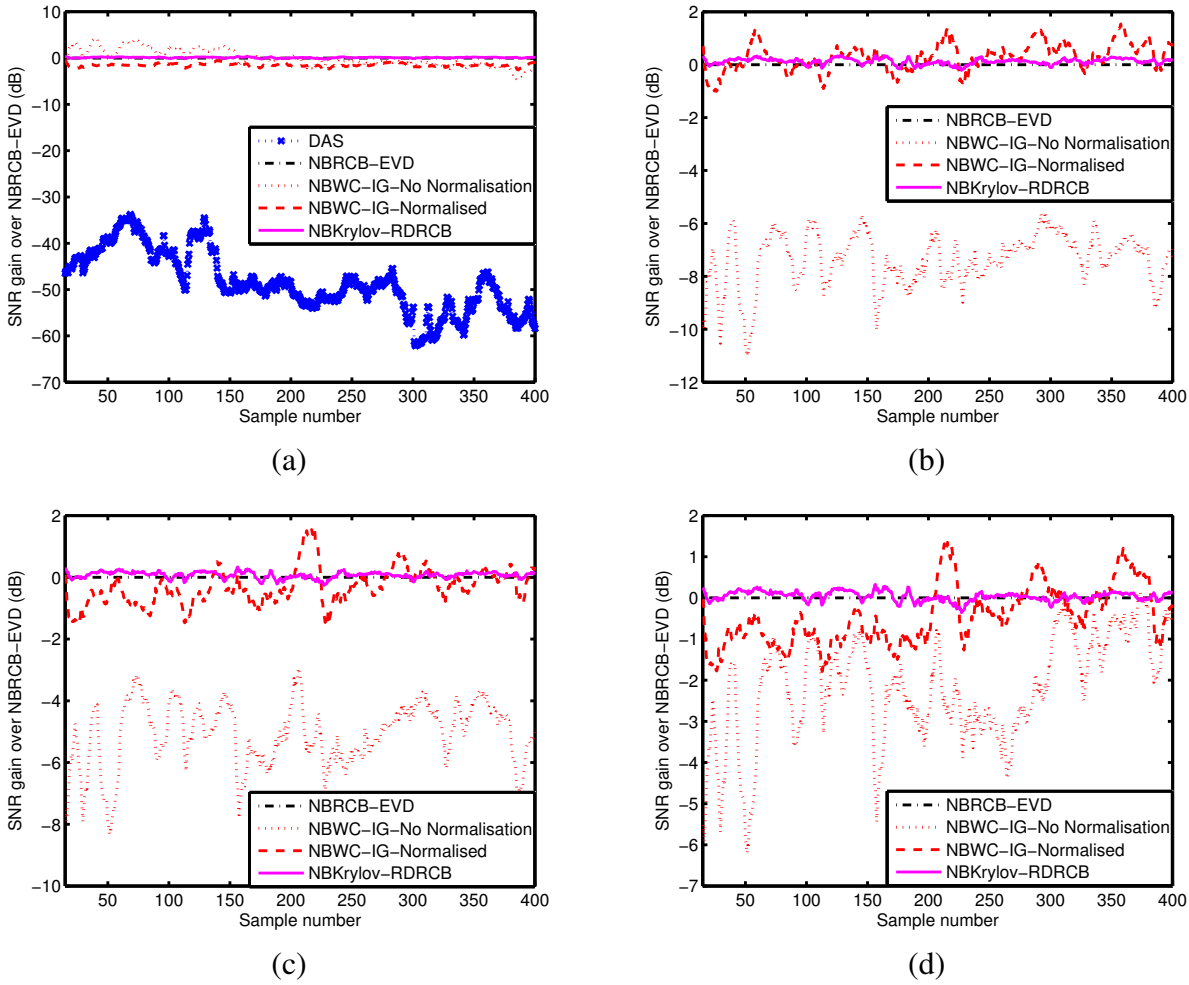

Fig. 7. For the strong controlled source run, SNR gain over NBRCB-EVD for (a) the strong controlled source and the lines at (b) 94, (c) 120, and (d) 140.

for the strong controlled acoustic source as the other, weaker, sources are not detectable with SDAS. For the strong controlled source, the adaptive algorithms give 30-55 dB higher SNR than SDAS, as the high SDAS sidelobes lead to an increase in the mean noise level and variance. Since the weaker lines are undetectable using SDAS, we can assume that their SDAS SNRs are less than 0 dB. From Figs. 6(b)–(d), it is clear then that the adaptive algorithms give at least between 15-45 dB gain over the SDAS for these weak sources. The curves showing the SNR gain over NBRCB-EVD indicate that NBWC-IG and NBKrylov-RDRCB are within ± 2 dB and ± 0.4 dB of NBRCB-EVD, respectively, indicating that both are good implementations of the NBRCB-EVD. The results in Figs. 6 and 7 also show the importance of normalizing the weights using (29) in NBWC-IG, as not normalizing the weights results in losses of up to approximately 10 dB on the weaker sources.

B. Acceleration Run

The second data set was recorded during an acceleration run, where the boat was initially at low speed and moved in a straight line accelerating to high speed over a period of 20 minutes. The data was recorded in the Alboran basin, during mid-winter, at nighttime, and at sea state condition 0. In Fig. 8(a), we plot the bearing-time record for two segments of the data. Samples 1-295 contain the bearing-time record for a continuous segment of data containing low to medium speeds. Samples 296 and above show the BTR for a continuous segment of data recorded at high speed. Figs. 8(b)–(d) illustrate azimuth spectra for low, medium, and high speeds. For the initial segment at low to medium speeds, Figs. 9(a) and (b) show

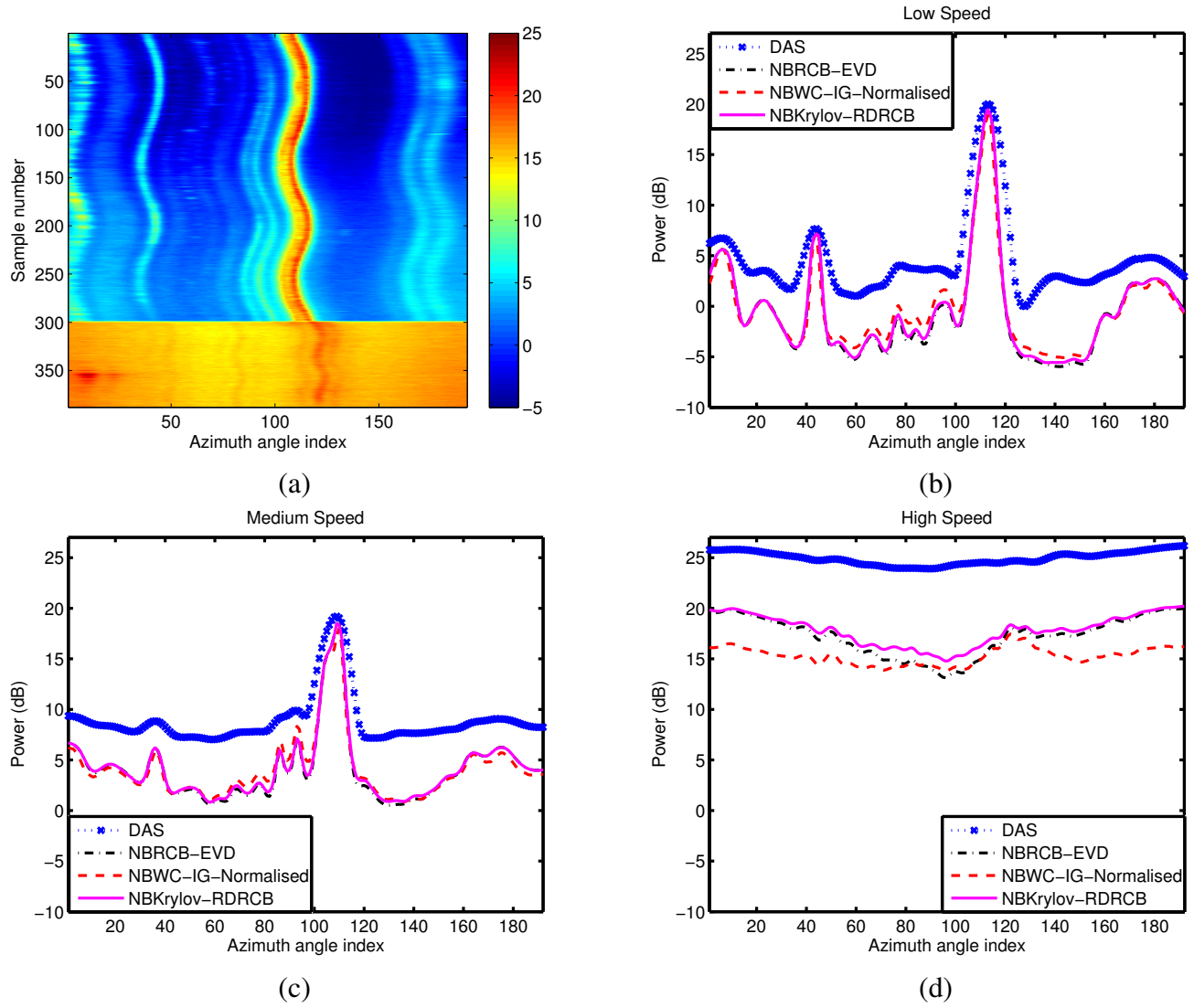

Fig. 8. For the acceleration run, (a) the NBWC-IG bearing time record and the azimuth spectra measured at (b) low, (c) medium and (d) high speed.

NBKrylov-RDRCB and SDAS BTRs. For the (deflection) SNR calculation, we assume that the beams at azimuth indices 130-150 contain noise only and we use a window length of $T = 10$ samples. For the weak source at around azimuth index 40, Figs. 9(c) and (d) illustrate the SNR and the SNR gain over NBRCB-EVD, respectively. The associated plots for the stronger source at around azimuth index 110 are shown in Figs. 9(e) and (f). Figs. 10 (a) and (b) show the NBWC-IG and SDAS BTRs for the continuous segment recorded at high speed. For the (deflection) SNR calculation, we assume that the beams at azimuth indices 60-70 only contain noise and use a window length of $T = 5$ samples. The results in Figs. 9 and 10 illustrate how significant gains over SDAS are achievable at all speeds. Similar to the results obtained on the strong controlled acoustic source run, NBRCB-EVD, NBKrylov-RDRCB and NBWC-IG perform similarly. At the highest speed there is deviation of the NBWC-IG spectrum from the NBRCB-EVD and NBKrylov-RDRCB spectra, which could be due to the ad hoc normalization. We remark that the sudden drop in (deflection) SNR for the conventional SDAS beamformer observed between samples 350 and 360 in Figs. 10 (c) and (d) is due to the appearance of a transient at around azimuth angle index 10, which can be seen, e.g., in Fig. 10 (a). The effect of this transient on the robust adaptive beamforming algorithms' SNRs is negligible.

To summarize, the experimental data results show that the efficient NBKrylov-RDRCB and NBWC-IG

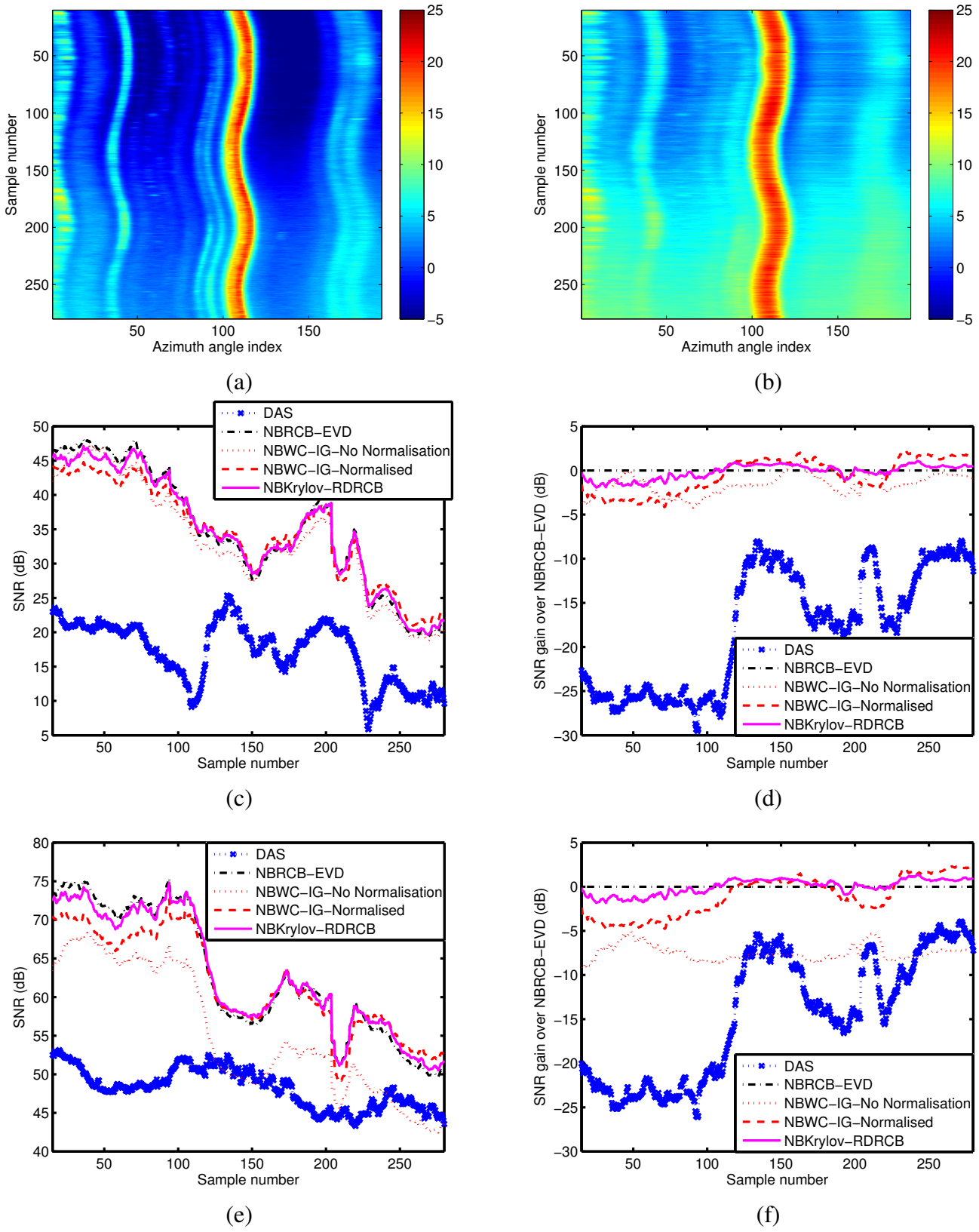

Fig. 9. For samples 10-280 of the acceleration run, BTRs for (a) NBKrylov-RDRCB and (b) SDAS; For the weak source at around azimuth index 40, (c) the SNRs and (d) the SNR gains over NBRCB-EVD; For the strong source at around azimuth index 110, (e) the SNRs and (f) the SNR gains over NBRCB-EVD

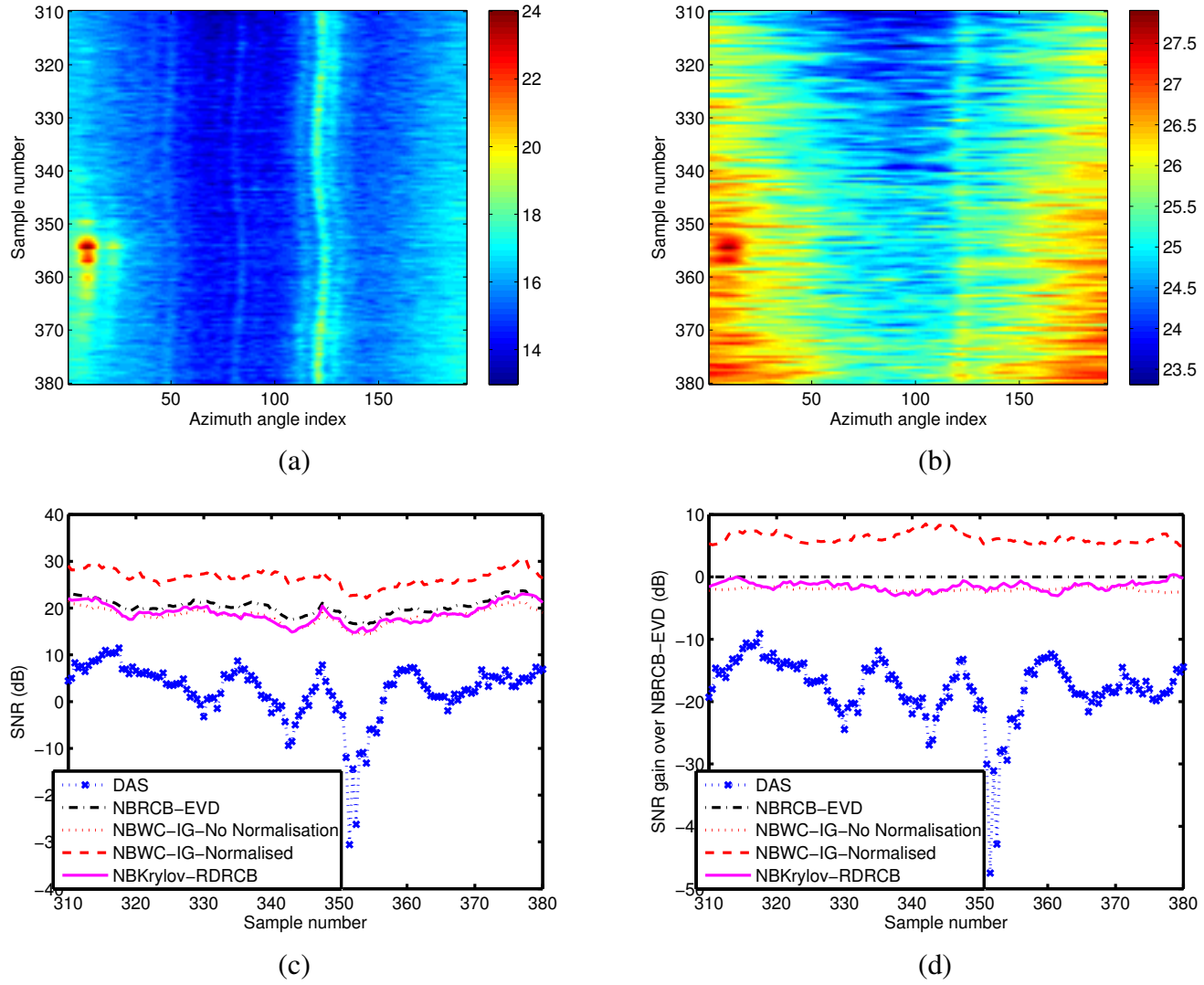

Fig. 10. For the fast section of the acceleration run, (a) the NBWC-IG BTR (b) the SDAS BTR, (c) the SNR as a function of sample number, and (f) the SNR gain over NBRCB-EVD.

algorithms performs similarly, in terms of (deflection) SNR, to the NBRCB-EVD. The NBKrylov-RDRCB is a closer representation to the NBRCB-EVD algorithm and also does not suffer when the environment changes rapidly.

XI. Conclusions

In this paper, we have examined five recent efficient (ellipsoidal uncertainty set based) robust adaptive beamforming algorithms. Using simulated data, we conclude that the Krylov reduced-dimension robust Capon beamformer implementation of the robust Capon beamformer and the iterative gradient based implementation of the worst-case robust adaptive beamformer performed the best for the considered passive sonar application, being around 20 times and 80 times less complex than the standard eigenvalue decomposition based RCB, respectively, on a 40 degrees of freedom adaptive array, and also more amenable to parallelization on, e.g., graphical processing units. We used these approaches to implement the recent NBRCB algorithm, terming the efficient implementations NBKrylov-RDRCB and NBWC-IG, and evaluated them on recorded passive sonar data with reasonable snapshot support. We found that NBKrylov-RDRCB performed more like NBRCB-EVD and was more robust to rapidly varying environments than the NBWC-IG, and that both performed significantly better than the conventional shaded

DAS beamformer. In summary, the efficient Krylov-RDRCB and WC-IG based implementations provide excellent performance at a significantly reduced computationally complexity.

XII. Acknowledgments

The authors are grateful to Dr. Nigel Parsons, Dr. Martin Green, Mr. Paul Bacchus, and Dr. Paul Gosling for interesting discussions and their continued support.

References

[1] H. L. V. Trees, *Detection, Estimation, and Modulation Theory, Part IV, Optimum Array Processing*. John Wiley and Sons, Inc., 2002.
[2] J. Li and P. Stoica, *Robust Adaptive Beamforming*. New York: Wiley, 2005.
[3] H. Cox, "Resolving power and sensitivity mismatch of optimum array processor," *Journal of the Acoustical Society of America*, vol. 54, no. 3, pp. 771–785, 1973.
[4] S. P. Applebaum and D. J. Chapman, "Adaptive Arrays with Main Beam Constraints," *IEEE Trans. Antennas Propag.*, vol. AP-24, no. 5, pp. 650–662, Sep. 1976.
[5] M. H. Er and A. Cantoni, "Derivative Constraints for Broad-band Element Space Antenna Array Processors," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-31, pp. 1378–1393, December 1983.
[6] K. M. Buckley and L. J. Griffiths, "An Adaptive Generalized Sidelobe Canceller with Derivative Constraints," *IEEE Trans. Antennas Propag.*, vol. AP-34, pp. 311–319, Mar. 1986.
[7] M. H. Er and A. Cantoni, "A New Set of Linear Constraints for Broadband Time-Domain Element-Space Processors," *IEEE Trans. Antennas Propag.*, vol. AP-34, pp. 320–329, Mar. 1986.
[8] S. D. Somasundaram, "Linearly Constrained Robust Capon Beamforming," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5845–5856, Nov. 2012.
[9] B. D. Carlson, "Covariance Matrix Estimation Errors and Diagonal Loading in Adaptive Arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 397–401, Jul. 1988.
[10] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust Adaptive Beamforming," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 10, pp. 1365–1376, Oct. 1987.
[11] P. Stoica, Z. Wang, and J. Li, "Robust Capon Beamforming," *IEEE Signal Process. Lett.*, vol. 10, no. 6, pp. 172–175, June 2003.
[12] 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, July 2003.
[13] 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, February 2003.
[14] R. G. Lorenz and S. Boyd, "Robust Minimum Variance Beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684–1696, May 2005.
[15] J. Li, P. Stoica, and Z. Wang, "Doubly Constrained Robust Capon Beamformer," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2407–2423, Sep. 2004.
[16] S. D. Somasundaram and N. H. Parsons, "Evaluation of Robust Capon Beamforming for Passive Sonar," *IEEE J. Ocean. Eng.*, vol. 36, no. 4, pp. 686–695, Oct. 2011.
[17] 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.
[18] 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.
[19] ——, "Wideband Robust Capon Beamforming for Passive Sonar," *IEEE J. Ocean. Eng.*, vol. 38, no. 2, pp. 308–322, April 2013.
[20] A. B. Baggeroer and H. Cox, "Passive Sonar Limits Upon Nulling Multiple Moving Ships with Large Aperture Arrays," in *IEEE 33rd Asilomar Conf. Signals, Systems Comput.*, Pacific Grove, CA, October 1999, pp. 103–108.
[21] D. Feldman and L. J. Griffiths, "A Projection Approach for Robust Adaptive Beamforming," *IEEE Trans. Signal Process.*, vol. 42, no. 4, pp. 867–876, April 1994.
[22] D. A. Gray, "Formulation of the Maximum Signal-to-Noise Ratio Array Processor in Beam Space," *J. Acoust. Soc. Am.*, vol. 72, no. 4, pp. 1195–1201, Oct. 1982.
[23] N. L. Owsley, "Systolic Array Adaptive Beamforming," in *Naval Underwater Systems Center Technical Report, NUSC 7981*, Apr. 1987.
[24] D. N. Swingler, "A Low-Complexity MVDR Beamformer for Use with Short Observation Times," *IEEE Trans. Signal Process.*, vol. 47, no. 4, pp. 1154–1160, Apr. 1999.
[25] H. Cox and H. Lai, "Sub-Aperture Beam-Based Adaptive Beamforming for Large Dynamic Arrays," in *Proc. 38th Asilomar Conf. Signals, Systems, Comput.*, vol. 2, Nov. 2004, pp. 2355–2358.
[26] Y. Doisy, L. Deruaz, and R. Been, "Interference Suppression of Subarray Adaptive Beamforming in Presence of Sensor Dispersions," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4195–4212, Aug. 2010.
[27] N. Lee, L. M. Zurk, and J. Ward, "Evaluation of Reduced-Rank, Adaptive Matched Field Processing Algorithms for Passive Sonar Detection in a Shallow-Water Environment," in *Proc. 33rd Asilomar Conf. Signals, Systems, Comput.*, vol. 2, Pacific Grove, CA, Oct. 1999, pp. 876–880.
[28] S. D. Somasundaram and N. H. Parsons, "Robust Krylov-Subspace Methods for Passive Sonar Adaptive Beamforming," in *Proc. 3rd IMA Conference on Mathematics in Defence*, Malvern, UK, Oct. 24 2013.

[29] S. D. Somasundaram, P. Li, N. H. Parsons, and R. C. de Lamare, "Reduced-Dimension Robust Capon Beamforming Using Krylov-Subspace Techniques," to appear in *IEEE Trans. Aerosp. Electron. Syst.*

[30] D. A. Abraham and N. L. Owsley, "Beamforming with Dominant Mode Rejection," in *Proc. IEEE OCEANS*, Washington, DC, Sep. 1990, pp. 470–475.

[31] T. R. Messerschmitt and R. A. Gramann, "Evaluation of the Dominant Mode Rejection Beamformer Using Reduced Integration Times," *IEEE J. Ocean. Eng.*, vol. 22, no. 2, pp. 385–392, April 1997.

[32] H. Song, W. A. Kuperman, W. S. Hodgkiss, P. Gerstoft, and J. S. Kim, "Null-Broadening with Snapshot-Deficient Covariance Matrices in Passive Sonar," *IEEE J. Ocean. Eng.*, vol. 28, no. 2, pp. 250–261, April 2003.

[33] 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.

[34] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," in *28th IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Hong Kong, April 2003, pp. 337–349.

[35] B. Jiang, "Low-Complexity Implementation for Worst-Case Optimization-Based Robust Adaptive Beamforming," in *Proc. 5th IEEE Sensor Array and Multichannel Signal Processing Workshop*, Darmstadt, Germany, July 21-23 2008, pp. 311–314.

[36] A. E.-Keyi, T. Kirubarajan, and A. B. Gershman, "Robust Adaptive Beamforming Based on the Kalman Filter," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3032–3041, Aug. 2005.

[37] A. Elnashar, "Efficient Implementation of Robust Adaptive Beamforming Based on Worst-Case Performance Optimization," *IET Signal Processing*, vol. 2, no. 4, pp. 381–393, Dec. 2008.

[38] A. Elnashar, S. M. Elnoubi, and H. A. El-Mikati, "Further Study on Robust Adaptive Beamforming With Optimum Diagonal Loading," *IEEE Trans. Antennas Propag.*, vol. 54, no. 12, pp. 3647–3658, Dec. 2006.

[39] W. Zhang and S. Wu, "Low-Complexity Online Implementation of a Robust Capon Beamformer," in *Proc. Sensor Signal Processing for Defence*, London, U.K., Sep. 25-27 2012, pp. 1–4.

[40] A. J. A. P. S. D. Somasundaram, N. R. Butt and L. Hart, "Computationally Efficient Robust Adaptive Beamforming for Passive Sonar," in *Proc. Underwater Defence Technology*, Liverpool, U.K., Jun. 2014.

[41] S. D. Somasundaram and N. H. Parsons, "Using Robust Capon Beamforming to Minimise the Impact of Angle of Arrival Errors in 2-D Arrays," in *IEEE Workshop on Statistical Signal Processing*, 2009.

[42] M. R. Hestenes and E. Stiefel, "Methods of Conjugate Gradients for Solving Linear Systems," *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409–436, Dec. 1952.

[43] 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.

[44] F. J. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform," *Proceedings of the IEEE*, vol. 66, no. 1, pp. 51–83, Jan. 1978.