Source Localization and Sensing: A Nonparametric Iterative Adaptive Approach Based on Weighted Least Squares

Tarik Yardibi, *Student Member, IEEE,* Jian Li*, *Fellow, IEEE,*
Petre Stoica, *Fellow, IEEE,* Ming Xue, *Student Member, IEEE,*
Arthur B. Baggeroer, *Fellow, IEEE*

Abstract

Array processing is widely used in sensing applications for estimating the locations and waveforms of the sources in a given field. In the absence of a large number of snapshots, which is the case in numerous practical applications such as underwater array processing, it becomes challenging to estimate the source parameters accurately. This paper presents a nonparametric and hyperparameter free weighted least squares-based iterative adaptive approach for amplitude and phase estimation (IAA-APES) in array processing. IAA-APES can work well with few snapshots (even one), uncorrelated, partially correlated, and coherent sources, and arbitrary array geometries. IAA-APES is extended to give sparse results via a model-order selection tool, the Bayesian information criterion (BIC). Moreover, it is shown that further

Manuscript submitted December 2007, revised July 2008 and November 2008. This work was supported in part by the Office of Naval Research (ONR) under Grants No. N00014-07-1-0193, N00014-07-1-0293 and N00014-01-1-0257, the Army Research Office (ARO) under Grant No. W911NF-07-1-0450, the National Aeronautics and Space Administration (NASA) under Grant No. NNX07AO15A, the National Science Foundation (NSF) under Grants No. CCF-0634786, ECS-0621879 and ECS-0729727, the Swedish Research Council (VR) and the European Research Council (ERC). Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

Tarik Yardibi, Jian Li and Ming Xue are with the Dept. of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611, USA. Jian Li performed this work while on sabbatical leave at MIT, Cambridge, MA.

Petre Stoica is with the Dept. of Information Technology, Uppsala University, Uppsala, Sweden.

Arthur B. Baggeroer is with the Depts. of Mechanical Engineering & Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

*Please address all correspondence to Jian Li. Phone: (352) 392-2642; Fax: (352) 392-0044; Email: li@dsp.ufl.edu.

improvements in resolution and accuracy can be achieved by applying the parametric RELAX algorithm to refine the IAA-APES&BIC estimates if desired. IAA-APES can also be applied to active sensing applications including single-input single-output (SISO) radar/sonar range-Doppler imaging and multi-input single-output (MISO) channel estimation for communications. Simulation results are presented to evaluate the performance of IAA-APES for all these applications, and IAA-APES is shown to outperform a number of existing approaches.

Index Terms

Passive or active sensing, Array processing, Range-Doppler imaging, Channel estimation, Underwater acoustic communications, Sparse signal representation, Few snapshots, Direction-of-arrival (DOA) estimation, Spectrum estimation.

I. INTRODUCTION

THE goal of array processing is to estimate the locations and waveforms of sources by combining the received data from multiple sensors so that the desired signal is enhanced, while the unwanted signals such as interference and noise are suppressed. In active sensing applications such as radar/sonar range-Doppler imaging, the aim is to find targets present in a region of interest. For channel estimation in communications, the aim is to estimate the non-zero channel taps and their Doppler shifts, which are then fed to the subsequent equalizer for symbol detection. Low numbers of snapshots and low signal-to-noise ratios (SNR) are among the many challenges that array processing systems frequently face. Another challenge is the presence of nearby sources, in terms of location or Doppler, since closely spaced sources are harder to discriminate. We consider herein both passive array processing and active sensing.

The most basic approach to array processing is the classical delay-and-sum (DAS) method, in which the received signal from each sensor is weighted and delayed so as to focus on different points in space. However, this method suffers from low resolution and high sidelobe levels. There is a vast amount of literature on methods that provide superior performance over the DAS approach when certain assumptions are met [1]. The well-known standard Capon beamformer (SCB) [2] and multiple signal classification (MUSIC) [3], [4] methods provide superresolution when the sources are uncorrelated and the number of snapshots is high. Many extensions to these methods have been proposed to deal with modelling errors such as steering vector mismatches,

see, e.g., [5]-[9]. However, none of these methods is able to cope with very low snapshot numbers, coherent or highly correlated sources, and severe noise.

Only a few snapshots are available when the environment being sensed by the array is stationary for a short duration of time. Moreover, to avoid smearing, i.e., losing resolution because of wide main beamwidths in the array response, averaging can only be done over a small bandwidth [10], [11]. Therefore, the number of available snapshots, which is directly related to the time-bandwidth product, can be very small, sometimes as small as 3, for applications such as underwater array processing. Furthermore, as discussed later, the data models for single-input single-output (SISO) radar/sonar range-Doppler imaging and multi-input single-output (MISO) channel estimation in communication problems are similar to the model used in array processing with a single snapshot and an arbitrary array geometry (dictated by the probing waveforms).

The array processing problem has been carried into the sparse signal representation area by noticing that the number of actual sources is usually much smaller than the number of potential source points that can be considered [12]-[29]. Sparsity-based techniques have also been used in spectral estimation [30], image processing, and array design [31]-[33], among many other application areas. Sparse signal representation algorithms can deal with a few snapshots (even one). However, they may require large computation times and the fine tuning of one or more hyperparameters.

This paper addresses application areas ranging from passive source localization to active radar/sonar range-Doppler imaging and channel estimation for communications. Section II considers passive sensing applications and describes a weighted least squares-based iterative adaptive approach for amplitude and phase estimation (IAA-APES). The algorithm is named IAA-APES herein since its derivation resembles that of the amplitude and phase estimation (APES) algorithm [34]-[36]; see Section II-C. This name also distinguishes it from the maximum likelihood-based iterative adaptive approach (IAA-ML) discussed in the appendix to further motivate IAA-APES. Using the Bayesian information criterion (BIC) [37], [38], IAA-APES can be extended to yield point source estimates (this approach is referred to as IAA-APES&BIC). Next, a parametric relaxation-based cyclic approach, namely RELAX [39], [40], will be discussed as a way to further refine the results of IAA-APES&BIC (this approach is referred to as IAA-APES&RELAX). Section III presents the data models for the aforementioned active sensing applications and emphasizes the similarities to the passive sensing case. IAA-APES will be

TABLE I

NOTATION USED IN THE TEXT.

$\|\cdot\|_0$	ℓ_0-norm
$\|\cdot\|_1$	ℓ_1-norm
$\|\cdot\|_2$	ℓ_2-norm
$\|\cdot\|_{\mathrm{F}}$	Frobenius norm of a matrix
$\odot$	the Hadamard (elementwise) matrix product
$\mathrm{tr}(\cdot)$	trace of a matrix
$(\cdot)^T$	transpose of a vector or matrix
$(\cdot)^H$	conjugate transpose of a vector or matrix

evaluated via comprehensive simulations in Section IV, and its performance will be compared with that of a number of existing approaches.

Notation: We denote vectors and matrices by boldface lowercase and boldface uppercase letters, respectively. The kth component of a vector $\mathbf{x}$ is written as x_k. The kth diagonal element of a matrix $\mathbf{P}$ is written as P_k. See Table I for other symbols and their meanings.

II. PASSIVE SENSING

In passive sensing applications such as aeroacoustic noise measurements [41] and underwater acoustic measurements [42], an array of sensors is used to estimate the desired source characteristics. This section first introduces the data model for such applications and then presents the IAA-APES algorithm.

A. Data Model

Consider the wavefield generated by K sources located at $\boldsymbol{\theta}$, where $\boldsymbol{\theta} \triangleq [\theta_1, \theta_2, \ldots, \theta_K]$ and θ_k[1] is the location parameter of the kth signal, $k = 1, \ldots, K$. In the narrowband multi-snapshot case, the $M \times 1$ array output vector of an M element array in the presence of additive noise

[1]With a slight abuse of notation, we do not use bold font for θ_k, $k = 1, \ldots, K$, which might be multi-dimensional, for simplicity.

can be represented as [5], [43]

$$\mathbf{y}(n) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(n) + \mathbf{e}(n), \quad n = 1, \ldots, N, \tag{1}$$

where N is the number of snapshots, $\mathbf{A}(\boldsymbol{\theta})$ is the $M \times K$ steering matrix defined as $\mathbf{A}(\boldsymbol{\theta}) \triangleq [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \ldots, \mathbf{a}(\theta_K)]$, and $\mathbf{s}(n) \triangleq [s_1(n), s_2(n), \ldots, s_K(n)]^T$, $n = 1, \ldots, N$, is the source waveform vector at time n.

The array steering matrix has different expressions, depending on the array geometry and on whether the source is in the near-field or far-field of the array. For instance, the steering vector corresponding to the kth source for a far-field linear array (where θ_k represents the impinging angle of source k in this case) is given by

$$\mathbf{a}(\theta_k) = \left[e^{-j\frac{2\pi f}{c_0} x_1 \cos(\theta_k)}, \ldots, e^{-j\frac{2\pi f}{c_0} x_M \cos(\theta_k)} \right]^T, \tag{2}$$

where f is the center frequency, c_0 is the wave propagation velocity, and x_m is the position of the mth sensor, $m = 1, \ldots, M$; see Figure 1. Note that for a planar array or near-field sources, only the expressions for the steering vectors have to be modified; the algorithms that will be presented can be applied without any modifications, since $\mathbf{a}(\theta)$ is assumed to be a known function of θ. The number of sources, K, is usually unknown; hence, in this paper, K is considered to be the number of scanning points in the region. In other words, every point of a predefined grid that covers the region of interest is considered as a potential source whose power will be estimated. Consequently, K will be much larger than the actual number of sources present, and only a few components of $\{\mathbf{s}(n)\}$ will be non-zero. This is the main reason why sparse algorithms can be used in array processing applications.

B. Related Work

We will focus our attention on array processing algorithms exploiting sparsity, which have gained noticeable interest recently. Sparse signal representation aims at finding the sparsest $\mathbf{s}$ such that $\mathbf{y} = \mathbf{A}\mathbf{s}$ is satisfied, i.e., to minimize $\|\mathbf{s}\|_0$ such that $\mathbf{y} = \mathbf{A}\mathbf{s}$, where $\mathbf{A}$ is known and $\mathbf{y}$ is measured. The problem in its original form is a combinatorial problem and is NP hard, making it impractical [12]. Fortunately, when $\mathbf{s}$ is sufficiently sparse [12]-[15], $\|\mathbf{s}\|_0$ can be replaced by $\|\mathbf{s}\|_1$, which leads to a convex optimization problem that can be solved much more easily using, for instance, the least absolute shrinkage and selection operator (LASSO) [16] or basis

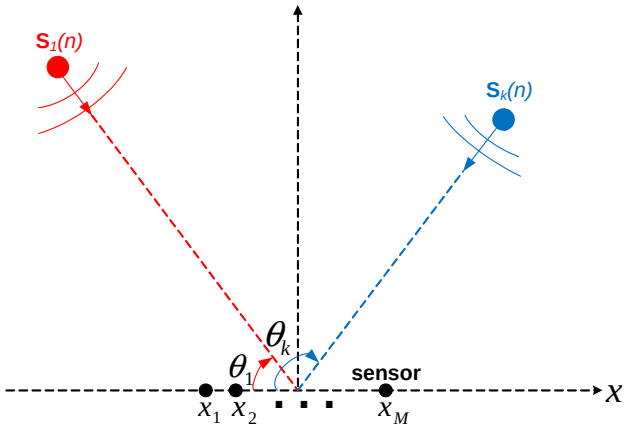


Fig. 1. Far-field linear array.

pursuit (BP) [17] algorithms. Alternatively, the focal underdetermined system solution (FOCUSS) algorithm [18], which is derived using Lagrange multipliers, can be used to iteratively solve the sparse problem. A Bayesian approach, such as sparse Bayesian learning (SBL) [19], [20] or the approach in [21], can also be used to estimate $\mathbf{s}$. The algorithms mentioned up to here are for the single-snapshot case only. The extensions of FOCUSS and SBL to the multiple-snapshot case are M-FOCUSS [22] and M-SBL [23], respectively. Another algorithm, the ℓ_1-SVD (singular value decomposition) algorithm [24], [25] is similar to BP or LASSO, but can work with multiple snapshots. (For the single-snapshot case, ℓ_1-SVD becomes a LASSO and BP type of method.) M-FOCUSS requires the tuning of two hyperparameters, which might affect the performance of the algorithm significantly. ℓ_1-SVD requires the tuning of a hyperparameter and an estimate for the number of sources. Moreover, implementing ℓ_1-SVD requires a convex optimization software such as SeDuMi [44]. M-SBL does not require any hyperparameters. However, M-SBL converges quite slowly in its original form [19], [20], [23].

Besides the above algorithms, which will be the focus of our attention in the numerical examples, there are other sparsity based approaches worth mentioning. [26] adds an additional spatial sparsity regularizing term (an ℓ_2-norm constraint) to the ℓ_1-norm constraint and minimizes a cost function similar to that of ℓ_1-SVD. However, this method has two hyperparameters, assumes that the source waveforms can be represented by a sparse basis, and has high computational complexity [26]. Fuchs [27], [28] uses a sparsity-constrained deconvolution approach that assumes the sources are uncorrelated and the number of snapshots is large. The sparsity-

constrained solution is obtained with a LASSO or BP type of algorithm. [29] introduces two hyperparameter free deconvolution algorithms exploiting sparsity: a sparsity-based extension to the deconvolution approach for the mapping of acoustic sources (DAMAS) [45], which is similar to [27], [28] and widely used in practice, and a sparsity based covariance-matrix fitting approach. Extensions to the correlated source case are also provided. However, the methods in [29] are based on the sample covariance matrix and hence do not work well with few snapshots.

C. IAA-APES

IAA-APES is a data-dependent nonparametric algorithm based on a weighted least squares (WLS) approach. Let $\mathbf{P}$ be a $K \times K$ diagonal matrix whose diagonal contains the power at each angle on the scanning grid. Then, $\mathbf{P}$ can be expressed as:

$$P_k = \frac{1}{N} \sum_{n=1}^{N} |s_k(n)|^2, \quad k = 1, \ldots, K. \tag{3}$$

Furthermore, define the interference (signals at angles other than the angle of current interest θ_k) and noise covariance matrix $\mathbf{Q}(\theta_k)$ to be

$$\mathbf{Q}(\theta_k) = \mathbf{R} - P_k \mathbf{a}(\theta_k)\mathbf{a}^H(\theta_k), \tag{4}$$

where $\mathbf{R} \triangleq \mathbf{A}(\boldsymbol{\theta})\mathbf{P}\mathbf{A}^H(\boldsymbol{\theta})$. Then, the WLS cost function is given by (see, e.g., [34]-[36], [43])

$$\sum_{n=1}^{N} \|\mathbf{y}(n) - s_k(n)\mathbf{a}(\theta_k)\|^2_{\mathbf{Q}^{-1}(\theta_k)}, \tag{5}$$

where $\|\mathbf{x}\|^2_{\mathbf{Q}^{-1}(\theta_k)} \triangleq \mathbf{x}^H \mathbf{Q}^{-1}(\theta_k)\mathbf{x}$ and $s_k(n)$ represents the signal waveform at angle θ_k and at time n. Minimizing (5) with respect to $s_k(n)$, $n = 1, \ldots, N$, yields

$$\hat{s}_k(n) = \frac{\mathbf{a}^H(\theta_k)\mathbf{Q}^{-1}(\theta_k)\mathbf{y}(n)}{\mathbf{a}^H(\theta_k)\mathbf{Q}^{-1}(\theta_k)\mathbf{a}(\theta_k)}, \quad n = 1, \ldots, N. \tag{6}$$

This looks like the result that would be obtained by employing APES [34]-[36], but it is actually different than APES since APES obtains $\mathbf{Q}(\theta_k)$ from the data by forming subapertures, while IAA-APES computes $\mathbf{Q}(\theta_k)$ as in (4). Moreover, IAA-APES is iterative, but APES is not, and APES cannot be used with arbitrary array geometries.

Using (4) and the matrix inversion lemma, (6) can be written as

$$\hat{s}_k(n) = \frac{\mathbf{a}^H(\theta_k)\mathbf{R}^{-1}\mathbf{y}(n)}{\mathbf{a}^H(\theta_k)\mathbf{R}^{-1}\mathbf{a}(\theta_k)}, \quad n = 1, \ldots, N. \tag{7}$$

This avoids the computation of $\mathbf{Q}^{-1}(\theta_k)$ for each scanning point, i.e., K times. Moreover, $\{\hat{s}_k(n)\}$ can be computed in parallel for each scanning point, which makes IAA-APES amenable to implementation on parallel hardware. IAA-APES is summarized in Table II. Since IAA-APES requires $\mathbf{R}$, which itself depends on the unknown signal powers, it has to be implemented as an iteration. The initialization is done by a standard DAS beamformer. Our empirical experience is that IAA-APES does not provide significant improvements in performance after about 15 iterations. In IAA-APES, $\mathbf{P}$ and hence $\mathbf{R}$ are obtained from the signal estimates of the previous iteration and not from the snapshots as in conventional adaptive beamforming algorithms such as SCB, which fails to work properly with coherent or highly correlated sources, or few snapshots [43]. The appendix provides an alternative derivation of IAA-APES based on the maximum likelihood (ML) principle. IAA-APES is shown to be an approximation of the IAA-ML algorithm, which is locally convergent due to cyclically maximizing the likelihood function. Hence, the analysis in the appendix provides an approximate calculation showing the local convergence of IAA-APES.

Because $\mathbf{P}$ is assumed to be a diagonal matrix, one degree of freedom (DOF) is needed to suppress one interfering source. This may lead to a larger than necessary reduction of the number of degrees of freedom when some of the interfering sources are coherent since the cancellation of multiple coherent interfering sources would require only one DOF if the correct structure of $\mathbf{P}$ were known. However, we do not assume that the true structure of $\mathbf{P}$ is known. Moreover, it is the diagonal structure of $\mathbf{P}$ assumed by IAA-APES that makes the algorithm work properly even for low number-of-snapshot cases and coherent sources.

D. IAA-APES&BIC

In many applications, it is desirable to obtain point estimates rather than a continuous spatial estimate. To achieve this sparsity, we incorporate a model-order selection tool, i.e., the Bayesian information criterion (BIC) [37], [38], into IAA-APES. Let $\mathcal{P}$ denote a set containing the indices of the peaks selected from the IAA-APES spatial power spectrum estimate. Also, let $\mathcal{I}$ denote the set of the indices of the peaks selected by the BIC algorithm so far. The IAA-APES&BIC algorithm works as follows: first, the peak, from the set $\mathcal{P}$, giving the minimum BIC is selected. Then, the second peak, from the set $\mathcal{P}-\mathcal{I}$, which together with the first peak

TABLE II

THE IAA-APES ALGORITHM.

$\hat{P}_k = \frac{1}{(\mathbf{a}^H(\theta_k)\mathbf{a}(\theta_k))^2 N} \sum_{n=1}^{N} |\mathbf{a}^H(\theta_k)\mathbf{y}(n)|^2, \quad k = 1, \ldots, K$

repeat

$\mathbf{R} = \mathbf{A}(\boldsymbol{\theta})\hat{\mathbf{P}}\mathbf{A}^H(\boldsymbol{\theta})$

for $k = 1, \ldots, K$

$\hat{s}_k(n) = \dfrac{\mathbf{a}^H(\theta_k)\mathbf{R}^{-1}\mathbf{y}(n)}{\mathbf{a}^H(\theta_k)\mathbf{R}^{-1}\mathbf{a}(\theta_k)}, \; n = 1, \ldots, N$

$\hat{P}_k = \frac{1}{N} \sum_{n=1}^{N} |\hat{s}_k(n)|^2$

end for

until (convergence)

TABLE III

THE IAA-APES&BIC ALGORITHM.

$\mathcal{P}$: Set of peaks obtained from IAA-APES

$\mathcal{I} = \emptyset$, $\eta = 1$, quit $= 0$, $\text{BIC}^{\text{old}} = \infty$

repeat

$i' = \text{argmin}_{i \in \mathcal{P} - \mathcal{I}} \; \text{BIC}_i(\eta)$

if $\text{BIC}_{i'}(\eta) < \text{BIC}^{\text{old}}$

$\mathcal{I} = \{\mathcal{I}\,,\, i'\}$

$\text{BIC}^{\text{old}} = \text{BIC}_{i'}(\eta)$

$\eta = \eta + 1$

else quit $= 1$

until (quit $= 1$)

gives the minimum BIC, is selected, and so on, until the BIC value does not decrease anymore.[2] The IAA-APES&BIC algorithm is summarized in Table III. $\text{BIC}_i(\eta)$ is calculated as follows (see [38])

[2]Alternatively, the largest η peaks can be selected so that when the $(\eta+1)$st largest peak is added to $\mathcal{I}$, the BIC value does not decrease. This simplified version gives similar results to the one described above in our numerical examples.

$$\text{BIC}_i(\eta) = 2MN \ln \left(\sum_{n=1}^{N} \left\| \mathbf{y}(n) - \sum_{j \in \{\mathcal{I} \cup i\}} \mathbf{a}(\theta_j) \hat{\mathbf{s}}_j(n) \right\|_2^2 \right) + 3\eta \ln(2MN), \tag{8}$$

where $\eta = |\mathcal{I}| + 1$, $|\mathcal{I}|$ denotes the size of the set $\mathcal{I}$, i is the index of the current peak under consideration, and $\{\hat{\mathbf{s}}_j(n)\}_{n=1}^{N}$ is the IAA-APES signal waveform estimate corresponding to angle θ_j, $j \in \{\mathcal{I} \cup i\}$. Note that the second term on the right side of (8) does not matter to peak selection; it matters only when (8) is used to select the number of peaks to retain.

E. IAA-APES&RELAX

RELAX [39], [40] is a parametric cyclic algorithm that requires an estimate of the number of sources in the field. The results of IAA-APES&BIC can be used to provide a good initial estimate for the last step of RELAX. This approach is outlined in Table IV. The RELAX iterations can be terminated when the norm of the difference between two consecutive estimates is smaller than a certain threshold (5×10^{-4} for the examples considered herein). The maximization step of the algorithm can be implemented without much computational effort by means of a fine search only around the peak values of the IAA-APES&BIC result.[3] RELAX can be useful in estimating off-grid sources accurately and for further improving the IAA-APES waveform and angle estimates.

III. ACTIVE SENSING

In active sensing applications, besides the receiver, there are also one or more transmitters. This section first investigates the radar/sonar range-Doppler imaging problem for a SISO system. Then, the channel estimation problem for MISO communications will be discussed.

A. Range-Doppler Imaging

Pulse compression refers to the process of transmitting a modulated pulse and matched filtering the returned signal, which arrives at the antenna altered by complex coefficients that bear target information [48].

[3]The fine search can be implemented efficiently using the fast Fourier transform (FFT) [39], [40] or a derivative-free uphill search method such as the Nelder-Mead algorithm [46], [47] Note that the latter method is available in the MATLAB optimization toolbox with the name of "fminsearch."

TABLE IV

THE IAA-APES&RELAX ALGORITHM.

$\boldsymbol{\theta}'$: Locations of the peaks obtained from IAA-APES&BIC

K' : Number of peaks obtained from IAA-APES&BIC

$\{\hat{s}_k(n)\}$: Corresponding waveforms obtained from IAA-APES&BIC

repeat

for $k = 1, \ldots, K'$

$$\mathbf{y}_k(n) = \mathbf{y}(n) - \sum_{i=1, i\neq k}^{K'} \mathbf{a}(\theta_i')\hat{s}_i(n), \quad n = 1, \ldots, N$$

$$\hat{\theta}_k' = \text{argmax}_{\theta' \in \mathbb{R}} \sum_{n=1}^{N} |\mathbf{a}^H(\theta')\mathbf{y}_k(n)|^2$$

$$\hat{s}_k(n) = \tfrac{1}{M}\mathbf{a}^H(\hat{\theta}_k')\mathbf{y}_k(n), \quad n = 1, \ldots, N$$

end for

until (convergence)

---

1) Data Model: Consider a range-Doppler imaging radar/sonar with a transmitted pulse

$$\tilde{\mathbf{s}} = [\tilde{s}(1), \tilde{s}(2), \ldots, \tilde{s}(M)]^T, \tag{9}$$

where M is the pulse length. Let

$$\mathbf{s}(\omega_l) = \tilde{\mathbf{s}} \odot \mathbf{d}(\omega_l) \tag{10}$$

be the Doppler shifted signal, where

$$\mathbf{d}(\omega_l) = [1, e^{j\omega_l}, \cdots, e^{j(M-1)\omega_l}]^T, \quad l = 1, \ldots, L, \tag{11}$$

the Doppler interval of interest is divided into L bins, and the Doppler frequency for the lth Doppler bin is denoted as ω_l. Then, the M samples of the received signal that is temporally aligned with the return from the range bin of current interest r (see Figure 2) can be represented by

$$\mathbf{y}_r = \sum_{l=1}^{L} \alpha_{r,l}\mathbf{s}(\omega_l) + \sum_{\substack{m=-M+1 \\ m\neq 0}}^{M-1} \sum_{l=1}^{L} \alpha_{r+m,l}\mathbf{J}_m\mathbf{s}(\omega_l) + \mathbf{e}_r, \tag{12}$$

for $r = 1, \ldots, R$, where $\alpha_{r,l}$, which is proportional to the complex "voltage" radar-cross section (RCS) of the corresponding target, denotes the complex amplitude of the returned signal from the

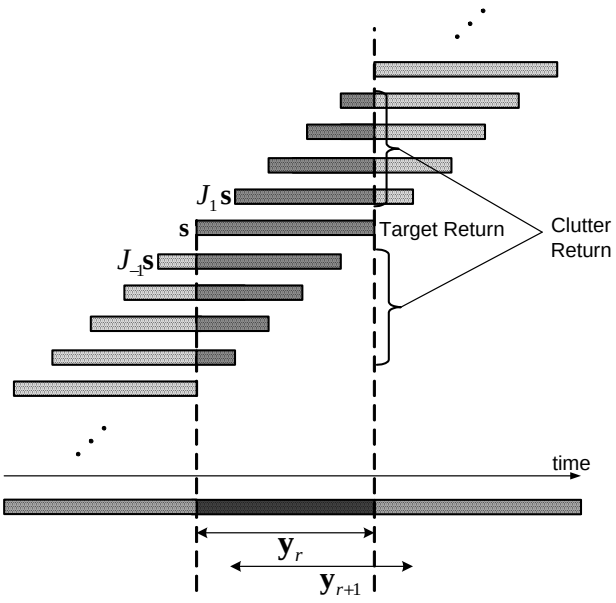


Fig. 2. Pulse compression for radar/sonar range-Doppler imaging.

range bin of current interest r and the lth Doppler bin, $\{\alpha_{m+r,l}\}$ denote the complex amplitudes of the returned signals from the adjacent range bins, and $\mathbf{e}_r$ denotes the noise. The reflections from nearby range bins are considered to be clutter. The $M \times M$ shift matrix $\mathbf{J}_m$ takes into account the fact that the clutter returns from adjacent range bins need different propagation times to reach the radar/sonar receiver:

$$\mathbf{J}_m = \begin{bmatrix} \longleftarrow m \longrightarrow & & & & \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & & \ddots & \vdots \\ 0 & & \ddots & & 1 \\ \vdots & \ddots & & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} = \mathbf{J}_{-m}^T, \tag{13}$$

for $m = 0, \ldots, M-1$. Eq. (12) can be written as

$$\mathbf{y}_r = \mathbf{S}\boldsymbol{\alpha}_r + \tilde{\mathbf{e}}_r, \tag{14}$$

where

$$\mathbf{S} = [\mathbf{s}(\omega_1), \mathbf{s}(\omega_2), \ldots, \mathbf{s}(\omega_L)], \tag{15}$$

$$\boldsymbol{\alpha}_r = [\alpha_{r,1}, \alpha_{r,2}, \ldots, \alpha_{r,L}]^T, \tag{16}$$

and

$$\tilde{\mathbf{e}}_r = \sum_{\substack{m=-M+1 \\ m\neq 0}}^{M-1} \sum_{l=1}^{L} \alpha_{r+m,l} \mathbf{J}_m \mathbf{s}(\omega_l) + \mathbf{e}_r. \tag{17}$$

Note that (12) is similar to the data model for passive sensing arrays, see (1), but with a single snapshot. The IAA-APES estimate at iteration number $i \in \{1, 2, \ldots\}$ becomes,

$$\hat{\alpha}_{r,l}^{(i)} = \frac{\mathbf{s}^H(\omega_l)\mathbf{R}_{(i-1)}^{-1}(r)\mathbf{y}_r}{\mathbf{s}^H(\omega_l)\mathbf{R}_{(i-1)}^{-1}(r)\mathbf{s}(\omega_l)}, \tag{18}$$

for $l = 1, \ldots, L$, $r = 1, \ldots, R$, where

$$\mathbf{R}_{(i-1)}(r) = \sum_{m=-M+1}^{M-1} \sum_{l=1}^{L} |\hat{\alpha}_{r+m,l}^{(i-1)}|^2 \mathbf{J}_m \mathbf{s}(\omega_l)\mathbf{s}^H(\omega_l)\mathbf{J}_m^T. \tag{19}$$

Here, IAA-APES is applied in a slightly different manner than in the passive array processing case. When $\boldsymbol{\alpha}_r$ is to be estimated, the previous values of $\{\boldsymbol{\alpha}_{r+m}\}_{m=-M+1}^{m=M-1}$ are used to estimate $\mathbf{R}(r)$. However, only $\boldsymbol{\alpha}_r$ is updated using $\mathbf{R}(r)$ and $\mathbf{y}_r$. When all $\{\boldsymbol{\alpha}_r\}_{r=1}^{R}$ are updated in this way, we advance to the next IAA-APES iteration. IAA-APES is initialized by matched filtering, matched to each range and Doppler bin, to obtain the initial $\{\alpha_{r,l}\}$.

2) Related Work: Matched filtering has been widely used for pulse compression since it gives optimal signal-power-to-output-power performance in the presence of a single target and white noise. However, in practical radar/sonar systems, matched filter performance is far from desirable since the pulse compression problem is usually clutter limited rather than noise limited (see, e.g., [49], [50]). Many data independent (see, e.g., [49]-[54]) and data-adaptive (see, e.g., [48], [55]) approaches have been proposed to achieve improved pulse compression.

Data independent approaches can be designed off-line and hence are convenient for real-time implementations in practical systems. Although receive filters based on data-independent instrumental variables (IV) can be used to achieve excellent pulse compression for the negligible Doppler case, their performance is unsatisfactory in the non-negligible Doppler case because of high sidelobe level problems [49]. On the other hand, data-adaptive approaches result in better performance, but at the cost of implementation complexity. Two important data-adaptive

methods for pulse compression are the adaptive pulse compression (APC) [48] and the Doppler-compensated APC (DC-APC) [55] algorithms. APC is an iterative minimum mean-squared error (MMSE) based data adaptive approach and DC-APC is the extension of APC to the non-negligible Doppler case. However, an IV approach [49] may be preferred to APC since the sidelobe level of the former method can be made arbitrarily low and the filter coefficients can be computed off-line, whereas APC updates the filter coefficients iteratively and adaptively. In the non-negligible Doppler case, however, adaptive approaches can perform much better. DC-APC is purported to work with at most one target per range bin, and the method used in [55] to estimate the Doppler values is not appealing from a performance viewpoint. On the other hand, IAA-APES can work with multiple Doppler targets located at the same range bin and it estimates the Doppler values in a robust manner. Moreover, APC and DC-APC require the tuning of hyperparameters, whereas IAA-APES is hyperparameter free. Some other differences are discussed below.

DC-APC assumes that there is at most one target per range bin. Let $\{\alpha_{r-M+1}, \ldots, \alpha_{r+M-1}\}$ and $\{\tilde{\omega}_{r-M+1}, \ldots, \tilde{\omega}_{r+M-1}\}$ denote the complex amplitudes and the Doppler frequencies, respectively, of the targets in the corresponding range bin r. DC-APC estimates the target parameters iteratively as follows:

$$\alpha_r^{(i)} = |\alpha_r^{(i-1)}|^2 \mathbf{s}^H(\tilde{\omega}_r)\mathbf{R}_{(i-1)}^{-1}(r)\mathbf{y}_r, \tag{20}$$

for $r = 1, \ldots, R$, where

$$\mathbf{R}_{(i-1)}(r) = \sum_{m=-M+1}^{M-1} |\alpha_{r+m}^{(i-1)}|^2 \mathbf{J}_m \mathbf{s}(\tilde{\omega}_{r+m})\mathbf{s}^H(\tilde{\omega}_{r+m})\mathbf{J}_m^T + \mathbf{Q}_r, \tag{21}$$

$\mathbf{Q}_r$ is the true noise covariance matrix of $\mathbf{e}_r$, which is assumed to be known, and $i \in \{1, 2, \ldots\}$ represents the current iteration number. The initial estimates of $\{\alpha_{r+m}\}_{m=-M+1}^{M-1}$ are obtained by using a standard matched filter that neglects the Doppler effect [55]. DC-APC also requires the estimates $\{\tilde{\omega}_{r+m}\}_{m=-M+1}^{M-1}$ of the target Doppler frequencies at each iteration. The Doppler frequency estimation approaches suggested in [55] are ad-hoc and not very accurate, especially for large Doppler shifts. This limitation of DC-APC, though, can be easily corrected by replacing $\tilde{\omega}_r$ in (20) with all possible Doppler frequencies $\{\omega_l\}$ to deal with multiple targets per range bin and to form range-Doppler images.

The DC-APC iterations defined in (20)-(21) result in the numerical ill-conditioning of $\mathbf{R}_{(i-1)}(r)$. To mitigate this problem, [48] suggests using $|\alpha_r^{(i-1)}|^\delta$ in (20) instead of $|\alpha_r^{(i-1)}|^2$ and σ^δ instead

of σ^2, where $0 \leq \delta \leq 2$ and the noise $\mathbf{e}_r$ is assumed to be white with a known variance σ^2. This approach requires the delicate tuning of δ (at each iteration), but [48], [55] do not provide a clear guideline on how to do this. We remark that if the noise covariance matrix $\mathbf{Q}_r$ in (21) is set to zero, then APC/DC-APC becomes *identical* to FOCUSS with the sparsity parameter $p = 0$ and the regularization parameter $\lambda = 0$ (see, e.g., (16) in [22]). In addition, with the introduction of δ, APC/DC-APC is still identical to FOCUSS, now with $1 - p/2 = \delta/2$ and $\lambda = \sigma^\delta$. We also note that both APC/DC-APC and FOCUSS are related to the approaches in [21], [30], [56]. It is interesting to note that many different ideas lead to the same result.

IAA-APES assumes that multiple targets can exist within the same range bin and $\{\alpha_{r,l}\}$ is calculated as in (18), where the term in the denominator can be viewed as the *current* estimate of $|\alpha_r^{(i-1)}|^2$ used in (20). This would be best estimate for $|\alpha_r^{(i-1)}|^2$ obtained by applying the minimum variance distortionless criterion, had the true covariance matrix been known [5]. One more advantage of IAA-APES is that the parameter estimates are inherently unbiased for the signal of interest, whereas this is not the case in DC-APC (or FOCUSS). Furthermore, in IAA-APES, the statistical properties of $\mathbf{e}_r$ are assumed unknown and taken into account implicitly as false targets in the current and adjacent range and Doppler bins; see (19). No matrix inversion problems were encountered with IAA-APES during our numerical simulations because of the unbiasedness property of IAA-APES.

B. Channel Estimation

The purpose of channel estimation in communications is to provide the subsequent equalizer with an accurate channel estimate so that the transmitted signals can be recovered successfully at the receiver side [57], [58]. SISO channel estimation for communications problem is similar to SISO radar/sonar range-Doppler imaging with the only difference being that the delay for the former is due to single-trip propagation, while the delay for the latter results from round-trip propagation.

1) Data Model: For the MISO channel estimation problem with I transmitters and a single receiver, let

$$\tilde{\mathbf{s}}_i = [\tilde{s}_i(1), \tilde{s}_i(2), \ldots, \tilde{s}_i(M)]^T, \quad i = 1, \ldots, I \tag{22}$$

denote the ith transmitted pulse and let L denote the total number of Doppler bins. Then, similar to (12), the received signal that is temporally aligned with the return from the rth tap,

$r = 1, \ldots, R$, with R denoting the total number of taps of the channel, can be written as:

$$\mathbf{y}_r = \sum_{i=1}^{I} \left\{ \sum_{l=1}^{L} \alpha_{r,l}(i)\mathbf{s}_i(\omega_l) + \sum_{\substack{m=-M+1 \\ m\neq 0}}^{M-1} \sum_{l=1}^{L} \alpha_{r+m,l}(i)\mathbf{J}_m\mathbf{s}_i(\omega_l) \right\} + \mathbf{e}_r, \tag{23}$$

where $\mathbf{J}_m$ was defined in (13),

$$\mathbf{s}_i(\omega_l) = \tilde{\mathbf{s}}_i \odot \mathbf{d}(\omega_l), \tag{24}$$

and $\mathbf{d}(\omega_l)$ was defined in (11). When determining the parameters of the ith channel $\{\alpha_{r,l}(i)\}$, the term $\sum_{l=1}^{L} \alpha_{r,l}(i)\mathbf{s}_i(\omega_l)$ is the signal term and all other terms in (23) are considered as clutter and noise. Consequently, IAA-APES can be applied directly to (23) to estimate $\{\alpha_{r,l}(i)\}$. Like in radar/sonar range compression problems, to determine $\{\alpha_{r,l}(i)\}$ with high accuracy, the transmitted pulses need to have both good auto and cross-correlation properties [59], [60].

2) Related Work: Most of the demodulation algorithms used in practice rely heavily on the accurate estimation of the channel impulse response. As in many other applications, sparse signal estimation approaches have also been proposed in this context. The main motivation for this is that underwater communications [61]-[63] and wireless channels are appropriately modelled as sparse channels consisting of only a few non-zero taps [64]. The existing approaches we will evaluate for sparse channel estimation, besides IAA-APES, include the matching pursuit (MP), orthogonal matching pursuit (OMP) [65]-[67], and least squares matching pursuit (LSMP) [68] algorithms, which have been used for sparse channel estimation and equalization in many applications [69]-[72]. It is difficult to determine the stopping criterion when using matching pursuit algorithms and user intervention is needed.

IV. Numerical Examples

We evaluate the performance of IAA-APES and compare it with various alternative methods in this section. We first focus on passive sensing applications and then shift our attention to active sensing applications.

A. Passive Sensing Examples

This subsection investigates the performance of IAA-APES, M-FOCUSS, M-SBL and ℓ_1-SVD for various passive sensing scenarios. Unless noted otherwise, M-FOCUSS is implemented by setting the sparsity parameter $p = 0.8$ and fine tuning the regularization parameter λ to get the

best results.[4] ℓ_1-SVD is implemented by assuming that the number of sources is known a priori and by fine tuning the hyperparameter. Since the fine tuning of the hyperparameters assume knowledge of the true source parameters, the so obtained results of M-FOCUSS and ℓ_1-SVD are impractical. M-SBL is implemented by using the alternative update method for the parameters as described in [19], [20], [23]. For all approaches considered, the scanning grid is uniform in the range from 1° to 180° with 1° increment between adjacent grid points, unless noted otherwise.

We consider a uniform linear array with $M = 12$ sensors and half-wavelength interelement spacing. The far-field narrowband signal waveforms and the additive noise signals are assumed to be circularly symmetric independent identically distributed (i.i.d.) complex Gaussian random processes with zero mean and variance σ^2, which is varied to obtain various SNR values. Furthermore, each signal waveform is normalized such that $\frac{1}{N}\sum_{n=1}^{N}|s_k(n)|^2 = P_k$, $k = 1, \ldots, K_0$, for a given P_k value, where K_0 denotes the true number of sources. SNR is defined as $10\log_{10}(P_k/\sigma^2)$, $k = 1, \ldots, K_0$, in decibels (dB), where σ^2 is the noise variance. According to the comments made in the Introduction, we consider very low snapshot cases, viz. $N = 3$ and $N = 1$.

First, we consider three uncorrelated sources at 60°, 82°, and 90° with 5 dB, 10 dB, and 10 dB powers, respectively, and with $N = 3$. The noise power is 0 dB, which results in a minimum SNR of 5 dB. Figure 3 shows the power and location estimates of the algorithms. The circles and the vertical dotted lines that align with these circles represent the true source locations and powers, and the results of 10 Monte-Carlo trials are shown in each plot. DAS clearly suffers from smearing and leakage. IAA-APES provides a much better result than DAS with low sidelobes and peaks at the true source locations. Moreover, IAA-APES&RELAX indicates the number of sources, their locations and powers accurately. As observed in Figure 3(d)-(f), the sparse algorithms encounter source splitting, location bias, and power underestimation problems. Also, it is hard to tell how many sources are present by solely using the spatial estimates of the sparse algorithms.

Next, Figure 4 considers three coherent sources at 60°, 80° and 90° with 10 dB power each and with $N = 3$. The source waveforms are assumed to be identical for all three coherent sources. The noise power is 0 dB, resulting in a 10 dB SNR. Similar to Figure 3, the circles and the

[4]See Section V.E in [22] Also, note that the M-FOCUSS algorithm used here is referred to as regularized M-FOCUSS in [22] since $\lambda \neq 0$. We refer to the approach as M-FOCUSS for simplicity.

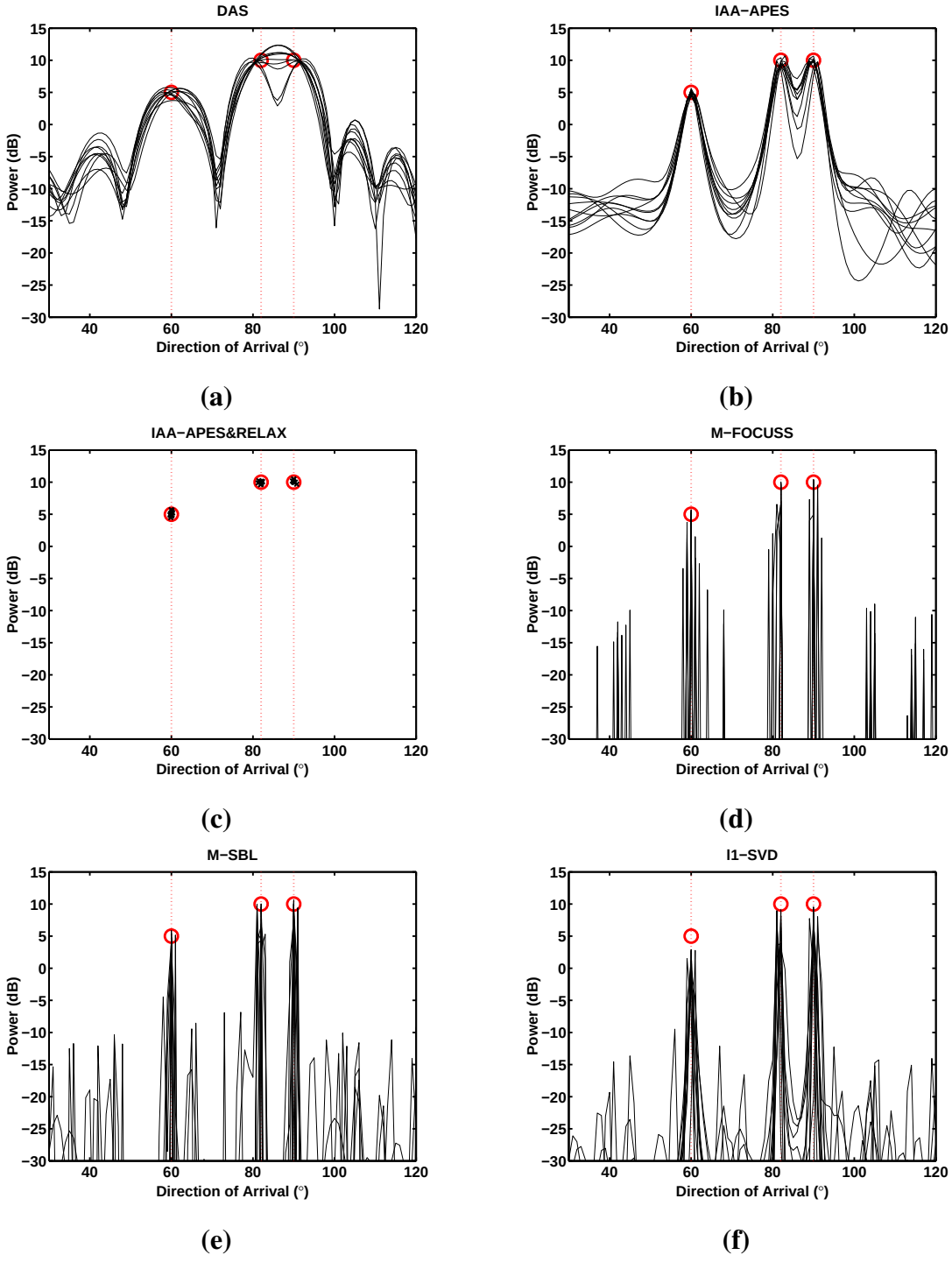


Fig. 3. Three uncorrelated sources at 60° (5 dB power), 82° (10 dB power), and 90° (10 dB power), as represented by the circles and vertical dotted lines in each plot. $N = 3$, the noise power is 0 dB (which results in a minimum SNR of 5 dB), and 10 Monte-Carlo trials are shown. (a) DAS spatial estimate, (b) IAA-APES spatial estimate, (c) IAA-APES&RELAX point source estimates, (d) M-FOCUSS spatial estimate, (e) M-SBL spatial estimate, and (f) ℓ_1-SVD spatial estimate.

vertical dotted lines represent the true source locations and the powers, and the results of 10 Monte-Carlo trials are shown in each plot. We observe that IAA-APES is able to resolve the sources successfully and IAA-APES&RELAX provides accurate point estimates. On the other hand, DAS fails to resolve the two closely spaced sources. The performances of the sparse algorithms are similar to those of the previous example.

Finally, Figure 5(a) compares the total mean-squared-error (MSE) (sum of each individual MSE) of the angle estimates of each algorithm with the Cramer-Rao bound (CRB) [5], and Figure 5(b) compares the total angle estimation bias (sum of each individual bias' modulus) of each algorithm for varying SNR. Two uncorrelated sources are placed at 77.51° and 90.51°, and N is set to 1. (The angle values are picked so that they are not on the size 1° grid used by IAA-APES.) To calculate the MSE and bias, we consider only the signals with the two largest powers as the estimated signals for FOCUSS and ℓ_1-SVD. We run the sparse algorithms with a fine grid of size 0.01° around the true source locations. The sparse algorithms are run with $K = 950$ scanning points, whereas IAA-APES is run with the same resolution as before, i.e., $K = 180$, and IAA-APES&RELAX is used to estimate the off-grid sources. M-SBL results are not shown because of the excessive computation time needed. Each point in Figure 5 is the average of 100 Monte-Carlo trials. We observe that IAA-APES&RELAX has both better variance and bias characteristics than the other methods for lower SNR.

Note that FOCUSS and ℓ_1-SVD both have two hyperparameters, and their selection affects the performances of these two algorithms significantly. Moreover, a different parameter setting should be used depending on the SNR, N, and the source structure, i.e., the number of sources, source spacing, source power levels, and correlation levels. We were able to tune the parameters relatively easily in our simulation scenarios, but when the problem dimensions are large and there is no prior knowledge of the scenario, it becomes difficult to find good hyperparameters. M-SBL, on the other hand, does not require any hyperparameters, but it takes the longest time to converge. These are the main reasons why BP or LASSO, which are the single-snapshot counterpart of ℓ_1-SVD, and SBL and FOCUSS, which are the single snapshot counterparts of M-FOCUSS and M-SBL, are not considered in the active sensing examples below.

Assuming $K \gg M$, the complexity of each IAA-APES iteration is $\mathcal{O}(M^2K)$. The complexity of the BIC extension is negligible compared to that of IAA-APES and the complexity of the RELAX extension depends on how many sources IAA-APES&BIC determines, the termination

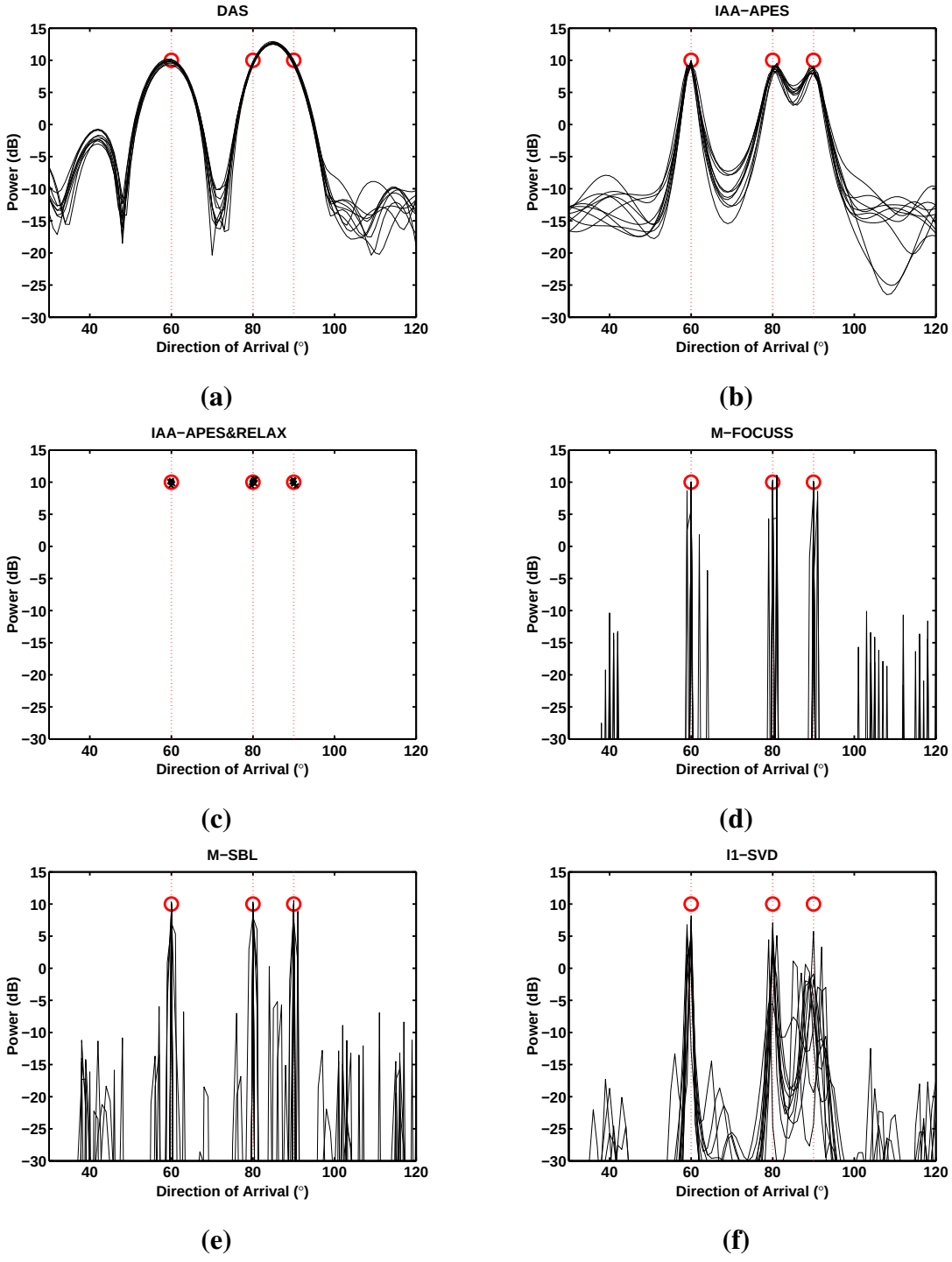
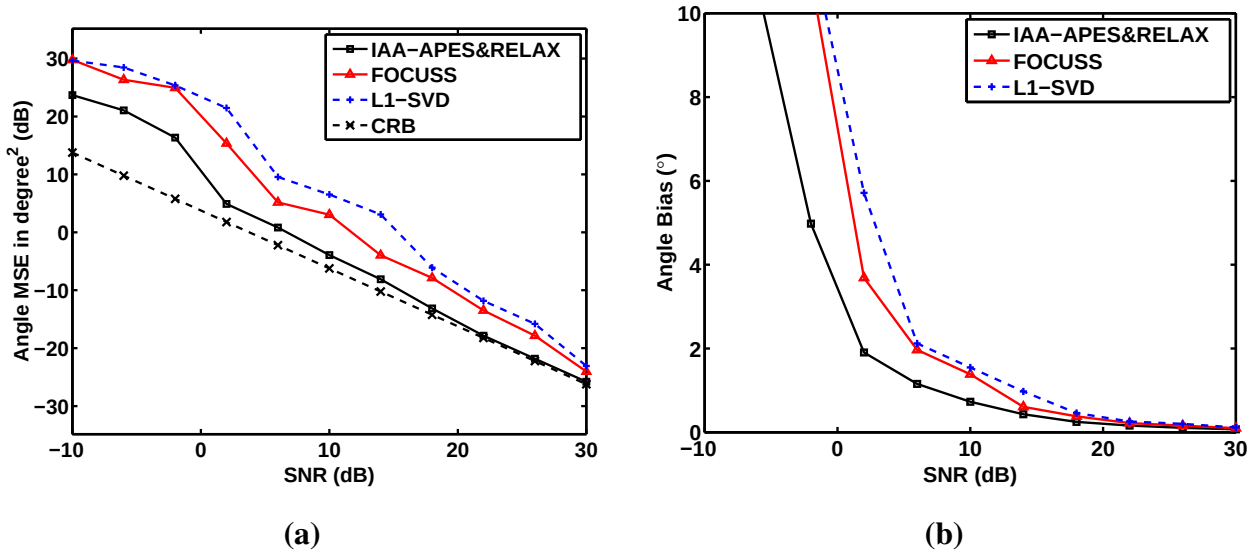


Fig. 4. Three coherent sources at 60°, 80°, and 90°, each with 10 dB power, as represented by the circles and vertical dotted lines in each plot. $N = 3$, the noise power is 0 dB (SNR=10 dB), and 10 Monte-Carlo trials are shown. (a) DAS spatial estimate, (b) IAA-APES spatial estimate, (c) IAA-APES&RELAX point source estimates, (d) M-FOCUSS spatial estimate, (e) M-SBL spatial estimate, and (f) ℓ_1-SVD spatial estimate.



(a) **(b)**

Fig. 5. Two uncorrelated sources at 77.51° and 90.51° with $N = 1$. (a) Total angle estimation MSE in dB together with the CRB, and the (b) Total angle estimation bias versus SNR. Each point is the average of 100 Monte-Carlo trials. Recall that FOCUSS is the single snapshot version of M-FOCUSS.

condition, and the method chosen for the maximization step. In our experiments, RELAX usually converged faster than the time needed for IAA-APES to converge. The complexities of M-FOCUSS and M-SBL are also $\mathcal{O}(M^2K)$ per iteration [23]. The complexity of ℓ_1-SVD, on the other hand, is $\mathcal{O}(K^3K_{\text{svd}}^3)$ [25] where K_{svd} is the estimate of the number of sources. In our simulations, IAA-APES was always faster than M-FOCUSS, M-SBL and ℓ_1-SVD, especially when the SNR was low. Note that the actual number of iterations required for convergence depends heavily on the specific scenario. For example, for the scenario considered in Figure 5 at SNR = 2 dB, the times required by FOCUSS and ℓ_1-SVD are 3.9 and 5.6 seconds, respectively, whereas IAA-APES&RELAX requires 0.1 seconds on average.[5]

B. Active Sensing Examples

This section considers examples for the SISO radar/sonar and the MISO channel estimation problems.

[5]The timing values in seconds are given as an example. The convergence times may vary depending on many factors such as how the algorithms are implemented, the specific hardware, etc. ℓ_1-SVD has been implemented using the software described in [73] and [44].

1) SISO Range-Doppler Imaging Example: We now evaluate the performance of IAA-APES for SISO range-Doppler imaging and compare it with that of the matched filtering, MP, OMP and LSMP. We use a 30-element P3 code for the transmitted pulse, i.e., $\tilde{s}(m) = e^{j(m-1)^2\pi/M}$, $m = 1, \ldots, M$, where $M = 30$ [55], [74]. The Doppler shift is expressed as $\Phi_l = \omega_l M(180^\circ/\pi)$, $l = 1, \ldots, L$, which is the total phase shift (in degrees) for the transmitted pulse duration.

We consider three moving targets with 5 dB power and Doppler shifts of -30°, -25° and 15° and six moving targets with 25 dB power and Doppler shifts of -70°, -55°, -10°, 10°, 20° and 60°, as shown in Figure 6(a). The background noise is assumed to be a circularly symmetric i.i.d. complex Gaussian random process with mean zero and a variance of 0 dB, the number of range bins is set to 100, and the number of Doppler bins is set to 37. The resulting minimum SNR is 5 dB. Figure 6 shows that the matched filter smears the targets in both the Doppler and range domains significantly. Figures 6(c)-(e) show that MP, OMP and LSMP results are not satisfactory in this example. The MP, OMP and LSMP algorithms were all terminated manually to give the best performance.[6] Figure 6(f) shows that IAA-APES provides a much more useful result, where the target locations and Doppler frequencies are easily observable, than the matched filter. IAA-APES&BIC can estimate the strong targets accurately and the weak targets reasonably well, as shown in Figure 6(g). IAA-APES&RELAX helps improve the results of IAA-APES&BIC even further, as shown in Figure 6(h). In Figures 6(c)-(e) and Figures 6(g)-(h), the cross marks represent the targets selected by the algorithms and the circles represent the ground truth. The numbers shown in Figures 6(g)-(h) are the power estimates obtained by the corresponding algorithms. Recall that BIC selects only the dominant components of the IAA-APES spatial estimate, but does not modify either the range-Doppler values or the power levels estimated by IAA-APES, whereas RELAX improves upon the IAA-APES estimates.

2) Channel Estimation: The performance of MP, OMP, LSMP and IAA-APES will also be investigated for the channel estimation problem encountered in communications. We consider six transmitters, viz. $I = 6$, and a single receiver. The probing pulses used in the simulations are obtained by using a cyclic algorithm to ensure both good auto and cross-correlation properties; see [59], [60] for details. The sampling frequency is assumed to be 24 kHz, M is chosen to be

[6]See [70] for a complexity analysis of MP, OMP and LSMP. In our examples, MP and OMP took less time than IAA-APES, whereas LSMP took longer than IAA-APES to converge.

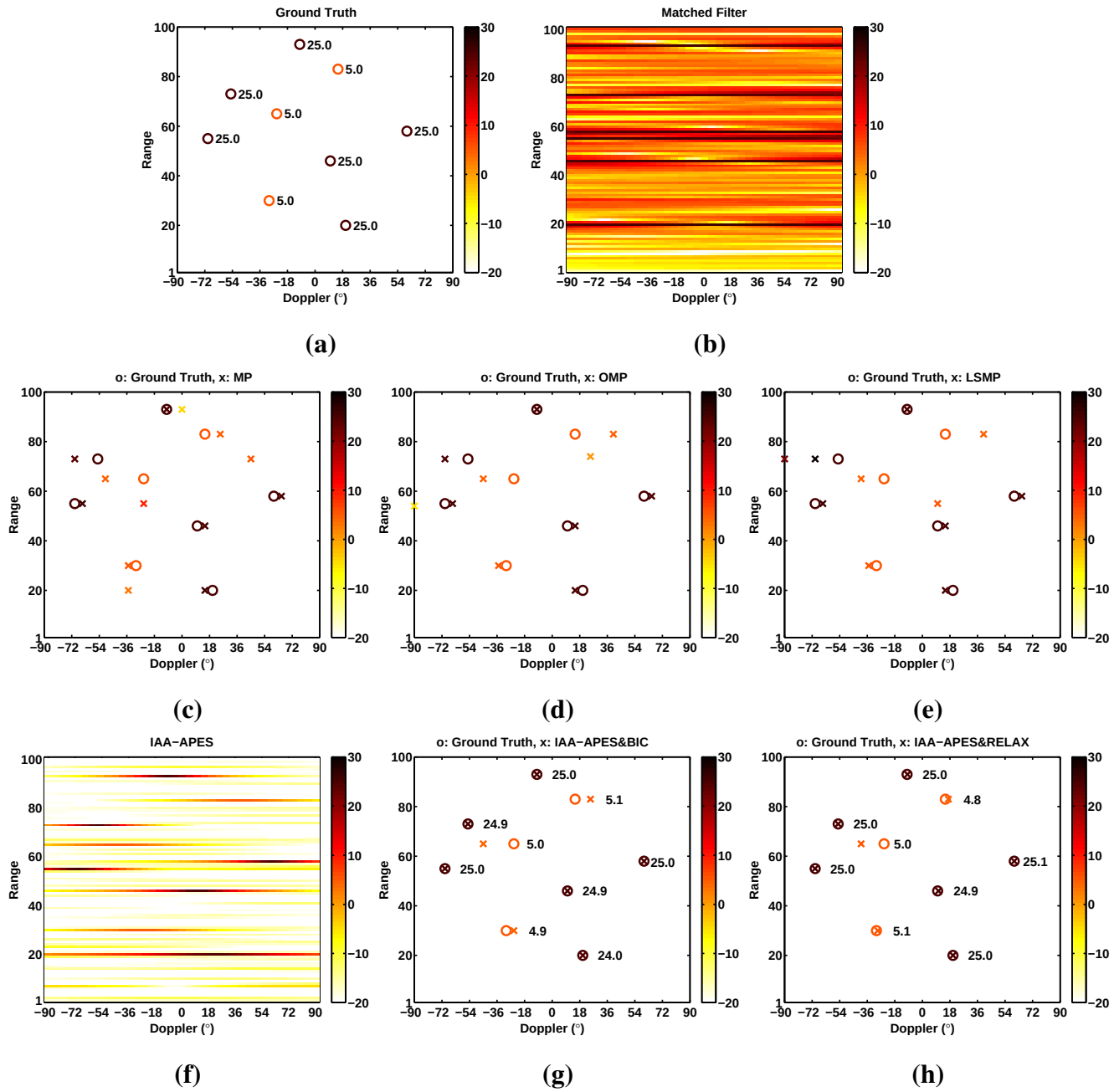


Fig. 6. SISO range-Doppler imaging with three 5 dB and six 25 dB targets, indicated by circles. The noise power is 0 dB, resulting in a minimum SNR of 5 dB. (a) The ground truth with power levels, (b) Matched filter, (c) MP and the ground truth, (d) OMP and the ground truth, (e) LSMP and the ground truth, (f) IAA-APES, (g) IAA-APES&BIC and the ground truth, and (h) IAA-APES&RELAX and the ground truth. All power levels are in dB.

1024 (which yields a 42.7 ms pulse duratio), the number of delay taps are set at 80 to yield a 3.3 ms of delay span and 13 Doppler bins are used to cover the range from -30 Hz to 30 Hz with 5° resolution. The six channels are assumed to be independent of each other and are simulated as circularly symmetric i.i.d. complex Gaussian random variables with mean zero and a variance decreasing exponentially with increasing delay, specifically $\sigma_k^2 = e^{-0.03(k-1)}$, where σ_k^2 is the variance of the channel tap with delay index k, $k = 1, \ldots, 80$. The noise is simulated as a circularly symmetric i.i.d. complex Gaussian random process with mean zero and a -20 dB variance. The locations of the non-zero taps of the channels are also simulated randomly and by trying to mimic practical channel functions encountered in underwater communications, see, e.g., [69], [70]. The power levels of the channel taps were in the range from -30 dB to 0 dB, which results in a minimum SNR of -10 dB. Figure 7 illustrates the performances of the algorithms for the first channel; IAA-APES&BIC provides the best result. The results for the other channels, which are not shown for conciseness, are similar to that for the first channel.

V. CONCLUSIONS

This paper has presented an iterative adaptive approach for amplitude and phase estimation (IAA-APES) in array processing applications. IAA-APES is a nonparametric, hyperparameter free algorithm that is designed to work under severe snapshot limitations and for uncorrelated, partially correlated, and coherent sources, as well as for arrays with arbitrary geometries. Because of the similarities between many active sensing applications and passive array processing, IAA-APES can be applied to these cases as well without any essential modifications. The Bayesian information criterion (BIC) can be used in conjunction with the IAA-APES algorithm to yield sparse solutions, which are desirable in many applications. Furthermore, the application of the parametric RELAX algorithm to the IAA-APES&BIC results provides further performance improvement. Simulations showed that IAA-APES and its variations outperform the most prominent methods in the literature in the corresponding application areas. IAA-APES is believed to be a viable candidate for practical applications, since it does not require any hyperparameters, has a simple formulation, provides superresolution, facilitates parallel processing, and shows good performance.

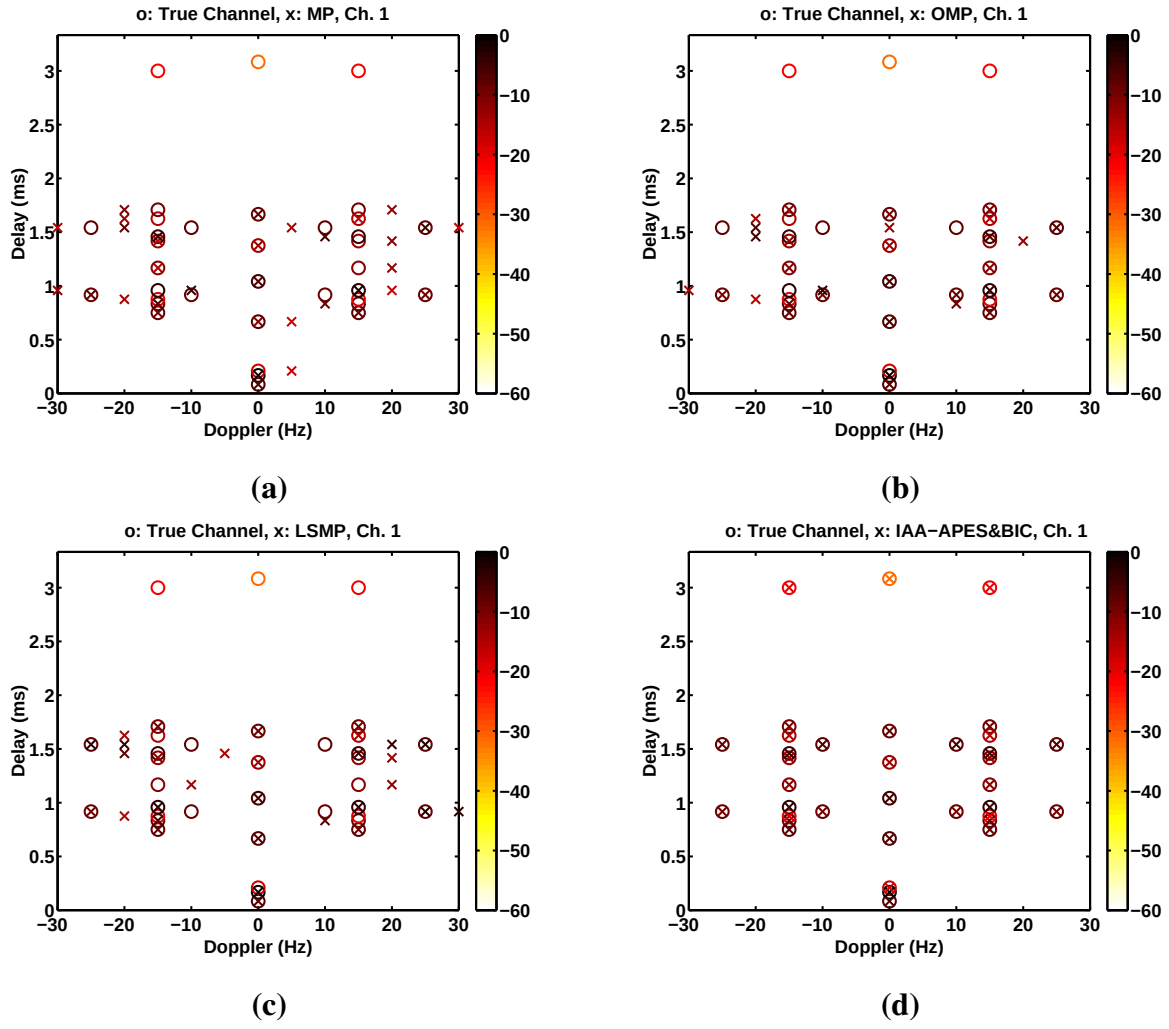


Fig. 7. MISO channel estimation with six transmitters and one receiver. Channel taps are indicated by circles. The channel tap power levels are in the range of -30 dB to 0 dB, and the noise level is -20 dB, which results in a -10 dB minimum SNR. Estimates for the first channel via (a) MP, (b) OMP, (c) LSMP, and (d) IAA-APES&BIC. All power levels are in dB.

APPENDIX A
AN APPROXIMATE ML INTERPRETATION OF IAA-APES

We derive a locally convergent ML-based iterative adaptive approach called IAA-ML. IAA-ML is similar to IAA-APES in nature, and we show that IAA-APES can be obtained as a simplified version of IAA-ML.

The negative log-likelihood function of $\{\mathbf{y}(n)\}_{n=1}^{N}$ can be represented as

$$\ln|\mathbf{R}| + \mathrm{tr}(\mathbf{R}^{-1}\mathbf{\Gamma}), \tag{25}$$

where $\mathbf{\Gamma} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{y}(n)\mathbf{y}^H(n)$ is the so-called sample covariance matrix. It is assumed that the received signal is a multivariate complex Gaussian random vector with zero-mean and covariance matrix $\mathbf{R}$ and that the snapshots are i.i.d. [5]. Minimizing (25) is equivalent to minimizing the Stein's loss, $-\ln|\mathbf{R}^{-1}\mathbf{\Gamma}| + \mathrm{tr}(\mathbf{R}^{-1}\mathbf{\Gamma})$ [75], [76], which stresses the covariance fitting character of the cost function, but without the i.i.d. Gaussian assumptions needed by the ML loss function.

Assume that $\mathbf{Q}(\theta_k)$ is known and that the signal power at θ_k is to be estimated. Using the fact that $|\mathbf{I}+\mathbf{AB}| = |\mathbf{I}+\mathbf{BA}|$ and the matrix inversion lemma together with (4), it can be shown that minimizing (25) with respect to P_k is equivalent to minimizing

$$f(P_k) \triangleq \ln(1+P_k\mathbf{a}_k^H\mathbf{Q}_k^{-1}\mathbf{a}_k) - \frac{P_k\mathbf{a}_k^H\mathbf{Q}_k^{-1}\mathbf{\Gamma}\mathbf{Q}_k^{-1}\mathbf{a}_k}{1+P_k\mathbf{a}_k^H\mathbf{Q}_k^{-1}\mathbf{a}_k}, \tag{26}$$

where $\mathbf{Q}(\theta_k)$ and $\mathbf{a}(\theta_k)$ have been replaced by $\mathbf{Q}_k$ and $\mathbf{a}_k$, respectively, for notational simplicity. Setting the first derivative of (26) with respect to P_k to zero, i.e., $f'(\tilde{P}_k) = 0$, gives

$$\tilde{P}_k = \frac{\mathbf{a}_k^H\mathbf{Q}_k^{-1}(\mathbf{\Gamma}-\mathbf{Q}_k)\mathbf{Q}_k^{-1}\mathbf{a}_k}{(\mathbf{a}_k^H\mathbf{Q}_k^{-1}\mathbf{a}_k)^2}. \tag{27}$$

The second derivative of (26) with respect to P_k is

$$f''(P_k) = \frac{-(\mathbf{a}_k^H\mathbf{Q}_k^{-1}\mathbf{a}_k)^2 + 2\frac{\mathbf{a}_k^H\mathbf{Q}_k^{-1}\mathbf{\Gamma}\mathbf{Q}_k^{-1}\mathbf{a}_k}{(1+P_k\mathbf{a}_k^H\mathbf{Q}_k^{-1}\mathbf{a}_k)}\mathbf{a}_k^H\mathbf{Q}_k^{-1}\mathbf{a}_k}{(1+P_k\mathbf{a}_k^H\mathbf{Q}_k^{-1}\mathbf{a}_k)^2},$$

and hence $f''(\tilde{P}_k)$ is

$$f''(\tilde{P}_k) = \frac{(\mathbf{a}_k^H\mathbf{Q}_k^{-1}\mathbf{a}_k)^2}{(1+\hat{P}_k\mathbf{a}_k^H\mathbf{Q}_k^{-1}\mathbf{a}_k)^2}, \tag{28}$$

which is strictly positive. This means that $\tilde{P}_k$ is the unique minimizer of $f(P_k)$. In principle, $\tilde{P}_k$ may be negative. However, since P_k represents power, it should be nonnegative. The minimizer of $f(P_k)$ subject to the constraint $P_k \geq 0$ is

$$\hat{P}_k = \max(0, \tilde{P}_k), \tag{29}$$

since $\tilde{P}_k$ is the unique minimizer of $f(P_k)$ (when $\tilde{P}_k < 0$, $f(0) \leq f(P_k)$ for $\forall P_k \geq 0$).

By using the matrix inversion lemma in (27) to replace $\mathbf{Q}_k$ by $\mathbf{R}$, we get

$$\hat{P}_k = \max\left(0, P_k + \frac{\mathbf{a}_k^H\mathbf{R}^{-1}(\mathbf{\Gamma}-\mathbf{R})\mathbf{R}^{-1}\mathbf{a}_k}{(\mathbf{a}_k^H\mathbf{R}^{-1}\mathbf{a}_k)^2}\right). \tag{30}$$

Computing $\hat{P}_k$ requires knowledge of P_k and $\mathbf{R}$ (recall that $\mathbf{R} = \mathbf{A}(\boldsymbol{\theta})\mathbf{P}\mathbf{A}^H(\boldsymbol{\theta})$, where $\mathbf{P}$ is a diagonal matrix with $\{P_k\}$ on its diagonal). Therefore, the algorithm must be implemented iteratively; the initialization is done with DAS. IAA-ML will be locally convergent if $\mathbf{R}$ is

recalculated after each P_k is updated because of the cyclical maximization of the likelihood function.

Note that (30) can be written as (assuming $\tilde{P}_k \geq 0$)

$$\hat{P}_k = \frac{\mathbf{a}_k^H \mathbf{R}^{-1} \mathbf{\Gamma} \mathbf{R}^{-1} \mathbf{a}_k}{(\mathbf{a}_k^H \mathbf{R}^{-1} \mathbf{a}_k)^2} + P_k - \frac{1}{(\mathbf{a}_k^H \mathbf{R}^{-1} \mathbf{a}_k)}. \tag{31}$$

By the properties of SCB, $P_k \approx 1/(\mathbf{a}_k^H \mathbf{R}^{-1} \mathbf{a}_k)$ [6], [43]. Hence, an approximate solution of (31) is

$$\hat{P}_k = \frac{\mathbf{a}_k^H \mathbf{R}^{-1} \mathbf{\Gamma} \mathbf{R}^{-1} \mathbf{a}_k}{(\mathbf{a}_k^H \mathbf{R}^{-1} \mathbf{a}_k)^2}, \quad k = 1, \ldots, K. \tag{32}$$

Iterating this equation by building $\mathbf{R}$ from the latest estimate of $\{P_k\}$, we get IAA-APES. This approximation has two advantages. First, the $\hat{P}_k$ in (32) is guaranteed to be nonnegative and it alleviates the need for the procedure in (29). Secondly, when $\{P_k\}$ are accurate, the difference $P_k - 1/(\mathbf{a}_k^H \mathbf{R}^{-1} \mathbf{a}_k)$ in (31) is small, as discussed above. When $\{P_k\}$ are inaccurate, however, this difference may not be small, and forcing the difference to zero may provide a better estimate of P_k. (IAA-ML tends to work well when the snapshot number is large. It is not considered in the numerical examples because of our focus on cases with few snapshots.) Since IAA-APES can be obtained as a close approximation to IAA-ML, which is locally convergent, IAA-APES is expected to also enjoy local convergence. We have never come across an example where IAA-APES did not converge; however, the search for a convergence proof could be a useful direction for future work.

ACKNOWLEDGMENTS

We thank Professor Aaron Lanterman and the reviewers for their useful comments on and the revisions of a former version of this paper.

REFERENCES

[1] J. Li and P. Stoica, Eds., *Robust Adaptive Beamforming*. New York, NY: John Wiley & Sons, 2005.

[2] J. Capon, "High resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, pp. 1408–1418, August 1969.

[3] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. AP-34, no. 3, pp. 276–280, March 1986.

[4] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-37, no. 5, pp. 720–741, May 1989.

[5] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. New York, NY: John Wiley & Sons, 2002.

[6] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Transactions on Signal Processing*, vol. 51, no. 7, pp. 1702–1715, July 2003.

[7] A. L. Swindlehurst and T. Kailath, "A performance analysis of subspace-based methods in the presence of model errors, Part I: The MUSIC algorithm," *IEEE Transactions on Signal Processing*, vol. 40, no. 7, pp. 1758–1773, July 1992.

[8] P. Stoica, Z. Wang, and J. Li, "Extended derivations of MUSIC in the presence of steering vector errors," *IEEE Transactions on Signal Processing*, vol. 53, no. 3, pp. 1209–1211, March 2005.

[9] J. Li, P. Stoica, and Z. Wang, "Doubly constrained robust Capon beamformer," *IEEE Transactions on Signal Processing*, vol. 52, no. 9, pp. 2407–2423, September 2004.

[10] A. B. Baggeroer and H. Cox, "Passive sonar limits upon nulling multiple moving ships with large aperture arrays," *33th Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 103–108, 1999.

[11] A. L. Kraay and A. B. Baggeroer, "A physically constrained maximum-likelihood method for snapshot-deficient adaptive array processing," *IEEE Transactions on Signal Processing*, vol. 55, pp. 4048–4063, 2007.

[12] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization," *Proceedings of the National Academy of Sciences*, vol. 100, no. 5, pp. 2197–2202, March 2003.

[13] J. A. Tropp, "Just relax: Convex programming methods for identifying sparse signals," *IEEE Transactions on Information Theory*, vol. 51, no. 3, pp. 1030–1051, March 2006.

[14] J. J. Fuchs, "On sparse representations in arbitrary redundant bases," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 1341–1344, 2004.

[15] ——, "Recovery of exact sparse representations in the presence of bounded noise," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3601–3608, 2005.

[16] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society*, vol. 58, no. 1, pp. 267–288, 1996.

[17] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

[18] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 600–616, 1997.

[19] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, pp. 211–244, 2001.

[20] D. P. Wipf and B. D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Transactions on Signal Processing*, vol. 52, no. 8, pp. 2153–2164, 2004.

[21] M. A. T. Figueiredo, "Adaptive sparseness for supervised learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1150–1159, 2003.

[22] S. F. Cotter, B. D. Rao, E. Kjersti, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2477–2488, 2005.

[23] D. P. Wipf and B. D. Rao, "An empirical Bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3704–3716, 2007.

[24] D. M. Malioutov, "A sparse signal reconstruction perspective for source localization with sensor arrays," Master's thesis, MIT, July 2003.

[25] D. M. Malioutov, M. Çetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 3010–3022, 2005.

[26] D. Model and M. Zibulevsky, "Signal reconstruction in sensor arrays using sparse representations," *Signal Processing*, vol. 86, no. 3, pp. 624–638, 2006.

[27] J. J. Fuchs, "Linear programming in spectral estimation. Application to array processing," in *IEEE Proceedings of the Acoustics, Speech, and Signal Processing*, 1996, pp. 3161–3164.

[28] ——, "On the application of the global matched filter to DOA estimation with uniform circular arrays," *IEEE Transactions on Signal Processing*, vol. 49, no. 4, pp. 702–709, 2001.

[29] T. Yardibi, J. Li, P. Stoica, and L. N. Cattafesta, "Sparsity constrained deconvolution approaches for acoustic source mapping," *Journal of the Acoustical Society of America*, vol. 123, no. 5, pp. 2631–2642, May 2008.

[30] M. D. Sacchi, T. J. Ulrych, and C. J. Walker, "Interpolation and extrapolation using a high-resolution discrete fourier transform," *IEEE Transactions on Signal Processing*, vol. 46, no. 1, pp. 31–38, January 1998.

[31] B. Jeffs, "Sparse inverse solution methods for signal and image processing applications," *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 3, pp. 1885–1888, May 1998.

[32] M. Ting, R. Raich, and A. Hero, "Sparse image reconstruction using sparse priors," *IEEE International Conference on Image Processing,* Atlanta, GA, October 2006.

[33] R. Raich and A. Hero, "Sparse image reconstruction for partially known blur functions," *IEEE International Conference on Image Processing,* Atlanta, GA, October 2006.

[34] J. Li and P. Stoica, "An adaptive filtering approach to spectral estimation and SAR imaging," *IEEE Transactions on Signal Processing*, vol. 44, no. 6, pp. 1469–1484, June 1996.

[35] P. Stoica, H. Li, and J. Li, "A new derivation of the APES filter," *IEEE Signal Processing Letters*, vol. 6, no. 8, pp. 205–206, August 1999.

[36] P. Stoica, A. Jakobsson, and J. Li, "Capon, APES and matched-filterbank spectral estimation," *Signal Processing*, vol. 66, no. 1, pp. 45–59, April 1998.

[37] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist.*, vol. 6, no. 2, pp. 461–464, 1978.

[38] P. Stoica and Y. Selén, "Model-order selection: A review of information criterion rules," *IEEE Signal Processing Magazine*, vol. 21, no. 4, pp. 36–47, July 2004.

[39] J. Li and P. Stoica, "Efficient mixed-spectrum estimation with applications to target feature extraction," *IEEE Transactions on Signal Processing*, vol. 44, pp. 281–295, February 1996.

[40] J. Li, P. Stoica, and D. Zheng, "Angle and waveform estimation via RELAX," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, pp. 1077–1087, July 1997.

[41] W. M. Humphreys, Jr., T. F. Brooks, W. W. Hunter, Jr., and K. R. Meadows, "Design and use of microphone directional arrays for aeroacoustic measurements," *AIAA Paper 98-0471, AIAA, 36th Aerospace Sciences Meeting and Exhibit, Reno, NV*, January 1998.

[42] A. B. Baggeroer, "Sonar arrays and array processing," *AIP Conf. Proc., Review of progress in quantitative nondestructive evaluation*, vol. 1, pp. 3–24, April 2005.

[43] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ: Prentice-Hall, 2005.

[44] Y. Labit, D. Peaucelle, and D. Henrion, "SEDUMI interface 1.02 - A tool for solving LMI problems with SEDUMI," *IEEE International Symposium on Computer, Glasgow, U.K.*, pp. 272–277, September 2002.

[45] T. F. Brooks and W. M. Humphreys, "A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays," *Journal of Sound and Vibration*, vol. 294, pp. 856–879, Jul. 2006.

[46] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Computer Journal*, vol. 7, pp. 308–313, 1965.

[47] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the Nelder-Mead simplex method in low dimensions," *SIAM Journal of Optimization*, vol. 9, no. 1, pp. 112–147, 1998.

[48] S. D. Blunt and K. Gerlach, "Adaptive pulse compression via MMSE estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 2, pp. 572–584, April 2006.

[49] P. Stoica, J. Li, and M. Xue, "Transmit codes and receive filters for pulse compression radar systems," accepted to *IEEE Signal Processing Magazine*, November 2008.

[50] ——, "On sequences with good correlation properties: A new perspective," *2007 IEEE Information Theory Workshop on Information Theory for Wireless Networks,* Bergen, Norway, July 2007.

[51] C. Stutt and L. Spafford, "A "best" mismatched filter response for radar clutter discrimination," *IEEE Transactions on Information Theory*, vol. 14, no. 2, pp. 280–287, March 1968.

[52] M. H. Ackroyd and F. Ghani, "Optimum mismatched filters for sidelobe suppression," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 9, no. 2, pp. 214–218, March 1973.

[53] C. Nunn, "Constrained optimization applied to pulse compression codes, and filters," *IEEE International Radar Conference,* Arlington, VA, USA, pp. 190–194, 9-12 May 2005.

[54] N. Levanon, "Cross-correlation of long binary signals with longer mismatched filters," *IEE Proceedings - Radar, Sonar, and Navigation*, vol. 152, no. 6, pp. 377–382, December 2005.

[55] S. D. Blunt, K. J. Smith, and K. Gerlach, "Doppler-compensated adaptive pulse compression," *IEEE Conference on Radar,* Verona, NY, USA, pp. 114–119, 24-27 April 2006.

[56] J. J. Fuchs, "Convergence of a sparse representations algorithm applicable to real or complex data," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 598–605, 2007.

[57] J. C. Preisig and G. B. Deane, "Surface wave focusing and acoustic communications in the surf zone," *Journal of the Acoustical Society of America*, vol. 116, no. 4, pp. 2067–2080, October 2004.

[58] J. C. Preisig, "Performance analysis of adaptive equalization for coherent acoustic communications in the time-varying ocean environment," *Journal of the Acoustical Society of America*, vol. 118, no. 1, pp. 263–278, July 2005.

[59] J. Li, X. Zheng, and P. Stoica, "MIMO SAR imaging: Signal synthesis and receiver design," *The 2nd International Workshop on Computational Advances in Multi-Sensor Adaptive Processing* St Thomas, US Virgin Islands, 2007.

[60] J. Li, P. Stoica, and X. Zheng, "Signal synthesis and receiver design for MIMO radar imaging," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3959–3968, August 2008.

[61] M. Kocic, D. Brady, and M. Stojanovic, "Sparse equalization for real-time digital underwater acoustic communications," *IEEE/MTS Oceans Conference*, vol. 3, pp. 1417–1422, October 1995.

[62] M. Stojanovic, "Recent advances in high-speed underwater acoustic communications," *IEEE Journal of Oceanic Engineering*, vol. 21, no. 2, pp. 125–136, April 1996.

[63] M. Stojanovic, L. Freitag, and M. Johnson, "Channel-estimation-based adaptive equalization of underwater acoustic signals," *IEEE/MTS Oceans Conference*, vol. 2, pp. 985–990, September 1999.

[64] C. Carbonelli, S. Vedantam, and U. Mitra, "Sparse channel estimation with zero tap detection," *IEEE Transactions on Wireless Communications*, vol. 6, no. 5, pp. 1743–1753, 2007.

[65] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.

[66] B. K. Natarajan, "Sparse approximation solutions to linear systems," *SIAM Journal of Computing*, vol. 24, pp. 227–234, 1995.

[67] S. F. Cotter, R. Adler, R. D. Rao, and K. Kreutz-Delgado, "Forward sequential algorithms for best basis selection," *IEE Proceedings of Vision, Image and Signal Processing*, vol. 146, no. 5, pp. 235–244, October 1999.

[68] W. Li, "Estimation and tracking of rapidly time-varying broadband acoustic communication channels," *PhD thesis*, Massachusetts Institute of Technology, Cambridge, MA, 2005.

[69] W. Li and J. C. Preisig, "Estimation and equalization of rapidly varying sparse acoustic communication channels," *IEEE/MTS Oceans Conference*, pp. 1–6, September 2006.

[70] ——, "Estimation of rapidly time-varying sparse channels," *IEEE Journal of Oceanic Engineering*, vol. 32, no. 4, pp. 927–939, 2007.

[71] S. F. Cotter and B. D. Rao, "The adaptive matching pursuit algorithm for estimation and equalization of sparse time-varying channels," *25th Asilomar Conference on Signals, Systems and Computers,* Pacific Grove, CA, vol. 2, pp. 1772–1776, October 2000.

[72] ——, "Sparse channel estimation via matching pursuit with application to equalization," *IEEE Transactions on Communications*, vol. 50, no. 3, pp. 374–377, March 2002.

[73] J. Löfberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," *The 2004 IEEE International Symposium on Computer Aided Control Systems Design*, pp. 284–289, September 2004.

[74] J. G. Proakis, *Digital Communications*. McGraw-Hill Inc., fourth edition, 2001.

[75] W. James and C. Stein, "Estimation with quadratic loss," *In Proc. Fourth Berkeley Symp. Math. Statist. Prob., University of California Press*, vol. 1, pp. 361–380, 1961.

[76] D. K. Dey and C. Srinivasan, "Estimation of a covariance matrix under Stein's loss," *The Annals of Statistics*, vol. 13, no. 4, pp. 1581–1591, December 1985.