Circuits Syst Signal Process (2015) 34:2363–2379
DOI 10.1007/s00034-014-9966-3

Covariance-Based DoA Estimation in a Krylov Subspace

Tadeu N. Ferreira · Marcello L. R. de Campos · Sergio L. Netto

Received: 4 August 2014 / Revised: 26 December 2014 / Accepted: 27 December 2014 /
Published online: 8 January 2015

Abstract Covariance-based DoA estimation (CB-DoA) algorithms represent lower computational complexity alternatives to the traditional ESPRIT approach. This paper investigates CB-DoA using Krylov-subspace techniques (including Arnoldi's and Lanczos' updates) with respect to the resulting computational cost and estimation error performance. The proposed modifications also allow an automatic estimation of the number of sources. Computational analyses performed for the resulting CB-DoA algorithm indicate cost savings above 60 % in comparison with the standard CB-DoA implementation, which already represents a 20 % improvement upon its ESPRIT counterpart, at an equivalent mean squared error level, as verified in numerical simulations.

Keywords Direction of arrival · Eigendecomposition · Antenna array

1 Introduction

Antenna arrays have brought advantages over single-antenna communication scenarios, such as increase in channel capacity or more effective noise mitigation [7],

T. N. Ferreira (✉)
Telecommunication Engineering Department, Fluminense Federal University,
Niterói, RJ 24210-240, Brazil
e-mail: tadeu_ferreira@id.uff.br

M. L. R. de Campos · S. L. Netto
Electrical Engineering Program - COPPE, Federal University of Rio de Janeiro,
Rio de Janeiro, RJ 21941-972, Brazil
e-mail: mcampos@ieee.org

S. L. Netto
e-mail: sergioln@smt.ufrj.br

providing either higher data rates for multiplexing transmission or better interference mitigation for diversity transmission.

In antenna-array scenarios, the spatial diversity provides to the receiver the capacity of finding the location of a given source. When the assumptions of coplanar transmitting sources and receiving array, isotropic propagation medium, and a receiving array in the far-field of the sources are valid, the source-localization problem can be mapped onto finding the direction of arrival (DoA).

For DoA finding, the ESPRIT algorithm (estimation of parameters via rotational invariance techniques) [13] has become one of the standard DoA algorithms. ESPRIT has been tested in wireless communication scenarios [1], due to its intrinsic reduced computational cost in comparison with previous algorithms, such as MUSIC [14], without degradation of its high resolution. The lower computational cost provided by ESPRIT results from taking advantage of the invariance constraints imposed upon the receiving array geometry, which must present uniformly spaced pairs (doublets) of sensors [13].

Covariance-based DoA (CB-DoA) estimation is derived for the same geometric constraints as ESPRIT, achieving equivalent performance in terms of mean squared error (MSE), with a reduced computational cost, as detailed in [4].

Krylov-subspace techniques are applied to several signal processing scenarios, especially for large and sparse systems, in order to decrease the overall computational complexity [3].

This article investigates the effect of incorporating Krylov-subspace techniques on the standard CB-DoA framework, which substitutes the traditional eigenvalue/eigenvector decomposition (EVD). Some of the alternatives to the traditional EVD implementations give rise to a regular structure in the eigenvector matrix whose columns span a Krylov subspace [18]. A popular iterative Krylov-subspace method for performing the eigendecomposition of Hermitian matrices is the Lanczos method [8]. The application of Lanczos iterations to perform ESPRIT's first EVD led to the so-called fast subspace decomposition (FSD)-ESPRIT algorithm [11], with reduced computational burden and a reliable strategy for automatically estimating the number of sources. For non-Hermitian matrices, the Arnoldi's iterations [2,8,15,18] are an alternative to the standard EVD. Other schemes, such as [9,15], also apply additional Krylov-subspace techniques to decrease the complexity of the DoA estimation algorithm. Recent works on applications of Krylov-subspace algorithms on signal processing and communication scenarios include [3,16].

For this article, the ESPRIT algorithm has been chosen as a performance benchmark, since it allows the same degree of freedom on the design of the receiving array geometry as that of the CB-DoA algorithm. Algorithms such as Root-MUSIC, Unitary ESPRIT [10], or Real CB-DoA [5] are left out of the scope of this work, due to their additional restrictions and requirements in the array geometry when compared to the proposed algorithm. The computational complexity of the resulting CB-DoA algorithm is compared to that of the original CB-DoA algorithm as well as that of the FSD-ESPRIT [11]. When the simplified CB-DoA algorithm is compared to other algorithms that use different basis for performing the EVD of the covariance matrix, such as the one in [9], the proposed simplified CB-DoA algorithm uses smaller structures and takes advantage of the structure of tridiagonal and Hessenberg matrices to

improve computational complexity. The algorithm proposed here presents an MSE equivalent to the FSD-ESPRIT algorithm with a significant reduction in complexity.

In addition to presenting the new algorithm, this paper also stresses important differences among the standard and simplified CB-DoA algorithms as well as different ESPRIT versions. The paper is organized as follows: In Sect. 2, the basic setup and notation for the DoA problem are presented and the standard CB-DoA algorithm is described. All modifications included by the proposed CB-DoA algorithm based on Krylov-space techniques are then considered in Sect. 3. An analysis on the reduction of computational operations achieved by the proposed techniques is provided in Sect. 4, whereas Sect. 5 considers the MSE performance of the standard and simplified CB-DoA algorithms and their ESPRIT counterparts. Section 6 summarizes the main contributions of this article.

2 CB-DoA Estimation

2.1 Antenna-Array Scenario

Consider an environment with M sources and N receiving sensors and define $s_m(k)$ as the kth sample of the mth source signal, $x_i(k)$ as the signal at the ith receiving antenna, and $a_i(\theta_m)$ as the ith antenna gain in the direction θ_m of the mth source. In the digital domain, one may then write that

$$x_i(k) = \sum_{m=0}^{M-1} s_m(k) a_i(\theta_m) + n_i(k), \tag{1}$$

for $i \in 0, 1, \ldots, N-1$, where $n_i(k)$ is an additive white Gaussian noise (AWGN) acquired at the ith antenna.

By arranging all data in a matrix form,

$$\mathbf{s}(k) = \left[s_0(k)\ s_1(k)\ \ldots\ s_{M-1}(k)\right]^{\mathrm{T}}, \tag{2}$$

$$\mathbf{x}(k) = \left[x_0(k)\ x_1(k)\ \ldots\ x_{N-1}(k)\right]^{\mathrm{T}}, \tag{3}$$

$$\mathbf{n}(k) = \left[n_0(k)\ n_1(k)\ \ldots\ n_{N-1}(k)\right]^{\mathrm{T}}, \tag{4}$$

$$\mathbf{A} = \begin{bmatrix} a_0(\theta_0) & a_0(\theta_1) & \cdots & a_0(\theta_{M-1}) \\ a_1(\theta_0) & a_1(\theta_1) & \cdots & a_1(\theta_{M-1}) \\ \vdots & \vdots & \ddots & \vdots \\ a_{N-1}(\theta_0) & a_{N-1}(\theta_1) & \cdots & a_{N-1}(\theta_{M-1}) \end{bmatrix}, \tag{5}$$

where $\mathbf{A}$ contains the directional gains of each of the antennas in the receiving array. Equation (1) can be rewritten as

$$\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) + \mathbf{n}(k). \tag{6}$$

If the receiving sensors satisfy the translational invariance constraints [6,13], they can be divided in two sub-arrays, comprised of P antenna doublets, and displaced by δ: one sub-array containing the initial points and the other containing the final points in each antenna pair. Hence, the received signal may be redefined as $x_{b,i}(k)$, referring to the bth sub-array, $b = 0, 1$, and the ith doublet, such that

$$x_{0,i}(k) = \sum_{m=0}^{M-1} s_m(k) a_i(\theta_m) + n_{0,i}(k), \tag{7}$$

$$x_{1,i}(k) = \sum_{m=0}^{M-1} s_m(k) \phi_m a_i(\theta_m) + n_{1,i}(k), \tag{8}$$

for $0 \leq i < P$, with

$$\phi_m = \mathrm{e}^{\frac{\jmath\omega\delta}{c}\sin\theta_m}, \tag{9}$$

where ω represents the carrier angular frequency, c is the speed of light, and δ is the distance between the antennas belonging to a doublet pair [13].

2.2 The Standard CB-DoA Algorithm

Consider that the transmitted signals have zero mean, then a covariance model for the received signal $\mathbf{x}(k)$ can be expressed by

$$\mathbf{R}_x = E[\mathbf{x}(k)\mathbf{x}^H(k)] = \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \mathbf{R}_n, \tag{10}$$

where $\mathbf{R}_s = E[\mathbf{s}(k)\mathbf{s}^H(k)]$ represents the covariance matrix for the transmitted signal and $\mathbf{R}_n$ is the noise covariance matrix. If all the source and noise components are uncorrelated and if noise is spatially independent and identically distributed, then $\mathbf{R}_s$ is diagonal and $\mathbf{R}_n = \sigma^2\mathbf{I}$. Hence, one may consider the matrices

$$\begin{aligned}\mathbf{R}_{00} &= \mathbf{J}_0^{P,N}\mathbf{R}_x(\mathbf{J}_0^{P,N})^T \\ &= \mathbf{J}_0^{P,N}\mathbf{A}\mathbf{R}_s\mathbf{A}^H(\mathbf{J}_0^{P,N})^T + \sigma^2\mathbf{I}, \end{aligned}\tag{11}$$

$$\begin{aligned}\mathbf{R}_{01} &= \mathbf{J}_0^{P,N}\mathbf{R}_x(\mathbf{J}_1^{P,N})^T \\ &= \mathbf{J}_0^{P,N}\mathbf{A}\mathbf{R}_s\boldsymbol{\Phi}^H\mathbf{A}^H(\mathbf{J}_0^{P,N})^T + \sigma^2\mathbf{I}, \end{aligned}\tag{12}$$

where $\mathbf{J}_0^{P,N}$ and $\mathbf{J}_1^{P,N}$ represent two auxiliary matrices which select the first and last P rows, respectively, of an $N \times 1$ vector, and $\boldsymbol{\Phi}$ is a diagonal matrix comprised of $\phi_0, \phi_1, \ldots, \phi_{M-1}$.

By performing an EVD on $\mathbf{R}_{00}$, such that $\mathrm{EVD}(\mathbf{R}_{00}) = \mathbf{U}\boldsymbol{\Lambda}^2\mathbf{U}^H$, one determines the diagonal matrix $\boldsymbol{\Lambda}^2$ containing all eigenvalues of $\mathbf{R}_{00}$ with the corresponding eigenvectors placed on the columns of $\mathbf{U}$. The smallest $(P - M)$ eigenvalues of $\mathbf{R}_{00}$ should be equal to the noise variance σ^2. The remaining M largest eigenvalues form the $M \times M$ diagonal matrix $\boldsymbol{\Lambda}_s$, and the corresponding eigenvectors, associated with

the signal subspace [13], can be arranged in a columnwise manner on matrix $\mathbf{U}_s$, allowing one to define the auxiliary matrix $\mathbf{F} = \mathbf{U}_s^H \mathbf{\Lambda}_s^{-1}$.

By applying $\mathbf{F}$ on both sides of the matrix pencil $(\mathbf{R}_{01} - \hat{\sigma}^2\mathbf{I})$, where $\hat{\sigma}^2$ is the smallest eigenvalue of $\mathbf{\Lambda}$, one gets

$$\mathbf{\Psi} = \mathbf{F}(\mathbf{R}_{01} - \hat{\sigma}^2\mathbf{I})\mathbf{F}^H, \tag{13}$$

such that [17]

$$\text{EVD}(\mathbf{\Psi}) = \mathbf{B}\hat{\mathbf{\Phi}}\mathbf{B}^{-1}, \tag{14}$$

which yields the DoA estimates θ_m from the DoA mapping given by Eq. (9).

The CB-DoA algorithm requires 1 Hermitian and 1 non-Hermitian EVD. In comparison, the total least squares (TLS) version of ESPRIT requires 2 non-Hermitian EVDs (an $\mathcal{O}(25n^3)$ operation according to Golub and van Loan [8]) and 1 Hermitian EVD (an $\mathcal{O}(n^2)$ operation [8]). Despite being less complex than the TLS-ESPRIT, the CB-DoA still requires 2 EVD operations, whose complexity may be significantly reduced when computed through Krylov-subspace techniques, as investigated below.

3 The Simplified CB-DoA Algorithm

In this section, some fundamental principles underlying a Krylov subspace are described. Techniques using Krylov subspaces are used here for transforming Hermitian covariance matrices into tridiagonal ones and non-Hermitian cross-covariance matrices into Hessenberg ones, which allow computing EVDs with smaller cost and subsequently the CB-DoA algorithm with lower complexity.

Consider a subspace $\mathcal{S}$ and a linear transformation $\mathbf{L} : \mathcal{S} \rightarrow \mathcal{S}$. When a vector $\mathbf{b}_0$ is repeatedly left-multiplied by $\mathbf{L}$, the resulting sequence of vectors $\{\mathbf{b}_0, \mathbf{L}\mathbf{b}_0, \mathbf{L}^2\mathbf{b}_0, \cdots\} \in \mathcal{S}$ spans a Krylov subspace. The matrix whose columns are formed by this sequence of vectors is known as the Krylov matrix.

In [8], it is shown that the QR decomposition of a Krylov matrix $\mathbf{K} = \mathbf{Q}\mathbf{R}_K$ also leads to the decomposition $\mathbf{K} = \mathbf{Q}\mathbf{H}\mathbf{Q}^T$, where $\mathbf{H}$ is an upper Hessenberg matrix. By exploiting these relationships, one is able to obtain EVDs with a smaller cost, for a given matrix, as discussed below.

3.1 Lanczos' Updates

In this subsection, the Lanczos' updates are used for transforming the Hermitian matrix $\mathbf{R}_{00}$ in an Hermitian tridiagonal matrix, in order to decrease the burden of computing its eigendecomposition.

As presented in [12], when the Krylov matrix $\mathbf{K}$ is Hermitian, then $\mathbf{H}$ is tridiagonal and also Hermitian. Consider now the Hermitian tridiagonal matrix $\mathbf{T}$ generated by the QR decomposition of the Hermitian matrix $\mathbf{R}_{00} = \mathbf{Q}\mathbf{T}\mathbf{Q}^H$, where $\mathbf{Q}$ is a unitary matrix,

Table 1 Lanczos' update operations

while $(\beta_k \neq 0)$
$\mathbf{q}_{k+1} = \mathbf{r}_k/\beta_k$;
$k = k + 1$;
$\alpha_k = \mathbf{q}_k^H \mathbf{R}_{00}\mathbf{q}_k$;
$\mathbf{r}_k = (\mathbf{A} - \alpha_k\mathbf{I})\mathbf{q}_k - \beta_{k-1}\mathbf{q}_{k-1}$;
$\beta_k = \|\mathbf{r}_k\|_2$;
end

$$\mathbf{T} = \begin{bmatrix} \alpha_1 & \beta_1^* & 0 & \dots & 0 & 0 \\ \beta_1 & \alpha_2 & \beta_2^* & \dots & 0 & 0 \\ 0 & \beta_2 & \alpha_3 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \alpha_{N-1} & \beta_{N-1}^* \\ 0 & 0 & 0 & \dots & \beta_{N-1} & \alpha_N \end{bmatrix}. \tag{15}$$

Since $\mathbf{R}_{00}\mathbf{Q} = \mathbf{QT}$, and by considering $\mathbf{q}_i$ as the ith column of $\mathbf{Q}$, then

$$\mathbf{R}_{00}\mathbf{q}_i = \beta_{i-1}^*\mathbf{q}_{i-1} + \alpha_i\mathbf{q}_i + \beta_i\mathbf{q}_{i+1}, \tag{16}$$

which is the basis for the Lanczos' updates.

In fact, by pre-multiplying (16) by $\mathbf{q}_i^H$ and by considering the orthonormality of $\mathbf{q}_i$, then $\alpha = \mathbf{q}_i^H\mathbf{R}_{00}\mathbf{q}_i$. The iterations are performed until the criterion $\mathbf{r}_i = [(\mathbf{R}_{00} - \alpha_i\mathbf{I})\mathbf{q}_i - \beta_{i-1}\mathbf{q}_{i-1}] \approx \mathbf{0}$ is matched. While the stop criterion is not reached, β_i is updated as $\|\mathbf{r}_i\|_2$ and $\mathbf{q}_{i+1} = \mathbf{r}_i/\beta_i$. After all the α_i and β_i are found, matrix $\mathbf{T}$ is then determined from the Hermitian version of the Lanczos' updates, as summarized in Table 1.

3.2 Bisection Method

Upon determining the tridiagonal matrix $\mathbf{T}$, a low complexity algorithm may be employed to find its eigenvalues. Consider $\mathbf{T}_r$ as the $r \times r$ leading principal submatrix generated from $\mathbf{T}$, i.e., by removing the last $(N-r)$ rows and columns from $\mathbf{T}$. Let $p_r(x) = \det(\mathbf{T}_r - x\mathbf{I})$ be an auxiliary polynomial. Whenever x is an eigenvalue of $\mathbf{T}_r$, then $p_r(x) = 0$ and it can be shown that [8]

$$p_r(x) = (\alpha_r - x)p_{r-1}(x) - \beta_{r-1}^2 p_{r-2}(x), \tag{17}$$

with α_r and β_r as defined in Eq. (15).

The bisection method performs an eigenvalue search based on the idea represented in Fig. 1. Consider that two points y and z are chosen such that $p_r(y)p_r(z) < 0$. This relation indicates that there is a point between y and z which is a root for the

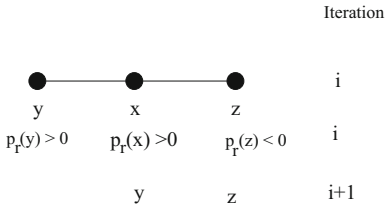

Fig. 1 Bisection method for finding eigenvalues

polynomial p_r and also an eigenvalue of $\mathbf{T}_r$. Then, a point x is chosen midway from y to z, and the polynomial value at x is analyzed.

Whenever $p_r(x)p_r(y) > 0$, then we set $y = x$ in the next iteration, and the process is repeated until the distance between y and z is smaller than a predetermined value.

By using the Gershgorin theorem of sensitivity, it can be proved that, for the ith largest eigenvalue λ_i [8],

$$\min_{1\le j\le N}[\alpha_j - |\beta_j| - |\beta_{j-1}|] \le \lambda_i \le \max_{1\le j\le N}[\alpha_j + |\beta_j| + |\beta_{j-1}|]. \tag{18}$$

Hence, the bisection iterations can be used for finding each eigenvalue, where y and z are initially chosen as $\min_{1\le j\le N}[\alpha_j - |\beta_j| - |\beta_{j-1}|]$ and $\max_{1\le j\le N}[\alpha_j + |\beta_j| + |\beta_{j-1}|]$, respectively.

At each iteration, the middle point $x = (y+z)/2$ is calculated. If $p_r(x) \ge (N-k)$, i.e., if the number of sign inversions is greater than $(N-k)$, then the operation is repeated with x in the role of z, that is, the segment from x to y is considered. When $p_r(x) < (N-k)$, then x is used in the role of y and the segment between z and x is employed.

Finally, the eigenvectors of $\mathbf{R}_{00}$ are found by using the inverse power iterations [8]. The characteristic equation $\mathbf{R}_{00}\mathbf{v}_i = \lambda_i\mathbf{I}\mathbf{v}_i$ of eigendecomposition is used. When $(\mathbf{R}_{00} - \lambda_i\mathbf{I})^{-1}$ is right-multiplied by a vector, it generates a vector in the direction of the eigenvector $\mathbf{v}_i$, associated with λ_i [8]. This multiplication is iteratively performed on the vectors which progressively converge to $\mathbf{v}_i$.

3.3 Arnoldi's Updates

In this subsection, the non-Hermitian matrix $\mathbf{R}_{01}$ is transformed into a Hessenberg matrix in order to perform a lower-complexity eigendecomposition. Another decomposition $\mathbf{R}_{01} = \mathbf{Q}_{n+1}\mathbf{H}_n\mathbf{Q}_n^H$ is performed on $\mathbf{R}_{01}$, where $\mathbf{Q}_n$ represents an $N \times n$ matrix obtained by extracting the first n columns of $\mathbf{Q}$ and $\mathbf{H}_n$ is a Hessenberg matrix containing the first $(n+1)$ rows and n columns of $\mathbf{H}$ such that

$$\mathbf{H}_n = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1n} \\ h_{21} & h_{22} & \dots & h_{2n} \\ 0 & h_{32} & \dots & h_{3n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & h_{n+1,n} \end{bmatrix}. \tag{19}$$

Table 2 Arnoldi's update operations

while $(\beta_k \neq 0)$
  $\mathbf{q}_{k+1} = \mathbf{r}_k / h_{k+1,k}$;
  $k = k + 1$;
  $\mathbf{r}_k = \mathbf{R}_{01}\mathbf{q}_k$;
  for $i = 1 : k$,
    $h_{i,k} = \mathbf{q}_i^H \mathbf{r}_k$;
    $\mathbf{r}_k = \mathbf{r}_k - h_{i,k}\mathbf{q}_i$;
  end
  $h_{k+1,k} = \|\mathbf{r}_k\|_2$;
end

If q_i denotes the ith column of $\mathbf{Q}$, then the following equations arise:

$$\begin{cases} \mathbf{R}_{01}\mathbf{q}_1 = h_{11}\mathbf{q}_1 + h_{21}\mathbf{q}_2 \\ \mathbf{R}_{01}\mathbf{q}_2 = h_{12}\mathbf{q}_1 + h_{22}\mathbf{q}_2 + h_{32}\mathbf{q}_3 \\ \quad \vdots \\ \mathbf{R}_{01}\mathbf{q}_n = h_{1n}\mathbf{q}_1 + \ldots + h_{nn}\mathbf{q}_n + h_{n+1,n}\mathbf{q}_{n+1} \end{cases} . \tag{20}$$

By exploiting the mutual orthogonality among the $\mathbf{q}_n$ vectors, a Gram-Schmidt-like algorithm, known as Arnoldi's updates, can be used to solve (20), yielding the h_{ij} entries of matrix $\mathbf{H}_n$ [18]. Arnoldi's updates transform any matrix $\mathbf{R}_{01}$ into a Hessenberg matrix $\mathbf{H}$, as given in Eq. (19), through the steps summarized in Table 2.

3.4 Eigendecomposition of a Hessenberg Matrix

In this subsection, we address two techniques for exploiting the Hessenberg structure of a matrix, in order to obtain its eigenvalues in a simplified manner.

After processing matrix $\mathbf{R}_{01}$ through the Arnoldi's updates, a Hessenberg matrix is generated. A Schur decomposition can then be performed on the Hessenberg matrix, which generates an upper triangular matrix $\mathbf{T}_S$, with the eigenvalues of the original matrix $\mathbf{R}_{01}$ in its main diagonal [8].

Alternatively, a QZ method may be performed for finding the eigenvalues of $\mathbf{R}_{01}$. This approach leads to a generalized Schur decomposition (GSD) [8], which states that if $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ complex matrices, then a pair of unitary matrices $\mathbf{Q}$ and $\mathbf{Z}$ generate $\mathbf{Q}^H\mathbf{A}\mathbf{Z} = \mathbf{T}$ and $\mathbf{Q}^H\mathbf{B}\mathbf{Z} = \mathbf{S}$, where $\mathbf{T}$ and $\mathbf{S}$ are both upper triangular. In our modeling, $\mathbf{B}$ is an identity matrix, which simplifies the problem to an EVD of $\mathbf{A}$. However, since $\mathbf{A} = \mathbf{\Psi}$ is still non-Hermitian and tridiagonal, a Hessenberg-triangular reduction (HTR) [8] operation must be performed instead of a simple QR decomposition. The HTR is an iterative process whose steps transform $\mathbf{A}$ progressively in a Hessenberg form, while $\mathbf{B}$ is transformed into an upper triangular matrix. That is the principle of the QZ method, which computes matrices $\mathbf{Q}$ and $\mathbf{Z}$ reducing the pair $(\mathbf{A}, \mathbf{B})$ to a pair of upper Hessenberg-upper triangular matrices. It is implemented by using a QR function and Givens rotations, as well as a QZ step inside a loop [8].

3.5 Detection of the Number of Sources

In the traditional ESPRIT and CB-DoA algorithms, the number of sources is estimated in separate stages, using, for example, the Akaike's information criterion (AIC) [19], which estimates the number of sources $\hat{M}$ as the value of k which minimizes:

$$\mathrm{AIC}(k) = -2\log\left(\frac{\prod_{i=k+1}^{N}\lambda_i^{\frac{1}{N-k}}}{\frac{1}{N-k}\sum_{i=k+1}^{N}\lambda_i}\right) + 2k(2N-k). \tag{21}$$

The simplified CB-DoA algorithm, however, employs the FSD algorithm [11], which estimates $\hat{M}$ as the value of k that minimizes

$$\mathrm{FSD}_M(k) = \log\left[\frac{\sqrt{\frac{1}{N-k}\left(\|\mathbf{R}_{00}\|^2 - \sum_{i=1}^{k}\lambda_i^2\right)}}{\frac{1}{N-k}\left(\mathrm{Tr}(\mathbf{R}_{00}) - \sum_{i=1}^{k}\lambda_i\right)}\right], \tag{22}$$

where $\mathrm{Tr}(\mathbf{X})$ denotes the trace of $\mathbf{X}$. As discussed in [11], the FSD estimation is more robust and less computationally complex than the AIC.

4 Computational Complexity Analysis

A block diagram describing the general structure of the standard and the simplified CB-DoA algorithms is depicted in Fig. 2. In addition, the simplified CB-DoA algorithm is summarized in Table 3.

The computational costs of all operations required by the standard and simplified versions of the CB-DoA algorithm are indicated in Table 4.

Although Hu and Kailath [11] state that the computational complexity for estimating the number of sources is $\mathcal{O}(n^2)$, it considers the method as computed simultaneously with the eigenvalues of the Lanczos loop. We have computed it after the loop, using $\|\mathbf{R}_{00}\| = \lambda_1$, where λ_1 is its largest eigenvalue, as well as $\mathrm{Tr}(\mathbf{R}_{00})$ equal to the sum of the eigenvalues, previously computed inside the Lanczos loop. By using such methods for computing $\|\mathbf{R}_{00}\|$ and $\mathrm{Tr}(\mathbf{R}_{00})$, the complexity of FSD-based evaluation of the number of sources reduces to $\mathcal{O}(n)$.

The FSD algorithm reduces the computational burden on both the EVD of $\mathbf{R}_{00}$ and the estimation of the number of sources. From Table 4, the reduction factor γ in the computational complexity for both versions of the CB-DoA algorithm is given by

$$\gamma = \left(1 - \frac{\frac{4M^3}{3} + M^2 + 2P^2 + M\log M + 3M^2P + 2P + M}{25M^3 + P^3 + 3M^2P + 2P + M^3}\right)100\,\%. \tag{23}$$

For instance, when $M = 1$ and $P = 2$, the reduction factor is $\gamma \approx 13\,\%$. By increasing P, the reduction factor also increases, approaching $\gamma = (1 - 1/P)100\,\%$

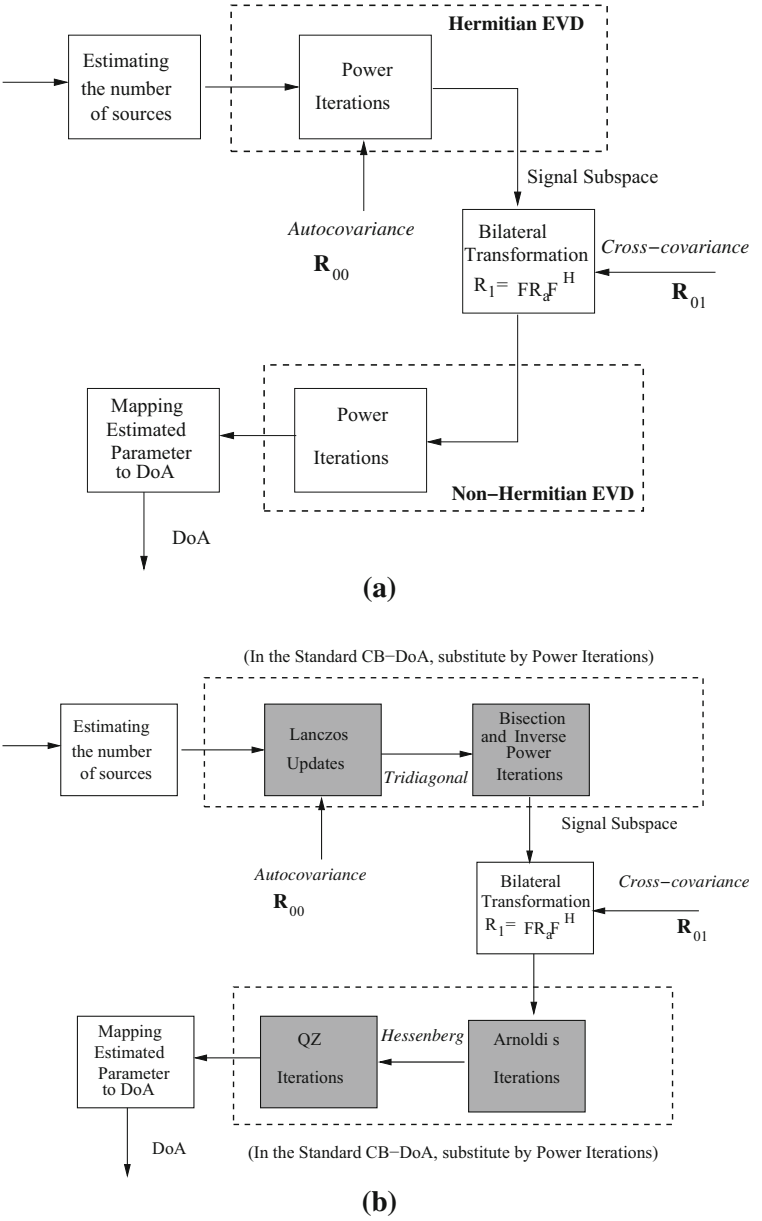

Fig. 2 Block diagrams for the **a** CB-DoA and **b** simplified CB-DoA algorithms. *Gray blocks* only appear in the simplified version, as the CB-DoA employs basic power iterations

in the limiting case of $M \ll P$. Such behavior is illustrated in Fig. 3 when $M = 2$ and $3 \leq P \leq 14$.

In Fig. 4, both $2 \leq M \leq 6$ and $3 \leq P \leq 14$ are varied for the computation of the reduction factor.

Table 3 Summary of operations for the simplified CB-DoA algorithm

$[\mathbf{L}, \mathbf{T}] = \text{Lanczos}(\mathbf{R}_{00})$

$\hat{M} = \text{FSD}_M(\mathbf{R}_{00})$

$\hat{\sigma}_N^2 = \text{Bisection}(\mathbf{L}, \mathbf{T}, \hat{M})$

$\mathbf{U}_s = \text{Inverse}_{\text{Power}}(\mathbf{L}, \mathbf{T}, \hat{M})$

$\mathbf{F} = \mathbf{\Sigma}_s^{-1}\mathbf{U}_s^H$

$\mathbf{R}_a = \hat{\mathbf{R}}_{01} - \hat{\sigma}^2\mathbf{I}$

$\mathbf{R}_1 = \mathbf{F}\mathbf{R}_a\mathbf{F}^H$

$[\mathbf{B}, \mathbf{H}] = \text{Arnoldi}(\mathbf{R}_1)$

$[\mathbf{\Phi}^H] = \text{Schur}(\mathbf{H})$ **or** $[\mathbf{\Phi}^H] = \text{QZ}(\mathbf{H})$

Table 4 Computational complexity of all operations required by the standard (bold), the simplified (italics) CB-DoA, and by both algorithms (bold italics) [8]

Operation (type)	Complexity	Flops
EVD (non-Hermit.)	$\mathcal{O}(25n^3)$	$25M^3$
EVD Arnoldi + Schur or QZ	$\mathcal{O}(4n^3/3 + n^2)$	$4M^3/3 + M^2$
EVD (Hermit.)	$\mathcal{O}(n^3)$	P^3
EVD Lanczos	$\mathcal{O}(2n^2 + n\log n)$	$2P^2 + M\log M$
Inversion (diag.)	$\mathcal{O}(n)$	P
Multiplication	$\mathcal{O}(n^3)$	$3 \times (3M^2P)$
Subtraction (diag.)	$\mathcal{O}(n)$	P
Source # (AIC)	$\mathcal{O}(n^3)$	M^3
Source # (FSD)	$\mathcal{O}(n^2)$	M^2

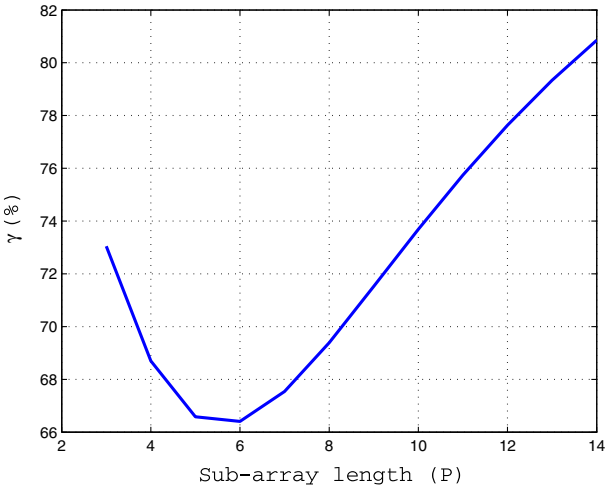

Fig. 3 Reduction factor γ for a fixed number of sources $M = 2$ and a varying sub-array dimension $3 \leq P \leq 14$

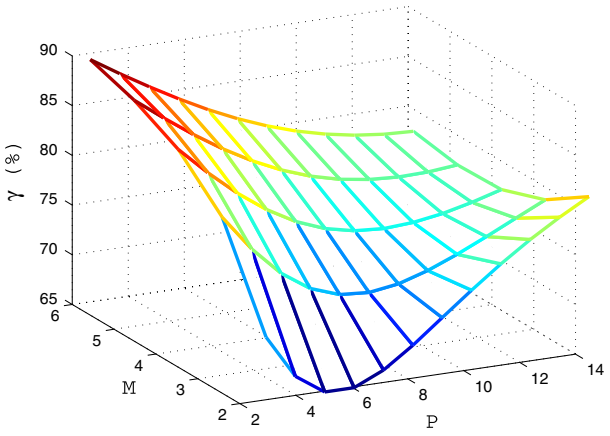

Fig. 4 Reduction factor γ for a varying number of sources $2 \leq M \leq 6$ and a varying sub-array dimension $3 \leq P \leq 14$

The associated computational complexity for the FSD-ESPRIT [11] algorithm, which incorporates Krylov-space techniques, is $(13P^3/3+5M^3/3+MN^2+2MNP+N^2+M)$. When considering the simplest ($M = 2$, $N = 4$, $P = 3$) DoA scenario, the simplified CB-DoA represents a 69 % computational saving against the FSD-ESPRIT algorithm. Even larger savings are achieved with the simplified CD-DoA algorithm in more practical cases when larger values of N and P are employed.

5 Simulations

In this section, we verify that the reduction in the computational complexity, as quantified in Sect. 4, does not imply performance degradation on the CB-DoA algorithm in terms of MSE. Our first DoA estimation scenario consists of $M = 2$ sources (both transmitting with the same power), positioned at $\theta_1 = 2°$ and $\theta_2 = 7°$, and the receiving array as shown in Fig. 5. When all $N = 12$ receiving antennas are grouped into $P = 6$ doublets, according to Fig. 3, simplified CB-DoA (with either Schur decomposition or the QZ iterations) presents a computational reduction factor of $\gamma = 66.4\,\%$ in comparison with the FSD-ESPRIT. If only the $N = 8$ black and gray antennas in Fig. 5 are employed with $P = 4$, one gets a constant-displacement scenario suited for the CB-DoA approach. Those geometrical constraints are less restrictive than either the uniform linear array imposed by Root-MUSIC or the centro-Hermitian configuration required by real-only transform algorithms, such as Unitary ESPRIT [10]. This ($M = 2$, $N = 8$, $P = 4$) case corresponds to a computational reduction factor of $\gamma = 68.8\,\%$ for the simplified CB-DoA algorithm against its standard version, as estimated by Eq. (23), and an 80.4 % computation saving in comparison with the FSD-ESPRIT.

In each run, 8,000 random values are transmitted, generated by a normalized Gaussian distribution. The resulting average MSE, over 400 Monte Carlo runs, as determined by

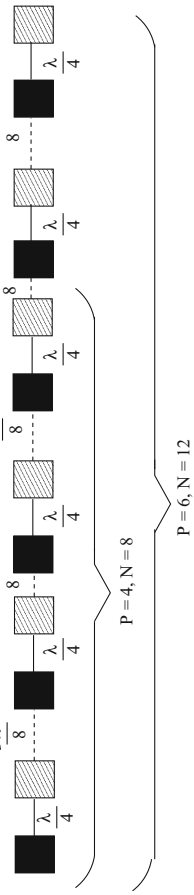

Fig. 5 Linear array with $\lambda/4$ space between antennas of the same doublet. Antennas in *black* belong to the first sub-array, and antennas in *white with diagonal thin lines* are in the second sub-array

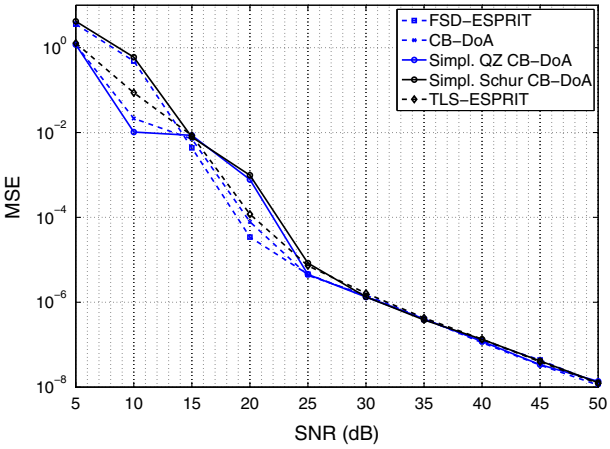

Fig. 6 MSE performance of standard and simplified CB-DoA algorithms as well as the FSD-ESPRIT with bisection, for the ($M = 2$, $N = 8$, $P = 4$) scenario, and $\lambda = 0.2$ m

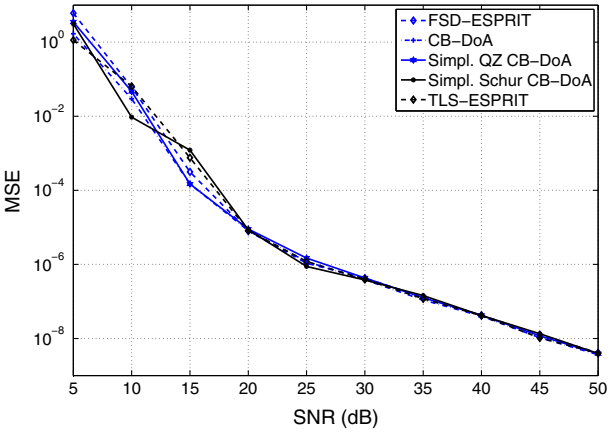

Fig. 7 MSE performance of standard and simplified CB-DoA algorithms as well as the FSD-ESPRIT with bisection, for the ($M = 2$, $N = 12$, $P = 6$) scenario, and $\lambda = 0.2$ m

$$\mathrm{MSE} = \frac{1}{800} \sum_{q=1}^{400} \sum_{m=1}^{2} |\theta_{m,q} - \hat{\theta}_{m,q}|^2, \tag{24}$$

is depicted in Fig. 6 for $P = 4$ antennas in each sub-array and in Fig. 7 for $P = 6$ antennas in each sub-array. In these sets of simulations, the standard and simplified CB-DoA versions were simulated as well as the FSD-ESPRIT algorithm using the bisection approach and the classical TLS-ESPRIT, for several levels of signal-to-noise ratio (SNR). In Figs. 6 and 7, transmission is performed at frequency 1.5 GHz, corresponding to $\lambda = 0.2$ m, which is a common wireless personal communication scenario.

In Fig. 8, transmission is performed for $P = 4$ antennas in each sub-array at a frequency 12 GHz (i.e., $\lambda = 0.025$ m), corresponding to a K_u-band satellite scenario.

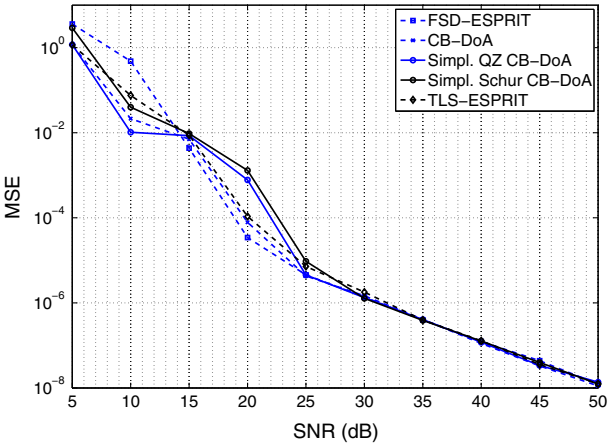

Fig. 8 MSE performance of standard and simplified CB-DoA algorithms as well as the FSD-ESPRIT with bisection, for the ($M = 2$, $N = 8$, $P = 4$) scenario, and $\lambda = 0.025$ m

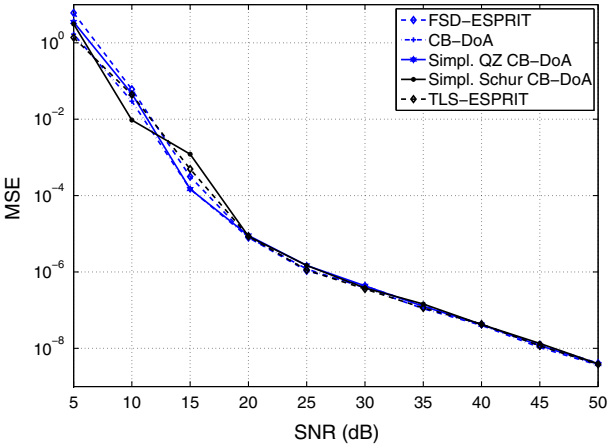

Fig. 9 MSE performance of standard and simplified CB-DoA algorithms as well as the FSD-ESPRIT with bisection, for the ($M = 2$, $N = 12$, $P = 6$) scenario, and $\lambda = 0.025$ m

In Fig. 9, transmission is also performed at a frequency 12 GHz, for $P = 6$ antennas in each sub-array.

From these sets of simulations, one readily observes a similar MSE performance for all the algorithms for the whole SNR range and all (M, N, P, λ) scenarios considered here.

Simulations have also been performed using both AIC and FSD methods for estimating the number of sources in conjunction with the CB-DoA algorithm. Results indicate a similar performance of both criteria for the entire $5 \leq$ SNR [dB] ≤ 50 range, which agrees to the FSD analysis presented in [11], with both methods correctly identifying $M = 2$ sources in all cases of SNR ≥ 15 dB.

6 Conclusion

A novel version for the CB-DoA estimation algorithm was considered using Krylov-subspace techniques, improving the computational complexity without increasing estimation error. In this proposal, the standard eigendecomposition operations were replaced by alternative algorithms for Hermitian (Lanczos' updates followed by bisection and inverse power) as well as non-Hermitian (Arnoldi's updates and Schur/QZ decomposition) matrices. The novel CB-DoA version also uses an alternative method for estimating the number of sources. Complexity analyses were performed indicating a reduction of at least 60 % on the computational burden of the simplified CB-DoA algorithm in comparison with the FSD-ESPRIT (which also incorporates similar Krylov-space techniques) and advantages over the standard CB-DoA algorithm. Numerical simulations validate the computational analysis and indicate similar MSE performances for the simplified CB-DoA as well as the standard CB-DoA and the FSD-ESPRIT algorithms in all the scenarios tested.

Acknowledgments The authors thank the Brazilian National Council for Research and Development (CNPq), the Brazilian Ministry of Education's CAPES, and the State of Rio de Janeiro Funding Support for Research (FAPERJ).

References

1. K. Almidfa, G.V. Tsoulos, A. Nix, Performance analysis of ESPRIT, TLS-ESPRIT and unitary ESPRIT algorithms for DoA estimation in a WCDMA mobile system, in *Proceedings of International Conference on 3G Mobile Communication Technologies* (2000), pp. 200–203
2. W.E. Arnoldi, The principle of minimized iterations in the solution of the matrix eigenvalue problem. Q. Appl. Math. **9**, 17–29 (1951)
3. S. Chouvardas, K. Slavakis, S. Theodoridis, Trading off complexity with communication costs in distributed adaptive learning via Krylov subspaces for dimensionality reduction. IEEE J. Sel. Top. Signal Process. **7**(2), 257–273 (2013)
4. T.N. Ferreira, S.L. Netto, P.S.R. Diniz, Direction-of-arrival estimation using a low-complexity covariance approach. IEEE Trans. Aerosp. Electron. Syst. **48**(3), 1924–1934 (2012)
5. T.N. Ferreira, S.L. Netto, P.S.R. Diniz, Covariance-based direction-of-arrival estimation with real structures. IEEE Signal Process. Lett. **15**, 757–760 (2008)
6. T.N. Ferreira, S.L. Netto, P.S.R. Diniz, Low complexity covariance-based DoA estimation algorithm, in *Proceedings of the 15th European Signal Processing Conference*, Poznan, Poland (2007), pp. 100–104
7. L.C. Godara, Applications of antenna arrays to mobile communications I: Performance improvement, feasibility and system considerations. Proc. IEEE **85**(7), 1031–1060 (1997)
8. G.H. Golub, C.F. van Loan, *Matrix Computations*, 3rd edn. (Johns Hopkins University Press, Baltimore, 1996)
9. R. Grover, D.A. Pados, M.J. Medley, Subspace direction finding with an auxiliary-vector basis. IEEE Trans. Signal Process. **55**(2), 758–763 (2007)
10. M. Haardt, J.A. Nossek, Unitary ESPRIT: how to obtain increased estimation accuracy with a reduced computational burden. IEEE Trans. Signal Process. **43**(5), 1232–1242 (1995)
11. G. Hu, T. Kailath, Fast subspace decomposition. IEEE Trans. Signal Process. **42**(3), 539–551 (1994)
12. S. Qiao, G. Liu, W. Xu, Block Lanczos tridiagonalization of complex symmetric matrices, in *Advanced Signal Processing Algorithms, Architectures, and Implementations XV, Proceedings of the SPIE*, vol. 5910 ed. by F.T. Luk (2005), pp. 285–295
13. R. Roy, T. Kailath, ESPRIT—estimation of parameters via rotational invariance techniques. IEEE Trans. Acoust. Speech Signal Process. **37**(7), 984–995 (1989)
14. R.O. Schmidt, Multiple emitter location and signal parameter estimation. IEEE Trans. Antennas Propag. **AP–34**(3), 276–280 (1986)

15. J. Steinwandt, R.C. de Lamare, M. Haardt, Beamspace direction finding based on the conjugate gradient and the auxiliary vector filtering algorithms. Signal Process. **93**, 641–651 (2012)
16. T. Tazeen, A. Kansal, Low-complexity equalization of OFDM systems using Krylov subspace methods over doubly selective channels. Int. J. Emerg. Res. Manag. Technol. **3**(8), 64–70 (2014)
17. L. Tong, G. Xu, T. Kailath, Blind identification and equalization based on second-order statistics: a time-domain approach. IEEE Trans. Inf. Theory **40**(2), 340–349 (1994)
18. L.N. Trefethen, D. Bau III, *Numerical Linear Algebra* (SIAM, Philadelphia, 1997)
19. M. Wax, T. Kailath, Detection of signals by information theoretic criteria. IEEE Trans. Acoust. Speech Signal Process. **33**(2), 387–392 (1985)