Signal Parameter Estimation via the Cayley–Hamilton Constraint

Prashanth Hande and Lang Tong

Abstract—The estimation of signal parameters via rotational invariance tachnique (ESPRIT) algorithm has proved to be a robust solution to exploit single invariance structure in array elements. However, ESPRIT is not flexible to incorporate multiple invariances. A closed form solution to the problem of estimating eigenvalues in multiple invariant structures is presented in this paper. It is shown that the algorithm is comparable in complexity to ESPRIT but performs better than ESPRIT, especially when only a limited number of signal samples are available. In addition, the algorithm is shown to be feasible even for single array elements sharing multiple invariance.

Index Terms—Estimation of signal parameters via rotational invariance tachnique (ESPRIT), multiple invariance, signal parameter estimation via Cayley–Hamilton constraint (SPECC), subspace methods.

I. INTRODUCTION

THE estimation of signal parameters via rotational invariance tachnique (ESPRIT) [1] algorithm has proved to be a powerful technique for the estimation of unknown parameters by exploitation of the rotational invariance property in the eigenstructure of the measured data. Originally developed for the estimation of the directions-of-arrivals (DOAs), ESPRIT finds widespread applications in radar, sonar, global positioning systems (GPS), wireless communications, and even image analysis. Conventionally, ESPRIT obtains closed-form estimates of the DOAs of narrowband signals and results in robust performance by exploiting single invariance structures. Unfortunately, ESPRIT is not flexible enough to incorporate multiple invariances in the data. In [2], Swindlehurst *et al.* proposed the multiple invariance ESPRIT (MI-ESPRIT), which provided better estimates. The main drawback of MI-ESPRIT is that the algorithm involves a multidimensional minimization.

In this paper, we state and prove a theorem based on the Cayley–Hamilton theorem for square matrices and propose an algorithm for parameter estimation based on this result. Referred to as the signal parameter estimation via Cayley–Hamilton constraint (SPECC), our algorithm provides a closed-form estimation that exploits multiple invariance structures.

II. PROBLEM FORMULATION

Consider the reception of a narrowband signal vector $\mathbf{s}(t) \in \mathcal{C}^d$ at an array of M sensors. The measured data at the sensors is to be used to estimate the DOA of each of the d sources. The response at the sensor array represented by $\mathbf{z}(t) \in \mathcal{C}^M$ is given by

$$\mathbf{z}(t) = \mathbf{G}\mathbf{s}(t) + \mathbf{n}(t), \quad \mathbf{G} = [\mathbf{a}(\theta_1) \quad \cdots \quad \mathbf{a}(\theta_d)] \quad (1)$$

where $\mathbf{a}(\theta)$ is the array response vector to a unit amplitude narrowband signal in the direction θ, and $\mathbf{n}(t)$ is additive white Gaussian noise. In this paper, we assume a uniformly spaced linear array although the proposed algorithm can be extended to a more general setting [3].

Consider selection matrices $\mathbf{J}_0, \mathbf{J}_1, \ldots, \mathbf{J}_p$ of size $m \times M$. As in [2], the selection matrices each pick m out of the M sensors in the uniform linear array and introduce the following rotational invariance structure

$$\mathbf{J}_0\mathbf{G} = \mathbf{A}, \quad \mathbf{J}_1\mathbf{G} = \mathbf{A}\mathbf{\Phi}, \quad \mathbf{J}_p\mathbf{G} = \mathbf{A}\mathbf{\Phi}^p. \quad (2)$$

Here, $\mathbf{\Phi}$ is a unitary diagonal matrix with diagonal elements ϕ_i given by

$$\phi_i = \exp\{-j2\pi\Delta\sin(\theta_i)/\lambda\}, \quad i = 1, \ldots, d \quad (3)$$

where λ is the wavelength and $\Delta \sim \lambda$ is the spacing between array elements. ESPRIT exploits this structure for $p = 1$, and estimates ϕ_i, based on which, the DOAs θ_i can be estimated. The complex vector space $\mathcal{C}^M$ of the received signal $\mathbf{z}(t)$ needs to be separated into the signal and noise subspaces which are orthogonal to each other. This is accomplished by considering the eigen-decomposition of the covariance matrix $\mathbf{R}_{ZZ}$ of the received signal. We have

$$\mathbf{R}_{ZZ} = E[\mathbf{z}(t)\mathbf{z}^H(t)] = \mathbf{G}\mathbf{R}_{SS}\mathbf{G}^H + \sigma^2\mathbf{I} \quad (4)$$

where σ^2 is the noise variance. We assume that the covariance matrix $\mathbf{R}_{SS}$ of the signal vector is of full rank d (no fully correlated signals) and the columns of $\mathbf{G}$ are linearly independent. The eigen-decomposition of the covariance matrix results in

$$\mathbf{R}_{ZZ} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^H = \mathbf{U}_S\mathbf{\Sigma}_S\mathbf{U}_S^H + \sigma^2\mathbf{U}_N\mathbf{U}_N^H \quad (5)$$

where $\mathbf{U}_S = [\mathbf{e}_1 \cdots \mathbf{e}_d]$, $\mathbf{U}_N = [\mathbf{e}_{d+1} \cdots \mathbf{e}_M]$ represent the signal space and noise space, respectively. Subspace techniques are based on the observation that $\text{span}\{\mathbf{U}_S\} = \text{span}\{\mathbf{G}\}$. This implies the existence of a $d \times d$ full rank matrix $\mathbf{T}$, satisfying $\mathbf{U}_S = \mathbf{G}\mathbf{T}$. The signal subspace corresponding to the subarrays are defined by $\{\mathbf{U}_i = \mathbf{J}_i\mathbf{U}_S, i = 0, \ldots, p\}$. The subspaces are related as

$$\mathbf{U}_i = \mathbf{A}\mathbf{\Phi}^i\mathbf{T} = \mathbf{B}\mathbf{\Psi}^i, \quad i = 0, \ldots, p \quad (6)$$

where $\mathbf{B} = \mathbf{AT}$ and $\mathbf{\Psi} = \mathbf{T}^{-1}\mathbf{\Phi T}$. In the presence of noise and a finite data set, we can only obtain estimates $\hat{\mathbf{U}}_i$ of $\mathbf{U}_i$, and the multiple invariance problem translates to

$$\min_{\mathbf{\Psi},\mathbf{B}} \left\| \left(\begin{bmatrix} \hat{\mathbf{U}}_0 \\ \hat{\mathbf{U}}_1 \\ \vdots \\ \hat{\mathbf{U}}_p \end{bmatrix} - \begin{bmatrix} \mathbf{B} \\ \mathbf{B\Psi} \\ \vdots \\ \mathbf{B\Psi}^p \end{bmatrix} \right) \right\|_F \tag{7}$$

where the notation $\|\cdot\|_F$ indicates Frobenius norm. The problem here is to obtain estimates of the eigenvalues of $\mathbf{\Psi}$, which are the diagonal elements of $\mathbf{\Phi}$. When $p = 1$, the problem reduces to ESPRIT, and the single invariance structure can be exploited by a least squares approach [1] or a total least squares approach [4] to estimate the eigenvalues of $\mathbf{\Psi}$. However, for multiple invariance, no closed form solution exists and the proposed algorithm in [2] is to perform a Newton search with the single invariance ESPRIT solution as the starting point.

III. THE CAYLEY–HAMILTON CONSTRAINT

An alternate solution to the problem is proposed in this section by reposing the problem in a different way. The solution is based on the following observation. Let $p(\lambda) = \sum_{i=0}^d b_i \lambda^i$ be the characteristic polynomial of the $d \times d$ square matrix $\mathbf{\Psi}$. The Cayley–Hamilton theorem enforces the following constraint on $\mathbf{\Psi}$:

$$p(\mathbf{\Psi}) = b_0 \mathbf{I}_d + b_1 \mathbf{\Psi} + \cdots + b_d \Psi^d = \mathbf{0} \tag{8}$$

where $\mathbf{I}_d$ is the identity matrix of size d. This, along with (6), translates to the following constraint on the signal space

$$b_0 \mathbf{U}_0 + \cdots + b_d \mathbf{U}_d = \mathbf{0}. \tag{9}$$

Since the eigenvalues of $\mathbf{\Psi}$ can be obtained as roots of $p(\lambda)$ if the coefficients $\{b_i, i = 0, \ldots, d\}$ are known, it remains to be shown that given $\mathbf{U}_i$, the characteristic polynomial is uniquely determined from (9).

Theorem 1: Let $\mathbf{U}_i$, $\mathbf{B}$ and $\mathbf{\Psi}$ be defined as in (6). Let $p(\lambda)$ be the characteristic polynomial of $\mathbf{\Psi}$. Let $\mathcal{Z}(g)$ be the notation to denote the set of roots of a polynomial $g(\lambda)$. So the set $\mathcal{Z}(p) = \{\lambda_1, \ldots, \lambda_d\}$ is the set of eigenvalues of $\mathbf{\Psi}$. Consider the minimization

$$\{a_i\} = \arg \min_{\alpha_i, \sum \alpha_i^2 = 1} \|\alpha_0 \mathbf{U}_0 + \cdots + \alpha_d \mathbf{U}_d\|_F. \tag{10}$$

Let $a(\lambda) = a_0 + a_1\lambda + \cdots + a_d\lambda^d$. Assume that $\mathbf{\Psi}$ is diagonalizable as $\mathbf{\Psi} = \mathbf{MDM}^{-1}$. If in addition, $\mathbf{\Psi}$ has distinct eigenvalues and $\mathbf{BM}$ has no zero columns, then $\mathcal{Z}(p) = \mathcal{Z}(a)$.

Remarks:

1) The condition $\sum \alpha_i^2 = 1$ is required to avoid the trivial solution $a_i = 0$ to the minimization problem. It can be replaced by other similar conditions without affecting the conclusions of the theorem since we are concerned about the roots of the polynomial rather than the polynomial coefficients.

2) $\mathcal{Z}(p) = \mathcal{Z}(a)$ tells us that (9) implies (8). This can be expected to be true if $\mathbf{B}$ is invertible. However, the theorem presents a much weaker condition under which this

holds. The invertibilty of $\mathbf{B}$ is not required. It suffices to have the row space of $\mathbf{B}$ orthogonal to no eigen-vectors of $\mathbf{\Psi}$, which translates to the no zero column condition on $\mathbf{BM}$.

3) This implies that the theorem holds even when the rank of the matrix $\mathbf{B}$ is less than d, the number of sources. This removes the ESPRIT restriction that the invariance be present among at least d subarray elements and makes it possible to exploit invariances even among subarrays with single elements.

Proof: Let $p(\lambda) = \sum_{i=0}^d b_i \lambda^i$. Since $\{a_0, \ldots, a_d\}$ is a minimizer of (10), we have

$$\|a_0 \mathbf{U}_0 + a_1 \mathbf{U}_1 + \cdots + a_d \mathbf{U_d}\|_F$$
$$\leq \|b_0 \mathbf{U}_0 + b_1 \mathbf{U}_1 + \cdots + b_d \mathbf{U_d}\|_F = \mathbf{0}$$

where the equality follows from (9). Since the norm is nonnegative, we have

$$\|a_0 \mathbf{U}_0 + a_1 \mathbf{U}_1 + \cdots + a_d \mathbf{U_d}\|$$
$$= \|a_0 \mathbf{B} + a_1 \mathbf{B\Psi} + \cdots + a_d \mathbf{B\Psi}^d\| = \mathbf{0}$$

which implies

$$\mathbf{B}(a_0 \mathbf{I}_d + a_1 \mathbf{\Psi} + \cdots + a_d \mathbf{\Psi}^d) = \mathbf{0}. \tag{11}$$

If $\mathbf{B}$ is invertible, it immediately follows that $a(\lambda)$ is an annihilating polynomial of $\mathbf{\Psi}$. An annihilating polynomial of a square matrix is simply a polynomial that reduces the matrix to zero ([5], p. 221). It can be shown that the invertibility condition is not a requirement. To see that, diagonalize $\mathbf{\Psi}$ in (11) to obtain

$$\mathbf{BM}(a_0\mathbf{I}_d + a_1\mathbf{D} + \cdots + a_d\mathbf{D}^d)\mathbf{M}^{-1} = \mathbf{0}.$$

Since $\mathbf{M}$ is full rank, we can write this as

$$\mathbf{BM\tilde{D}} = \mathbf{0}$$

where $\tilde{\mathbf{D}} = (a_0\mathbf{I}_d + a_1\mathbf{D} + \cdots + a_d\mathbf{D}^d)$ is also diagonal. If $\mathbf{BM} = [\mathbf{c}_1 \cdots \mathbf{c}_d]$ and $\{\beta_i, i = 1, \ldots, d\}$ are the diagonal elements of $\tilde{\mathbf{D}}$, then we have

$$[\beta_1 \mathbf{c}_1 \quad \cdots \quad \beta_d \mathbf{c}_d] = \mathbf{0} \tag{12}$$

which indicates that if none of the columns of $\mathbf{BM}$ are zero, then we have

$$\tilde{\mathbf{D}} = (a_0\mathbf{I}_d + a_1\mathbf{D} + \cdots + a_d\mathbf{D}^d) = \mathbf{0}. \tag{13}$$

From this, we conclude that $a(\lambda)$ is an annihilating polynomial of $\mathbf{\Psi}$. Since $\mathbf{\Psi}$ has distinct eigenvalues, $a(\lambda)$ is the minimal polynomial of $\mathbf{\Psi}$ where the minimal polynomial is defined as the annihilating polynomial of minimal degree. Notice that the characteristic polynomial is also an annihilating polynomial of the matrix. Also notice that $a(\lambda)$ is of the same degree as the characteristic polynomial $p(\lambda)$. Since the minimal polynomial is unique (p. 240, [5]), it follows that $a(\lambda)$ and $p(\lambda)$ are the same up to a constant. This concludes the proof of the theorem. $\square$

IV. SPECC ALGORITHM AND RESULTS

A description of the SPECC algorithm based on the theorem just stated is provided in this section. We solve for the following minimization:

$$\{a_0, \ldots, a_d\} = \min_{\{\alpha_0, \ldots, \alpha_d\}} \|\alpha_0 \hat{\mathbf{U}}_0 + \cdots + \alpha_d \hat{\mathbf{U}}_d\|_F. \quad (14)$$

The minimization leads to a unique solution up to a constant. The roots of the resulting polynomial $a(x) = a_0 + a_1 x + \cdots + a_d x^d$ represented by $\{\hat{\phi}_i, i = 1, \ldots, d\}$ are approximations to $\{\phi_i, i = 1, \ldots, d\}$, the eigenvalues of $\mathbf{\Psi}$. The DOAs $\{\theta_i\}$ can be estimated now just as in the ESPRIT case. Note that (14) involves minimization of a term quadratic in the parameters, and can be accomplished by a singular value decomposition (SVD). The polynomial rooting involved in the SPECC algorithm suggests similarities to the root-MUSIC algorithm. However, SPECC exploits a structure similar to the one that ESPRIT exploits and provides a solution comparable in complexity to the ESPRIT algorithm.

Before we compare SPECC with ESPRIT, it would be prudent to list the possible advantages and disadvantages. The advantage of SPECC is its ability to exploit multiple invariances. In addition, SPECC can exploit invariances even among single element subarrays whereas ESPRIT demands at least d elements in the two subarrays that share the invariance. A possible disadvantage is that SPECC assumes the availability of d invariances. If the subarrays are constructed in such a way that there is only a single invariance, SPECC can not be applied.

To compare the two, we assumed an array of $M = 12$ equispaced sensor elements. Two sources ($d = 2$) at $\theta_1 = 10°$ and $\theta_2 = 13°$ were assumed. The array elements were assumed to be separated by $\Delta = \lambda/2$. Labeling the array elements as $\{1, 2, 3 \ldots, 12\}$, we exploit the following invariance in the two algorithms.

- ESPRIT: $\{1, 2, \ldots, 11\}$ and $\{2, 3, \ldots, 12\}$.
- SPECC: $\{1, 2, 3, 4\}$, $\{5, 6, 7, 8\}$, $\{9, 10, 11, 12\}$.

The invariance structure exploited for ESPRIT is the one with the best performance among the structures considered in [2].

Fig. 1 shows a comparison of the two schemes by plotting the root mean square values for each source against SNR. The SNR we consider is the one at each array element due to all sources. The comparison was done by considering ten snapshots of the received signal at each array element. The estimation of the signal subspaces can be expected to be quite inaccurate with just ten snapshots of each signal. The plot shows that SPECC is more robust against inaccurate subspace estimation compared to ESPRIT. The effect is pronounced at low SNRs and brings out limitations of the ESPRIT algorithm.

As the number of snapshots considered increases, better performance can be expected by the ESPRIT algorithm. Fig. 2 shows comparison of the two algorithms for different numbers of snapshots. The simulation was done at an SNR of 0 dB. Only with an increased number of snapshots does the performance of ESPRIT approach that of SPECC. A detailed comparison including comparisons with the MI-ESPRIT algorithm is available at [6].
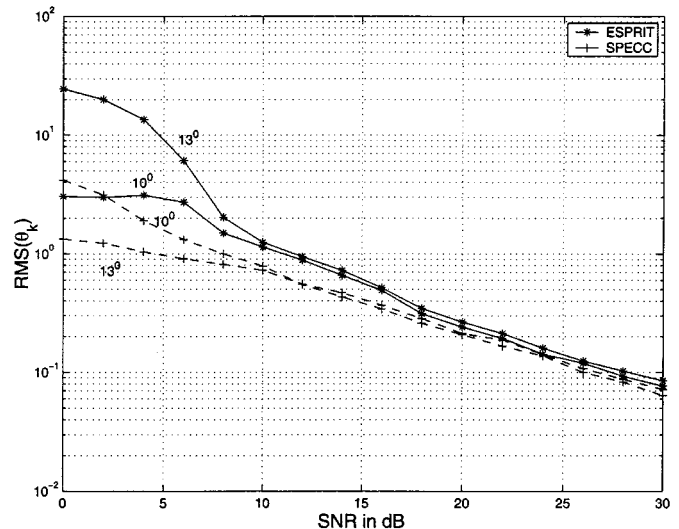

Fig. 1. Comparison of SPECC with ESPRIT: Plot of root mean square error (RMSE) for directions in degrees against SNR. $\theta = [10°\backslash 13°]$, ten snapshots.
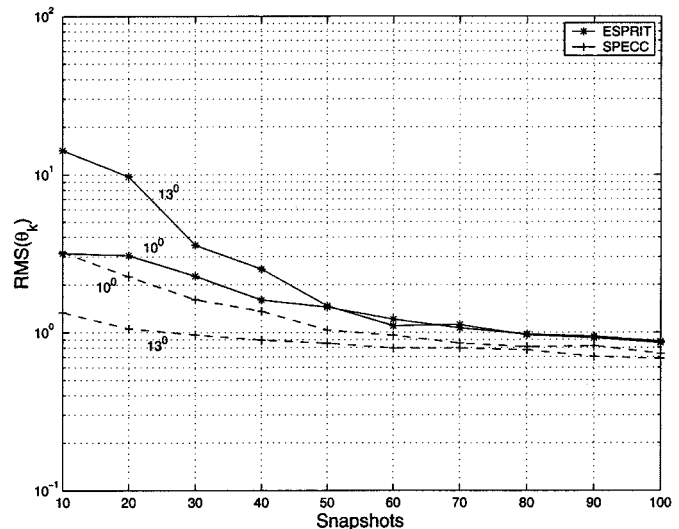

Fig. 2. Comparison of SPECC with ESPRIT for different number of snapshots, $\theta = [10°\backslash 13°]$, SNR = 0 dB.

V. CONCLUSION

The rotational invariance in an array structure has been exploited by invoking a linear constraint on the subspaces in the form of Cayley–Hamilton theorem. It is shown that the constraint leads to a closed form algorithm to exploit multiple invariance for estimation of parameters. Although SPECC has been formulated for the DOA estimation problem in this paper, it applies to a broader class of problems. Specifically, the technique can be applied to estimate the eigenvalues of $\mathbf{\Psi}$ from an invariant structure of the kind

$$\mathbf{U}_i = \mathbf{B} \mathbf{\Psi}^{k_i} \quad (15)$$

where $\{k_1, k_2, \ldots\}$ are positive integers that satisfy certain conditions. The robustness of SPECC makes it particularly attractive to those applications where there is a large number of invariances but only a small number of samples (snapshots) avail-

able. One such application is in the estimation of delays for a frequency hopping system using samples from the header part of data packets [6].

ACKNOWLEDGMENT

The authors wish to thank Dr. A. Swami for his comments and suggestions.

REFERENCES

[1] R. Roy, A. Paulraj, and T. Kailath, "ESPRIT—A subspace rotation approach to estimation of parameters of cisoids in noise," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-34, pp. 1340–1342, Oct. 1986.

[2] A. L. Swindlehurst, B. Ottersten, R. Roy, and T. Kailath, "Multiple invariance ESPRIT," *IEEE Trans. Signal Processing*, vol. SP-40, pp. 867–881, Apr. 1992.

[3] P. H. Hande, "Channel estimation for frequency hopping systems," M.S. thesis, Cornell Univ., Ithaca, NY, May 2000.

[4] R. Roy and T. Kailath, "ESPRIT—Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 37, pp. 984–995, July 1989.

[5] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*.   New York: Academic, 1985.

[6] P. H. Hande, L. Tong, and A. Swami, "Channel estimation for frequency hopping systems via multiple invariances," in *IEEE Wireless Communication and Networking Conf. 2000*, Chicago, IL, Sept. 2000.

[7] P. H. Hande and L. Tong, "Signal parameter estimation via the Cayley–Hamilton constraint," Tech. Rep, Cornell Univ., Ithaca, NY, Dec. 2000, Available online: http://www.ee.cornell.edu/~ltong/papers/HandeTongRpt.pdf.