

# SUBSPACE AVERAGING FOR SOURCE ENUMERATION IN LARGE ARRAYS

*I. Santamaria*<sup>1</sup>, *D. Ramírez*<sup>2,3</sup> and *L. L. Scharf*<sup>4</sup>

<sup>1</sup> Dept. of Communications Engineering, University of Cantabria, Spain

<sup>2</sup> Dept. of Signal Theory and Communications, University Carlos III of Madrid, Spain

<sup>3</sup> Gregorio Marañón Health Research Institute, Madrid, Spain

<sup>4</sup> Dept. of Mathematics, Colorado State University, Fort Collins, CO, USA

## ABSTRACT

Subspace averaging is proposed and examined as a method of enumerating sources in large linear arrays, under conditions of low sample support. The key idea is to exploit shift invariance as a way of extracting many subspaces, which may then be approximated by a single extrinsic average. An automatic order determination rule for this extrinsic average is then the rule for determining the number of sources. Experimental results are presented for cases where the number of array snapshots is roughly half the number of array elements, and sources are well separated with respect to the Rayleigh limit.

**Index Terms**— Array processing, Grassmann manifold, model order estimation, source enumeration, subspace averaging.

## 1. INTRODUCTION

Estimating the number of source signals received by a sensor array, which is called the source enumeration problem, is a classic problem in array signal processing, radar, and wireless communications [1, 2]. Approaches to this problem are based on information theoretic criteria, such as AIC [3,4], MDL [4,5], and BIC [6], all of which are functions of the eigenvalues of the sample covariance matrix.

A research trend for future wireless communication systems is the use of large antenna arrays with up to a few hundred antennas (e.g. massive MIMO systems) [7]. These large-scale arrays typically operate in the small sample regime for which the number of snapshots,  $N$ , is significantly smaller than the number of sensors,  $M$ . In this regime, most conventional order estimation or source enumeration techniques, which require a full-rank sample covariance matrix, provide poor results.

Using recent results from random matrix theory, Nadakuditi and Edelman proposed a version of the AIC in [8]. This work exploits the distribution of the moments of the eigenvalues of large sample covariance matrices, and introduces the concept of “effective number of identifiable signals”, which explains the undetectability of weak or closely-spaced signals using eigenvalues of the sample covariance

matrix in the small sample regime. Random matrix results were also used in [9], to derive an asymptotically consistent estimator of the sample covariance matrix with shrinkage of eigenvalues, followed by use of the minimum description length (MDL) criterion.

Another notable work on source enumeration using short data records is the exponentially embedded families (EEF) criterion [10, 11]. This approach applies a mutiparameter exponential embedding for all the models. From this embedding, the method constructs a likelihood function using sufficient statistics for the problem, and then chooses the model with maximum likelihood. From a different perspective, the exponential fitting test (EFT) is based on the observation that the profile of the ordered noise eigenvalues can be well approximated by a decaying exponential [12, 13]. The EFT method, which performs particularly well when both the number of sensors and snapshots is low, detects a break between the profile of population noise eigenvalues and the theoretical exponential profile.

In the work reported here, we assume the signal is received by a uniform linear array (ULA) for which a property called shift invariance holds [14–16]. This property means that, in a noiseless case and with  $K$  noncoherent sources, all subsets of  $L > K$  consecutive rows of the signal covariance matrix span the same  $K$ -dimensional subspace. In the noisy case the extracted subspaces are not shift-invariant, but we can apply a recently proposed technique for subspace averaging that also provides an estimate for the order of the optimal average subspace [17, 18]. The number of subspaces to average increases with the number of sensors, which makes the proposed technique competitive for scenarios with short data records and very large arrays, as our simulation results show.

**Notation.** In this paper we use  $\langle \mathbf{A} \rangle$  to denote a subspace of  $\mathbb{C}^n$  (a point on the complex Grassmann manifold), whereas  $\mathbf{A}$  is used to denote a matrix whose columns form a basis for that subspace. The superscripts  $(\cdot)^T$  and  $(\cdot)^H$  denote transpose and Hermitian, respectively. The expectation operator will be denoted by  $E[\cdot]$ , and  $\mathbf{x} \sim \mathcal{CN}_M(\mathbf{0}, \mathbf{R})$  indicates that  $\mathbf{x}$  is an  $M$ -dimensional complex circular Gaussian random vector of zero mean and covariance matrix  $\mathbf{R}$ .

## 2. SOURCE ENUMERATION IN LARGE ARRAYS

### 2.1. Problem Statement

Let us consider  $K$  narrowband signals impinging on a large, uniform, half-wavelength linear array with  $M$  antennas. The received signal is

$$\mathbf{x}[n] = [\mathbf{a}(\theta_1) \quad \cdots \quad \mathbf{a}(\theta_K)] \mathbf{s}[n] + \mathbf{e}[n] = \mathbf{A}\mathbf{s}[n] + \mathbf{e}[n], \quad (1)$$

The work of I. Santamaria has been partially supported by the Ministerio de Economía y Competitividad (MINECO) of Spain, and AEI/FEDER funds of the E.U., under grant TEC2016-75067-C4-4-R (CARMEN). The work of D. Ramírez has been partly supported by Ministerio de Economía of Spain under projects: OTOSIS (TEC2013-41718-R) and the COMONSENS Network (TEC2015-69648-REDC), by the Ministerio de Economía of Spain jointly with the European Commission (ERDF) under projects ADVENTURE (TEC2015-69868-C2-1-R) and CAIMAN (TEC2017-86921-C2-2-R), and by the Comunidad de Madrid under project CASI-CAM-CM (S2013/ICE-2845). The work of L. L. Scharf was supported by the National Science Foundation (NSF) under grant CCF-1018472.

where  $\mathbf{a}(\theta_k) = [1 \ e^{-j\theta_k} \ e^{-j\theta_k(M-1)}]^T$  is the  $M \times 1$  complex array response for the  $k$ th source,  $s_k$ , with unknown direction-of-arrival (DOA)  $\theta_k$ . The signal and noise vectors are modeled as  $\mathbf{s}[n] \sim \mathcal{CN}_K(\mathbf{0}, \mathbf{S})$  and  $\mathbf{e}[n] \sim \mathcal{CN}_M(\mathbf{0}, \sigma^2 \mathbf{I})$ , respectively. The dimensions are these:  $\mathbf{x} \in \mathbb{C}^M$ ,  $\mathbf{A} \in \mathbb{C}^{M \times K}$ ,  $\mathbf{s} \in \mathbb{C}^K$ , and  $\mathbf{e} \in \mathbb{C}^M$ . From the signal model (1), the theoretical covariance matrix is

$$\mathbf{R} = E[\mathbf{x}[n]\mathbf{x}^H[n]] = \mathbf{A}\mathbf{S}\mathbf{A}^H + \sigma^2 \mathbf{I}. \quad (2)$$

We assume there are  $N$  snapshots collected in the data matrix  $\mathbf{X} = [\mathbf{x}[1] \ \dots \ \mathbf{x}[N]]$ . The source enumeration problem consists of estimating  $K$  from  $\mathbf{X}$ .

## 2.2. Prior Work

In this subsection we review some representative methods for order estimation in the small sample regime that will be used for comparison. All methods are functions of the eigenvalues  $\lambda_1 \geq \dots \geq \lambda_M$  of the sample covariance matrix,

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{n=1}^N \mathbf{x}[n]\mathbf{x}^H[n] = \frac{1}{N} \mathbf{X}\mathbf{X}^H.$$

**LS-MDL:** The standard MDL method proposed by Wax and Kailath in [4], based on a fundamental result of Anderson [19], is

$$\hat{k}_{MDL} = \underset{0 \leq k \leq M-1}{\operatorname{argmin}} (M-k)N \log \left( \frac{a(k)}{g(k)} \right) + \frac{1}{2}k(2M-k) \log N, \quad (3)$$

where  $a(k)$  and  $g(k)$  are the geometric and the arithmetic mean, respectively, of the  $M-k$  smallest eigenvalues of  $\hat{\mathbf{R}}$ . When the number of snapshots is smaller than the number of sensors or antennas ( $N < M$ ), the sample covariance becomes rank-deficient and (3) can not be applied directly. The LS-MDL method proposed by Huang and So in [9] replaces the noise eigenvalues  $\lambda_i$  in the MDL criterion by a linear shrinkage (LS), calculated as

$$\rho_i^{(k)} = \beta^{(k)} a(k) + (1 - \beta^{(k)}) \lambda_i, \quad i = k+1, \dots, M,$$

where  $\beta^{(k)} = \min(1, \alpha^{(k)})$ , with

$$\alpha^{(k)} = \frac{\sum_{i=k+1}^M \lambda_i^2 + (M-k)^2 a(k)^2}{(N+1) \left( \sum_{i=k+1}^M \lambda_i^2 - (M-k) a(k)^2 \right)}.$$

**NE:** The method proposed by Nadakuditi and Edelman in [8], which we refer to as the NE criterion, is given by

$$\hat{k}_{NE} = \underset{0 \leq k \leq M-1}{\operatorname{argmin}} \left\{ \frac{1}{2} \left( \frac{N t_k}{M} \right)^2 \right\} + 2(k+1),$$

where

$$t_k = \left[ \frac{\sum_{i=k+1}^M \lambda_i^2}{a(k)^2 (M-k)} - \left( 1 + \frac{M}{N} \right) \right] M.$$

## 3. ORDER ESTIMATION BY SUBSPACE AVERAGING

As an alternative to the previous model order estimation techniques in the small sample regime, we propose a subspace averaging (SA) method, which exploits the shift invariant structure of uniform linear arrays in combination with the recently proposed subspace averaging technique of [17]. This method is similar to the use of subspace averaging for direction of arrival estimation in [20–22].

### 3.1. Subspace Shifting Invariance

When uniform linear arrays are used, a property called shift invariance holds, which forms the basis of the ESPRIT method [14, 15] and its many variants. Let  $\mathbf{A}_p$  be the  $L \times K$  matrix with rows  $p, \dots, p+L-1$  extracted from the steering matrix  $\mathbf{A}$ . In Matlab notation  $\mathbf{A}_p = \mathbf{A}(p:p+L-1, :)$ . Then, from (1) it is readily verified that

$$\mathbf{A}_p \operatorname{diag}(e^{-j\theta_1}, \dots, e^{-j\theta_K}) = \mathbf{A}_{p+1},$$

which is the shift invariance property. In this way,  $\mathbf{A}_p$  and  $\mathbf{A}_{p+1}$  are related by a nonsingular rotation matrix,

$$\mathbf{D} = \operatorname{diag}(e^{-j\theta_1}, \dots, e^{-j\theta_K}),$$

and therefore they span the same subspace  $\langle \mathbf{A}_p \rangle = \langle \mathbf{A}_{p+1} \rangle$  with  $\dim(\langle \mathbf{A}_p \rangle) = K < L$ . In ESPRIT, two sub-arrays of dimension  $L = M-1$  are considered, and thus we have  $\mathbf{A}_1 \mathbf{D} = \mathbf{A}_2$ , where  $\mathbf{A}_1$  and  $\mathbf{A}_2$  select, respectively, the first and the last  $M-1$  rows of  $\mathbf{A}$ .

There is an interesting characterization of the shift invariance property. Let  $\mathbf{x}_p[n]$  be an  $L \times 1$  vector containing the observations acquired by sensors  $p, \dots, p+L-1$  of  $\mathbf{x}[n]$ , and let  $S^r$  denote a shift operator, so that  $S^r \mathbf{x}_p[n] = \mathbf{x}_{p+r}[n]$ . Then, in the noise-free model  $\mathbf{x}_p[n] = \mathbf{A}_p \mathbf{s}[n]$ , this shift invariance produces

$$S^r \mathbf{x}_p[n] = S^r \mathbf{A}_p \mathbf{s}[n] = \mathbf{A}_{p+r} \mathbf{s}[n] = \mathbf{A}_p \mathbf{D}^r \mathbf{s}[n].$$

The source signal  $\mathbf{D}^r \mathbf{s}[n]$  is distributed as  $\mathbf{s}[n]$  is distributed, provided  $\mathbf{s}[n]$  is complex normal with covariance matrix  $\sigma_s^2 \mathbf{I}$ , making  $\mathbf{D}$  a measure-preserving transformation. So shift on  $\mathbf{x}_p[n]$  is measure-preserving on  $\mathbf{s}$ . This property leaves second-order matrices invariant.

Following similar arguments, it is possible to show that, when the signal covariance matrix in (2) is  $\mathbf{S} = \sigma_s^2 \mathbf{I}$ , the  $K$  principal left singular vectors of  $\mathbf{R}$  are also shift-invariant [15, 16]. When noise is present, however, the shift-invariance property does not hold for the principal eigenvectors extracted from the sample covariance matrix. The Optimal Subspace Estimation (OSE) technique proposed by Vaccaro *et al.* obtains an improved estimate of the signal subspace with the required structure (up to the first order) [16]. The OSE algorithm, which requires the number of sources  $K$  to be known, has recently been used with the subspace averaging of [17] to improve DOA estimation [20–22].

### 3.2. Order Selection by Subspace Averaging

In this section we briefly review the order estimation rule described in [17]. Let  $\langle \mathbf{G}_p \rangle$  be a subspace of dimension  $K$  in an ambient space of dimension  $L$ , and let  $\mathbf{G}_p \in \mathbb{C}^{L \times K}$  be a matrix whose columns form a unitary basis for  $\langle \mathbf{G}_p \rangle$ . The orthogonal projection matrix onto the subspace is  $\mathbf{P}_p = \mathbf{G}_p \mathbf{G}_p^H$ , which is a unique representation for a subspace in its Grassmann manifold. Suppose now that we have a collection of  $P$  subspaces  $\langle \mathbf{G}_p \rangle$ ,  $p = 1, \dots, P$ , and we want to determine the dimension of the optimal average subspace according to an extrinsic distance measure, which is given by the Frobenius norm of the difference between the respective projection matrices [23, 24].

**SA:** The subspace averaging method proposed in [17] (see also [25]) first computes the average projection matrix as

$$\bar{\mathbf{P}} = \frac{1}{P} \sum_{p=1}^P \mathbf{P}_p.$$

---

**Algorithm 1:** Subspace averaging criterion.
 

---

**Input:**  $\hat{\mathbf{R}}, L = \lfloor 3M/4 \rfloor$  and  $k_{max}$ ;  
**Output:** Order estimate  $\hat{k}_{SA}$   
**for**  $p = 1, \dots, P$  **do**  
     Extract  $\hat{\mathbf{R}}_p$  from  $\hat{\mathbf{R}}$  and obtain  $\hat{\mathbf{R}}_p = \mathbf{U}_p \boldsymbol{\Sigma}_p \mathbf{U}_p^H$   
     Compute the projection matrix  $\mathbf{P}_p = \mathbf{G}_p \mathbf{G}_p^H$ , where  
      $\mathbf{G}_p = [\mathbf{u}_{p,1}, \dots, \mathbf{u}_{p,k_{max}}]$   
     Compute  $\bar{\mathbf{P}}$  and its eigenvalues  $(\sigma_1, \dots, \sigma_L)$   
     Estimate  $\hat{k}_{SA}$  as the largest  $k$  such that  $\sigma_k > \frac{1}{2}$

---

Writing the eigenvalue decomposition (EVD) of the average projection matrix as  $\bar{\mathbf{P}} = \mathbf{F} \boldsymbol{\Sigma} \mathbf{F}^H$ , where  $\boldsymbol{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_L)$  with  $1 \geq \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_L$ , the subspace order selection rule that minimizes the extrinsic distance is [17]

$$\hat{k}_{SA} = \underset{0 \leq k \leq L}{\text{argmin}} \sum_{i=1}^k (1 - \sigma_i) + \sum_{k+1}^L \sigma_i,$$

which amounts to selecting the largest  $k$  such that  $\sigma_k > \frac{1}{2}$ .

### 3.3. Application of the Proposed Method

From the  $L \times 1$  ( $L > K$ ) sub-array snapshots  $\mathbf{x}_p[n]$  we can estimate an  $L \times L$  sample covariance as

$$\hat{\mathbf{R}}_p = \frac{1}{N} \sum_{n=1}^N \mathbf{x}_p[n] \mathbf{x}_p^H[n].$$

Each  $\hat{\mathbf{R}}_p$  block corresponds to an  $L \times L$  submatrix of the full sample covariance  $\hat{\mathbf{R}}$  extracted along its diagonal, that is,  $\hat{\mathbf{R}}_p = \hat{\mathbf{R}}(p : p + L - 1, p : p + L - 1)$ .

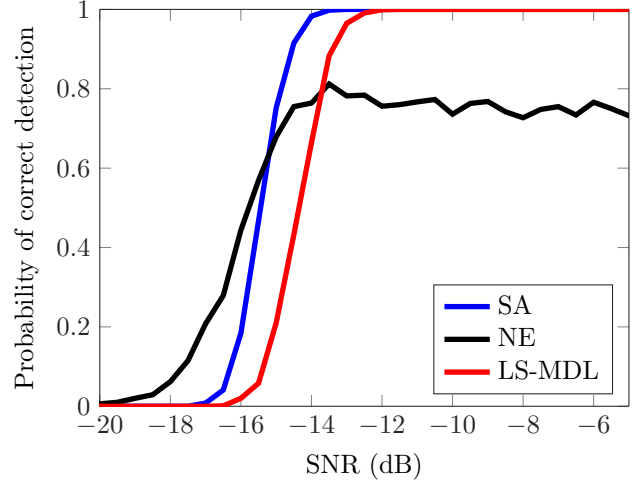
Due to the shift invariance property of uniform linear arrays, the noiseless signal subspaces of the theoretical  $\mathbf{R}_p$  are identical. Since there are  $M$  sensors and we extract  $L$ -dimensional subarrays, there are  $P = M - L + 1$  different submatrices  $\hat{\mathbf{R}}_p, p = 1, \dots, P$ . For each  $\hat{\mathbf{R}}_p$  we compute its EVD  $\hat{\mathbf{R}}_p = \mathbf{U}_p \boldsymbol{\Sigma}_p \mathbf{U}_p^H$ , and then extract a subspace formed by the  $k_{max}$  principal eigenvectors

$$\mathbf{G}_p = [\mathbf{u}_{p,1} \quad \dots \quad \mathbf{u}_{p,k_{max}}] \in \mathbb{C}^{L \times k_{max}},$$

where  $k_{max} < \min(L, N)$  is an overestimate of the maximum number of sources that we expect in our problem. From the collection of subspaces with unitary basis  $\mathbf{G}_p, p = 1, \dots, P$ , the order estimation method described in Sec. 3.2 can be applied. A summary of the proposed algorithm is shown in Algorithm 1.

Notice that the only parameters in the method are the dimension of the subarrays,  $L$ , and the dimension of the extracted subspaces,  $k_{max}$ . For the results in this paper, we have used  $L = \lfloor 3M/4 \rfloor$ , which seems to provide good performance for a wide range of values. In fact, as long as each extracted subspace contains a large common portion of the signal subspace and (more or less) independent portions of the noise subspace, then, the averaging procedure enhances signal coordinates while averaging out noise coordinates.

A final comment regarding the computational cost is in order. While the computational complexity of the order estimation methods reviewed in Sec. 2.2 is roughly  $\mathcal{O}(M^3)$  due to the EVD of  $\hat{\mathbf{R}}$ , the proposed SA technique requires  $(P + 1)\mathcal{O}(L^3)$  operations since we need to perform  $P + 1$  EVDs of the  $L \times L$  matrices  $\bar{\mathbf{P}}$  and  $\hat{\mathbf{R}}_p, p = 1, \dots, P$ .



**Fig. 1.** Probability of correct detection vs. SNR for all methods when  $K = 3, \Delta_\theta = 10^\circ, M = 100$ , and  $N = 60$ .

## 4. SIMULATION RESULTS

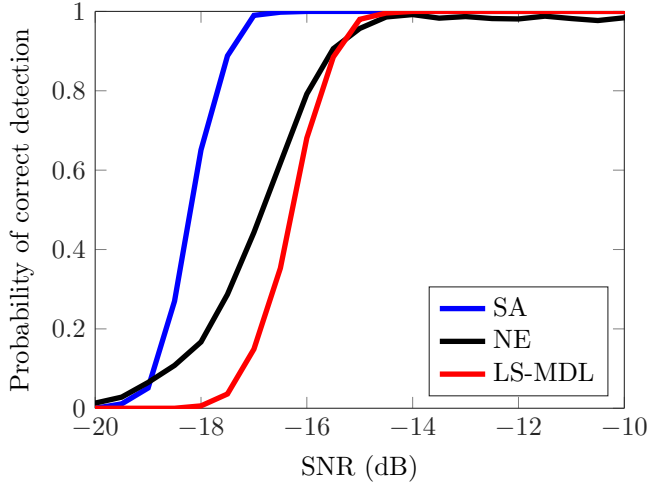
In this section we compare the performance of the proposed subspace averaging (SA) technique with LS-MDL and NE. Notice that neither LS-MDL nor NE exploits the shift invariance for this problem.

We consider a scenario with  $K$  narrowband incoherent unit-power signals and DOAs separated  $\Delta_\theta$  impinging on a uniform linear array with  $M$  antennas and half-wavelength element separation. For almost all examples,  $M = 100$  antenna elements, for which the Rayleigh limit to resolution is approximately 1.2 degrees. For all experiments in this section, the SA method extracts subspaces of ambient dimension  $L = \lfloor 3M/4 \rfloor$ , and the maximum number of sources is fixed at  $k_{max} = 7$  for Experiments 1-4 and at  $k_{max} = 15$  for Experiment 5. In other words, the subspaces to average are points on the Grassmann manifold  $\mathbb{G}(k_{max}, L)$ .

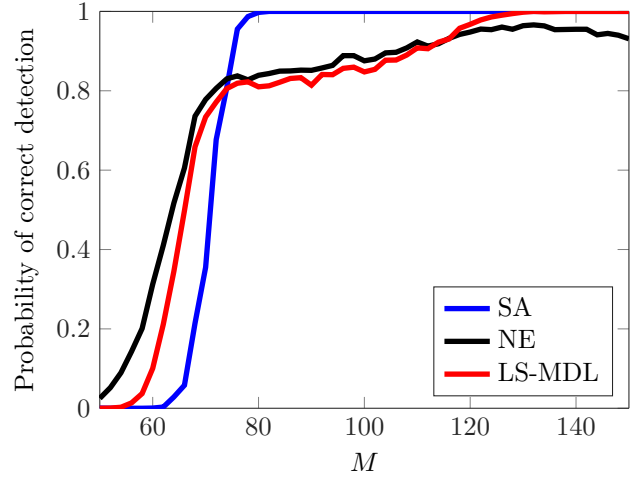
**Experiment 1:** In the first example we consider an array with  $M = 100$  antennas receiving  $K = 3$  sources separated  $\Delta_\theta = 10^\circ$ , and  $N = 60$  snapshots, thus yielding a rank-deficient sample covariance matrix. Fig. 1 shows the probability of correct detection vs. the signal-to-noise-ratio (SNR) for all methods. For this scenario with a large array, very few snapshots, and widely separated sources, the SA method outperforms LS-MDL. For very low SNR, the NE performs well, but it frequently overestimates the number of sources for higher values of the SNR.

**Experiment 2:** Increasing the number of snapshots to  $N = 150$  and keeping fixed the rest of the parameters, we obtain the results shown in Fig. 2. The advantages of the SA method become more evident for this scenario.

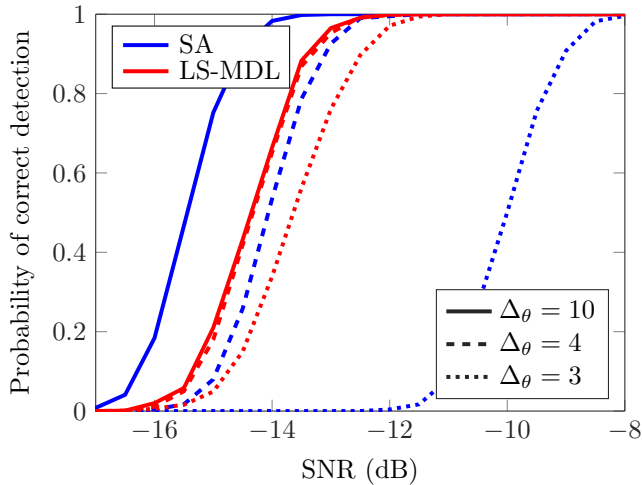
**Experiment 3:** To analyze the impact of the separation between sources, we consider again a scenario with  $M = 100$  antennas,  $N = 60$  snapshots,  $K = 3$  sources. But now sources are separated at angles of  $\Delta_\theta = \{10^\circ, 4^\circ, 3^\circ\}$ . We restrict the comparison to the LS-MDL criterion. The results are shown in Fig. 3. When the separation between sources is small (high-resolution scenario), the SA method typically underestimates the true number of sources, whereas the LS-MDL criterion provides a more accurate estimate. Clearly, the smaller the resolution  $\Delta_\theta$ , the higher the collinearity between the steering vectors. In this situation, the SA method, as



**Fig. 2.** Probability of correct detection vs. SNR for all methods when  $K = 3$ ,  $\Delta_\theta = 10^\circ$ ,  $M = 100$ , and  $N = 150$ .



**Fig. 4.** Probability of correct detection vs. number of antennas for the SA, NE and LS-MDL methods when  $K = 3$ ,  $\Delta_\theta = 5^\circ$ ,  $N = 150$ , and SNR is  $-15$  dB.

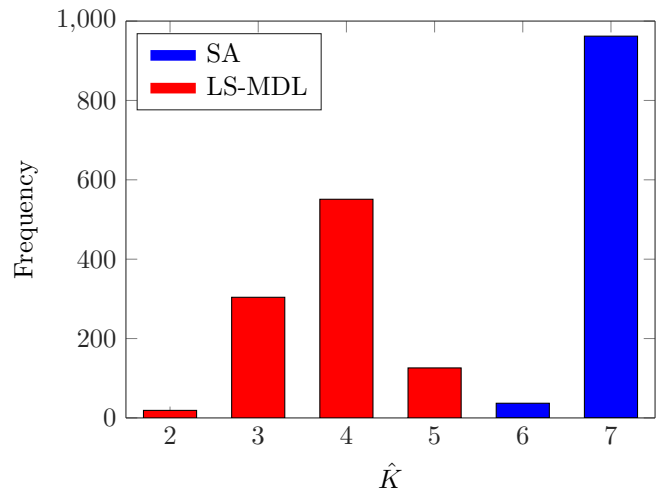


**Fig. 3.** Probability of correct detection vs. SNR for the SA and LS-MDL methods when  $K = 3$ ,  $\Delta_\theta = \{10^\circ, 4^\circ, 3^\circ\}$ , and  $N = 60$ .

configured here, tends to underestimate the true dimension of the central subspace. This problem can be mitigated by adjusting the subspace ambient dimension  $L = \lfloor 3M/4 \rfloor$ . This is a trade-off that will be analyzed in future work.

**Experiment 4.** In Fig. 4 we compare the performance of the SA, the NE and the LS-MDL criteria for an increasing number of antennas when  $N = 150$  snapshots,  $K = 3$  sources separated  $\Delta_\theta = 5^\circ$ , and SNR =  $-15$  dB. Notice that since we are using for the SA criterion a fixed ambient dimension  $L = \lfloor 3M/4 \rfloor$ , the number of subspaces to average  $P = M - L + 1$  increases with the number of antennas, which is beneficial for the SA method.

**Experiment 5.** In the final experiment, we consider a scenario with  $K = 7$  sources widely separated at  $\Delta_\theta = 10^\circ$ ,  $M = 100$  antennas,  $N = 60$  snapshots and SNR =  $-15$  dB. Fig. 5 depicts the histogram of the estimated order for the SA and the LS-MDL methods. In this scenario, the LS-MDL underestimates the true value, whereas the proposed SA criterion does not.



**Fig. 5.** Histogram of the estimated order for SA and LS-MDL when  $K = 7$ ,  $\Delta_\theta = 10^\circ$ , SNR =  $-15$  dB,  $M = 100$  and  $N = 60$ .

## 5. CONCLUSIONS

This paper addresses the problem of source enumeration in conditions of low sample support. Subspace averaging, based on shift-invariance in linear, equally-spaced arrays, is used as a method to enhance signal coordinates and mitigate the effects of low sample support. Experimental results for subspace averaging are promising for enumerating sources that are widely separated with respect to the Rayleigh limit. The problem of enumerating sources closely spaced with respect to the Rayleigh limit, in conditions of low sample support and low SNR remains open. In fact, the open question is: “at what sample supports and SNRs *can* closely spaced sources be enumerated? This is a question of performance bounding and threshold effects for source enumeration.

## 6. REFERENCES

- [1] H. V. Trees, *Detection, Estimation and Modulation Theory Part IV: Optimum Array Processing*, New York, NJ, Wiley, 2002.
- [2] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*, Reading, MA, Addison-Wesley, 1991.
- [3] H. Akaike, “A new look at the statistical model identification,” *IEEE Trans. Autom. Control*, vol. 19, no. 6, pp. 716–723, 1974.
- [4] M. Wax and T. Kailath, “Detection of signals by information theoretic criteria,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 2, pp. 387–392, 1985.
- [5] J. Rissanen, “Modeling by shortest data description,” *Automatica*, vol. 14, pp. 465–471, 1978.
- [6] Z. Lu and A. M. Zoubir, “Generalized Bayesian information criterion for source enumeration in array processing,” *IEEE Trans. Signal Process.*, vol. 61, no. 6, pp. 1470–1480, 2013.
- [7] F. Rusek, *et al.*, “Scaling up MIMO: Opportunities and challenges with very large arrays,” *IEEE Signal Process. Mag.*, vol. 40, no. 1, pp. 40–60, 2013.
- [8] R. R. Nadakuditi and A. Edelman, “Sample eigenvalue based detection of high-dimensional signals in white noise with relatively few samples,” *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2625–2638, 2008.
- [9] L. Huang and H. C. So, “Source enumeration via MDL criterion based on linear shrinkage estimation of noise subspace covariance matrix,” *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4806–4821, 2013.
- [10] S. Kay, “Exponentially embedded families: New approaches to model order estimation,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 1, pp. 333–345, 2005.
- [11] C. Xu and S. Kay, “Source enumeration via the EEF criterion,” *IEEE Signal Process. Letters*, vol. 15, pp. 569–572, 2008.
- [12] A. Quinlan, J. P. Barbot, P. Larzabal, and M. Haardt, “Model order selection for short data: an exponential fitting test (EFT),” *EURASIP J. Adv. Signal Process.*, 2007.
- [13] J. P. C. L. da Costa, M. Haardt, and F. Römer, G. Del Galdo, “Enhanced model order estimation using higher-order arrays,” *Proc. 40th Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Nov. 2007.
- [14] A. Paulraj, R. Roy, and T. Kailath, “A subspace rotation approach to signal parameter estimation,” *Proc. of the IEEE*, vol. 74, pp. 1044–1045, 1986.
- [15] R. Roy, and T. Kailath, “ESPRIT- Estimation of signal parameters via rotational invariance techniques,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, 1989.
- [16] F. Li, and R. J. Vaccaro, “Unified analysis for DOA estimation algorithms in array signal processing,” *Signal Process.*, vol. 25, pp. 147–169, 1991.
- [17] I. Santamaria, L. L. Scharf, C. Peterson, M. Kirby, and J. Francos, “An order fitting rule for optimal subspace averaging,” *Proc. IEEE Work. Stat. Signal Process. (SSP)*, Palma de Mallorca, June, 2016.
- [18] Z. Yavo, J. M. Francos, I. Santamaria, and L. L. Scharf, “Estimating the mean manifold of a deformable object from noisy observations,” *Proc. IEEE Image Video and Multidimensional Signal Process. Work. (IVMSP)*, Bordeaux, France, July, 2016.
- [19] T. W. Anderson, “Asymptotic theory for principal component analysis,” *Ann. J. Math. Stat.*, vol. 34, pp. 122–148, 1963.
- [20] R. J. Vaccaro, “The role of subspace estimation in sensor array signal processing,” *Proc. 2017 Underwater Acoust. Signal Process. Work.*, University of Rhode Island, Oct. 4-7, 2017.
- [21] T. A. Palka, “Bounds and algorithms for subspace estimation on Riemannian quotient submanifolds,” Ph.D. Dissertation, University of Rhode Island, 2016.
- [22] T. A. Palka and R. J. Vaccaro, “A differential geometry-based framework for constrained subspace estimation,” *Proc. 2017 Underwater Acoust. Signal Process. Work.*, University of Rhode Island, Oct. 4-7, 2017.
- [23] A. Srivastava and E. Klassen, “Monte Carlo extrinsic estimators of manifold-valued parameters,” *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 299–308, Aug. 2002.
- [24] A. Edelman, T. Arias, and S. T. Smith, “The geometry of algorithms with orthogonality constraints,” *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 2, pp. 303–353, 1998.
- [25] F. Hlawatsch and W. Kozek, “Time-frequency projection filters and time-frequency signal expansions,” *IEEE Trans. Signal Process.*, vol. 42, no. 12, pp. 3321–3334, Dec. 1994.