SPARSE SUBSPACE AVERAGING FOR ORDER ESTIMATION

Vaibhav Garg[1], David Ramírez[2,3] and Ignacio Santamaria[1]

[1]Dept. of Communications Engineering, Universidad de Cantabria, Spain
[2] Dept. of Signal Theory and Communications, Universidad Carlos III de Madrid, Spain
[3] Gregorio Marañón Health Research Institute, Spain
Email: vaibhav.garg@unican.es, david.ramirez@uc3m.es, i.santamaria@unican.es

ABSTRACT

This paper addresses the problem of source enumeration for arbitrary geometry arrays in the presence of spatially correlated noise. The method combines a sparse reconstruction (SR) step with a subspace averaging (SA) approach, and hence it is named sparse subspace averaging (SSA). In the first step, each received snapshot is approximated by a sparse linear combination of the rest of snapshots. The SR problem is regularized by the logarithm-based surrogate of the ℓ_0-norm and solved using a majorization-minimization approach. Based on the SR solution, a sampling mechanism is proposed in the second step to generate a collection of subspaces, all of which approximately span the same signal subspace. Finally, the dimension of the average of this collection of subspaces provides a robust estimate for the number of sources. Our simulation results show that SSA provides robust order estimates under a variety of noise models.

Index Terms— Array processing, source enumeration, sparse representation, subspace averaging.

1. INTRODUCTION

Detecting the number of signals received by an array of sensors is a crucial problem in many applications, such as wireless communications, radar, and biomedical and geophysical signal processing [1,2]. In array processing, this order estimation problem is known as source enumeration and boils down to the estimation of the array manifold dimension. Most common order estimation approaches are based on information-theoretic criteria [3–6], which minimize functions of the eigenvalues of the sample covariance matrix (SCM) penalized by a term that quantifies the model complexity. Many variants of these methods, based on random matrix theory, have been introduced to address the problem in the small sample regime where the number of observations is on the order of or smaller than the number of sensors [7–9]. These methods, however, are typically designed for spatially white noise and do not perform well under non-white or correlated noise. Source enumeration techniques when the noise samples have an unknown covariance matrix have recently been proposed in [10, 11], but these methods require that the noise be sufficiently weaker than the signal.

Recently, [12, 13] proposed a source enumeration technique for white noise based on subspace averaging (SA). The method was developed for uniform linear arrays (ULAs), as it exploits the shift invariance property (SIP) to generate multiple subspaces. In particular, the SIP allows us to extract submatrices from the SCM that span the same signal subspace in noiseless conditions. Then, averaging these subspaces is a mechanism that enhances the signal subspace, while cancelling out the noise subspace. In the end, the dimension of the average subspace yields an estimate the number of sources. In [14], we extended the technique to account for independent but not identically distributed (non-iid) noise models. However, this method still exploits the SIP, and it is therefore only applicable to ULAs. Moreover it entails a large computational cost.

This paper extends the SA enumeration technique to arrays with arbitrary geometry, referred to here as non-uniform linear arrays (NULAs), and to noise with arbitrary covariance matrix. The method first approximates each snapshot by a sparse linear combination of the rest of the received snapshots. In a noiseless situation, the number of nonzero coefficients of the sparse expansion would directly reveal the signal subspace dimension. In a noisy situation, the sparse expansions can be used to generate a collection of subspaces that, when averaged, will provide the order estimate using the SA method in [13]. This method, which combines sparse reconstruction with subspace averaging, is termed sparse subspace averaging (SSA).

The sparse reconstruction (SR) problem is solved by generalizing to complex-valued signals the well-known logarithm-based surrogate of the ℓ_0-norm [15]. Using the majorization-minimization (MM) framework [16], we show that the SR problem amounts to solving a reweighted regularized least squares (LS) problem. Based on the SR solution, a sampling mechanism is proposed to generate a collection of random subspaces, all of which approximately span the same signal subspace. Finally, an average or central subspace is estimated as in [13], and the dimension of this average provides a robust estimate for the number of sources impinging on the array. Sparse reconstruction approaches have been used before in array signal processing, but mainly for direction-of-arrival (DOA) estimation problems [17–19]. Differently from these works, in this paper we use sparse reconstruction methods as a means to generate multiple approximations to the signal subspace that, when averaged, provide accurate estimates of the underlying signal subspace rank.

Our simulation results show that SSA provides robust order estimates under a variety of noise models, ranging from uncorrelated noise with different variances to noise with arbitrary covariance matrices. Moreover, unlike the method in [13], SSA can be applied to any array geometry, not necessarily a ULA.

2. PROBLEM STATEMENT

Let us consider K narrowband signals impinging on an array of arbitrary geometry composed of M antennas. The received signal is

$$\mathbf{x}[n] = \mathbf{H}\mathbf{s}[n] + \mathbf{e}[n], \qquad (1)$$

where $\mathbf{H}$ is the $M \times K$ unknown multiple input multiple output (MIMO) channel. The signals are assumed to be uncorrelated and are modelled as $\mathbf{s}[n] \sim \mathcal{CN}_K(\mathbf{0}, \boldsymbol{\Psi})$, where $\boldsymbol{\Psi}$ is a diagonal matrix and the noise is $\mathbf{e}[n] \sim \mathcal{CN}_M(\mathbf{0}, \mathbf{R}_n)$, where $\mathbf{R}_n$ is the noise covariance matrix. In this paper, we do not assume any particular structure (scale identity, diagonal) for the noise covariance matrix $\mathbf{R}_n$.

From (1), the covariance matrix of the received signal is

$$\mathbf{R} = \mathbb{E}\left[\mathbf{x}[n]\mathbf{x}^H[n]\right] = \mathbf{R}_s + \mathbf{R}_n,$$

where $\mathbf{R}_s = \mathbf{H}\boldsymbol{\Psi}\mathbf{H}^H$ is the signal covariance matrix. Note that since the MIMO channel $\mathbf{H}$ is unknown, we can assume without loss of generality that $\boldsymbol{\Psi} = \mathbf{I}$.

If N snapshots are collected in the data matrix $\mathbf{X} = [\mathbf{x}[1], \ldots, \mathbf{x}[N]]$, then the sample covariance matrix is

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

and the source enumeration or order estimation problem consists in estimating K from $\mathbf{X}$ or $\hat{\mathbf{R}}$.

3. SPARSE SUBSPACE AVERAGING

The key idea of subspace averaging for order estimation is to average subspaces that contain correlated (or ideally identical) versions of the signal subspace, but uncorrelated portions of the noise subspace. For ULAs, these subspaces can be obtained from consecutive subarrays by exploiting the shift invariance property [12,13,20]. When the MIMO channel $\mathbf{H}$ has no particular structure, however, a different approach to generate these subspaces is needed. The SSA method generates the subspaces from the sparse reconstructions of each snapshot, as we describe in the following sections.

3.1. Sparse Representation

In a noiseless situation, a basis for the K-dimensional signal subspace can be constructed from K randomly chosen columns of the data matrix $\mathbf{X}$. This means that each snapshot can be represented as linear combination of other K snapshots that lie in the same subspace. Note that the K snapshots will be linearly independent with probability one. In a noisy situation, such an expansion will not exist in general, but still we can find a sparse representation of each snapshot in terms of the rest of snapshots. The snapshots selected by the sparse reconstruction algorithm will allow us to build (or generate) approximate basis for the signal subspace. Note also that the number of sources is typically much smaller than the number of snapshots $K \ll N$. This is the idea behind the sparse reconstruction stage in SSA. Similar ideas have successfully been applied to sparse subspace clustering problems in [21].

The sparse expansion for each snapshot is obtained by solving

$$\min_{\boldsymbol{\alpha}_n} \|\boldsymbol{\alpha}_n\|_0 \quad \text{subject to} \quad \mathbf{x}[n] = \mathbf{X}_n \boldsymbol{\alpha}_n \qquad (2)$$

Algorithm 1: Sparse Reconstruction Algorithm
Input: $\mathbf{X}_n$, $\mathbf{x}[n]$, λ, δ **Output:** $\hat{\boldsymbol{\alpha}}_n$ **Initialization**: $\hat{\boldsymbol{\alpha}}_n = \mathbf{X}_n^+ \mathbf{x}[n]$, where $(\cdot)^+$ is the Moore-Penrose pseudoinverse   **REPEAT**   Compute $\mathbf{D}$ using (5)   Compute $\hat{\boldsymbol{\alpha}}_n$ using (6)   **Until** Convergence = true

where $\mathbf{X}_n$ is the data matrix $\mathbf{X}$ after removing the nth column/snapshot and $\boldsymbol{\alpha}_n = [\alpha_{n,1}, \ldots, \alpha_{n,N-1}]^T$ are the coefficients of the sparsest linear combination obtained for the nth snapshot. However, the optimization problem in (2) is NP-hard. One well-known solution to this problem is to replace the ℓ_0-norm with the ℓ_1-norm [22], which can be solved by the least absolute shrinkage and selection operator (LASSO) optimization algorithm [23]. Nevertheless, the ℓ_1-norm surrogate function can sometimes perform poorly and other proxies for the ℓ_0-norm are preferred, such as the logarithm of the absolute value [15]. Moreover, the equality constraint in (2) is not adequate for low signal-to-noise-ratio (SNR) scenarios. Hence, in this paper the SR will be obtained as the solution to

$$\min_{\boldsymbol{\alpha}_n} \|\mathbf{x}[n] - \mathbf{X}_n \boldsymbol{\alpha}_n\|_2^2 + \lambda \sum_{i=1}^{N-1} \log(|\alpha_{n,i}| + \delta) \qquad (3)$$

which is the Lagrangian form of (2) where we have substituted the ℓ_0-norm by the log-surrogate and the equality by an inequality. We propose to solve (3) by using an MM-based approach [16]. The majorizer of the cost function in (3) is based on a first-order Taylor series of the logarithm, that is,

$$\sum_{i=1}^{N-1} \log(|\alpha_{n,i}| + \delta) = \sum_{i=1}^{N-1} \log\left(\sqrt{|\alpha_{n,i}|^2} + \delta\right)$$

$$\leq \sum_{i=1}^{N-1} \frac{1}{2\left|\alpha_{n,i}^{(t)}\right|^2 + 2\delta\left|\alpha_{n,i}^{(t)}\right|} \left(\left|\alpha_{n,i}\right|^2 - \left|\alpha_{n,i}^{(t)}\right|^2\right)$$

where $\alpha_{n,i}^{(t)}$ is the ith component of the solution at the tth iteration. Hence, the solution at the $(t+1)$th iteration is computed from

$$\min_{\boldsymbol{\alpha}_n} \|\mathbf{x}[n] - \mathbf{X}_n \boldsymbol{\alpha}_n\|_2^2 + \lambda \boldsymbol{\alpha}_n^H \mathbf{D} \boldsymbol{\alpha}_n, \qquad (4)$$

where $\mathbf{D} = \text{diag}(d_1, \ldots, d_N)$, with

$$d_i = \frac{1}{2\left|\alpha_i^{(t)}\right|^2 + 2\delta\left|\alpha_i^{(t)}\right| + \delta}, \qquad (5)$$

where a small value δ has been added to the denominator to avoid divisions by zero. Now, we can find a closed-form solution to (4) by taking the derivative and setting it to zero, which yields

$$\boldsymbol{\alpha}_n^{(t+1)} = \left(\mathbf{X}_n^H \mathbf{X}_n + \lambda \mathbf{D}\right)^{-1} \mathbf{X}_n^H \mathbf{x}[n]. \qquad (6)$$

The sparse reconstruction algorithm is shown in Algorithm 1.

3.2. Random Generation of Subspaces

In this section we develop a method to randomly generate subspaces based on the SR obtained in Section 3.1, which follows the lines

---

Algorithm 2: Generation of a Random Subspace

Input: $\mathbf{X}_n, \boldsymbol{\alpha}_n, \mu$
Initialization: $\tilde{\mathbf{X}}_n = \mathbf{X}_n$, $\mathbf{A}_{nt} = \emptyset$, and $L = N - 1$
Compute $\boldsymbol{\gamma}_n$ from (7)
Compute k_{max} as the number of $|\alpha_{i,n}|$ greater than
$\quad \mu \max_i(|\alpha_{i,n}|)$
Compute $\bar{\gamma}_n = \frac{1}{2} \sum_{i=1}^{N-1} \gamma_{n,i}$
while rank$(\mathbf{A}_{nt}) \leq k_{max}$ *or* $\sum_i \gamma_{n,i} \geq \bar{\gamma}_n$ **do**
$\quad$ /* Sample from $\mathcal{D}(\tilde{\mathbf{X}}_n, \boldsymbol{\gamma}_n)$ */
$\quad$ $\mathbf{g} = [g_1, \ldots, g_L]^T$, with $g_i \sim \mathcal{U}(0, 1)$
$\quad$ $\mathcal{I} = \{i \mid g_i \leq \gamma_{n,i}\}$ and $\overline{\mathcal{I}} = \{i \mid g_i > \gamma_{n,i}\}$
$\quad$ $\mathbf{G} = \tilde{\mathbf{X}}_n(:, \mathcal{I})$
$\quad$ /* Append new submatrix */
$\quad$ $\mathbf{A}_{nt} = \begin{bmatrix} \mathbf{A}_{nt} & \mathbf{G} \end{bmatrix}$
$\quad$ /* Eliminate selected directions */
$\quad$ $\tilde{\mathbf{X}}_n = \tilde{\mathbf{X}}_n(:, \overline{\mathcal{I}})$, $\boldsymbol{\gamma}_n = \boldsymbol{\gamma}_n(\overline{\mathcal{I}})$, and $L = |\overline{\mathcal{I}}|$
Output: Projection matrix onto the subspace
$\quad \mathbf{P}_{nt} = \mathbf{A}_{nt}(\mathbf{A}_{nt}^H \mathbf{A}_{nt})^{-1} \mathbf{A}_{nt}^H$

---

in [13]. In noisy conditions, it is important for the SA method that the extracted subspaces contain a large common portion of the signal subspace and independent portions of the noise subspace. This translates into an averaging procedure that enhances signal coordinates while averaging out noise coordinates. Concretely, the subspace associated to the nth snapshot is constructed by randomly selecting columns from $\mathbf{X}_n$, with probabilities proportional to the sparse coefficients $\boldsymbol{\alpha}_n$ as

$$\gamma_{n,i} = \frac{|\alpha_{n,i}|}{\sum_{j=1}^{N-1} |\alpha_{n,j}|}. \tag{7}$$

This process can be interpreted as sampling from a discrete distribution that we denote as $\mathcal{D}(\mathbf{X}_n, \boldsymbol{\gamma}_n)$, where $\boldsymbol{\gamma}_n = [\gamma_{n,1}, \ldots, \gamma_{n,N-1}]^T$ in (7) are the concentration parameters of the distribution. To shed some light on this distribution, let us explain the experiment that determines $\mathcal{D}$. Let us take for example the snapshot $n = N$ and let $\mathbf{X}_N = [\mathbf{x}[1], \ldots, \mathbf{x}[N-1]]$ be the matrix with snapshots $n = 1, \ldots, N - 1$. Solving the SR problem, we get the coefficients $\boldsymbol{\alpha}_N$ of the sparse expansion of $\mathbf{x}[N]$ in terms of the rest of snapshots. The sparse coefficients, when properly normalized, give us the probabilities $(\gamma_{N,1}, \ldots, \gamma_{N,N-1})$. The sampling experiment to generate a random subspace is as follows. Draw 1 includes $\mathbf{x}[1]$ with probability $\gamma_{N,1}$, and excludes it with probability $(1 - \gamma_{N,1})$; draw 2 includes $\mathbf{x}[2]$ with probability $\gamma_{N,2}$, and excludes it with probability $(1 - \gamma_{N,2})$; and so on. With the choice for $\mathcal{D}(\mathbf{X}_n, \boldsymbol{\gamma}_n)$, the probability of picking the ith snapshot from X_n is proportional to $|\alpha_{n,i}|$. Therefore, the snapshots associated to the higher values of $|\alpha_{n,i}|$ should be chosen more often. For more details, on this interpretation we refer the reader to [13].

Finally, the subspaces are the column spaces of the matrices iteratively constructed as shown in Algorithm 2. This algorithm considers two thresholds: 1) k_{max}, which is the maximum dimension of the signal subspace (in practice an overestimate of K), and is selected as the number of $|\alpha_{i,n}|$ greater than $\mu \max_i(|\alpha_{i,n}|)$, with $0 < \mu < 1$; and 2) $\bar{\gamma}_n$, which is a minimum value for the probabilities in $\boldsymbol{\gamma}_n$, and is chosen as $\bar{\gamma}_n = \frac{1}{2} \sum_{i=1}^{N-1} \gamma_{n,i}$.

Algorithm 3: Sparse Subspace Averaging

Input: $\mathbf{X}, d_{max}$
Output: Order estimate $\hat{K}_{SSA}$
for $n = 1, \ldots, N$ **do**
$\quad$ Find $\mathbf{X}_n$ by removing the nth column from $\mathbf{X}$
$\quad$ Compute $\hat{\boldsymbol{\alpha}}_n$ using Algorithm 1
$\quad$ Generate T random projection matrices using
$\quad\quad$ Algorithm 2
Compute $\overline{\mathbf{P}}$ using (8)
Obtain $\hat{K}_{SSA}$ using (9)

3.3. Subspace Averaging (SA) Criterion

The random generation procedure can be repeated T times to generate T subspaces for each value of $\boldsymbol{\alpha}_n, n = 1, \ldots, N$. Therefore, we get a total of NT orthogonal projection matrices to be used in the SA procedure. The average projection matrix is

$$\overline{\mathbf{P}} = \frac{1}{NT} \sum_{n=1}^{N} \sum_{t=1}^{T} \mathbf{P}_{nt}, \tag{8}$$

Once we get $\overline{\mathbf{P}}$, its eigenvalues $1 \geq k_1 \geq k_2 \geq \cdots \geq k_{d_{max}}$ are used to determine number of sources as

$$\hat{K}_{SSA} = \operatorname*{argmax}_{1 \leq k \leq d_{max}} k_k - k_{k+1}. \tag{9}$$

Basically, the criterion in (9) detects the gap between the signal and noise eigenvalues. It is noted here that to avoid numerical issues and to save computational cost, we are considering only the largest d_{max} eigenvalues of $\overline{\mathbf{P}}$, where $d_{max} \gg K$ is an overestimation of K. A summary of the proposed method is shown in Algorithm 3.

4. SIMULATION RESULTS

In this section, the performance of the proposed method is compared with some representative methods for different noise models. As a figure of merit we use the probability of correct detection (P_D), which is the probability that $\hat{K}_{SSA} = K$. For all simulations we assume that K uncorrelated narrowband equal-power signals are impinging on a NULA with M antennas. The signal-to-noise-ratio is defined as SNR $= 10 \log \frac{\text{tr}(\mathbf{R}_s)}{\text{tr}(\mathbf{R}_n)}$, and 2000 Monte Carlo simulations are averaged for all the results. We found that $\lambda = \frac{M}{2N}$, $T = 20$, $\mu = 0.1$, and $d_{max} = \frac{M}{5}$ provide in general good performance over a wide range of scenarios. Since δ is a small constant that is used to avoid numerical issues, we select a $\delta = 10^{-15}$.

In the first experiment, we assume that the noise is uncorrelated across antennas with different variances at each sensor. Therefore, the noise covariance matrix is $\mathbf{R}_n = \text{diag}\left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_M^2\right)$, where σ_m^2 is the noise variance at the mth sensor. The noise variances are modeled as uniformly distributed independent random variables: $\sigma_m^2 \sim \mathcal{U}[\sigma^2(1 - \epsilon), \sigma^2(1 + \epsilon)]$, where σ^2 is a common noise variance and $0 \leq \epsilon \leq 1$ allows us to control the spatial non-whiteness of the noise. Note that noise would be white for $\epsilon = 0$. As methods for comparison, we select LSMDL [8] and BIC [9], which are designed for white noise, the MDL criterion adapted to uncorrelated noises [14, 24–26] (labeled in the figures as MDL-unc), and the linear-regression method designed for colored noise in [10]. Fig. 1 shows the P_D vs. ϵ for $M = 70$, $N = 100$, $K = 5$, and SNR
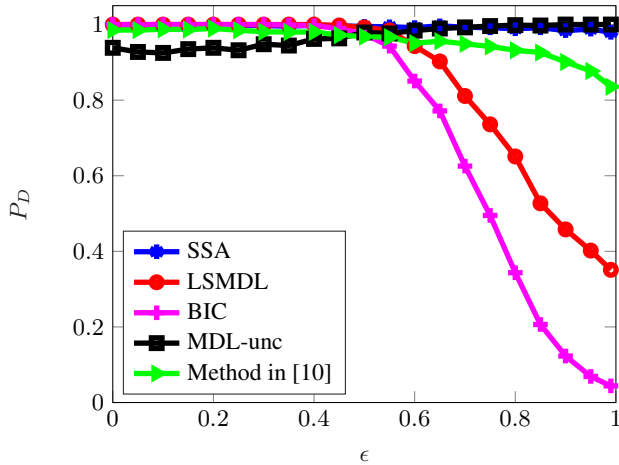
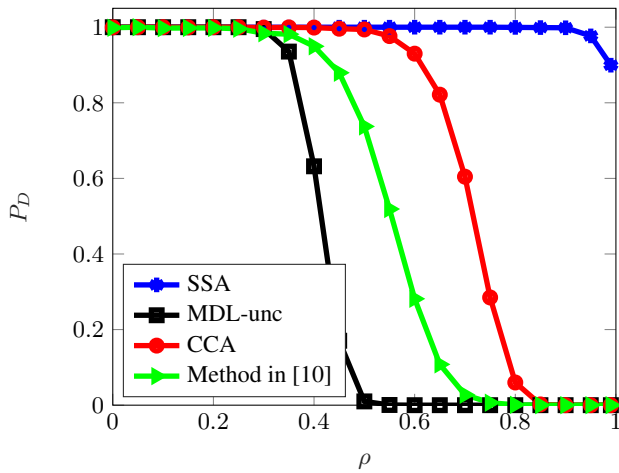**Fig. 1**: P_D vs. ϵ for diagonal noise in a scenario with $M = 70$, $N = 100$, $K = 5$, and SNR $= -2$ dB.

Fig. 3: P_D vs. N for an arbitrary noise model in a scenario with $M = 100$, $K = 4$, and SNR $= 15$ dB.

Fig. 2: P_D vs. ρ for exponentially correlated noise in a scenario with $M = 20$, $N = 80$, $K = 3$, and SNR $= 10$ dB.

All other methods fail when ρ increases.

In the last experiment, we consider Gaussian noise with arbitrary unknown covariance matrix, that is, the noise covariance matrix is drawn from a complex Wishart distribution with M degrees of freedom and scale matrix proportional to $\mathbf{I}$. The scenario has $M = 100$ antennas, $K = 4$ sources, and the number of snapshots N varies upto 100, so that we are in the small sample regime. The SNR for this experiment is chosen as SNR $= 15$ dB, because when the noise covariance matrix is arbitrary, the SNR must be sufficiently high for any method to work. Since CCA and method in [10] do not usually provide good results with few snapshots, we are considering $N \geq 15$ for these methods, however, lower values of N are considered for SSA. As it is shown in Fig. 3, CCA does not perform well for this noise model in the small sample regime, while the method in [10] provides good results for sufficiently large N. Finally, SSA performs satisfactorily even when N/M is around 0.1.

5. CONCLUSIONS

In this work we have presented an order estimation technique for arbitrary geometry arrays and noise with unknown spatial correlation. The method borrows from subspace averaging techniques for source enumeration, originally proposed for uniform linear arrays and white noise. The key idea to generate a collection of subspaces to be averaged is obtaining a sparse representation of each snapshot as a linear combination of the others. To perform this sparse recovery (SR) problem, we proposed a generalization of the log-surrogate of the ℓ_0-norm, which is solved using the majorization-minimization approach. Then, based on the sparse coefficients of the reconstruction, a sampling mechanism is presented to obtain projection matrices that share a large common portion of the signal subspace. Finally, the eigenvalues of an average projection matrix are then used to estimate the number of sources. It is illustrated by some simulation examples that the proposed sparse subspace averaging (SSA) method performs robustly for a wide range of noise models in the small sample regime.

$= -2$ dB. In this example SSA provides robust and accurate detection results over the entire range of variation of ϵ. Since LSMDL and BIC are originally designed for white noise, their performances degrade for higher values of ϵ. The performance of the linear regression method in [10] also degrades for higher values of ϵ. Finally, MDL-unc provides good results for higher values of ϵ, but its performance degrades when the noise is nearly spatially white.

In the second experiment, we consider an exponentially correlated noise model whose noise covariance matrix is defined as $[\mathbf{R}_n]_{i,j} = \rho^{|i-j|}$, where $0 \leq \rho \leq 1$ is the correlation coefficient. Fig. 2 shows P_D vs. ρ for an array with $M = 20$ antennas, $N = 80$ snapshots, $K = 3$ sources, and SNR $= 10$ dB. For this noise model, we are comparing the results of MDL-unc, the method in [10] and the method in [27], which applies a rank-reduction preprocessing step followed by an order estimation criterion based on canonical correlation analysis (CCA). As Fig. 2 suggests, the proposed SSA method again provides a robust solution for the whole range of ρ.
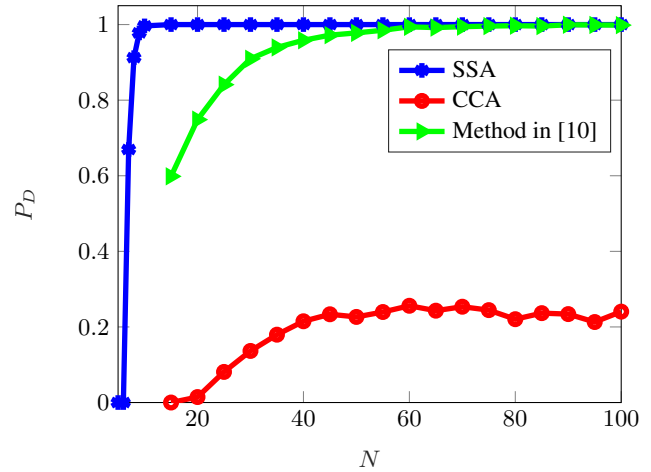
6. REFERENCES

[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] J. Rissanen, "Modeling by shortest data description," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 14, pp. 465–471, 1978.

[5] 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.

[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] 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.

[8] 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.

[9] L. Huang, Y. Xiao, H. C. So, and J.-K. Zhang, "Bayesian information criterion for source enumeration in large-scale adaptive antenna array," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3018–3032, 2016.

[10] A. Eguizabal, C. Lameiro, D. Ramírez, and P. J. Schreier, "Source enumeration in the presence of colored noise," *IEEE Signal Process. Letters*, vol. 26, pp. 475–479, 2019.

[11] M. Wax and A. Adler, "Detection of the number of signals by signal subspace matching," *IEEE Trans. Signal Process.*, vol. 69, pp. 973–985, 2021.

[12] I. Santamaria, D. Ramírez, and L. L. Scharf, "Subspace averaging for source enumeration in large arrays," in *Proc. IEEE Work. Stat. Signal Process. (SSP)*, (Freiburg, Germany), Jun. 2018.

[13] V. Garg, I. Santamaria, D. Ramírez, and L. L. Scharf, "Subspace averaging and order determination for source enumeration," *IEEE Trans. Signal Process.*, vol. 67, pp. 3028–3041, 2019.

[14] V. Garg and I. Santamaria, "Source enumeration in non-white noise and small sample size via subspace averaging," in *2019 27th European Signal Processing Conference (EUSIPCO)*, (A Coruña, Spain), Sep. 2019.

[15] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted l_1 minimization," *Journal Fourier Anal. App.*, vol. 14, pp. 877–905, 2008.

[16] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, pp. 794–816, Feb. 2017.

[17] S. Li, S. Ma, W. He, and Y. Wang, "An improved sparse representation method for direction-of-arrival estimation in the presence of unknown mutual coupling," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (Guilin, China), pp. 676–679, Aug 2014.

[18] J. Yin and T. Chen, "Direction-of-arrival estimation using a sparse representation of array covariance vectors," *IEEE Transactions on Signal Processing*, vol. 59, pp. 4489–4493, Sept 2011.

[19] D. Malioutov, M. Cetin, and A. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Transactions on Signal Processing*, vol. 53, pp. 3010–3022, Aug 2005.

[20] I. Santamaria, L. L. Scharf, M. Kirby, C. Peterson, and J. Francos, "An order fitting rule for optimal subspace averaging," in *Proc. IEEE Work. Stat. Signal Process. (SSP)*, (Palma de Mallorca, Spain), Jun. 2016.

[21] E. Elhamifar and R. Vidal, "Sparse subspace clustering," in *IEEE Conf. Comp. Vision and Pattern Recog.*, (Miami, FL, USA), pp. 2790–2797, June 2009.

[22] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal l_1 norm solution is also the sparsest solution," *Communications on Pure and Applied Mathematics,*, vol. 59, pp. 797–829, Jun 2006.

[23] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society B*, vol. 58, pp. 267–288, 1996.

[24] K. G. Joreskog, "Some contributions to maximum likelihood factor analysis," *psychometrika*, vol. 32, pp. 443–482, 1967.

[25] D. Ramírez, G. Vázquez-Vilar, R. Lopez-Valcarce, J. Via, and I. Santamaria, "Detection of rank-P signals in cognitive radio networks with uncalibrated multiple antennas," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3764–3774, 2011.

[26] G. Vazquez-Vilar, D. Ramírez, R. Lopez-Valcarce, J. Via, and I. Santamaria, "Spatial rank estimation in cognitive radio networks with uncalibrated multiple antennas," in *4th International Conference on Cognitive Radio and Advanced Spectrum Management (cogART)*, (Barcelona, Spain), Oct. 2011.

[27] Y. Song, P. J. Schreier, D. Ramírez, and T. Hasija, "Canonical correlation analysis of high-dimensional data with very small sample support," *Signal Processing*, vol. 128, pp. 449–458, 2016.