

Generalized Minimum Noise Subspace For Array Processing

Viet-Dung Nguyen, Karim Abed-Meraim, Nguyen Linh-Trung, Rodolphe Weber

▶ **To cite this version:**

Viet-Dung Nguyen, Karim Abed-Meraim, Nguyen Linh-Trung, Rodolphe Weber. Generalized Minimum Noise Subspace For Array Processing. IEEE Transactions on Signal Processing, 2017, 64 (14), pp.3789 - 3802. 10.1109/TSP.2017.2695457 . hal-01295030v2

HAL Id: hal-01295030

https://inria.hal.science/hal-01295030v2

Submitted on 25 Jul 2016 (v2), last revised 8 Jun 2017 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Generalized Minimum Noise Subspace For Array Processing

Viet-Dung Nguyen, Karim Abed-Meraim, Nguyen Linh-Trung, and Rodolphe Weber

***Abstract*—Based on the minimum noise subspace (MNS) method previously introduced in the context of blind channel identification, generalized minimum noise subspace (GMNS) is proposed in this paper for array processing that generalizes MNS with respect to the availability of only a fixed number of parallel computing units. Different batch and adaptive algorithms are then introduced for fast and parallel computation of signal (principal) and noise (minor) subspaces. The computational complexity of GMNS and its related estimation accuracy are investigated by simulated experiments and a real-life experiment in radio astronomy. It is shown that GMNS represents an excellent trade-off between the computational gain and the subspace estimation accuracy, as compared to several standard subspace methods.**

***Index Terms*—Batch and adaptive algorithms, principal and minor subspace, MNS, GMNS, PCA, MCA, parallel computing, radio frequency interference (RFI) mitigation, radio astronomy.**

I. Introduction

PRINCIPAL subspace analysis (PSA) and minor subspace analysis (MSA) play important roles in many practical signal processing applications such as high resolution parameter estimation [1], blind source separation [2], and radio frequency interference mitigation [3]. Important problems closely related to PSA and MSA are principal component analysis (PCA) and minor component analysis (MCA) that usually require eigen-subspaces of the data covariance matrix [4].

For batch systems, standard subspace techniques based on singular value decomposition (SVD) or eigenvalue decomposition (EVD) are often applied. Although these techniques have high performance advantages, they face high computational complexity, generally of $O(n^3)$ operations, where n is the dimension of the observation vector or the number of sensors. This complexity causes a serious handicap for large-dimensional systems such as large sensor networks [5], massive multiple-input-multiple-output (MIMO) systems [6], or large antenna arrays like the square kilometer array (SKA) used in radio astronomy [7]. In addition, SVD-like methods are not suitable for adaptive subspace tracking because it requires repeated decompositions. In such a case, the use of distributed algorithms [8], parallel computation schemes [3], and fast adaptive techniques [9]–[11] becomes of high interest and generally leads to large gains in terms of computational complexity and memory requirements.

To concretely motivate our research, we now give an example about the LOw Frequency Array (LOFAR) telescope [12], which can be considered as a pathfinder for SKAs [7]. A LOFAR telescope provides a flexible way for all-sky monitoring. In such a system, to attain desired sensitivity (at least several orders of magnitude higher than most communication systems) and accurate spatial resolution, a huge amount of small phased arrays for future SKAs are used and distributed over a large area instead of a small number of big dishes. According to [7], several hundreds of mid- to high-frequency 15 m dishes will be located in South Africa and Africa, and several hundreds of thousands of low-frequency antennas will be located in Western Australia. Thus, it allows an observation of multiple directions at the same time on the sky at a reasonable cost. However, it also leads to challenging problems as follows. First, a massive amount of data, which are collected from some distributed stations and then transmitted to a data center, need to be processed. Second, due to high sensitivity, observed radio astronomical signals are very weak; they are typically 40 dB to over 100 dB weaker than signals from active services. If we suppose further that subspace approaches such as multiple signal classification (MUSIC) [13], estimation of signal parameters via rotational invariance techniques (ESPRIT) [14] and their variants are adopted, it is then problematic to deal with both computational complexity and "subspace fusion" of the massive data. The latter, similar to data fusion, is the process of integration of multiple subspaces representing the same real-world object into consistent, accurate, and useful representations of the "global" minor and principal subspaces.

From a technical point of view, suppose that we have at hand a parallel computing architecture with K computing units[1]. The question of interest in this paper is: How can we exploit this architecture to fuse subspaces, and then to achieve the "global" minor or principal subspaces, as well as to reduce the computational cost for extracting these subspaces? A simple and efficient subspace method called minimum noise subspace (MNS) was proposed in [15]. This method estimates the noise subspace via a set of noise vectors (a basis of the noise subspace) that can be computed *in parallel* from a set of tuples of system outputs. MNS is much more efficient because it avoids large-scale EVD/SVD computation in the standard subspace method. MNS has been applied to blind system identification [15], source localization [16], array

V. D. Nguyen, K. Abed-Meraim, and R. Weber are with PRISME laboratory, University of Orléans, 12 rue de Blois BP 6744, Orléans, France. Emails: {viet-dung.nguyen, karim.abed-meraim, rodolphe.weber}@univ-orleans.fr.

R. Weber is also with the Station de Radioastronomie de Nancay, Observatoire de Paris, CNRS/INSU, F-18330 Nancay, France.

N. Linh-Trung is with the University of Engineering and Technology, Vietnam National University Hanoi, 144 Xuan Thuy, Cau Giay, Hanoi, Vietnam. E-mail: linhtrung@vnu.edu.vn.

[1]A parallel computing architecture can be, but not limited to, a multi-core processor, a graphics processing unit (GPU) or a field-programmable gate array (FPGA) board.

calibration [16], multichannel blind image deconvolution [17], and adaptive subspace tracking [18]. However, the number of parallel computing units based on which MNS is implemented is the same as the dimension of the noise subspace and is, generally not equal to the actually available number of computing units (K) of the parallel computing architecture in use. In addition, the performance of MNS is degraded if the number of outputs is very large compared to that of the inputs.

In this paper, we are not only to generalize MNS from [15] in such a way that we can handle MNS computation with a given K number of parallel computing units but also to handle principal subspace computation that was not dealt with in [15]. The main contributions of this paper are summarized as follows.

1) Via the introduction of the concept of a generalized properly connected sequence (GPCS), we propose the GMNS concept and then several efficient GMNS-based algorithms for MSA and PSA, given a fixed number of parallel computing units. It is noted that we have disseminated partial results on GMNS in our conference contribution [19], and in this paper we provide detailed proofs and extensive experiments.
2) We then propose an algorithm to estimate the principal eigenpairs from the corresponding principal subspaces (i.e., solving the PCA problem) by solving a joint diagonalization problem.
3) We develop efficient adaptive versions of the above proposed GMNS-based batch algorithms for both principal subspace tracking (PST) and minor subspace tracking (MST) by integrating several existing (non-GMNS-based) subspace tracking algorithms in our parallel framework of GMNS. As will be shown, our GMNS-based adaptive algorithms have advantages in terms of computational complexity and convergence rate. It should be noted that some adaptive algorithms for MNS have already been proposed in [20] but they are limited to least minor subspace analysis.
4) We further propose efficient GMNS-based adaptive algorithms for principal eigenvector tracking (PET) from the corresponding GMNS-based PST algorithms. The performance of our algorithms is nearly identical to that of standard SVD-based algorithms and the computational complexity is lower.
5) We apply our GMNS-based PSA method to a real-life application– radio frequency interference (RFI) mitigation in radio astronomy.

The minor and principal subspaces can be exploited in different ways in different algorithms. For example, while MUSIC and its variants use the minor subspace for direction of arrival (DOA) estimation, ESPRIT and its variants use the principal subspace to achieve the desired result. In this paper, GMNS considers both minor and principal subspace estimation, whereas MNS only considers minor subspace estimation. The potential of GMNS are two-fold. First, GMNS provides a framework that allows the estimation of both minor and principal subspaces in parallel. Second, it can be used to fuse data from several data sources. In the practical application on RFI mitigation studied in Section VII, it will be shown that it is difficult or expensive to directly estimate the "global" principal subspace from the available data covariances of several stations. On another note, GMNS has recently been applied to parallelizable tensor decomposition [21].

The rest of the paper is organized as follows. The MNS concept and its implementation are briefly reviewed in Section II. The GMNS concept and the GMNS-based batch algorithms for MSA and PSA/PCA are proposed in Sections III and IV, respectively. GMNS-based adaptive algorithms are then developed in Section V, with some details on the computational complexity. The performance of the proposed GMNS-based batch and adaptive algorithms is presented in Sections VI and VII, including the real-life application on RFI mitigation.

II. Minimum Noise Subspace: A Review

Let us consider a general linear system with p inputs and n outputs ($p < n$), which obeys the following model for the input-output relationship:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \tag{1}$$

where $\mathbf{x}(t) \in \mathbb{C}^n$ is the observation vector, $\mathbf{A} \in \mathbb{C}^{n\times p}$ is the system matrix of full column rank, $\mathbf{s}(t) \in \mathbb{C}^p$ is the random source vector, and $\mathbf{n}(t) \in \mathbb{C}^p$ is the additive white noise vector with unknown variance σ^2. The data covariance matrix is then given by

$$\mathbf{R}_{xx} = E\{\mathbf{x}(t)\mathbf{x}(t)^H\} = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^H + \sigma^2\mathbf{I}, \tag{2}$$

where $\mathbf{R}_{ss}$ is the source covariance matrix of full rank.

It is of interest to compute the minor subspace of $\mathbf{R}_{ss}$ for that MNS was proposed in order to achieve fast computation in a parallel manner [15]. In particular, the concept of properly connected sequence (PCS) is defined and the system is reorganized into $n - p$ subsystems based on a selected PCS. Then, the minor subspace will be efficiently estimated by computing the least (noise) eigenvector, corresponding to the least eigenvalue, of each subsystem. The PCS concept is to guarantee that the noise vectors computed from the subsystems form a basis of the noise subspace. In the following, we will review the concept of PCS and the implementation of MNS.

A. Properly Connected Sequence

Denote the n system outputs by a set of members $m_1, m_2, \ldots, m_n$. Consider a sequence of $n - p$ subsets of outputs, wherein each subset i contains $p + 1$ members and is denoted by the $(p + 1)$-tuple $t_i = (m_{i_1}, m_{i_2}, \ldots, m_{i_{p+1}})$, $i = 1, \ldots, n - p$. This sequence is said to be *properly connected*, if the following conditions are satisfied:

$$\{m_{i_1}, \ldots, m_{i_p}\} \subset \{m_{j_k} \mid j < i, 1 \le k \le p + 1\},$$
$$m_{i_{p+1}} \notin \{m_{j_k} \mid j < i, 1 \le k \le p + 1\}.$$

These conditions mean that each tuple in the sequence has p members in common with its preceding tuples, along with another member not so. For instance, if we consider a system

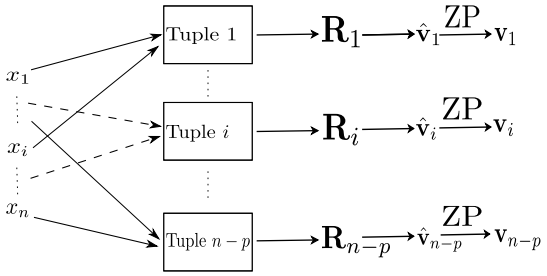


Fig. 1. MNS implementation.

with $p = 2$ inputs and $n = 6$ outputs, the following sequence of $n - p = 4$ tuples is a PCS:

$$t_1 = (m_1, m_2, m_3); t_2 = (m_1, m_2, m_4);$$
$$t_3 = (m_2, m_3, m_5); t_4 = (m_2, m_5, m_6).$$

B. MNS Implementation

Now, given a PCS of $n-p$ tuples of system outputs, the (partial) observation vector $\mathbf{x}_i(t) = [x_{i_1}(t), x_{i_2}(t), \ldots, x_{i_{p+1}}(t)]^T$ of $p + 1$ outputs corresponding to the i-th tuple has the following structure:

$$\mathbf{x}_i = \mathbf{A}_i\mathbf{s}(t) + \mathbf{n}_i(t), \tag{3}$$

and its covariance matrix is given by

$$\mathbf{R}_i = E\{\mathbf{x}_i(t)\mathbf{x}_i^H(t)\} = \mathbf{A}_i\mathbf{R}_{ss}\mathbf{A}_i^H + \sigma^2\mathbf{I}, \tag{4}$$

where $\mathbf{A}_i$ is the response matrix of the i-th subsystem and $\mathbf{n}_i(t)$ is the corresponding additive white noise. Each subsystem in (4) has a noise subspace of minimum dimension (i.e., equal to one), suggesting the naming of MNS.

From each $\mathbf{R}_i$, a noise vector $\mathbf{v}_i$ is constructed by first computing the least eigenvector $\hat{\mathbf{v}}_i$ of $\mathbf{R}_i$, and then zero-padding (ZP) it according to

$$\mathbf{v}_i(j) = \begin{cases} 0, & \text{if the } j\text{-th system output does} \\ & \text{not belong to the } i\text{-th tuple,} \\ \hat{\mathbf{v}}_i(j'), & \text{if the } j\text{-th system output is the} \\ & j'\text{-th entry of } i\text{-th tuple,} \end{cases} \tag{5}$$

for $1 \le j \le n$.

In practice, $\mathbf{R}_i$ is estimated by the sample average as $\hat{\mathbf{R}}_i = \frac{1}{T}\sum_t \mathbf{x}_i(t)\mathbf{x}_i^H(t)$, with T being the sample size. It is proved in [15] and [16] that the resulting set of noise vectors $\{\mathbf{v}_i\}$, for $1 \le i \le n - p$, forms a basis of the noise subspace. Figure 1 illustrates the MNS implementation.

III. Minor Subspace Analysis using GMNS

As stated in Section I, the objective of this paper is to exploit the available parallel computing architecture, when having access to exactly K computing units, in order to reduce the computational cost for extracting the minor or principal subspaces of $\mathbf{R}_{xx}$. We approach this by generalizing MNS.

It is observed that each noise vector in MNS is estimated by the use of a minimum number of system outputs (i.e., $p + 1$) which might lead to a non-negligible performance loss if $n \gg p$. In addition, to achieve the parallel computation of the noise vectors, $(n - p)$ computing units are needed, a number which depends on the impinging source number p and is usually a non-controllable system parameter.

Next, we propose a GMNS-based method for MSA, hereafter referred to as GMNS-MSA, to overcome the above-mentioned shortcomings. Given K computing units, we need to compute the $(n - p)$ noise vectors.

Let us write $n - p = dK + r$, where d and r are integers and $0 \le r < K$, and for simplicity we assume that $r = 0$, i.e., $(n-p)/K$ is integer-valued. Now, we propose a concept of generalized properly connected sequence (GPCS), which generalizes the PCS concept used in MNS.

Definition 1. *A GPCS is a sequence of K $(p+d)$-tuples $t_i = (m_{i_1}, \ldots, m_{i_{p+d}})$ for $1 \le i \le K$ that satisfies the following conditions:*

$$\{m_{i_1}, \ldots, m_{i_p}\} \subset \{m_{j_k} \mid j < i, 1 \le k \le p + d\}, \tag{6}$$
$$\{m_{i_{p+1}}, \ldots, m_{i_{p+d}}\} \notin \{m_{j_k} \mid j < i, 1 \le k \le p + d\}. \tag{7}$$

In other words, each tuple in the sequence has p members in common with its preceding tuples along with d other members not so.

Given a GPCS of K $(p+d)$-tuples of the outputs, the noise vectors are computed as follows. First, for each i-th subsystem, we compute the covariance matrix $\mathbf{R}_i$ of size $(p+d)\times(p+d)$, and hence its d least eigenvectors, represented by matrix $\hat{\mathbf{V}}_i$. Then, we construct the desired noise submatrix $\mathbf{V}_i$ according to the following rule[2]:

$$\mathbf{V}_i(j,:) = \begin{cases} 0, & \text{if the } j\text{-th system output does} \\ & \text{not belong to the } i\text{-th tuple,} \\ \hat{\mathbf{V}}_i(j',:), & \text{if the } j\text{-th system output is the} \\ & j'\text{-th entry of } i\text{-th tuple.} \end{cases} \tag{8}$$

Finally, we concatenate the K submatrices $\mathbf{V}_i$, $1 \le i \le K$, to have the global noise matrix $\mathbf{V} = [\mathbf{V}_1, \ldots, \mathbf{V}_K]$ whose columns form a basis of the noise subspace under the conditions given in the following theorem, with proof given in Appendix A.

Theorem 1. *Under the assumption that every p rows of $\mathbf{A}$ are linearly independent, the noise matrix $\mathbf{V}$ has full column rank (i.e., $\mathrm{rank}(\mathbf{V}) = n - p$) and hence its columns span the desired noise subspace of the data covariance matrix $\mathbf{R}_{xx}$.*

The GMNS-MSA method is illustrated in Figure 2 and its implementation is summarized in Table I.

The main advantage of GMNS-MSA is the reduction of the computational cost in comparison with the standard method. In particular, GMNS-MSA requires $O((p+(n-p)/K)^2T)$ flops for the computation of the subsystem covariance matrices plus $O((p + (n - p)/K)^2(n - p)/K)$ flops for the estimation of the least eigenvectors. Whereas, the standard method requires $O(n^2T)$ flops for the estimation of the global covariance matrix plus $O(n^2(n-p))$ flops for the extraction of the noise vectors. Obviously, if $n \gg p$, the overall cost is reduced by almost a factor of K^2 for the covariance matrix estimation and a factor of K^3 for the noise subspace estimation.

[2]In the sequel, $\mathbf{V}(j,:)$ refers to j-th row vector of $\mathbf{V}$.

TABLE I
SUMMARY OF GMNS FOR MSA

Input: Observed data $\mathbf{X}$

1) Extract subsystems $\mathbf{X}_i$ as described in III
2) Compute covariance matrix of each subsystem: $\mathbf{R}_i = \frac{1}{T}\mathbf{X}_i\mathbf{X}_i^H$
3) Extract d least eigenvectors of $\mathbf{R}_i$: $\hat{\mathbf{V}}_i = \text{eig}(\mathbf{R}_i)$
4) Construct desired noise matrix:

$$\mathbf{V}_i(j,:) = \begin{cases} 0, & \text{system output } j \notin \text{tuple } i \\ \hat{\mathbf{V}}_i(j',:), & \text{system output } j = \text{entry } j' \text{ of tuple } i \end{cases}$$

Output: Minor subspace weight matrix: $\mathbf{V} = [\mathbf{V}_1, \ldots, \mathbf{V}_K]$

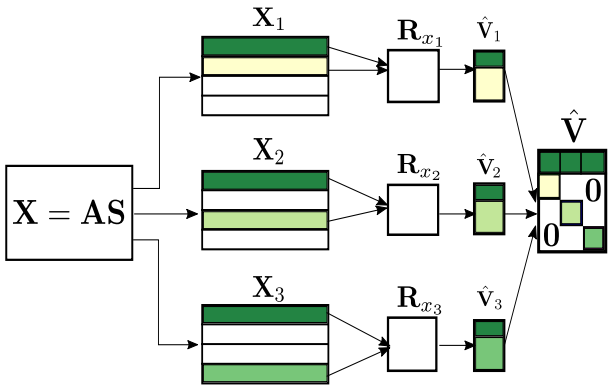


Fig. 2. GMNS for MSA, with $K = 3$ subsystems. The green (bold) part represents the p outputs shared by three subsystems.

Following are some remarks on the GMNS-MSA method.

Remark 1. Generally, $(n-p)/K$ is non integer-valued; that is, $0 < r < K$ for r in $n - p = dK + r$. In that case, we will use r tuples of length $p + d'$, with $d' = d + 1$, and $K - r$ tuples of length $p + d$.

Remark 2. It should be noted that the GPCS concept is just a practical way to guarantee that the obtained vectors form a basis of the desired subspace. In other words, GPCS does not represent necessary conditions to meet but only sufficient conditions.

Remark 3. For large dimensional systems when $n \gg p$, using only $p+1$ system outputs as in the original MNS to compute a noise vector may result in non-negligible performance loss. Now if $\lfloor (n-p)/K \rfloor = d$ is relatively large, we will instead use $p + d + 1$ outputs to estimate a given noise vector which improves its estimation accuracy, as later shown in Figure 5. It also means that we need $K < (n-d)/p$ so that the size of subsystems is larger than p.

Remark 4. When $p \ll n$, one way to estimate the minor subspace is to first estimate the signal subspace (with a low cost) and then obtain the noise subspace as its orthogonal complement. However, it is quite expensive to compute the orthogonal complement subspace; i.e., $O(p^2 n)$. In addition, extraction of an orthogonal basis of the noise subspace would require a further cost.

Remark 5. So far, we have assumed that p is known in advance. When this assumption is violated, many subspace algorithms still work well in practice if we replace p with an overestimated value, for example in the MUSIC algorithm. In such a case, the value used by the algorithm can be just a guess (overestimated value) of the exact value p.

Remark 6. In [22], a fast subspace estimation method was proposed, exploiting the spatial whiteness of the additive noise. In this paper, we exploit this property together with parallel computing to achieve a much higher computational gain.

IV. PRINCIPAL SUBSPACE ANALYSIS AND PRINCIPAL COMPONENT ANALYSIS USING GMNS

The original MNS was dedicated to MSA and above we have proposed a GMNS-based method for MSA. In this section, we propose a GMNS-based method for PSA using K subsystems in a parallel manner. This method is hereafter referred to as GMNS-PSA. In particular, we proposed algorithms for overlapping and non-overlapping subsystems respectively. In addition, we extend the method for PCA, which is hereafter referred to as GMNS-PCA.

A. Principal Subspace Analysis using GMNS

1) Subsystems without overlapping parts:

Let us assume that we have a large dimensional system such that $l = n/K > p$ and, for simplicity, l is integer-valued. We divide the n system outputs into K subsystems of length l each represented by

$$(m_{(i-1)l+1}, \ldots, m_{il}), \quad i = 1, \ldots, K.$$

Now, for each subsystem i, we compute the corresponding covariance matrix $\mathbf{R}_i$ and its principal subspace matrix $\mathbf{W}_i = \mathbf{A}_i\mathbf{Q}_i$, where $\mathbf{Q}_i$ is an unknown nonsingular matrix of size $p \times p$.

To have a global estimate of the signal subspace (i.e., a matrix $\mathbf{W} = \mathbf{AQ}$ of size $n \times p$ where $\mathbf{Q}$ is any $p \times p$ nonsingular matrix), we need to get rid of the unknown matrices $\mathbf{Q}_i$. For that, we exploit the fact that all subsystems receive the same source $\mathbf{S}$ of size $p \times T$, that is,

$$\mathbf{X}_i = \mathbf{A}_i\mathbf{S} + \mathbf{N}_i, \quad i = 1, \ldots, K, \tag{9}$$

where $\mathbf{S} = [\mathbf{s}(1), \cdots, \mathbf{s}(T)]$ and $\mathbf{N}_i$ is noise affecting the i-th subsystem. Let us define

$$\mathbf{S}_i = \mathbf{W}_i^{\#}\mathbf{X}_i, \tag{10}$$

where $\#$ denotes the pseudo-inverse operator[3]. Then, thanks to (9), we have

$$\mathbf{S}_i = \mathbf{Q}_i^{-1}\mathbf{S} + \mathbf{W}_i^{\#}\mathbf{N}_i, \quad i = 1, \ldots, K. \tag{11}$$

In the noiseless case, it can be shown that

$$\mathbf{S}_i = \mathbf{T}_i\mathbf{S}_1,$$

where $\mathbf{T}_i = \mathbf{Q}_i^{-1}\mathbf{Q}_1$. Matrix $\mathbf{T}_i$ can be estimated by solving the following least square problem:

$$\min_{\mathbf{T}_i} \| \mathbf{S}_i - \mathbf{T}_i\mathbf{S}_1 \|_2^2 .$$

[3]Most subspace estimation methods compute an orthonormal basis of the desired subspace (see, e.g., [23]) in which case we have $\mathbf{W}_i^{\#} = \mathbf{W}_i^H$.

Its solution is given by

$$\hat{\mathbf{T}}_i = \mathbf{S}_i \mathbf{S}_1^{\#}. \tag{12}$$

Finally, the principal subspace weight matrix is obtained as

$$\mathbf{W} = \left[\mathbf{W}_1^T, (\mathbf{W}_2\mathbf{T}_2)^T, \ldots, (\mathbf{W}_K\mathbf{T}_K)^T\right]^T \approx \mathbf{A}\mathbf{Q}_1. \tag{13}$$

In the noisy case, the estimate of the principal subspace weight matrix in (13) is *biased* due to the effect of the noise term on the estimation of $\mathbf{T}_i$ in (12). In fact, (12) can be rewritten as

$$\hat{\mathbf{T}}_i = \left(\frac{\mathbf{S}_i\mathbf{S}_1^H}{T}\right)\left(\frac{\mathbf{S}_1\mathbf{S}_1^H}{T}\right)^{-1} \tag{14}$$

$$\approx (\mathbf{Q}_i^{-1}\hat{\mathbf{R}}_{ss}\mathbf{Q}_1^{-H})(\mathbf{Q}_1^{-1}\hat{\mathbf{R}}_{ss}\mathbf{Q}_1^{-H} + \sigma^2\mathbf{I})^{-1}. \tag{15}$$

Here we have substituted (11) into (14), and used the facts that the subsystems are non-overlapping, their noise terms are uncorrelated (spatially white noise assumption) and $\mathbf{W}_i$ are unitary matrices that leads to $E[\mathbf{W}_1^H\mathbf{n}_1(t)\mathbf{n}_1^H(t)\mathbf{W}_1] = \sigma^2\mathbf{I}$ and, hence, $\mathbf{W}_1^H(\mathbf{N}_1\mathbf{N}_1^H/T)\mathbf{W}_1 \approx \sigma^2\mathbf{I}$. Moreover, we remind that the signal and the noise are uncorrelated, i.e., $E[\mathbf{S}\mathbf{N}_i^H] = 0$ and hence $\mathbf{S}\mathbf{N}_i^H/T \approx 0$. Because of the additive term $\sigma^2\mathbf{I}$, $\hat{\mathbf{T}}_i$ deviates from its desired value and leads to an estimation bias for the global weight matrix, especially, at low signal-to-noise ratio (SNR).

To overcome this problem, we replace the previous estimate of $\mathbf{T}_i$ by the following asymptotically unbiased estimate:

$$\tilde{\mathbf{T}}_i = \left(\frac{\mathbf{S}_i\mathbf{S}_1^H}{T}\right)\left(\frac{\mathbf{S}_1\mathbf{S}_1^H}{T} - \hat{\sigma}^2\mathbf{I}\right)^{-1}$$

$$= (\mathbf{W}_i^H\mathbf{R}_{i,1}\mathbf{W}_1)(\mathbf{W}_1^H\mathbf{R}_1\mathbf{W}_1 - \hat{\sigma}^2\mathbf{I})^{-1}, \tag{16}$$

where

$$\mathbf{R}_{i,1} = E[\mathbf{x}_i(t)\mathbf{x}_1^H(t)] \tag{17}$$

$$\hat{\sigma}^2 = [\mathrm{Tr}(\mathbf{R}_1) - \mathrm{Tr}(\mathbf{W}_1^H\mathbf{R}_1\mathbf{W}_1)]/(l-p). \tag{18}$$

To obtain the noise power estimate in (18), we first obtain the EVD of $\mathbf{R}_1$ as

$$\mathbf{R}_1 = \mathbf{A}_1\mathbf{R}_{ss}\mathbf{A}_1^H + \sigma^2\mathbf{I}$$

$$= \mathbf{U}_s(\mathbf{\Lambda}_s + \sigma^2\mathbf{I})\mathbf{U}_s^H + \sigma^2\mathbf{U}_n\mathbf{U}_n^H, \tag{19}$$

in which the columns of $\mathbf{U}_s$ span the signal subspace of $\mathbf{R}_1$ and those of $\mathbf{U}_n$ span the noise subspace of $\mathbf{R}_1$. From (19) and using $\mathbf{W}_1^H\mathbf{U}_n = \mathbf{0}$, we then have

$$\mathbf{W}_1^H\mathbf{R}_1\mathbf{W}_1 = \mathbf{W}_1^H\mathbf{U}_s(\mathbf{\Lambda}_s + \sigma^2\mathbf{I})\mathbf{U}_s^H\mathbf{W}_1. \tag{20}$$

Since $\mathbf{W}_1^H\mathbf{U}_s$ is unitary, we have $\mathrm{Tr}(\mathbf{R}_1 - \mathbf{W}_1^H\mathbf{R}_1\mathbf{W}_1) = (l-p)\sigma^2$, and thus (18).

The above biased and unbiased algorithms are referred to as GMNS-N-PSA (N stands for non-overlapping) and GMNS-NU-PSA (NU stands for non-overlapping and unbiased), respectively. They are summarized in Table II. Figure 3 illustrates the steps described in this section.

TABLE II
SUMMARY OF GMNS FOR PSA: NON-OVERLAPPING AND OVERLAPPING SUBSYSTEMS

Input: Captured data $\mathbf{X}$

1) Extract non-overlapping subsystems $\mathbf{X}_i$ as described in IV-A1, or overlapping ones as in IV-A2
2) Compute covariance matrix of each subsystem: $\mathbf{R}_i = \frac{1}{T}\mathbf{X}_i\mathbf{X}_i^H$.
3) Extract p principal eigenvectors of $\mathbf{R}_i$: $\mathbf{W}_i = \mathrm{eig}(\mathbf{R}_i, p)$

4*) *Non-overlapping*:
Compute $\mathbf{T}_i$
- Biased: $\mathbf{S}_i = \mathbf{W}_i^{\#}\mathbf{X}_i, \quad \tilde{\mathbf{T}}_i = \mathbf{S}_i\mathbf{S}_1^{\#}$.
- Unbiased: $\hat{\sigma}^2 = [\mathrm{Tr}(\mathbf{R}_1) - \mathrm{Tr}(\mathbf{W}_1^H\mathbf{R}_1\mathbf{W}_1)]/(l-p)$, $\tilde{\mathbf{T}}_i = \left(\frac{\mathbf{S}_i\mathbf{S}_1^H}{T}\right)\left(\frac{\mathbf{S}_1\mathbf{S}_1^H}{T} - \hat{\sigma}^2\mathbf{I}\right)^{-1}$.

4**) *Overlapping*:
Extract overlap: $\mathbf{W}_1^{\mathrm{olap}}(t)$, $\mathbf{W}_i^{\mathrm{olap}}(t)$
Compute: $\mathbf{T}_i = (\mathbf{W}_i^{\mathrm{olap}})^{\#}(\mathbf{W}_1^{\mathrm{olap}}), i = 2, \ldots, K$
Compute principal weight matrix: $\tilde{\mathbf{W}}_1 = \mathbf{W}_1$, $\tilde{\mathbf{W}}_i = \mathbf{W}_i\mathbf{T}_i$

Output: Principal subspace
- *Non-overlapping*: $\mathbf{W} = [\mathbf{W}_1^T, (\mathbf{W}_2\tilde{\mathbf{T}}_2)^T, \ldots, (\mathbf{W}_K\tilde{\mathbf{T}}_K)^T]^T$
- *Overlapping*: $\mathbf{W} = [\tilde{\mathbf{W}}_1^T, \tilde{\mathbf{W}}_2^T, \ldots, \tilde{\mathbf{W}}_K^T]^T$

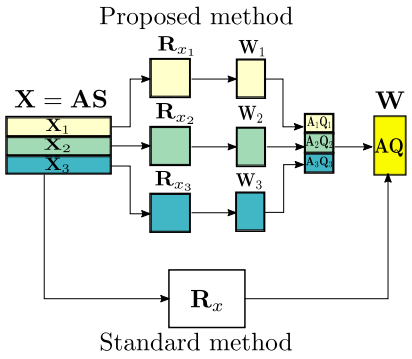


Fig. 3. GMNS for PSA with non-overlapping parts; $K = 3$.

2) Subsystems with overlapping parts:
In the above non-overlapping case, we have assumed that $n/K > p$. To relax this assumption and extend our method to cover also the case[4] where $n/K < p$, we consider here overlapping subsystems of size $p+q$ sharing q system outputs, and represented by the K tuples. For example, here, we choose all subsystems which overlap with the first one:

$$(m_1, \ldots, m_q, m_{q+1}, \ldots m_{p+q}),$$
$$(m_1, \ldots, m_q, m_{(i-1)d+1}, \ldots, m_{id}), \quad i = 2, \ldots, K.$$

In other words, the q first members of the first subsystem are the q first members of the i-th subsystem, for $i = 2, \ldots, K$. For simplicity, we assume that $d = (n-p)/K$ is integer-valued. Now, for each subsystem, we compute the covariance matrix $\mathbf{R}_i$ and its corresponding weight matrix $\mathbf{W}_i$ which can be written as

$$\mathbf{W}_i = \begin{bmatrix}\mathbf{W}_i^{\mathrm{olap}} \\ \mathbf{W}_i'\end{bmatrix} = \begin{bmatrix}\mathbf{A}_i^{\mathrm{olap}} \\ \mathbf{A}_i'\end{bmatrix}\mathbf{Q}_i. \tag{21}$$

[4]When using overlapping subsystems, we can deal with either $n/K \geq p$ or $n/K < p$, while in the non-overlapping case, it is necessary to have $n/K > p$.

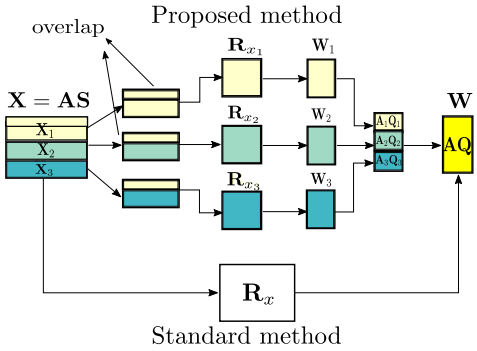


Fig. 4. GMNS for PSA with overlapping parts; $K = 3$.

To get rid of the matrices $\mathbf{Q}_i$, one exploits the overlap between the first subsystem and the i-th subsystem by assuming that any $p \times p$ submatrix of $\mathbf{A}$ has full rank. In that case, the global weight matrix is estimated as

$$\mathbf{W} = \left[\tilde{\mathbf{W}}_1^T, \tilde{\mathbf{W}}_2^T, \ldots, \tilde{\mathbf{W}}_K^T\right]^T, \tag{22}$$

where

$$\tilde{\mathbf{W}}_1 = \mathbf{W}_1, \tag{23}$$

$$\tilde{\mathbf{W}}_i = \mathbf{W}_i \mathbf{T}_i, \quad i = 2, \ldots, K, \tag{24}$$

$$\mathbf{T}_i = (\mathbf{W}_i^{\text{olap}})^{\#} (\mathbf{W}_1^{\text{olap}}). \tag{25}$$

This algorithm is referred to as GMNS-O-PSA (where O stands for overlapping) and is summarized in Table II and illustrated in Figure 4.

3) Complexity:
Similar to GMNS-MSA in Section III, the main advantage of the proposed GMNS-based PSA algorithms resides in its reduced computational cost. GMNS-NU-PSA costs $O((n/K)^2 p + p^2(n/K + p))$ flops for the computation of p signal subspace vectors and $O(2(n/K)^2 T)$ flops for the computation of the covariance matrices $\mathbf{R}_i$ and the correlation matrices $\mathbf{R}_{i,1}$, for $i = 1, \ldots, K$. This overall cost is approximately K^2 less than the cost of a direct computation of the signal subspace using the global covariance matrix, which takes $O(n^2(T + p))$ flops. GMNS-O-PSA costs $O((p+q)^2 p + p^2(2p+q))$ flops for the computation of p signal subspace vectors and $O((p+q)^2 T)$ flops for the parallel computation of the covariance matrices. If $n \gg p$, $n \gg K$ and $T \gg 1$, then $q \approx n/K$, and hence GMNS-O-PSA is slightly cheaper than GMNS-NU-PSA since it does not require computing the correlation matrices[5].

B. Principal Component Analysis using GMNS

Having estimated the principal subspace by the GMNS-PSA method as presented in Section IV-A, we can further extract the principal eigenvalues and eigenvectors of the covariance matrix $\mathbf{R}_{xx}$, resulting in a GMNS-based method for PCA. This method is hereafter referred to as GMNS-PCA and is described below.

[5]For certain applications, e.g., radio astronomy, the global covariance matrix is available (or already computed) for other needs. Hence, GMNS-NU-PSA becomes more advantageous than GMNS-O-PSA in term of computational complexity.

TABLE III
SUMMARY OF GMNS FOR PCA

Input: Principal subspace $\mathbf{W}$
1) Compute expression: $\mathbf{W}^{\#}\mathbf{R}_{xx}\mathbf{W}$
2) Solve diagonalization problem to find $\tilde{\mathbf{Q}}^{-1}$ and $\{\lambda_i\}$
3) Extract p principal eigenvectors: $\mathbf{U}_s = \mathbf{W}\tilde{\mathbf{Q}}^{-1}$

Output: Principal eigenpairs: $\mathbf{U}_s$ and $\{\lambda_i\}_{i=1:p}$

First, we note that

$$\mathbf{W} = \mathbf{AQ} = \mathbf{U}_s \tilde{\mathbf{Q}}, \tag{26}$$

where the columns of $\mathbf{U}_s$ are p principal eigenvectors of $\mathbf{R}_{xx}$, and $\tilde{\mathbf{Q}}$ is a non-singular matrix of size $p \times p$.

Now, consider the following expression:

$$\mathbf{W}^{\#}\mathbf{R}_{xx}\mathbf{W} = \tilde{\mathbf{Q}}^{-1}(\mathbf{U}_s^H \mathbf{R}_{xx} \mathbf{U}_s)\tilde{\mathbf{Q}} \tag{27}$$

$$= \tilde{\mathbf{Q}}^{-1} \begin{bmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_p \end{bmatrix} \tilde{\mathbf{Q}}, \tag{28}$$

where $\lambda_1, \ldots, \lambda_p$ are the principal eigenvalues of $\mathbf{R}_{xx}$.

Therefore, the principal eigenvectors and eigenvalues can be found by computing matrix $\tilde{\mathbf{Q}}$ that diagonalizes the matrix $\mathbf{W}^{\#}\mathbf{R}_{xx}\mathbf{W}$.

We note that since GMNS-PCA estimates the principal eigenvectors from the principal subspace obtained by GMNS-PSA, the eigenvectors are arranged in descending order of the corresponding eigenvalues, i.e., $\lambda_1 > \lambda_2 > \cdots > \lambda_p$. Therefore, the performance of principal eigenvector estimation depends on that of principal subspace estimation. The whole method is summarized in Table III.

V. Adaptive GMNS-based Algorithms

In this section, we are interested in estimating the minor and principal subspaces at each time index t from streaming observations $\{\mathbf{x}(t)\}_{t \geq 1}$.

A. Minor Subspace Tracking using GMNS

Thanks to the parallel structure of GMNS-MSA, the conversion from a batch system to an adaptive one is quite simple. This leads to our proposed GMNS-based method for MST, which is hereafter referred to as GMNS-MST.

In practice, to track the underlying minor subspace, we replace the SVD-based computation of the minor subspace from the correlation matrix (i.e., Step 2 in Table I) with any existing MST algorithms while keeping the remaining steps unchanged. In our GMNS-MST, we integrate the FOOja algorithm [24] and the FDPM algorithm [11].

It should be noted that the estimation of the covariance matrix (i.e., Step 1 in Table I) is not required in FOOja and FDPM. As a result, GMNS-MST provides a way to reduce the complexity of the algorithms by a factor of K. The implementation of GMNS-MST is summarized in Table IV.

Because of the way we construct the desired noise matrix (i.e., Step 3 in Table IV), the computational cost of GMNS-MST equals the cost of the tracking algorithm in use reduced

TABLE IV
SUMMARY OF GMNS FOR MST

Initialization
Input: Observed data at snapshot t: $\mathbf{x}(t)$
1) Extract subsystems $\mathbf{x}_i(t)$ as described in III
2) Track minor subspace of each subsystem using existing algorithms (e.g. FOOja, FDPM):
$$\hat{\mathbf{V}}_i(t) = \text{TrackingAlgorithm}(\mathbf{V}_i(t-1), \mathbf{x}_i(t))$$
3) Construct the desired noise matrix $\mathbf{V}_i(t)$:
$$\mathbf{V}_i(j,:) = \begin{cases} 0, & \text{if system output } j \notin \text{tuple } i \\ \hat{\mathbf{V}}_i(j',:), & \text{if system output } j = \text{entry } j' \text{ of tuple } i \end{cases}$$
Output: Minor subspace $\mathbf{V}(t) = [\mathbf{V}_1(t), \ldots, \mathbf{V}_K(t)]$

by a factor of K. For example, it costs $O(np/K)$ if FOOja or FDPM is used.

B. Principal Subspace Tracking using GMNS

By extending GMNS-PSA as given in Section IV-A to adaptive processing, we now propose a GMNS-based method for PST, which is hereafter referred to as GMNS-PST. Similar to GMNS-PSA, we deal with two cases whether we have the subsystems with or without overlapping parts.

1) Subsystems without overlapping parts:
We observe that matrix $\mathbf{T}_i$ can be expressed by

$$\mathbf{T}_i = \mathbf{W}_i^{\#}(\mathbf{X}_i\mathbf{X}_1^{\#})\mathbf{W}_1. \tag{29}$$

By this way, we can track the principal subspace $\mathbf{W}_i$ of each subsystem and then compute the global weight matrix $\mathbf{W}$ as in (13).

Because calculating the term $\mathbf{X}_i\mathbf{X}_1^{\#}$ in full scale is expensive and not suitable for adaptive processing, we propose to use a sliding window technique to overcome this problem. Denote by N the window size. At time instant t, the subsystem $\mathbf{X}_i(t)$ can be written as

$$\mathbf{X}_i(t) = [\mathbf{x}_i(t-(N-1)) \quad \cdots \quad \mathbf{x}_i(t)]. \tag{30}$$

In our GMNS-PST, we use the sliding window version of the orthogonal projection approximation subspace tracking (OPAST) algorithm [25].

Now, to update $\mathbf{T}_i(t)$ efficiently, we first rewrite the term $\mathbf{X}_i(t)\mathbf{X}_1^{\#}(t)$ as

$$\mathbf{X}_i(t)\mathbf{X}_1^{\#}(t) = \mathbf{P}_i(t)\mathbf{M}^{-1}(t), \tag{31}$$

where

$$\mathbf{P}_i(t) = \mathbf{X}_i(t)\mathbf{X}_1^H(t),$$
$$\mathbf{M}(t) = \mathbf{X}_1(t)\mathbf{X}_1^H(t).$$

Then, we obtain

$$\begin{aligned}\mathbf{P}_i(t) &= \sum_{\tau=0}^{N-1} \mathbf{x}_i(t-\tau)\mathbf{x}_1^H(t-\tau) \\ &= \mathbf{P}_i(t-1) + \mathbf{x}_i(t)\mathbf{x}_1^H(t) - \mathbf{x}_i(t-N)\mathbf{x}_1^H(t-N),\end{aligned} \tag{32}$$

$$\begin{aligned}\mathbf{M}(t) &= \sum_{\tau=0}^{N-1} \mathbf{x}_1(t-\tau)\mathbf{x}_1^H(t-\tau) \\ &= \mathbf{M}(t-1) + \mathbf{x}_1(t)\mathbf{x}_1^H(t) - \mathbf{x}_1(t-N)\mathbf{x}_1^H(t-N).\end{aligned} \tag{33}$$

Since $\mathbf{M}(t)$ has a rank-2 update structure, it can be efficiently inverted by applying the matrix inversion lemma to yield $\mathbf{M}^{-1}(t)$. However, substituting (32) and (33) into (31) and hence into (29) still includes expensive matrix-matrix multiplications. Fortunately, many fast tracking algorithms have a rank-1 update structure (e.g., the OPAST algorithm [10]) expressed in the form of

$$\mathbf{W}(t) = \mathbf{W}(t-1) + \mathbf{p}(t)\mathbf{q}^H(t). \tag{34}$$

Thus, we can use this observation to compute (29) recursively with only matrix-vector multiplications. To initialize $\mathbf{P}_i(0)$ and $\mathbf{M}(0)$, we can either choose them arbitrarily or use the N first snapshots in a batch way.

This algorithm is referred to as GMNS-N-PST and summarized in Table V.

Remark 7. Recall that in GMNS-PSA we have developed both biased and unbiased estimators of $\mathbf{T}_i$. However, in adaptive processing, we have observed that the performance improvement of PST in the unbiased case is negligible as compared to the biased case. Thus, we only presented the latter, as above, due to its simplicity.

2) Subsystems with overlapping parts:
Similar to modifying GMNS-MST to deal with the case where the subsystems have overlapping parts, at each time instant in the adaptive version of GMNS-O-PSA, we simply replace the SVD-based computation of the principal subspace from the correlation matrix (i.e., Steps 2 and 3 in Table II) with any existing PST algorithms while keeping the remaining steps unchanged. Again, a similar observation for efficiently computing $\mathbf{T}_i$ in the non-overlapping case can also be applied here. This algorithm is referred to as GMNS-O-PST and summarized in Table V.

C. Principal Eigenvector Tracking using GMNS

Here we present an adaptive version of GMNS-PCA, as described in Section IV-B, to track the principal eigenvectors from the estimated principal subspace. It is hereafter referred to as GMNS-PET.

According to (27), at time instant t, the following diagonalization is performed:

$$\mathbf{W}^{\#}(t)\mathbf{R}_{xx}(t)\mathbf{W}(t) = \tilde{\mathbf{Q}}^{-1}(t)\begin{bmatrix}\lambda_1(t) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_p(t)\end{bmatrix}\tilde{\mathbf{Q}}(t). \tag{35}$$

TABLE V
SUMMARY OF GMNS FOR PST

Initialize
Input: Observed data at snapshot t: $\mathbf{x}(t)$

1) Extract non-overlapping subsystems $\mathbf{x}_i(t)$ as described in IV-A1, or overlapping ones as described in IV-A2
2) Track principal subspace of each subsytem using existing algorithms (e.g. YAST, OPAST):

$$\mathbf{W}_i(t) = \mathrm{TrackingAlgorithm}(\mathbf{W}_i(t-1), \mathbf{x}_i(t))$$

3*) *Non-overlapping*:
Compute recursive updates of $\mathbf{P}_i(t)$ and $\mathbf{M}(t)^{-1}$
Compute matrix: $\mathbf{T}_i(t) = \mathbf{W}_i^{\#}(t)(\mathbf{P}_i(t)\mathbf{M}^{-1}(t))\mathbf{W}_1(t)$
Compute principal weight matrix:

$$\mathbf{W}(t) = \left[\mathbf{W}_1^T(t), (\mathbf{W}_2^T(t)\mathbf{T}_2^T(t)), \ldots, (\mathbf{W}_K^T(t)\mathbf{T}_K^T(t))\right]^T$$

3**) *Overlapping*:
Extract overlap: $\mathbf{W}_1^{\mathrm{olap}}(t)$, $\mathbf{W}_i^{\mathrm{olap}}(t)$
Compute matrix:

$$\mathbf{T}_i(t) = (\mathbf{W}_i^{\mathrm{olap}}(t))^{\#}(\mathbf{W}_1^{\mathrm{olap}}(t)), i = 2, \ldots, K$$

Compute principal weight matrix:

$$\tilde{\mathbf{W}}_i(t) = \mathbf{W}_i(t)\mathbf{T}_i(t), \quad \text{with } \tilde{\mathbf{W}}_1(t) = \mathbf{W}_1(t)$$

Output: Principal subspace $\mathbf{W}(t)$ and $\tilde{\mathbf{W}}_i(t)$

Then, the principal eigenvectors are found to be

$$\mathbf{U}_s(t) = \mathbf{W}(t)\tilde{\mathbf{Q}}^{-1}(t). \tag{36}$$

A naive implementation of the above estimation is not numerically efficient. Observed that (27) can be written as

$$\begin{aligned}\mathbf{W}^{\#}(t)\mathbf{R}_{xx}(t)\mathbf{W}(t) &\cong E\{\mathbf{W}^{\#}(t)\mathbf{x}(t)\mathbf{x}(t)^H\mathbf{W}(t)\} \\ &\cong E\{\tilde{\mathbf{y}}(t)\mathbf{y}^H(t)\},\end{aligned}$$

where

$$\tilde{\mathbf{y}}(t) = \mathbf{W}^{\#}(t)\mathbf{x}(t), \tag{37}$$
$$\mathbf{y}(t) = \mathbf{W}^H(t)\mathbf{x}(t). \tag{38}$$

The main cost thus comes from the calculation of the pseudo-inverse operator in (37), which is then expressed as

$$\tilde{\mathbf{y}}(t) = [\mathbf{W}(t)^H\mathbf{W}(t)]^{-1}\mathbf{W}(t)^H\mathbf{x}(t) = \mathbf{D}(t)\mathbf{y}(t),$$

where

$$\mathbf{D}(t) = [\mathbf{W}(t)^H\mathbf{W}(t)]^{-1} = \left[\mathbf{I} + \sum_{k=2}^{K}\mathbf{T}_k^H(t)\mathbf{T}_k(t)\right]^{-1}. \tag{39}$$

Equation (39) yields because of the fact that $\mathbf{W}_k(t)$ ($k = 1, \ldots, K$) are orthogonal, i.e., $\mathbf{W}_k^H(t)\mathbf{W}_k(t) = \mathbf{I}$. Thus, $\tilde{\mathbf{y}}(t)$ and $\mathbf{y}(t)$ can be obtained as

$$\tilde{\mathbf{y}}(t) = \sum_{k=1}^{K}\tilde{\mathbf{y}}_k(t), \tag{40}$$
$$\mathbf{y}(t) = \sum_{k=1}^{K}\mathbf{y}_k(t), \tag{41}$$

TABLE VI
SUMMARY OF GMNS FOR PET

Input: Principal subspaces $\mathbf{W}_k(t)$, $\mathbf{T}_k(t)$

1) Compute $\mathbf{D}(t) = \left[\mathbf{I} + \sum_{k=2}^{K}\mathbf{T}_k^H(t)\mathbf{T}_k(t)\right]^{-1}$
2) Compute $\mathbf{Z}(t)$ as in (44) or (45)
3) Compute $\tilde{\mathbf{Q}}^{-1}(t) = \mathrm{eig}(\mathbf{Z}(t))$
4) Extract p principal eigenvectors: $\mathbf{U}_s(t) = \mathbf{W}(t)\tilde{\mathbf{Q}}^{-1}(t)$

Output: Principal eigenpairs: $\mathbf{U}_s(t)$ and $\{\lambda_i(t)\}_{i=1:p}$

where

$$\tilde{\mathbf{y}}_k = \mathbf{D}(t)\mathbf{y}_k(t), \tag{42}$$
$$\mathbf{y}_k = \mathbf{W}_k^H(t)\mathbf{x}_k(t), \tag{43}$$

which can be implemented in parallel in each computing unit.

Then, according to (35), $\tilde{\mathbf{Q}}^{-1}(t)$ can be estimated by performing EVD of $\mathbf{Z}(t)$. Here, $\mathbf{Z}(t)$ is either obtained by

$$\begin{aligned}\mathbf{Z}(t) &= \sum_{\tau=0}^{L-1}\tilde{\mathbf{y}}(t-\tau)\mathbf{y}^H(t-\tau) \\ &= \mathbf{Z}(t-1) + \tilde{\mathbf{y}}(t)\mathbf{y}^H(t) - \tilde{\mathbf{y}}(t-L+1)\mathbf{y}^H(t-L+1),\end{aligned} \tag{44}$$

if a sliding window of length L is used, or by

$$\begin{aligned}\mathbf{Z}(t) &= \sum_{\tau=0}^{t}\beta^{t-\tau}\tilde{\mathbf{y}}(\tau)\mathbf{y}^H(\tau) \\ &= \beta\mathbf{Z}(t-1) + \tilde{\mathbf{y}}(t)\mathbf{y}^H(t),\end{aligned} \tag{45}$$

if the exponential window is used (β being the forgetting factor, $0 < \beta \leq 1$). Once $\tilde{\mathbf{Q}}^{-1}(t)$ has been estimated, the principal eigenvectors are obtained by (36), which can again be estimated in parallel. The algorithm costs $O(np^2/K) + O(p^3)$ and is summarized in Table VI.

VI. SIMULATED EXPERIMENTS

In this section, the performance of subspace estimation is studied by simulation. In all experiments, the system matrix $\mathbf{A}$ is randomly generated but kept fixed for all Monte Carlo runs. The sources are i.i.d. Gaussian processes of unit power. The noise is spatially white with variance of σ^2. The SNR is defined as

$$\mathrm{SNR} = 10\log_{10}\frac{\parallel \mathbf{A} \parallel^2}{\sigma^2}. \tag{46}$$

To assess the performance of the proposed algorithms, the following two criteria are used: subspace estimation performance (SEP) and eigenvector estimation performance (EEP). The lower the values of SEP or EEP, the better the performance.

The SEP is defined as

$$\mathrm{SEP}(t) = \frac{\mathrm{Tr}\{\mathbf{W}_i^H(t)(\mathbf{I} - \mathbf{W}_{\mathrm{ex}}(t)\mathbf{W}_{\mathrm{ex}}^H(t))\mathbf{W}_i(t)\}}{\mathrm{Tr}\{\mathbf{W}_i^H(t)(\mathbf{W}_{\mathrm{ex}}(t)\mathbf{W}_{\mathrm{ex}}^H(t))\mathbf{W}_i(t)\}}, \tag{47}$$

where $\mathbf{W}_i$ is the estimated subspace at the i-th run, and $\mathbf{W}_{\mathrm{ex}}$ is the exact subspace weight matrix computed by orthorgonalizing $\mathbf{A}$. In the case of a batch system, we can drop t from SEP.

TABLE VII
PARAMETERS USED IN OUR EXPERIMENTS

Experiment	Figure	n	p	K	T	q
1	5	30	2	4	500	N/A
2	5	30	6	4	500	N/A
3	6	30	1:10	4	500	N/A
4	6	30	1:10	4	500	N/A
5	7	48	2	4	500	4
6	7	48	6	4	500	8
7	8	48	1:10	4	500	8
8	8	48	1:10	4	500	8
9	9	48	2	4	500	4
10	9	48	6	4	500	8
11	10	30	2	4	4000	N/A
12	11	30	6	4	4000	N/A
13	12	48	2	4	3000	4
14	13	48	6	4	3000	8
15	14	48	2	4	3000	4
16	15	48	6	4	3000	8
17	16	48	2	4	20971	4

The EEP is defined as

$$\mathrm{EEP}(t) = \| \mathbf{U}(t) - \mathbf{U}_{\mathrm{ex}} \|_F^2, \tag{48}$$

where $\mathbf{U}(t)$ is the matrix of the estimated eigenvectors[6], and $\mathbf{U}_{\mathrm{ex}}$ is the matrix of the exact eigenvectors computed from the exact covariance matrix (i.e., noiseless case) using the full SVD algorithm. Similarly, we can drop t when considering a batch system.

In all cases, the results are reported by taking the average over 100 Monte Carlo runs. To assess the performance of the algorithms with respect to the number of sources, p, we present two different scenarios ($p = 2$ and $p = 6$) for relatively large dimensional systems. A summary of parameters used in the experiments is given in Table VII. For principal subspace estimation/tracking, parameters are chosen based on the configuration of one of real radio astronomy systems from which we collected data, as described in Section VII. For minor subspace estimation/tracking, the number of sensors is chosen randomly but in such a way that $(n-d)/p$ is integer-valued. The value of T is considered following the radio astronomy application (Section VII) where the sample size was large enough and may even exceed 100 times n.

A. Minor Subspace Analysis

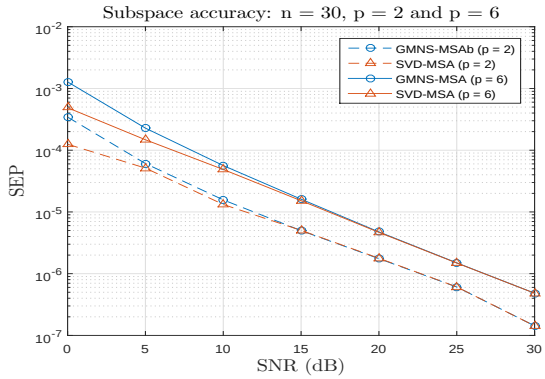
First, we assess the performance of GMNS-MSA against the standard MSA method using SVD (SVD-MSA) with both small ($p = 2$) and large ($p = 6$) numbers of sources. The results indicate that GMNS-MSA has performance close to SVD-MSA when $p = 2$, and it losses some accuracy at low SNR (i.e., SNR < 5 dB) when $p = 6$, as shown in Figure 5(a). However, in both experiments, the dominant cost of GMNS-MSA is reduced by a factor of K^2, as compared to SVD-MSA. The selected GPCS is given in Table VIII.

We further consider the effect of the number of sources on the performance of GMNS-MSA by fixing n and changing p. Moreover, we take into account two scenarios: low SNR (SNR $= 5$ dB) and high SNR (SNR $= 20$ dB). Figure 5(b) suggests that GMNS-MSA is robust and comparable with SVD-MSA in both scenarios.

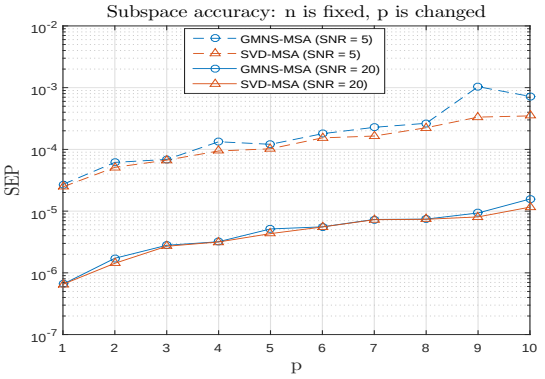
[6]To remove inherent ambiguities, the eigenvector norm is set to one and its first entry is chosen to be positive real-valued.

TABLE VIII
GPCS USED FOR FIRST EXPERIMENT

Tuples	Members
t_1	$(m_1, m_2, m_3, m_4, m_5, m_6, m_7, m_8, m_9)$
t_2	$(m_1, m_2, m_{10}, m_{11}, m_{12}, m_{13}, m_{14}, m_{15}, m_{16})$
t_3	$(m_1, m_2, m_{17}, m_{18}, m_{19}, m_{20}, m_{21}, m_{22}, m_{23})$
t_4	$(m_1, m_2, m_{24}, m_{25}, m_{26}, m_{27}, m_{28}, m_{29}, m_{30})$



(a) SEP vs. SNR

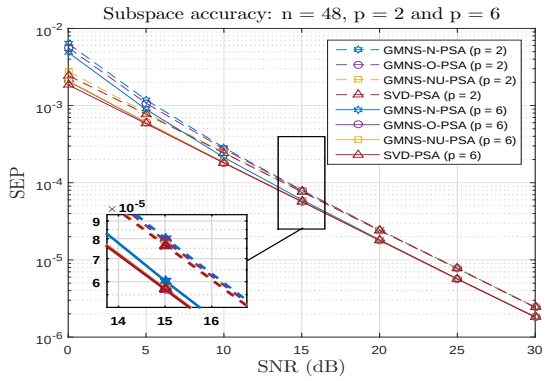


(b) SEP vs. p

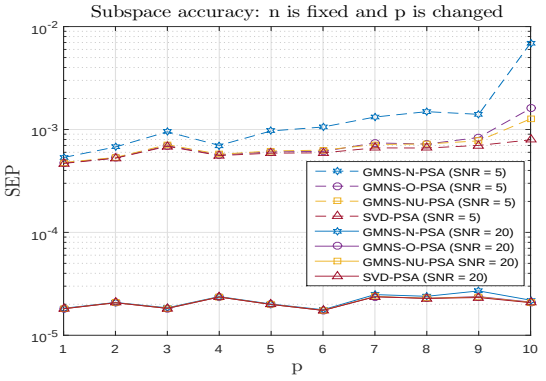
Fig. 5. Minor subspace estimation

B. Principal Subspace/Component Analysis

For PSA, we compare the performance of GMNS-N-PSA, GMNS-O-PSA and GMNS-NU-PSA with the standard PSA method using SVD (SVD-PSA), for $p = 2$ and $p = 6$ as shown in Figures 6(a). Among the four algorithms, GMNS-N-PSA has the lowest performance, as noticed in Section IV-A1. The three other methods reach the SVD, except at low SNRs. Additionally, we conduct a similar experiment as in the MSA case to evaluate the effect of p on the performance of GMNS-N-PSA, GMNS-O-PSA and GMNS-NU-PSA (Figure 6(b)). At low SNR, GMNS-N-PSA degrades the performance when p increases while GMNS-O-PSA, GMNS-NU-PSA and SVD-PSA are comparable. At high SNR, all algorithms are nearly identical in terms of estimation accuracy.



(a) SEP vs. SNR



(b) SEP vs. p

Fig. 6. Principal subspace estimation.

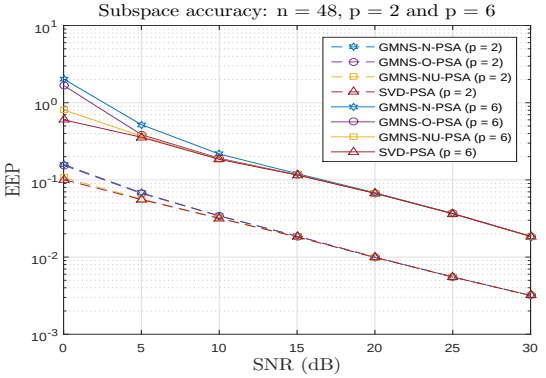
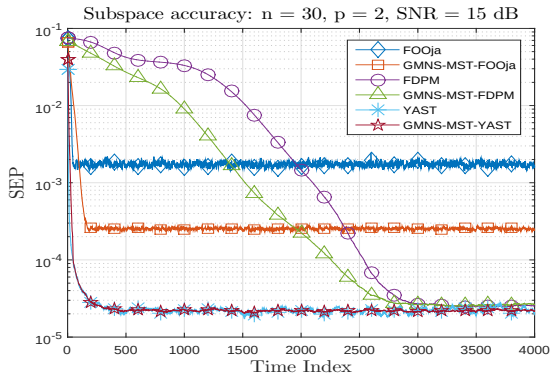


Fig. 7. Principal eigenvector estimation: EEP vs. SNR.

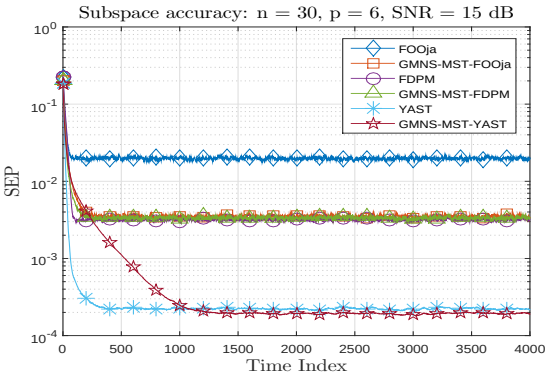
For PCA, the same observation is done as for PSA, as shown in Figures 7. Similar to GMNS-MSA, GMNS-PCA based on the three GMNS-MSA algorithms have the advantage of lower computational cost with a gain approximately being equal to K^2, as compared to the standard PCA method based on SVD.

C. Minor Subspace Tracking

For MST, we choose two low-cost algorithms, FOOja [24] and FDPM [11], and one moderate-cost one, YAST (yet another subspace tracking) [26], and compare them with the corresponding GMNS-based algorithms: GMNS-MST-FOOja, GMNS-MST-FDPM and GMNS-MST-YAST. All algorithms



(a) $p = 2$



(b) $p = 6$

Fig. 8. Minor subspace tracking.

run in a noisy environment with SNR $= 15$ dB. The performance results with respect to $p = 2$ and $p = 6$ are shown in Figure 8, respectively.

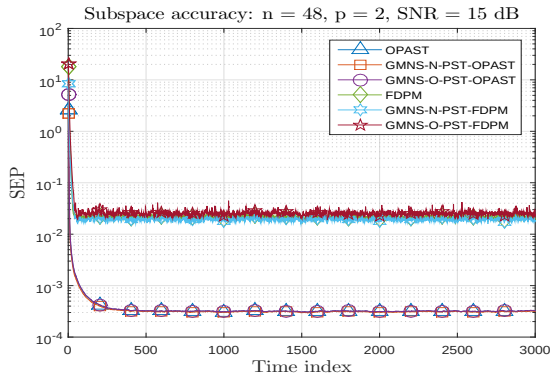
Interestingly, the performance of GMNS-MST-FOOja is better than FOOja even though its convergence rate is slower. The reason is that dividing data into small subsystems reduces the search space which then mitigates the local minima convergence problem and enhances the overall performance. An analogous observation can be seen for GMNS-MST-FDPM and FDPM. Better convergence rate and estimation accuracy are obtained by both YAST[7] and GMNS-MST-YAST but at the expense of higher computational complexity.

D. Principal Subspace Tracking

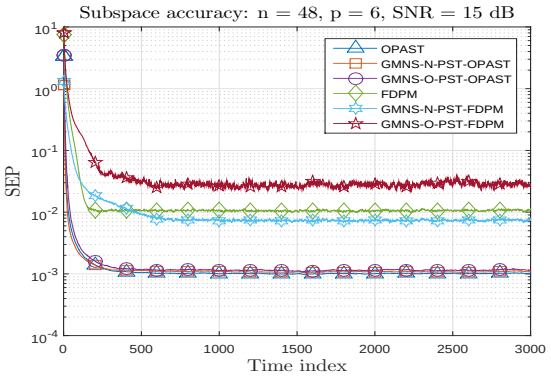
For PST, we compare OPAST and FDPM[8], with their corresponding GMNS-based algorithms: GMNS-N-PST-OPAST, GMNS-O-PST-OPAST, GMNS-N-PST-FDPM and GMNS-O-PST-FDPM. It can be seen from Figure 9 that the GMNS-based algorithms have the same performance as their original algorithms, but with a reduced cost. Also, we observe a clear advantage in favor of OPAST-based algorithms as compared to FDPM-based ones.

[7]YAST has complexity of $O(np^2)$ as presented in [26].

[8]The principal subspace can be obtained by FDPM by changing the sign of step size parameter.

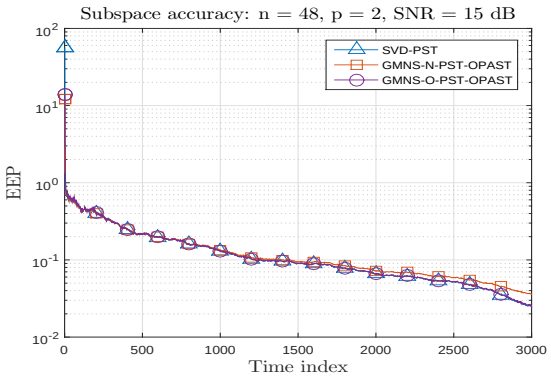


(a) $p = 2$

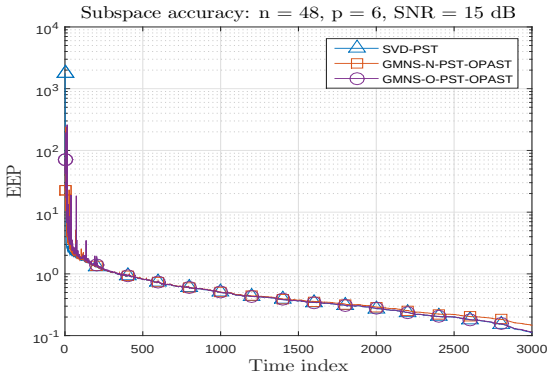


(b) $p = 6$

Fig. 9. Principal subspace tracking.



(a) $p = 2$



(b) $p = 6$

Fig. 10. Principal eigenvector tracking.

E. Principal Eigenvector Tracking

Since the performance of GMNS-based FDPM is worse than GMNS-based OPAST, as just shown above, we apply the GMNS-PET method in Section V-C using OPAST to track the principal eigenvectors, with two parallelized GMNS-based algorithms: GMNS-N-PST-OPAST and GMNS-O-PST-OPAST. Then, we compare their results with the stand SVD-based algorithm (SVD-PST). As shown in Figure 10, both GMNS-N-PST-OPAST and GMNS-O-PST-OPAST have the same performance as that of SVD-PST.

VII. Application to RFI Mitigation in Radio Astronomy

Now, we consider RFI mitigation which is a challenging problem in radio astronomy [27]. In general, RFI is difficult to void even with the spectrum being protected and the deployed area being relatively remote. The interference sources in radio astronomy can stem from various man-made wireless services such as mobile cellular telephone, global positioning system satellites, digital audio and video broadcasting, and so on.

The effect of RFI can be observed in Figure 11(b). In this example, the dataset comes from a single LOFAR station with $n = 48$ antennas, as shown in Figure 11(a). Usually, those 48 antenna outputs are beamformed in real time and the corresponding signal is sent to a central correlator for further radio astronomical processing with other remote stations.

In our experiment, a 5.2 ms signal of those 48 antennas has been stored on disk, and a 2 MHz band around a specific terrestrial RFI (a land mobile at 55 MHz) has been selected to produce a 48×48 covariance matrix. The different sky images presented in this section are derived from this matrix.

In Figure 11(b), while signals-of-interest (SOIs) are not really well-defined, a "strong" RFI appears at the horizon. This image with distortions and artifacts is referred to as "dirty image". RFI can lead to distorted data and unwanted artifacts, causing difficulty in astronomical observation.

To mitigate RFI, an efficient method is based on projection in which a key step is subspace estimation. However, as stated in the introduction, SVD- or EVD-based subspace estimation methods are quite expensive in terms of computational complexity. Here we will illustrate how we can apply the proposed GMNS-based methods to tackle this problem while still preserving the imaging quality.

The model under consideration can be presented as [28]

$$\mathbf{x}(t) = \mathbf{A}_c\mathbf{c}(t) + \mathbf{A}_r\mathbf{r}(t) + \mathbf{n}(t), \tag{49}$$

where $\mathbf{A}_c \in \mathbb{C}^{n\times m}$ is the cosmic source spatial signature matrix, $\mathbf{c} \in \mathbb{C}^m$ is the cosmic source signal vector, $\mathbf{A}_r \in \mathbb{C}^{n\times p}$ is the interference spatial signature matrix, $\mathbf{r}(t) \in \mathbb{C}^p$ is the RFI signal vector, and $\mathbf{n}(t) \in \mathbb{C}^n$ is the additive white noise vector with unknown variance σ^2.

Hence, we can estimate the data covariance matrix as

$$\mathbf{R}_{xx} = \mathbf{A}_c\mathbf{R}_{cc}\mathbf{A}_c^H + \mathbf{A}_r\mathbf{R}_{rr}\mathbf{A}_r^H + \sigma^2\mathbf{I}, \tag{50}$$

where $\mathbf{R}_{cc}$ and $\mathbf{R}_{rr}$ are cosmic and RFI covariance matrices, respectively. Here we have assumed that the cosmic sources, the RFIs and the system noise are uncorrelated. The cosmic sources are point sources because of relative distance between the source and the instrument.

1) Orthogonal projection based RFI mitigation algorithm: This method can be implemented by first computing an orthogonal projection matrix $\mathbf{P}_r$ and then applying it to a "dirty" covariance matrix to produce "clean" one (see [3], [29], [30] and references therein). In particular, the orthogonal projector is computed as

$$\mathbf{P}_r = \mathbf{I} - \mathbf{W}_r(\mathbf{W}_r^H\mathbf{W}_r)^{-1}\mathbf{W}_r^H, \tag{51}$$

where $\mathbf{W}_r$ is the estimated RFI principal subspace using SVD/EVD. We can estimate the "clean" covariance matrix as

$$\bar{\mathbf{R}} = \mathbf{P}_r\mathbf{R}\mathbf{P}_r^H. \tag{52}$$

In fact, we can apply the projection matrix at the pre-correlation stage (i.e., at antenna array output). However, because the data covariance matrix is produced by the radio astronomy system by default, the described method is preferred.

In Figure 11(c), the SVD-based subspace projector is applied according to (51). A subspace of dimension 48×2, corresponding to two principal eigenvalues, is selected. The Milky Way as well as Cassiopeia A and Cygnus A are now strongly visible and the land mobile signal has been mitigated.

2) Qualitative comparison:

In Section VI, we have compared the subspace estimation accuracy between the proposed algorithms and the corresponding SVD-based algorithms quantitatively via numerical simulation. Now we conduct a qualitative experiment using real-life data for further evaluation.

We replace the principal subspace in (51), estimated by the standard SVD-PSA method, with the principal subspace estimated by our GMNS-PSA method and build skymaps after RFI mitigation. Again, the subspace corresponding to two principal eigenvalues defines the RFI subspace ($p = 2$). As can be seen from Figures 11(d)–(f), SOIs are enhanced while RFIs are removed (only the strongest RFI is circled in these figures). From those figures, we can see that the imaging quality based on GMNS-PSA is comparable with that of SVD-PSA. Again, our main advantage is the fact that the cost is reduced by a factor of K^2 compared to SVD-PSA.

While this experiment is based on a significant but still limited number of antennas, our method is potential when massive data like in the SKA project [7] need to be processed or when data from some stations need to be fused.

VIII. Conclusions

In this paper, we have proposed a simple but efficient approach for estimation and tracking of the signal and noise subspaces. The different problems considered in this paper are quite common in many array processing applications and are known as the most expensive tasks in source localization and separation to the extent that many efficient spatial filtering methods have been disregarded in real-life applications which use large antenna arrays, e.g., RFI mitigation in radio astronomy [27], [28][9]. Our GMNS solution exploits the specific array processing model together with a parallel computing architecture to reduce the overall cost by a factor close to K^2, where K is the number of parallel computing units, for large dimensional systems. At the same time, it can be used to fuse data from a number of data sources. Several algorithmic versions of the GMNS have been developed and the performance was assessed via simulated and real-life experiments. The performance results showed that GMNS represents an excellent solution to deal with large size arrays when distributed resources or parallel computing units are available.

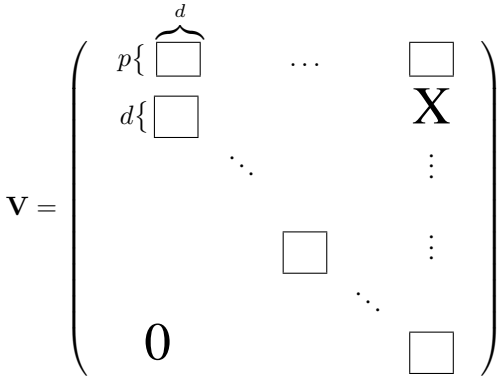


Fig. 12. Block diagonal structure of matrix $\mathbf{V}$.

Appendix A
Proof of Theorem 1

To prove Theorem 1, we need to show that the columns of $\mathbf{V}_i$ belong to the noise subspace and the columns of $\mathbf{V}$ form a vector basis of the noise subspace.

First, note that $\tilde{\mathbf{V}}_i^H\mathbf{A}_i = \mathbf{0}$ leads to $\mathbf{V}_i^H\mathbf{A} = \mathbf{0}$ because of the zero-padding procedure. Hence, the columns of $\mathbf{V}_i$ belong to the noise subspace.

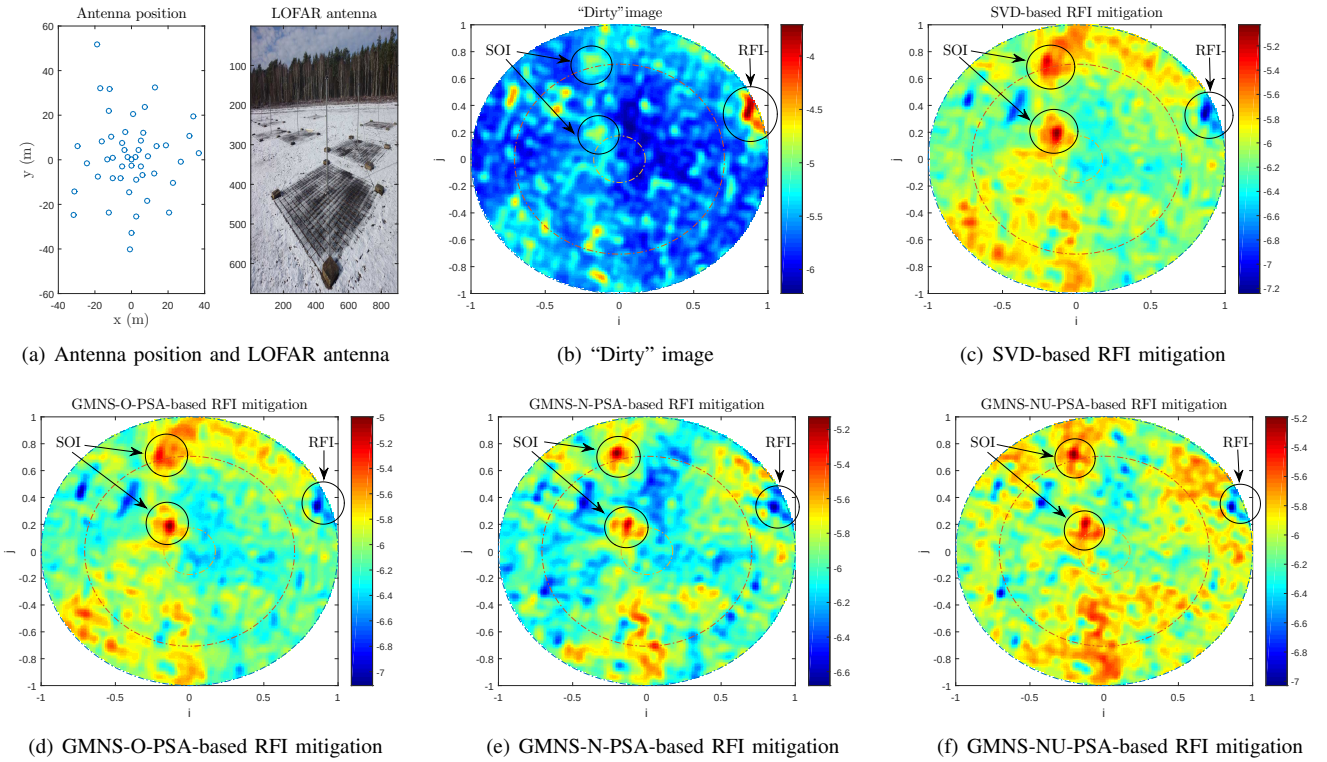
To prove that the columns of $\mathbf{V}$ form a vector basis, let us show that the noise matrix has (up to row permutation) a block diagonal structure as illustrated in Figure 12 with non-singular $d \times d$ diagonal blocks, which guarantee its full column rank. Indeed, according to the GPCS concept, the i-th $d \times d$ diagonal block represents the entries of $\mathbf{V}_i$ corresponding to the d system outputs not shared by the preceding subsystems (i.e., associated to tuples $1, 2, \ldots, i-1$). The block diagonal structure is then a direct consequence of the zero-padding technique used to build $\mathbf{V}_i$ from $\tilde{\mathbf{V}}_i$.

Let us prove now that the $d \times d$ diagonal blocks are non-singular. Indeed, if a given diagonal block matrix is singular, then there exists a noise vector with at most p non-zero entries; this is in contradiction with the assumption that any p rows of matrix $\mathbf{A}$ are linearly independent, as assumed in Theorem 1.

Acknowledgment

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 102.02-2015.32.

[9]Due to their high computational cost, efficient subspace-based RFI mitigation methods are replaced in practice by a simple RFI or 'No RFI' labeling method.



(a) Antenna position and LOFAR antenna (b) "Dirty" image (c) SVD-based RFI mitigation

(d) GMNS-O-PSA-based RFI mitigation (e) GMNS-N-PSA-based RFI mitigation (f) GMNS-NU-PSA-based RFI mitigation

Fig. 11. Image formation before (b) and after RFI mitigation (c-f). Qualitative comparison between SVD-based (c) and GMNS-based (d-f) subspace estimation on RFI mitigation in radio astronomy. Signals-Of-Interest (SOIs): Cassiopeia A and Cygnus A; Radio-Frequency-Interference (RFI): a land mobile signal. (Best view in color)

REFERENCES

[1] P. Stoica, R. Moses, *Spectral Analysis of Signals*. Prentice Hall, 2005.

[2] P. Comon and Ch. Jutten, *Handbook of Blind Source Separation: ICA and Applications*. Elsevier, 2010.

[3] G. Hellbourg, "RFI Spatial Processing for Modern Radio Telescopes," *PhD Thesis, Univ. of Orléans*, Jan. 2014.

[4] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2002.

[5] S. Haykin, K. J. Ray Liu, *Handbook on Array Processing and Sensor Networks*. Wiley-IEEE Press, 2010.

[6] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for Next Generation Wireless Systems," *IEEE Com. Mag.*, 2013.

[7] "Square Kilometre Array: Exploring the Universe with the world's largest radio telescope," *https://www.skatelescope.org/*.

[8] A. Bertrand and M. Moonen, "Distributed adaptive eigenvector estimation of the sensor signal covariance matrix in a fully connected sensor network," in *Proc. ICASSP*, May 2013, pp. 4236–4240.

[9] B. Yang, "Projection approximation subspace tracking," *Signal Processing, IEEE Transactions on*, vol. 43, no. 1, pp. 95–107, 1995.

[10] K. Abed-Meraim, A. Chkeif, and Y. Hua, "Fast orthonormal PAST algorithm," *IEEE Sig. Proc. Ltrs.*, vol. 7, no. 3, pp. 60–62, 2000.

[11] X. G. Doukopoulos and G. V. Moustakides, "Fast and stable subspace tracking," *IEEE Tr. Sig. Proc.*, vol. 56, no. 4, pp. 1452–1465, 2008.

[12] M. Van Haarlem, M. Wise, A. Gunst, G. Heald, J. McKean, J. Hessels, A. De Bruyn, R. Nijboer, J. Swinbank, R. Fallows *et al.*, "LOFAR: The low-frequency array," *Astronomy & Astrophysics*, vol. 556, p. A2, 2013.

[13] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Tr. Ant. Prop.*, vol. 34, no. 3, pp. 276–280, 1986.

[14] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Tr. Acous., Speech, Sig. Proc.*, vol. 37, no. 7, pp. 984–995, 1989.

[15] K. Abed-Meraim and Y. Hua, "Blind identification of multi-input multi-output system using minimum noise subspace," *IEEE Tr-SP*, no. 1, pp. 254–258, 1997.

[16] K. Abed-Meraim, Y. Hua and A. Belouchrani, "Minimum noise subspace: concepts and applications," in *Proc. ICICS*, 1997, pp. 118–121.

[17] W. Souidene, K. Abed-Meraim, and A. Beghdadi, "A new look to multichannel blind image deconvolution," *Image Processing, IEEE Transactions on*, vol. 18, no. 7, pp. 1487–1500, 2009.

[18] M. Thameri, K. Abed-Meraim and A. Belouchrani, "Minor subspace tracking using MNS technique," in *Proc. ICASSP*, 2012, pp. 2433–2436.

[19] V.-D. Nguyen, K. Abed-Meraim, N. Linh-Trung, and R. Weber, "Generalized MNS method for parallel minor and principal subspace analysis," in *Proc. EUSIPCO*. IEEE, 2014, pp. 2265–2269.

[20] M. Thameri, K. Abed-Meraim, and A. Belouchrani, "Minor subspace tracking using MNS technique," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, March 2012, pp. 2433–2436.

[21] V.-D. Nguyen, K. Abed-Meraim, and N. Linh-Trung, "Parallelizable PARAFAC decomposition of 3-way tensors," in *Proc. ICASSP*. IEEE, 2015, pp. 5505–5509.

[22] G. Xu and T. Kailath, "Fast subspace decomposition," *Signal Processing, IEEE Transactions on*, vol. 42, no. 3, pp. 539–551, 1994.

[23] G.H. Golub, C. F. Van Loan, *Matrix computations*. JHU Press, 2012.

[24] S. Bartelmaos and K. Abed-Meraim, "Principal and Minor Subspace Tracking: Algorithms and Stability Analysis," in *Proc. ICASSP*, vol. 3, May 2006, pp. III–III.

[25] R. Badeau, K. Abed-Meraim, G. Richard, and B. David, "Sliding window orthonormal PAST algorithm," in *Proc. ICASSP*, vol. 5, 2003, pp. 261–264.

[26] R. Badeau, G. Richard, and B. David, "Fast and stable YAST algorithm for principal and minor subspace tracking," *Signal Processing, IEEE Transactions on*, vol. 56, no. 8, pp. 3437–3446, 2008.

[27] R. Weber, G. Hellbourg, C. Dumez-Viou, A. Boonstra, S. Torchinsky, C. Capdessus, and K. Abed-Meraim, "RFI Mitigation in Radio Astronomy: an Overview," in *Les Journées Scientifiques d'URSI-France L'électromagnétisme,*, 2013.

[28] G. Hellbourg, R. Weber, K. Abed-Meraim, and A. Boonstra, "RFI spatial processing at Nancay observatory: Approaches and experiments," in *Proc. ICASSP*. IEEE, 2014, pp. 5387–5391.

[29] A.-J. Boonstra, *Radio frequency interference mitigation in radio astronomy*. TU Delft, Delft University of Technology, 2005.

[30] A.-J. van der Veen and S. J. Wijnholds, "Signal processing tools for radio astronomy," in *Handbook of Signal Processing Systems*. Springer, 2013, pp. 421–463.