

LOWER-RANK TENSOR APPROXIMATION AND MULTIWAY FILTERING*

DAMIEN MUTI[†], SALAH BOURENNANE[†], AND JULIEN MAROT[†]

Abstract. This paper presents some recent filtering methods based on the lower-rank tensor approximation approach for denoising tensor signals. In this approach, multicomponent data are represented by tensors, that is, multiway arrays, and the presented tensor filtering methods rely on multilinear algebra. First, the classical channel-by-channel SVD-based filtering method is overviewed. Then, an extension of the classical matrix filtering method is presented. It is based on the lower rank- (K_1, \dots, K_N) truncation of the higher order SVD which performs a multimode principal component analysis (PCA) and is implicitly developed for an additive white Gaussian noise. Two tensor filtering methods recently developed by the authors are also overviewed. The first method consists of an improvement of the multimode PCA-based tensor filtering in the case of an additive correlated Gaussian noise. This improvement is specially done thanks to the fourth order cumulant slice matrix. The second method consists of an extension of Wiener filtering for data tensors. The performances and comparative results between all these tensor filtering methods are presented for the cases of noise reduction in color images, multispectral images, and multicomponent seismic data.

Key words. multilinear algebra, tensor decomposition, multiway arrays, lower-rank approximation, filtering

AMS subject classifications. 15A69, 15A18, 49M27

DOI. 10.1137/060653263

1. Introduction. Tensor data modeling and tensor analysis have been improved and used in several application fields, such as quantum physics, economy, chemometrics, psychology, and data analysis. Nevertheless, only recent studies focus their interest on tensor methods in signal processing applications. Tensor formulation in signal processing has received great attention since the recent development of multicomponent sensors, especially in imagery (color or multispectral images, video, etc.) and seismic fields (antenna of sensors recording waves with polarization properties). Indeed, the digital data obtained from these sensors are fundamentally higher order tensor objects, that is, multiway arrays whose elements are accessed via more than two indexes. Each index is associated with a dimension of the tensor generally called “*n*th-mode” [13, 14, 28, 29].

In recent decades, the classical algebraic processing methods have been specifically developed for vector and matrix representations. They are usually based on the covariance matrix, the cross-spectral matrix, or, more recently, the higher order statistics. Their overall aim is classically to determine a subspace associated with the signal or the parameters to estimate. They mainly rely on three algebraic tools.

- (1) The singular value decomposition (SVD) [18], which is used in principal component analysis (PCA);
- (2) Penrose–Moore matrix inversion [18]; and
- (3) The matrix lower rank approximation, which, according to the Eckart–Young theorem [15], can be achieved thanks to a simple SVD truncation.

These methods have proved to be very efficient in several applications.

*Received by the editors March 1, 2006; accepted for publication (in revised form) by L. De Lathauwer September 10, 2007; published electronically September 25, 2008.

<http://www.siam.org/journals/simax/30-3/65326.html>

[†]Institut Fresnel / CNRS UMR 6133 - Domaine Universitaire de Saint Jérôme, F-13397 Marseille Cedex 20, France (damien.muti@fresnel.fr, salah.bourennane@fresnel.fr, julien.marot@fresnel.fr).

When dealing with multicomponent data represented as tensors, the classical processing techniques consist in rearranging or splitting the data set into matrices or vectors in order for the previously quoted classical algebraic processing methods to be applicable. The original data structure is then built anew, after processing.

In order to keep the data tensor as a whole entity, new signal processing methods have been proposed [35, 36, 37]. Hence, instead of adapting the data tensor to the classical matrix-based algebraic techniques (by rearrangement or splitting), these new methods propose to adapt their processing to the tensor structure of the multicomponent data. This new approach implicitly implies the use of multilinear algebra and mathematical tools that extend the SVD to tensors.

Two main tensor decomposition methods that generalize the matrix SVD have been initially developed to achieve a multimode PCA and recently used in tensor signal processing. They rely on two models, the TUCKER3 model and the PARAFAC model.

The TUCKER3 model [29, 48] was adopted in higher order SVD (HOSVD) [2, 13] and in lower rank- (K_1, \dots, K_N) tensor approximation [11, 14, 47]. We denote by HOSVD- (K_1, \dots, K_N) the truncation of HOSVD, performed with ranks (K_1, \dots, K_N) , in modes $1, \dots, N$, respectively. This model recently has been used as multimode PCA, in seismics for wave separation based on a subspace method, in image processing for face recognition and expression analysis [49, 52], and in noise filtering of color images [36].

The PARAFAC model and the CANDECOMP model were developed in [20] and [10], respectively. In [30] the link was set between CANDECOMP and PARAFAC models. The CANDECOMP/PARAFAC model, referred to as the CP model [25], has recently been applied to the food industry [9], array processing [45], and telecommunications [46].

These two decomposition methods differ in the tensor rank definition on which they are based. The HOSVD- (K_1, \dots, K_N) and the rank- (K_1, \dots, K_N) approximation rely on the n th-mode rank definition, that is, the rank of the tensor n th-mode flattening matrix [13, 14]. The rank- (K_1, \dots, K_N) approximation [14] relies on an optimization algorithm which is initialized by the HOSVD- (K_1, \dots, K_N) [13]. The rank- (K_1, \dots, K_N) approximation improves the approximation obtained with the HOSVD- (K_1, \dots, K_N) . It relies on the determination of the signal subspace in every n th-mode of the data tensor and copes with additive white Gaussian noise. The rank- (K_1, \dots, K_N) approximation provides the best approximation in the sense of least Frobenius norm of the difference between estimated and expected tensors. Nevertheless it assumes a noncorrelated Gaussian noise. To face the case of correlated Gaussian noise, a variant of rank- (K_1, \dots, K_N) approximation, based on fourth order cumulants, was proposed [39]. Indeed, as it is proved in [33], the fourth order cumulants of a Gaussian variable are null.

A tensor framework was employed by [12] to express the solution to the linear independent component analysis (ICA) problem which employs fourth order cumulants. The multilinear ICA (N-mode ICA) model [50, 51], which was developed for face recognition, encodes the fourth order cumulants for each of the n th-mode flattening matrices of the tensor.

The CP model relies on a canonical decomposition of a tensor into a summation of rank-one tensors and on the extension of the classical matrix rank. Details on the tensor ranks and orthogonal tensor decomposition can be found in [22, 27].

When the TUCKER3 model and the PARAFAC model are associated with an ALS loop, they are known respectively as the TUCKALS3 algorithm [29, 28] and

the PARAFAC ALS algorithm [30, 20]. Many recent studies have been conducted to improve the convergence of these algorithms [14, 26, 56, 44].

The goal of this paper is to present an overview of the principal results concerning this new approach of data tensor filtering. More details on the algorithms presented in this survey can be found in [35, 36, 38, 39]. These algorithms are analogous to multilinear ICA but were developed independently for image filtering. The presented algorithms are based on a signal subspace approach, so they are efficient when the noise components are uncorrelated, when the signal and the additive noise are uncorrelated, and when some rows or columns of the image are redundant. In this case it is possible to distinguish between a signal subspace and a noise subspace, as for the traditional SVD-based filtering and Wiener filtering algorithms. Wiener filtering requires prior knowledge on the expected noise-free signal or image. However, multiway filtering methods provide the following advantage over traditional filtering methods: by apprehending a multiway data set as a whole entity, they take into account the dependence between modes thanks to ALS algorithms. The goal of the paper is also to present some simulations and comparative results concerning color images and multicomponent seismic signal filtering.

The paper is organized as follows. Section 2 presents the tensor data and a short overview of its main properties. Section 3 introduces the tensor formulation of the classical noise-removal problem as well as some new tensor filtering notations. First, we explain how the channel-by-channel SVD-based method processes successively each component of the data tensor. Second, we consider two methods that take into account the relationships between each component of the considered tensor. These two methods are based on the n th-mode signal subspace. The first method for signal tensor estimation is based on multimode PCA achieved by rank- (K_1, \dots, K_N) approximation. The second method is a new tensor version of Wiener filtering. Section 4 presents some comparative results where the overviewed multiway filtering methods are applied to noise reduction in color images, denoising of multispectral images, and denoising of multicomponent seismic waves. Section 5 concludes the paper.

The following notation is used in the rest of the paper. Scalars are denoted by italic lowercase roman (a); vectors by boldface lowercase roman (\mathbf{a}); matrices by boldface uppercase roman (\mathbf{A}); and tensors by uppercase calligraphic (\mathcal{A}). We distinguish a random vector, like \mathbf{a} , from one of its realizations by using a supplementary index, like \mathbf{a}_i .

2. Tensor representation and properties. We define a tensor of order N as a multidimensional array whose entries are accessed via N indexes. A tensor is denoted by $\mathcal{A} \in \mathbb{R}^{I_1 \times \dots \times I_N}$, where each element is denoted by $a_{i_1 \dots i_N}$, and \mathbb{R} is the real manifold. Each dimension of a tensor is called n th-mode, where n refers to the n^{th} index. Figure 2.1 shows how a color image can be represented by a third order tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, where I_1 is the number of rows, I_2 is the number of columns, and I_3 is the number of color channels. In the case of a color image, we have $I_3 = 3$. Let us define $E^{(n)}$ as the n th-mode vector space of dimension I_n , associated with the n th-mode of tensor \mathcal{A} . By definition, $E^{(n)}$ is generated by the column vectors of the n th-mode flattening matrix. The n th-mode flattening matrix \mathbf{A}_n of tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times \dots \times I_N}$ is defined as a matrix from $\mathbb{R}^{I_n \times M_n}$, where

$$(2.1) \quad M_n = I_{n+1} I_{n+2} \cdots I_N I_1 I_2 \cdots I_{n-1}.$$

For example, when we consider a third order tensor, the definition of the matrix flattening involves the dimensions I_1 , I_2 , I_3 in a backward cyclic way [5, 13, 25].

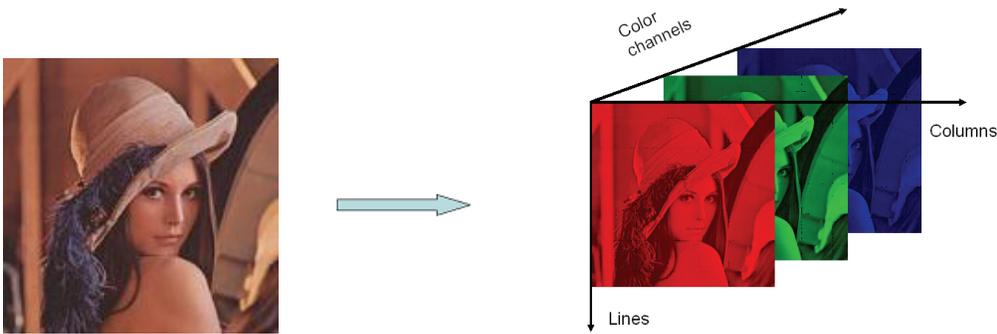


FIG. 2.1. *Lena standard color image and its tensor representation.*

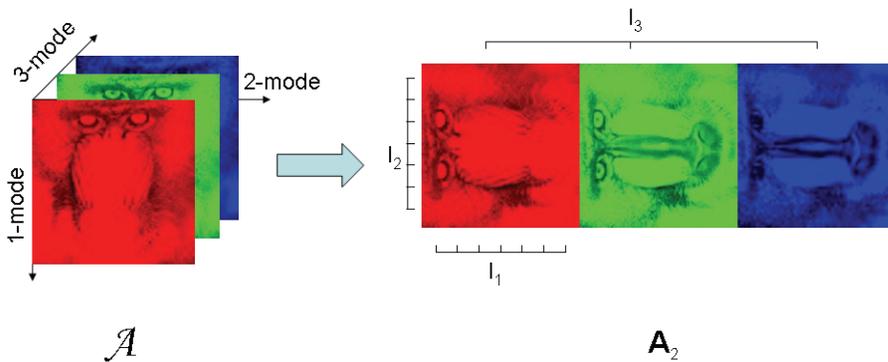


FIG. 2.2. *2nd-mode flattening of tensor \mathcal{A} : \mathbf{A}_2 .*

When dealing with a 1st-mode flattening of dimensionality $I_1 \times (I_2 I_3)$, we formally assume that the index i_2 varies more slowly than i_3 . For all $n = 1$ to 3, \mathbf{A}_n columns are the I_n -dimensional vectors obtained from \mathcal{A} by varying the index i_n from 1 to I_n and keeping the other indexes fixed. These vectors are called the n th-mode vectors of tensor \mathcal{A} . An illustration of the 2nd-mode flattening of a color image is presented in Figure 2.2.

In the following, we use the operator \times_n as the n th-mode product, which generalizes the matrix product to tensors. Given $\mathcal{A} \in \mathbb{R}^{I_1 \times \dots \times I_N}$ and a matrix $\mathbf{U} \in \mathbb{R}^{J_n \times I_n}$, the n th-mode product between tensor \mathcal{A} and matrix \mathbf{U} leads to the tensor $\mathcal{B} = \mathcal{A} \times_n \mathbf{U}$, which is a tensor of $\mathbb{R}^{I_1 \times \dots \times I_{n-1} \times J_n \times I_{n+1} \times \dots \times I_N}$, whose entries are given by

$$(2.2) \quad b_{i_1 \dots i_{n-1} j_n i_{n+1} \dots i_N} = \sum_{i_n=1}^{I_n} a_{i_1 \dots i_{n-1} i_n i_{n+1} \dots i_N} u_{j_n i_n}.$$

The next section presents the recent filtering methods for tensor data.

3. Tensor filtering problem formulation. The tensor data extend the classical vector data. The measurement of a multidimensional and multiway signal \mathcal{X} by multicomponent sensors with additive noise \mathcal{N} results in a data tensor \mathcal{R} such that

$$(3.1) \quad \mathcal{R} = \mathcal{X} + \mathcal{N}.$$

\mathcal{R} , \mathcal{X} , and \mathcal{N} are tensors of order N from $\mathbb{R}^{I_1 \times \dots \times I_N}$. Tensors \mathcal{N} and \mathcal{X} represent noise and signal parts of the data, respectively. The goal of this study is to estimate the expected signal \mathcal{X} thanks to a multidimensional filtering of the data [35, 36, 38, 39]:

$$(3.2) \quad \hat{\mathcal{X}} = \mathcal{R} \times_1 \mathbf{H}^{(1)} \times_2 \mathbf{H}^{(2)} \times_3 \dots \times_N \mathbf{H}^{(N)},$$

From a signal processing point of view, the n th-mode product is a n th-mode filtering of data tensor \mathcal{R} by n th-mode filter $\mathbf{H}^{(n)}$. Consequently, for all $n = 1$ to N , $\mathbf{H}^{(n)}$ is the n th-mode filter applied to the n th-mode of the data tensor \mathcal{R} .

In this paper we assume that the noise \mathcal{N} is independent from the signal \mathcal{X} and that the n th-mode rank K_n is smaller than the n th-mode dimension I_n ($K_n < I_n$, for all $n = 1$ to N). Then it is possible to extend the classical subspace approach to tensors by assuming that, whatever the n th-mode, the vector space $E^{(n)}$ is the direct sum of two orthogonal subspaces, namely, $E_1^{(n)}$ and $E_2^{(n)}$, which are defined as follows:

- $E_1^{(n)}$ is the subspace of dimension K_n , spanned by the K_n singular vectors associated with the K_n largest singular values of matrix \mathbf{X}_n ; $E_1^{(n)}$ is called the signal subspace [1, 33, 55, 54].
- $E_2^{(n)}$ is the subspace of dimension $I_n - K_n$, spanned by the $I_n - K_n$ singular vectors associated with the $I_n - K_n$ smallest singular values of matrix \mathbf{X}_n ; $E_2^{(n)}$ is called the noise subspace [1, 33, 55, 54].

The dimensions K_1, K_2, \dots, K_N can be estimated by means of the well-known Akaike information criterion (AIC) or Minimum description length (MDL) criteria [53], which are entropy-based information criteria. Hence, one way to estimate signal tensor \mathcal{X} from noisy data tensor \mathcal{R} is to estimate $E_1^{(n)}$ in every n th-mode of \mathcal{R} . The following section presents three tensor filtering methods based on n th-mode signal subspaces. The first method is an extension of classical matrix filtering algorithms. It consists of a channel-by-channel SVD-based filtering.

The second filtering method is based on multimode PCA achieved by rank- (K_1, \dots, K_N) approximation. Two algorithms are presented for this case. The first algorithm is implicitly developed for an additive *white* and Gaussian noise assumption, whereas the second algorithm represents an improvement of the first one in the case of a *correlated* Gaussian noise. This improvement is achieved thanks to higher order statistics.

The third method, the multiway Wiener filtering (Wmm- (K_1, \dots, K_N)), is an algorithm that extends the classical two-dimensional Wiener filtering to tensor data.

3.1. Channel-by-channel SVD-based filtering. The classical algebraical methods operate on two-dimensional data matrices and are based on the SVD [1, 3, 4] and on the Eckart–Young theorem concerning the best lower rank approximation of a matrix [15] in the least-squares sense.

In the first method, a preprocessing is applied to the multidimensional and multiway data. It consists in splitting data tensor \mathcal{R} , representing the noisy multicomponent image into two-dimensional “slice matrices” of data, each representing a specific channel. According to the classical signal subspace methods [8], the left and right signal subspaces, corresponding to, respectively, the column and the row vectors of each slice matrix, are simultaneously determined by processing the SVD of the matrix associated with the data of the slice matrix. Let us consider the slice matrix $\mathcal{R}(:, :, i_3, \dots, i_j, \dots, i_N)$ of data tensor \mathcal{R} . Projectors \mathbf{P} on the left signal subspace and \mathbf{Q} on the right signal subspace are built from, respectively, the left and the right singular vectors associated with the K largest singular values of $\mathcal{R}(:, :, i_3, \dots, i_j, \dots, i_N)$.

The parameter K simultaneously defines the dimensions of the left and right signal subspaces. Applying the projectors \mathbf{P} and \mathbf{Q} on the slice $\mathcal{R}(:, :, i_3, \dots, i_j, \dots, i_N)$ amounts to computing its best lower rank- K matrix approximation [15] in the least-squares sense.

The filtering of each slice matrix of data tensor \mathcal{R} separately is called in the following “channel-by-channel” SVD-based filtering of \mathcal{R} . It consists of a first way to estimate the signal tensor \mathcal{X} and can be summarized by the following steps:

1. input: data tensor \mathcal{R} , left and right signal subspace dimension K .
 - for $i_N = 1$ to I_N :
 - for $i_{N-1} = 1$ to I_{N-1} :
 - ⋮
 - for $i_4 = 1$ to I_4 :
 - for $i_3 = 1$ to I_3 :

- (a) calculate matrix $\mathcal{R}(:, :, i_3, \dots, i_j, \dots, i_N)$ SVD:

$$\mathcal{R}(:, :, i_3, \dots, i_j, \dots, i_N) = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T,$$

where $\mathbf{\Sigma}$ is the core matrix regrouping the singular values of the matrix $\mathcal{R}(:, :, i_3, \dots, i_j, \dots, i_N)$, and $\mathbf{U} = [\mathbf{u}_1 \dots \mathbf{u}_{I_1}]$ and $\mathbf{V} = [\mathbf{v}_1 \dots \mathbf{v}_{I_2}]$ are the matrices containing the left and right singular vectors defined respectively by \mathbf{u}_{i_1} and \mathbf{v}_{i_2} .

- (b) construct matrices $\mathbf{U}_K = [\mathbf{u}_1 \dots \mathbf{u}_K]$ and $\mathbf{V}_K = [\mathbf{v}_1 \dots \mathbf{v}_K]$ containing the K largest left and right eigenvectors of $\mathcal{R}(:, :, i_3, \dots, i_j, \dots, i_N)$;
- (c) compute the projector $\mathbf{P} = \mathbf{U}_K \mathbf{U}_K^T$ on the column signal subspace, and projector $\mathbf{Q} = \mathbf{V}_K \mathbf{V}_K^T$ on the row signal subspace.
- (d) compute the two-dimensional slice matrices of the estimated expected signal $\hat{\mathcal{X}}$:

$$\hat{\mathcal{X}}(:, :, i_3, \dots, i_j, \dots, i_N) = \mathbf{P} \mathcal{R}(:, :, i_3, \dots, i_j, \dots, i_N) \mathbf{Q}$$

2. output: estimated expected signal: $\hat{\mathcal{X}}$.

Channel-by-channel SVD-based filtering is based on a common efficient method but exhibits a major drawback: it does not take into account the relationships between the components of the processed tensor. Moreover, channel-by-channel SVD-based filtering is appropriate only on some conditions. For example, applying SVD-based filtering to an image is generally appropriate when the rows or columns of an image are redundant, that is, linearly dependent. In this case, the rank K of the image is equal to the number of linearly independent rows or columns. It is only in this case that it would be safe to throw out eigenvectors from $K + 1$ on. It is only in this special case that the noise subspace is orthogonal to the signal subspace. Otherwise, the noise simply increases the variance of the signal subspace and underestimating the signal subspace dimension would result in throwing out both signal and noise information. Thus, one would lose spatial resolution.

The next subsection presents a multiway filtering method that processes jointly, and not successively, each component of the data tensor.

3.2. Tensor filtering based on multimode PCA.

3.2.1. White decorrelated Gaussian noise and second-order-statistics-based method. Assuming that the dimension K_n of the signal subspace is known

for all $n = 1$ to N , one way to estimate the expected signal tensor \mathcal{X} from the noisy data tensor $\mathcal{R} = \mathcal{X} + \mathcal{N}$ is to orthogonally project, for every n th-mode, the vectors of tensor \mathcal{R} on the n th-mode signal subspace $E_1^{(n)}$ for all $n = 1$ to N . This statement is equivalent to replacing in (3.2) the filters $\mathbf{H}^{(n)}$ by the projectors $\mathbf{P}^{(n)}$ on the n th-mode signal subspace:

$$(3.3) \quad \widehat{\mathcal{X}} = \mathcal{R} \times_1 \mathbf{P}^{(1)} \times_2 \cdots \times_N \mathbf{P}^{(N)}.$$

In this last formulation, projectors $\mathbf{P}^{(n)}$ are estimated thanks to a multimode PCA applied to data tensor \mathcal{R} . This multimode PCA-based filtering generalizes the classical matrix filtering methods [16, 17, 21, 23, 24, 32] and implicitly supposes that the additive noise is *white* and *Gaussian*.

In the vector or matrix formulation, the definition of the projector on the signal subspace is based on the eigenvectors associated with the largest eigenvalues of the covariance matrix of the set of observation vectors. Hence, the determination of the signal subspace amounts to determine the best approximation (in the least-squares sense) of the observation matrix or the covariance matrix.

As an extension to the vector and matrix cases, in the tensor formulation, the projectors on the n th-mode vector spaces are determined by computing the rank- (K_1, \dots, K_N) approximation of \mathcal{R} in the least-squares sense. From a mathematical point of view, the rank- (K_1, \dots, K_N) approximation of \mathcal{R} is represented by tensor $\mathcal{B}^{K_1, \dots, K_N}$ which minimizes the quadratic tensor Frobenius norm $\|\mathcal{R} - \mathcal{B}\|^2$ subject to the condition that $\mathcal{B} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ is a rank- (K_1, \dots, K_N) tensor. The description of the TUCKALS3 algorithm used in rank- (K_1, \dots, K_N) approximation is provided in the following.

Rank- (K_1, \dots, K_N) approximation - TUCKALS3 algorithm.

1. **Input:** data tensor \mathcal{R} , and dimensions K_1, \dots, K_N of all n th-mode signal subspaces.
2. **Initialization** $k = 0$: For $n = 1$ to N , calculate the projectors $\mathbf{P}_0^{(n)}$ given by HOSVD- (K_1, \dots, K_N) :
 - (a) n th-mode flatten \mathcal{R} into matrix \mathbf{R}_n ;
 - (b) Compute the SVD of \mathbf{R}_n ;
 - (c) Compute matrix $\mathbf{U}_0^{(n)}$ formed by the K_n eigenvectors associated with the K_n largest singular values of \mathbf{R}_n . $\mathbf{U}_0^{(n)}$ is the initial matrix of the n th-mode signal subspace orthogonal basis vectors;
 - (d) Form the initial orthogonal projector $\mathbf{P}_0^{(n)} = \mathbf{U}_0^{(n)} \mathbf{U}_0^{(n)T}$ on the n th-mode signal subspace;
 - (e) Compute the HOSVD- (K_1, \dots, K_N) of tensor \mathcal{R} given by $\mathcal{B}_0 = \mathcal{R} \times_1 \mathbf{P}_0^{(1)} \times_2 \cdots \times_N \mathbf{P}_0^{(N)}$;
3. **ALS loop:** Repeat until convergence, that is, for example, while $\|\mathcal{B}_{k+1} - \mathcal{B}_k\|^2 > \epsilon$, $\epsilon > 0$ being a prior fixed threshold,
 - (a) For $n = 1$ to N :
 - i. Form $\mathcal{B}^{(n),k}$:

$$\mathcal{B}^{(n),k} = \mathcal{R} \times_1 \mathbf{P}_{k+1}^{(1)} \times_2 \cdots \times_{n-1} \mathbf{P}_{k+1}^{(n-1)} \times_{n+1} \mathbf{P}_k^{(n+1)} \times_{n+2} \cdots \times_N \mathbf{P}_k^{(N)};$$
 - ii. n th-mode flatten tensor $\mathcal{B}^{(n),k}$ into matrix $\mathbf{B}_n^{(n),k}$;
 - iii. Compute matrix $\mathbf{C}^{(n),k} = \mathbf{B}_n^{(n),k} \mathbf{R}_n^T$;
 - iv. Compute matrix $\mathbf{U}_{k+1}^{(n)}$ composed of the K_n eigenvectors associated

with the K_n largest eigenvalues of $\mathbf{C}^{(n),k}$. $\mathbf{U}_k^{(n)}$ is the matrix of the n th-mode signal subspace orthogonal basis vectors at the k^{th} iteration;

v. Compute $\mathbf{P}_{k+1}^{(n)} = \mathbf{U}_{k+1}^{(n)} \mathbf{U}_{k+1}^{(n)T}$;

(b) Compute $\mathcal{B}_{k+1} = \mathcal{R} \times_1 \mathbf{P}_{k+1}^{(1)} \times_2 \cdots \times_N \mathbf{P}_{k+1}^{(N)}$;

(c) Increment k .

4. **Output:** the estimated signal tensor is obtained through $\hat{\mathcal{X}} = \mathcal{R} \times_1 \mathbf{P}_{k_{stop}}^{(1)} \times_2 \cdots \times_N \mathbf{P}_{k_{stop}}^{(N)}$. $\hat{\mathcal{X}}$ is the rank- (K_1, \dots, K_N) approximation of \mathcal{R} , where k_{stop} is the index of the last iteration after the convergence of TUCKALS3 algorithm.

In this algorithm, the second order statistics come from the SVD of matrix \mathbf{R}_n at step 2(b), which is equivalent, up to $\frac{1}{M_n}$ multiplicative factor, to the estimation of tensor \mathcal{R} n th-mode vectors [39]. The definition of M_n is given in (2.1). In the same way, at step 3(a)iii, matrix $\mathbf{C}^{(n),k}$ is, up to $\frac{1}{M_n}$ multiplicative factor, the estimation of the covariance matrix between tensor \mathcal{R} and tensor $\mathcal{B}^{(n),k}$ n th-mode vectors. According to step 3(a)ii, $\mathcal{B}^{(n),k}$ represents data tensor \mathcal{R} filtered in every m th-mode but the n th-mode, by projection-filters $\mathbf{P}_l^{(m)}$, with $m \neq n$, $l = k$ if $m > n$ and $l = k + 1$ if $m < n$. TUCKALS3 algorithm has recently been used to process a multimode PCA in order to perform white noise removal in color images [36].

A good approximation of the rank- (K_1, \dots, K_N) approximation can simply be achieved by computing the HOSVD- (K_1, \dots, K_N) of tensor \mathcal{R} [14, 34]. Indeed, the HOSVD- (K_1, \dots, K_N) of \mathcal{R} consists of the initialization step of TUCKALS3 algorithm and hence can be considered as a suboptimal solution for the rank- (K_1, \dots, K_N) approximation of tensor \mathcal{R} [14]. This HOSVD-based technique has recently been used in [39] for denoising and source separation of multicomponent seismic waves.

3.2.2. Correlated Gaussian noise and higher-order-statistics-based method. In practice, the condition of noise whiteness is not always fulfilled. Hence, in the case of an additive *correlated* Gaussian noise, the TUCKALS3 algorithm is theoretically incapable of providing a good estimation of the n th-mode signal subspaces since it is based on second order moments. A classical means to remove the Gaussian (noise) components is to use the higher order statistics, and especially the higher order cumulants. The tensor framework has been used to compute the fourth order cumulants as a means of solving the ICA problem [12]. Vasilescu and Terzopoulos introduced a multilinear ICA (N-mode ICA) for face recognition, which encodes the higher order statistics associated with each mode of the tensor [50, 51]. The related methods are based on the well-known cumulant property stating that the higher order cumulants of a Gaussian variable are null [31, 33].

As a consequence, in the case of an additive *correlated* Gaussian noise, a recent study [39] has proposed to improve the multimode PCA-based filtering by incorporating into the TUCKALS3 algorithm the fourth order cumulants instead of the second order moments.

From a practical point of view, second order matrices $\mathbf{C}^{(n),0}$ and $\mathbf{C}^{(n),k}$ at steps 2(b) and 3(a)iii of the TUCKALS3 algorithm are replaced with the corresponding fourth order cumulants. In the following, we present only the details of the procedure for matrix $\mathbf{C}^{(n),k}$. Obtaining the details concerning $\mathbf{C}^{(n),0}$ is straightforward.

We assume that $\{\mathbf{r}_p^{(n)}, p = 1, \dots, M_n\}$ and $\{\mathbf{b}_p^{(n),k}, p = 1, \dots, M_n\}$ are the M_n realizations of two random vectors $\mathbf{r}^{(n)}$ and $\mathbf{b}^{(n),k}$. In practice, we take as the realizations of these two random vectors the n th-mode vectors of data tensors \mathcal{R} and $\mathcal{B}^{(n),k}$.

Matrix $\mathbf{C}^{(n),k}$ reads

$$(3.4) \quad \mathbf{C}^{(n),k} = \sum_{p=1}^{M_n} \mathbf{b}_p^{(n),k} \mathbf{r}_p^{(n)T}.$$

The fourth order cumulants associated with vectors $\mathbf{r}^{(n)}$ and $\mathbf{b}^{(n),k}$ are denoted by

$$(3.5) \quad \mathcal{C}^{(n),k} = \text{Cum}(\mathbf{b}^{(n),k}, \mathbf{b}^{(n),kT}, \mathbf{r}^{(n)}, \mathbf{r}^{(n)T}),$$

where $\text{Cum}(\cdot)$ denotes the cumulant operator. $\mathcal{C}^{(n),k}$ is a fourth order super-symmetric tensor from $\mathbb{R}^{I_n \times I_n \times I_n \times I_n}$, whose generic term for indexes (i_1, i_2, j_1, j_2) , for centered variables, is given by [19, 31]

$$(3.6) \quad \begin{aligned} (\mathcal{C}^{(n),k})_{i_1, i_2, j_1, j_2} &= \mathbb{E}[b_{i_1}^{(n),k} b_{i_2}^{(n),k} r_{j_1}^{(n)} r_{j_2}^{(n)}] \\ &\quad - \mathbb{E}[b_{i_1}^{(n),k} r_{j_1}^{(n)}] \mathbb{E}[b_{i_2}^{(n),k} r_{j_2}^{(n)}], -\mathbb{E}[b_{i_1}^{(n),k} r_{j_2}^{(n)}] \mathbb{E}[b_{i_2}^{(n),k} r_{j_1}^{(n)}] \end{aligned}$$

where $b_i^{(n),k}$ and $r_j^{(n)}$ are the i th and j th components of random vectors $\mathbf{b}^{(n),k}$ and $\mathbf{r}^{(n)}$, and $\mathbb{E}[\cdot]$ is the expectation operator. The practical estimation of $(\mathcal{C}^{(n),k})_{i_1, i_2, j_1, j_2}$ is given by

$$(3.7) \quad \begin{aligned} (\mathcal{C}^{(n),k})_{i_1, i_2, j_1, j_2} &= \frac{1}{M_n} \left(\sum_{p=1}^{M_n} (b_{i_1 p}^{(n),k} b_{i_2 p}^{(n),k} r_{j_1 p}^{(n)} r_{j_2 p}^{(n)}) \right) \\ &\quad - \frac{1}{M_n^2} \left(\sum_{p=1}^{M_n} (b_{i_1 p}^{(n),k} r_{j_1 p}^{(n)}) \right) \left(\sum_{p=1}^{M_n} (b_{i_2 p}^{(n),k} r_{j_2 p}^{(n)}) \right) \\ &\quad - \frac{1}{M_n^2} \left(\sum_{p=1}^{M_n} (b_{i_1 p}^{(n),k} r_{j_2 p}^{(n)}) \right) \left(\sum_{p=1}^{M_n} (b_{i_2 p}^{(n),k} r_{j_1 p}^{(n)}) \right). \end{aligned}$$

Here, $b_{ip}^{(n),k}$ and $r_{jp}^{(n)}$ are the elements at position (i, j) of tensors $\mathcal{B}^{(n),k}$ and \mathcal{R} n th-mode flattening matrices $\mathbf{B}_n^{(n),k}$ and \mathbf{R}_n .

In the classical TUCKALS3 algorithm, the K_n n th-mode signal subspace basis vectors, given by matrix $\mathbf{U}^{(n),k}$, are estimated by computing, at step 3a, the eigenvectors associated with the K_n largest eigenvalues of matrix $\mathbf{C}^{(n),k}$. This amounts to computing the best lower rank- K_n approximation of $\mathbf{C}^{(n),k}$. In [41] fourth order cumulants are used instead of the covariance matrix because of their ability to remove Gaussian noise. Indeed, the fourth order cumulants of Gaussian variables are null. Therefore, when dealing with an additive *correlated* Gaussian noise, we also use fourth order cumulants [39].

The main drawback of fourth order cumulants is the high computational load to build every fourth order cumulant tensor associated with the n th-mode of the data tensor. This computational load depends on the size of the data tensor \mathcal{R} , that is, the values of I_n , for all $n = 1$ to N . One way to reduce the computational load has been proposed in [39] and consists in using the fourth order cumulant slice matrix. The cumulant slice matrix has initially been introduced in array processing for source localization or directions-of-arrival (DOA) estimation [7, 55, 54]. In [19, 55, 54], it is proved that the signal subspace spanned by the eigenvectors associated with the largest eigenvalues of a cumulant slice matrix is the same as signal subspace obtained from the whole cumulant tensor defined in (3.5) [55, 54]. Therefore, we use only the

eigenvectors of one cumulant slice matrix in our algorithm (see step 2(a)iii because the other cumulant slice matrices provide redundant information. The use of the fourth order cumulant slice matrix provides a much faster algorithm [54]. In our application, the fourth order cumulant slice matrix $\mathbf{C}_q^{(n),k}$ can be defined, from (3.6), by fixing the q^{th} component of vector $\mathbf{b}^{(n),k}$ as follows:

$$(3.8) \quad \left(\mathbf{C}_q^{(n),k}\right)_{ij} = \mathbb{E} \left[(b_q^{(n),k})^2 r_i^{(n)} r_j^{(n)} \right] - 2\mathbb{E} \left[b_q^{(n),k} r_i^{(n)} \right] \mathbb{E} \left[b_q^{(n),k} r_j^{(n)} \right].$$

The practical estimation of $(\mathbf{C}_q^{(n),k})_{ij}$ can be given by

$$(3.9) \quad \left(\mathbf{C}_q^{(n),k}\right)_{ij} = \frac{1}{M_n} \left(\sum_{p=1}^{M_n} (b_{qp}^{(n),k})^2 r_{ip}^{(n)} r_{jp}^{(n)} \right) - \frac{2}{M_n^2} \left(\sum_{p=1}^{M_n} b_{qp}^{(n),k} r_{ip}^{(n)} \right) \left(\sum_{p=1}^{M_n} b_{qp}^{(n),k} r_{jp}^{(n)} \right),$$

where $b_{ij}^{(n),k}$ and $r_{ij}^{(n)}$ are, respectively, the elements at position (i, j) in the n th-mode flattening matrices $\mathbf{B}_n^{(n),k}$ and \mathbf{R}_n of tensors $\mathcal{B}^{(n),k}$ and \mathcal{R} .

As a consequence, in the case of an additive *correlated* Gaussian noise, the K_n n th-mode signal subspace basis vectors can now be estimated by computing matrix $\mathbf{C}_q^{(n),k}$ lower rank- K_n approximation. Then, the fourth order cumulant slice matrix-based multimode PCA-based filtering can be summarized as follows:

1. **Initialization** $k = 0$:

For all $n = 1$ to N , $\mathbf{P}_0^{(n)} = \mathbf{U}_0^{(n)} \mathbf{U}_0^{(n)T}$. $\mathbf{U}_0^{(n)}$ is the matrix of the K_n eigenvectors associated with the K_n largest eigenvalues of fourth order cumulant slice matrix $\mathbf{C}_q^{(n),0}$ of tensor \mathcal{R} n th-mode vectors.

2. **ALS loop:**

The steps (b) and (c) of the ALS loop are the same as in the algorithm “**rank- (K_1, \dots, K_N) approximation - TUCKALS3 algorithm**” described previously, and step (a) is replaced by

(a) For $n = 1$ to N :

i. $\mathcal{B}^{(n),k} = \mathcal{R} \times_1 \mathbf{P}_{k+1}^{(1)} \times_2 \dots \times_{n-1} \mathbf{P}_{k+1}^{(n-1)} \times_{n+1} \mathbf{P}_k^{(n+1)} \times_{n+2} \dots \times_N \mathbf{P}_k^{(N)}$;

ii. Compute cumulant slice matrix $\mathbf{C}_q^{(n),k}$ associated with the fourth order cumulants of tensors \mathcal{R} and $\mathcal{B}^{(n),k}$ n th-mode vectors. Every element of $\mathbf{C}_q^{(n),k}$ is given in (3.9);

iii. Process matrix $\mathbf{C}_q^{(n),k}$ eigenvalue decomposition (EVD) and put the K_n eigenvectors associated with the K_n largest eigenvalues into $\mathbf{U}_{k+1}^{(n)}$;

iv. Compute projector $\mathbf{P}_{k+1}^{(n)} = \mathbf{U}_{k+1}^{(n)} \mathbf{U}_{k+1}^{(n)T}$;

3. **Output:** $\hat{\mathcal{X}} = \mathcal{R} \times_1 \mathbf{P}_{k_{stop}}^{(1)} \times_2 \dots \times_N \mathbf{P}_{k_{stop}}^{(N)}$, with k_{stop} being the index of the last iteration after convergence of the algorithm.

It was experimentally shown in [39] that when the parameter q involved in $\mathbf{C}_q^{(n),k}$ is chosen properly, multimode PCA filtering based on fourth order cumulants (denoted by $\text{rank-}\mathcal{C}(K_1, \dots, K_N)$) and on fourth order cumulant slice matrix (denoted by $\text{rank-}\mathbf{C}_1(K_1, \dots, K_N)$) give sensibly the same performances in regard to noise reduction in color images and multicomponent seismic waves.

3.3. Multiway Wiener filtering. Let \mathbf{R}_n , \mathbf{X}_n , and \mathbf{N}_n be the n th-mode flattening matrices of tensors \mathcal{R} , \mathcal{X} , and \mathcal{N} , respectively.

In the previous subsection, the estimation of signal tensor \mathcal{X} has been performed by projecting noisy data tensor \mathcal{R} on each n th-mode signal subspace. The n th-mode

projectors have been estimated thanks to the use of multimode PCA achieved by rank- (K_1, \dots, K_N) approximation. Despite the good results given by this method, it is possible to improve the tensor filtering quality by determining n th-mode filters $\mathbf{H}^{(n)}$, $n = 1$ to N , in (3.2), which optimize an estimation criterion. The most classical method is to minimize the mean squared error between the expected signal tensor \mathcal{X} and the estimated signal tensor $\hat{\mathcal{X}}$ given in (3.2):

$$(3.10) \quad e(\mathbf{H}^{(1)}, \dots, \mathbf{H}^{(N)}) = \mathbb{E}[\|\mathcal{X} - \mathcal{R} \times_1 \mathbf{H}^{(1)} \times_2 \dots \times_N \mathbf{H}^{(N)}\|^2].$$

Due to the criterion which is minimized, filters $\mathbf{H}^{(n)}$, $n = 1$ to N , can be called “ n th-mode Wiener filters” [38].

According to the calculations presented in Appendix A, especially from (A.1) to (A.15), the minimization of (3.10) with respect to filter $\mathbf{H}^{(n)}$, for fixed $\mathbf{H}^{(m)}$, $m \neq n$, leads to the following expression of n th-mode Wiener filter:

$$(3.11) \quad \mathbf{H}^{(n)} = \gamma_{\mathbf{X}\mathbf{R}}^{(n)} \mathbf{\Gamma}_{\mathbf{R}\mathbf{R}}^{(n)-1},$$

where

$$(3.12) \quad \gamma_{\mathbf{X}\mathbf{R}}^{(n)} = \mathbb{E}[\mathbf{X}_n \mathbf{T}^{(n)} \mathbf{R}_n^T]$$

is the $\mathbf{T}^{(n)}$ -weighted covariance matrix between the random column vectors of signal \mathbf{X}_n and data \mathbf{R}_n , with

$$(3.13) \quad \mathbf{T}^{(n)} = \mathbf{H}^{(1)} \otimes \dots \otimes \mathbf{H}^{(n-1)} \otimes \mathbf{H}^{(n+1)} \otimes \dots \otimes \mathbf{H}^{(N)},$$

where \otimes stands for Kronecker product, and

$$(3.14) \quad \mathbf{\Gamma}_{\mathbf{R}\mathbf{R}}^{(n)} = \mathbb{E}[\mathbf{R}_n \mathbf{Q}^{(n)} \mathbf{R}_n^T]$$

is the $\mathbf{Q}^{(n)}$ -weighted covariance matrix of the data \mathbf{R}_n , with

$$(3.15) \quad \mathbf{Q}^{(n)} = \mathbf{T}^{(n)T} \mathbf{T}^{(n)}.$$

In order to obtain $\mathbf{H}^{(n)}$ through (3.11), we suppose that the filters $\{\mathbf{H}^{(m)}, m = 1 \text{ to } N, m \neq n\}$ are known. Data tensor \mathcal{R} is available, but signal tensor \mathcal{X} is unknown. So, only the term $\mathbf{\Gamma}_{\mathbf{R}\mathbf{R}}^{(n)}$ can be derived, and not the term $\gamma_{\mathbf{X}\mathbf{R}}^{(n)}$. Hence, some more assumptions on \mathcal{X} have to be made in order to overcome the indetermination over $\gamma_{\mathbf{X}\mathbf{R}}^{(n)}$ [35, 38]. In the one-dimensional case, a classical assumption is to consider that a signal vector is a weighted combination of the signal subspace basis vectors. In extension to the tensor case, [35, 38] have proposed considering that the n th-mode flattening matrix \mathbf{X}_n can be expressed as a weighted combination of K_n vectors from the n th-mode signal subspace $E_1^{(n)}$:

$$(3.16) \quad \mathbf{X}_n = \mathbf{V}_s^{(n)} \mathbf{O}^{(n)}$$

with $\mathbf{X}_n \in \mathbb{R}^{I_n \times M_n}$, and $\mathbf{V}_s^{(n)} \in \mathbb{R}^{I_n \times K_n}$ being the matrix containing the K_n orthonormal basis vectors of n th-mode signal subspace $E_1^{(n)}$. Matrix $\mathbf{O}^{(n)} \in \mathbb{R}^{K_n \times M_n}$ is a weight matrix and contains the whole information on expected signal tensor \mathcal{X} . This model implies that signal n th-mode flattening matrix \mathbf{X}_n is orthogonal to n th-mode

noise flattening matrix \mathbf{N}_n , since signal subspace $E_1^{(n)}$ and noise subspace $E_2^{(n)}$ are supposed mutually orthogonal.

Supposing that noise \mathcal{N} in (3.1) is white, Gaussian, and independent from signal \mathcal{X} , and introducing the signal model (3.16) in (3.11) leads to a computable expression of n th-mode Wiener filter $\mathbf{H}^{(n)}$ (see Appendix B),

$$(3.17) \quad \mathbf{H}^{(n)} = \mathbf{V}_s^{(n)} \gamma_{\mathbf{OO}}^{(n)} \mathbf{\Lambda}_{\mathbf{\Gamma s}}^{(n)-1} \mathbf{V}_s^{(n)T},$$

where $\gamma_{\mathbf{OO}}^{(n)} \mathbf{\Lambda}_{\mathbf{\Gamma s}}^{(n)-1}$ is a diagonal weight matrix given by

$$(3.18) \quad \gamma_{\mathbf{OO}}^{(n)} \mathbf{\Lambda}_{\mathbf{\Gamma s}}^{(n)-1} = \text{diag} \left[\frac{\beta_1}{\lambda_1^\Gamma}, \dots, \frac{\beta_{K_n}}{\lambda_{K_n}^\Gamma} \right],$$

where $\lambda_1^\Gamma, \dots, \lambda_{K_n}^\Gamma$ are the K_n largest eigenvalues of $\mathbf{Q}^{(n)}$ -weighted covariance matrix $\mathbf{\Gamma}_{\mathbf{RR}}^{(n)}$ (see (3.14)). Parameters $\beta_1, \dots, \beta_{K_n}$ depend on $\lambda_1^\gamma, \dots, \lambda_{K_n}^\gamma$, which are the K_n largest eigenvalues of $\mathbf{T}^{(n)}$ -weighted covariance matrix

$\gamma_{\mathbf{RR}}^{(n)} = \mathbb{E}[\mathbf{R}_n \mathbf{T}^{(n)} \mathbf{R}_n^T]$, according to the following relation:

$$(3.19) \quad \beta_{k_n} = \lambda_{k_n}^\gamma - \sigma_\Gamma^{(n)2} \quad \forall k_n = 1, \dots, K_n.$$

Superscript γ refers to the $\mathbf{T}^{(n)}$ -weighted covariance and subscript $\mathbf{\Gamma}$ to the $\mathbf{Q}^{(n)}$ -weighted covariance. $\sigma_\Gamma^{(n)2}$ is the degenerated eigenvalue of noise $\mathbf{T}^{(n)}$ -weighted covariance matrix $\gamma_{\mathbf{NN}}^{(n)} = \mathbb{E}[\mathbf{N}_n \mathbf{T}^{(n)} \mathbf{N}_n^T]$. Thanks to the additive noise and the signal independence assumptions, the $I_n - K_n$ smallest eigenvalues of $\gamma_{\mathbf{RR}}^{(n)}$ are equal to $\sigma_\Gamma^{(n)2}$ and thus can be estimated by the following relation:

$$(3.20) \quad \hat{\sigma}_\Gamma^{(n)2} = \frac{1}{I_n - K_n} \sum_{k_n=K_n+1}^{I_n} \lambda_{k_n}^\gamma.$$

In order to determine the n th-mode Wiener filters $\mathbf{H}^{(n)}$ that minimize the mean squared error (3.10), the alternating least squares (ALS) algorithm has been proposed in [35, 38]. It can be summarized in the following steps:

1. **Initialization** $k = 0$: $\mathcal{R}^0 = \mathcal{R} \Leftrightarrow \mathbf{H}_0^{(n)} = \mathbf{I}_{I_n}$, Identity matrix, $\forall n = 1 \dots N$.
2. **ALS loop**:

Repeat until convergence, that is, $\|\mathcal{R}^{k+1} - \mathcal{R}^k\|^2 < \epsilon$, with $\epsilon > 0$ prior fixed threshold,

(a) for $n = 1$ to N :

- i. Form $\mathcal{R}^{(n),k}$: $\mathcal{R}^{(n),k} = \mathcal{R} \times_1 \mathbf{H}_{k+1}^{(1)} \times_2 \dots \times_{n-1} \mathbf{H}_{k+1}^{(n-1)} \times_{n+1} \mathbf{H}_k^{(n+1)} \times_{n+2} \dots \times_N \mathbf{H}_k^{(N)}$;
- ii. Determine $\mathbf{H}_{k+1}^{(n)} = \arg \min_{\mathbf{Z}^{(n)}} \|\mathcal{X} - \mathcal{R}^{(n),k} \times_n \mathbf{Z}^{(n)}\|^2$

subject to $\mathbf{Z}^{(n)} \in \mathbb{R}^{I_n \times I_n}$ thanks to the following procedure:

- A. n th-mode flatten $\mathcal{R}^{(n),k}$ into $\mathbf{R}_n^{(n),k} = \mathbf{R}_n(\mathbf{H}_{k+1}^{(1)} \otimes \dots \otimes \mathbf{H}_{k+1}^{(n-1)} \otimes \mathbf{H}_k^{(n+1)} \otimes \dots \otimes \mathbf{H}_k^{(N)})^T$, and \mathcal{R} into \mathbf{R}_n ;
- B. Compute $\gamma_{\mathbf{RR}}^{(n)} = \mathbb{E}[\mathbf{R}_n \mathbf{R}_n^{(n),kT}]$,
- C. Determine $\lambda_1^\gamma, \dots, \lambda_{K_n}^\gamma$, the K_n largest eigenvalues of $\gamma_{\mathbf{RR}}^{(n)}$;

- D. For $k_n = 1$ to I_n , estimate $\sigma_\Gamma^{(n)2}$ thanks to (3.20) and for $k_n = 1$ to K_n , estimate β_{k_n} thanks to (3.19);
- E. Compute $\mathbf{\Gamma}_{\mathbf{RR}}^{(n)} = \mathbb{E}[\mathbf{R}_n^{(n),k} \mathbf{R}_n^{(n),kT}]$;
- F. Determine $\lambda_1^\Gamma, \dots, \lambda_{K_n}^\Gamma$, the K_n largest eigenvalues of $\mathbf{\Gamma}_{\mathbf{RR}}^{(n)}$;
- G. Determine $\mathbf{V}_s^{(n)}$, the matrix of the K_n eigenvectors associated with the K_n largest eigenvalues of $\mathbf{\Gamma}_{\mathbf{RR}}^{(n)}$;
- H. Compute the weight matrix $\gamma_{\mathbf{OO}}^{(n)} \mathbf{\Lambda}_{\mathbf{Ts}}^{(n)-1}$ given in (3.18);
- I. Compute $\mathbf{H}_{k+1}^{(n)}$, the n th-mode Wiener filter at the $(k + 1)$ th iteration, using (3.17);
- (b) Form $\mathcal{R}^{k+1} = \mathcal{R} \times_1 \mathbf{H}_{k+1}^{(1)} \times_2 \dots \times_N \mathbf{H}_{k+1}^{(N)}$;
- (c) Increment k ;
- 3. **output:** $\hat{\mathcal{X}} = \mathcal{R} \times_1 \mathbf{H}_{k_{stop}}^{(1)} \times_2 \dots \times_N \mathbf{H}_{k_{stop}}^{(N)}$, with k_{stop} being the last iteration after convergence of the algorithm.

In subsection 3.2, we presented the adaptation of multimode PCA to the case of a noncorrelated Gaussian noise, by using higher order statistics. In the same way, it is possible to use higher order statistics for multiway Wiener filtering. For this, one should replace step 2(a)iiB by step 2(a)ii of the ALS loop in subsection 3.2, and replace step 2(a)iiE by the computation of the cumulant slice $\mathbf{C}_q^{(n),k}$ associated with the fourth order cumulants of matrix $\mathbf{R}_n^{(n),k}$ and matrix $(\mathbf{R}_n^{(n),k})^T$. Elements of $\mathbf{C}_q^{(n),k}$ are given in (3.9).

4. Simulation results. In the following simulations, the channel-by-channel SVD-based filtering defined in subsection 3.1 and the rank- (K_1, \dots, K_N) approximation-based multiway and multidimensional filtering are applied to the denoising of color images and multispectral images and to the denoising of seismic signals. Color images, multispectral images, and seismic signals can be represented by a third order tensor from $\mathbb{R}^{I_1 \times I_2 \times I_3}$, where I_1 , I_2 , and I_3 take different values. In all these applications, the efficiency of denoising is tested in the presence of an additive Gaussian noise, either correlated or not.

A multidimensional and multiway white Gaussian noise \mathcal{N} which is added to signal tensor \mathcal{X} can be expressed as

$$(4.1) \quad \mathcal{N} = \alpha \cdot \mathcal{G},$$

where every element of $\mathcal{G} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ is an independent realization of a normalized centered Gaussian law and where α is a coefficient that permits to set the signal-to-noise ratio (SNR) in noisy data tensor \mathcal{R} .

When we process images impaired by correlated Gaussian noise, the noise which is added is a third order tensor defined by

$$(4.2) \quad \mathcal{N}^c = \mathcal{N} \times_1 \mathbf{W}^{(1)} \times_2 \mathbf{W}^{(2)} \times_3 \mathbf{W}^{(3)},$$

where every element of \mathcal{N} represents an independent realization of a white Gaussian noise and $\mathbf{W}^{(n)}$ is a weight matrix in the n th-mode, $n = 1, 2, 3$.

In order to evaluate the performances of the overviewed tensor signal processing methods, a particular performance criterion is employed as proposed in [38, 39].

4.1. Performance criterion. Following the representation of (3.1), the multiway noisy data tensor is expressed as $\mathcal{R} = \mathcal{X} + \mathcal{N}$, where \mathcal{X} is the expected signal

tensor and \mathcal{N} is the additive noise tensor. Let us define the SNR, in dB, in the noisy data tensor by

$$(4.3) \quad \text{SNR} = 10 \log \left(\frac{\|\mathcal{X}\|^2}{\|\mathcal{N}\|^2} \right).$$

In order to a posteriori verify the quality of the estimated signal tensor, we use the normalized quadratic error criterion (NQE) defined as follows:

$$(4.4) \quad \text{NQE}(\hat{\mathcal{X}}) = \frac{\|\hat{\mathcal{X}} - \mathcal{X}\|^2}{\|\mathcal{X}\|^2}.$$

The NQE criterion permits a quantitative comparison of the channel-by-channel SVD-based filtering and the rank- (K_1, K_2, K_3) approximation multiway and multidimensional filtering. Considering this criterion, we expect the rank- (K_1, K_2, K_3) approximation to give better results than the channel-by-channel SVD-based filtering method.

4.2. Denoising of color images. Denoising of color images has been studied in several works [6, 40, 43]. Some solutions have been brought from the field of wavelet processing, exhibiting good results in terms of output SNR. These studies concern only bidimensional data, whereas the methods that we compare are adapted to the processing of third order tensors as a whole, and in particular to three-channel images. We focus on subspace-based methods. We first consider the channel-by-channel SVD-based filtering, the rank- (K_1, K_2, K_3) approximation and multiway Wiener filtering (Wmm- (K_1, K_2, K_3)), applied to images impaired by an additive white Gaussian noise.

Then we present the results obtained with rank- (K_1, K_2, K_3) based on second order and higher order statistics, applied to images impaired by an additive correlated Gaussian noise. We compare the performances of the methods applied in this subsection in terms of denoising efficiency and computational load.

4.2.1. Denoising of a color image impaired by additive Gaussian noise.

Let us consider the “sailboat” standard color image of Figure 4.1(a) represented as a third order tensor $\mathcal{X} \in \mathbb{R}^{256 \times 256 \times 3}$. The ranks of the signal subspace for each mode are 30 for the 1st-mode, 30 for the 2nd-mode, and 2 for the 3rd-mode. This is fixed thanks to the following process. For Figure 4.1(a), we took the standard nonnoisy sailboat image and artificially reduced the ranks of the nonnoisy image, that is, we set the parameters (K_1, K_2, K_3) to $(30, 30, 2)$, thanks to the truncation of HOSVD. This ensures that, for each mode, the rank of the signal subspace is lower than the corresponding dimension. This also permits us to evaluate the performances of the filtering methods applied, independently from the accuracy of the estimation of the values of the ranks by MDL or AIC criterion.

Figure 4.1(b) shows the noisy image resulting from the impairment of Figure 4.1(a) and represented as $\mathcal{R} = \mathcal{X} + \mathcal{N}$. Third-order noise tensor \mathcal{N} is defined by relation (4.1) by choosing α such that, considering previous definition of (4.3), the SNR in the noisy image of Figure 4.1(b) is 8.1 dB. In these simulations, the value of the parameter K of channel-by-channel SVD-based filtering, the values of the dimensions of the row and column signal subspace are supposed to be known and fixed to 30. In the same way, parameters (K_1, K_2, K_3) of rank- (K_1, K_2, K_3) approximation are fixed to $(30, 30, 2)$.

The channel-by-channel SVD-based filtering of noisy image \mathcal{R} (see Figure 4.1(b)) yields the image of Figure 4.1(c), and rank- $(30, 30, 2)$ approximation of noisy data ten-

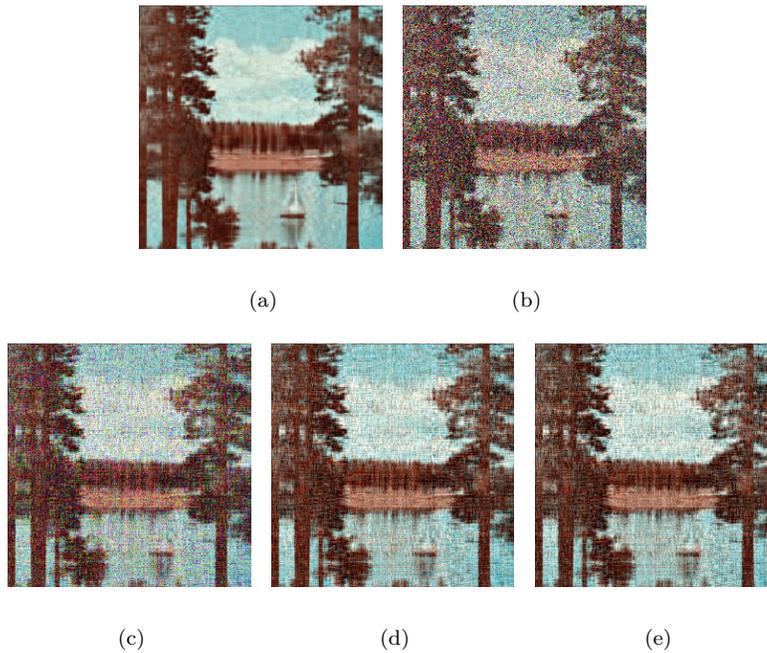


FIG. 4.1. (a) *Nonnoisy image.* (b) *Image to be processed, impaired by an additive white Gaussian noise, with SNR = 8.1 dB.* (c) *Channel-by-channel SVD-based filtering of parameter $K = 30$.* (d) *rank-(30, 30, 2) approximation.* (e) *Wmm-(30, 30, 2) filtering.*

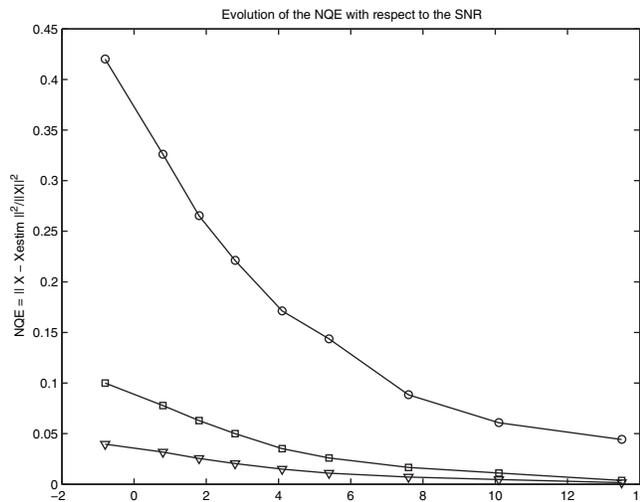


FIG. 4.2. *NQE evolution with respect to SNR (dB): channel-by-channel SVD-based filtering of parameter 30 (-o-), rank-(30, 30, 2) approximation (-□-), Wmm-(30, 30, 2) filtering (-▽-).*

sor \mathcal{R} yields the image of Figure 4.1(d). The NQE, defined in (4.4), permits a qualitative comparison between channel-by-channel SVD-based filtering and rank-(30, 30, 2) approximation. Figure 4.2, which presents the evolution of the NQE with respect to SNR varying from 3 dB to 18 dB, shows the NQE obtained with Wmm-(30, 30, 2) is lower than the NQE obtained with the filtering with rank-(30, 30, 2) approximation.

For this simulation, the rank- (K_1, K_2, K_3) approximation gives better results than channel-by-channel SVD-based filtering according to the NQE criterion. From the resulting image, presented on Figure 4.1(d), we note that dimension reduction leads to a loss of spatial resolution. However, the choice of a set of values K_1, K_2, K_3 which are small enough is the condition for an efficient noise reduction effect.

Therefore, a trade-off should be considered between noise reduction and detail preservation. This trade-off was discussed in [42]. We were interested in using the minimum description length (MDL) criterion [53], applied to the left singular values of the flattening matrices computed over the successive n th-modes. As a rule of thumb, the MDL criterion overestimates the value of parameters K_1, K_2 , and K_3 . This results in the preservation of the details in the processed image, at the expense of an efficient denoising.

Concerning the qualitative results obtained with this color image, we note that the intraclass variance of the pixel values of each component (or color mode) of the resulting image is lower for the image obtained with Wmm-(30, 30, 2) than for those images obtained with other methods applied in this subsection. This allows, for example, applying after denoising a high level classification method with a higher efficiency than when classification is applied after channel-by-channel SVD-based filtering or HOSVD-(30, 30, 2).

For the $256 \times 256 \times 3$ sailboat image of Figure 4.1, the computational times needed when Matlab programs are used on a 3 Ghz Pentium 4 processor running Windows are as follows. HOSVD-(30, 30, 2) lasts 1.61 seconds, the channel-by-channel SVD-based filtering lasts 1.94 seconds, the rank-(30, 30, 2) approximation run with 25 iterations lasts 54.1 seconds, and Wmm-(30, 30, 2) run with 25 iterations lasts 40.0 seconds.

The results presented in Figure 4.1 show that Wmm- (K_1, K_2, K_3) allows one to obtain better results in terms of NQE with a computational load which is lower than that of the rank- (K_1, K_2, K_3) approximation. In the next two examples we study the influence of the values of the n th-mode ranks. In the example of Figure 4.3 we set, in the same way as in the previous example, the ranks of the truncated image to $(30, 30, 3)$ (see Figure 4.3(a)). Note that $K_3 = I_3 = 3$. Thus the assumption $K_3 < I_3$ is not fulfilled. We aim at studying the behavior of the proposed tensor filtering algorithms when the color mode rank is equal to the color mode dimension ($K_3 = I_3$). The truncated image is impaired by a noncorrelated Gaussian noise such that SNR = 8.1 dB (see Figure 4.3(b)). The results obtained show that channel-by-channel Wiener-based filtering of parameter $K = 30$ (see Figure 4.3(c)) is outperformed by rank-(30, 30, 3) approximation (see Figure 4.3(d)) and Wmm-(30, 30, 3) (see Figure 4.3(e)). Indeed, the proposed tensor filtering algorithms rely on an ALS loop which permits us to take into account the relationships between the filters of each mode when multiway filters are used. In particular, concerning multiway Wiener filtering, it can be adapted to the case where it is applied with $K_3 = I_3$. For this, the weight matrix $\gamma_{\mathbf{O}\mathbf{O}\mathbf{r}_s}^{(3)} \mathbf{\Lambda}_{\mathbf{r}_s}^{(3)-1}$ of step 2a) of the multiway Wiener filtering algorithm presented in subsection 3.3 is set to identity. That is, $\mathbf{H}^{(3)}$ is replaced by $\mathbf{P}^{(3)}$. We adapted the algorithm in order to take into account the channel mode information for the computation of the two spatial filters thanks to the ALS loop.

This proves the interest of multiway filtering even in the case where the rank of the signal subspace along the third mode is equal to the number of channels.

In the example of Figure 4.4 we study the case where the ranks of the signal subspaces are underestimated for the spatial modes. Let us consider the ‘‘Mondriaan’’ standard color image of Figure 4.4 represented as a third order tensor $\mathcal{X} \in \mathbb{R}^{256 \times 256 \times 3}$. We set the ranks of the truncated image to $(150, 150, 3)$. The ranks along the spa-

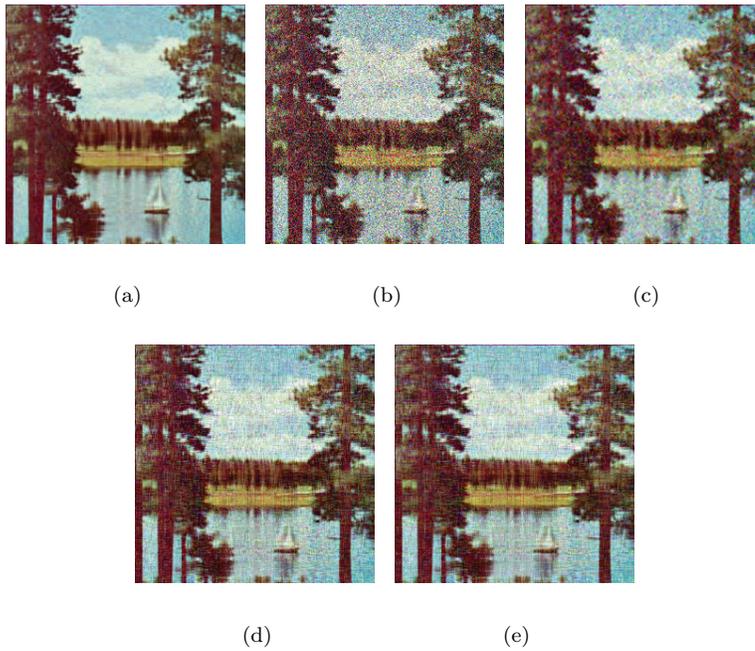


FIG. 4.3. (a) *Nonnoisy image.* (b) *Image to be processed, impaired by an additive white Gaussian noise, with SNR = 8.1 dB.* (c) *Channel-by-channel Wiener-based filtering of parameter $K = 30$.* (d) *rank-(30, 30, 3) approximation.* (e) *Wmm-(30, 30, 3) filtering.*

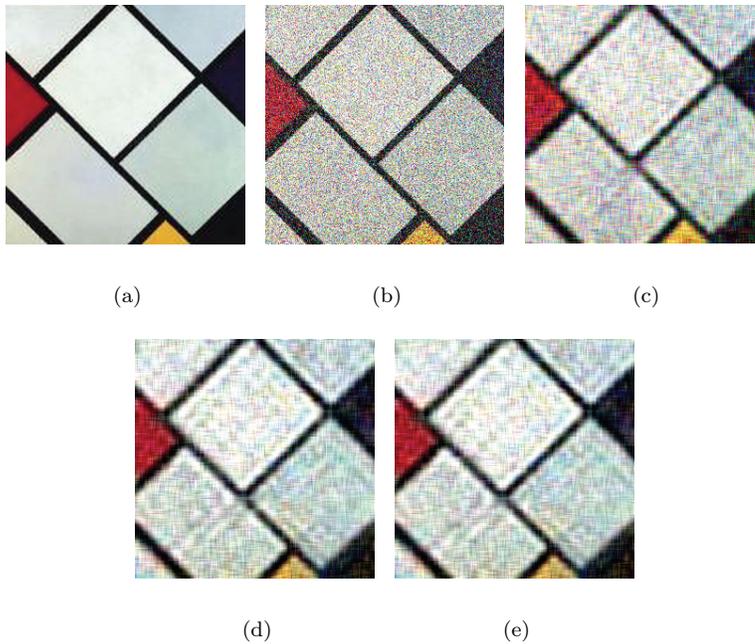


FIG. 4.4. (a) *Nonnoisy image.* (b) *Image to be processed, impaired by an additive white Gaussian noise, with SNR = 8.0 dB.* (c) *Channel-by-channel SVD-based filtering of parameter $K = 19$.* (d) *rank-(19, 19, 3) approximation.* (e) *Wmm-(19, 19, 3) filtering.*

tial modes will be fixed intentionally to a value which is smaller than 150 when the reviewed methods are applied. Figure 4.4(a) gives the nonnoisy image, Figure 4.4(b) shows the noisy image resulting from the impairment, with SNR= 8.0 dB, of the image of Figure 4.4(a). Figure 4.4(c) gives the result obtained with channel-by-channel SVD-based filtering of parameter $K = 19$. Figure 4.4(d) gives the result obtained with rank-(19, 19, 3) approximation, and Figure 4.4(e) gives the result obtained with Wmm-(19, 19, 3) filtering. Note that choosing $(K_1, K_2, K_3)=(19, 19, 3)$ results in throwing out both signal and noise information along the spatial modes, as the ranks of the noisy image are (150, 150, 3). Underestimating the ranks along the spatial modes induces some blurry effect in the result images: part of the spatial resolution is lost. The presented subspace-based algorithms perform well if there is a high level of redundancy in the column or row space or if the image exhibits many soft or blurry edges, and the n th-mode ranks are not underestimated.

4.2.2. HOSVD- (K_1, K_2, K_3) , rank- (K_1, K_2, K_3) approximation based on second order and higher order statistics, applied to an image impaired by an additive correlated Gaussian noise. The purpose here is to compare methods based on second order statistics with methods based on higher order statistics when an image is impaired by a correlated Gaussian noise. Figure 4.5 shows the results obtained with the HOSVD- (K_1, K_2, K_3) , and the rank- (K_1, K_2, K_3) approximation based on second order and higher order statistics, used for the denoising of an image impaired by an additive correlated Gaussian noise. We consider the nonnoisy image of Figure 4.5(a) whose ranks are fixed to (30, 30, 2): we artificially reduced the ranks of the nonnoisy image, that is, we set the parameters (K_1, K_2, K_3) to (30,30,2), thanks to the truncation of HOSVD. This image is impaired by a correlated Gaussian noise (see (4.2)). Figure 4.5(b) shows the noisy image. The result of HOSVD- (K_1, K_2, K_3) is given in Figure 4.5(c), and the result of rank- (K_1, K_2, K_3) approximation based on second order statistics is given in Figure 4.5(d), the result of

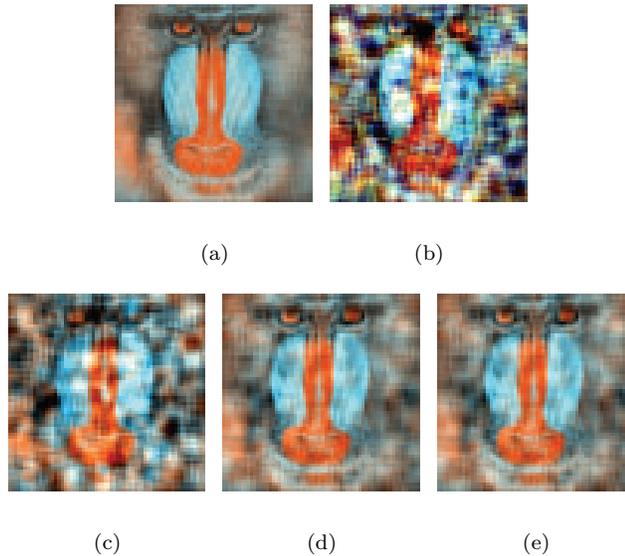


FIG. 4.5. (a) *Initial nonnoisy image.* (b) *Initial image with an additive correlated Gaussian noise, SNR = 2.48 dB.* (c) *HOSVD-(30, 30, 2).* (d) *rank-C(30, 30, 2) approximation.* (e) *rank- $C_1(30, 30, 2)$ approximation.*

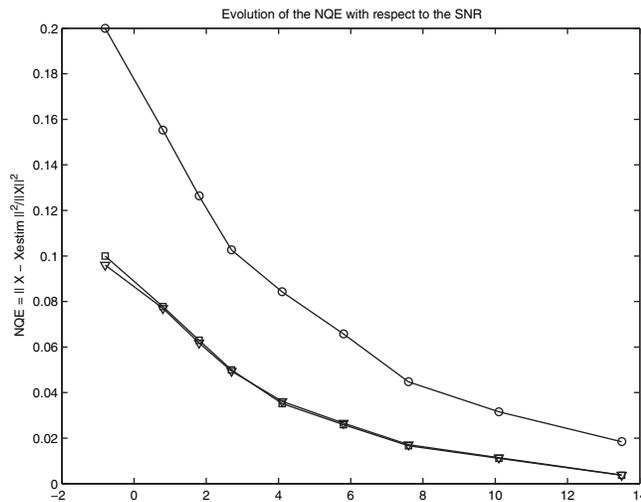


FIG. 4.6. Evolution of the NQE with respect to the SNR(dB) for each tensor filtering method: ○: HOSVD-(30, 30, 2); ▽: rank-C(30, 30, 2); □: rank-C₁(30, 30, 2).

rank- $\mathcal{C}(K_1, K_2, K_3)$ approximation based on higher order statistics is given in Figure 4.5(e). The evolution of the NQE with respect to the SNR for HOSVD-(30, 30, 2), rank- $\mathcal{C}(30, 30, 2)$ approximation based on fourth order cumulants, rank- $\mathbf{C}_1(30, 30, 2)$ approximation based on one slice of the fourth order cumulants is represented in Figure 4.6.

The main conclusion from Figure 4.5 is that the methods based on fourth order cumulants give similar visual results and better results than HOSVD-(30, 30, 2). Whatever the SNR, the methods based on fourth order cumulants give a lower NQE value than the methods based on second order statistics. The method based on fourth order cumulant slice matrix gives sensibly the same NQE values as the method based on fourth order cumulants.

For the $256 \times 256 \times 3$ baboon image of Figure 4.5, the computational times needed in the same conditions of processor and software as in previous subsection are the following: HOSVD-(30, 30, 2) lasts 1.61 seconds, rank- $\mathcal{C}(30, 30, 2)$ based filtering lasts 2h. 11 min. 40 seconds, rank- $\mathbf{C}_1(30, 30, 2)$ lasts 3 min. 50 seconds.

4.3. Denoising of multispectral images. The results obtained from the processing of a multispectral image composed of 72 rows, 160 columns and 100 spectral channels representing a truck are considered. This set of spectral images can be represented as a tensor $\mathcal{X} \in \mathbb{R}^{72 \times 160 \times 100}$. Images shown on Figures 4.7(a) to 4.7(e) represent channels 30 to 34 of the multispectral image. To evaluate the performances of the reviewed methods, some signal-independent white Gaussian noise \mathcal{N} is added to \mathcal{X} and results in noisy tensor $\mathcal{R} = \mathcal{X} + \mathcal{N}$. Channels 30 to 34 of noisy multispectral image represented as \mathcal{R} are shown in Figures 4.7(f) to 4.7(j), and correspond to a noise impairment level SNR = -1 dB. Figures 4.7(k) to 4.7(o) represent channels 30 to 34 of the multispectral image obtained by applying channel-by-channel-based SVD-filtering to noisy image \mathcal{R} . Finally, Figures 4.7(p) to 4.7(t) represent channels 30 to 34 of the multispectral image obtained after applying rank-(30,30,30) approximation to noisy image \mathcal{R} . This last simulation clearly shows that the rank-(30,30,30) approximation-based filtering gives better results than channel-by-channel SVD-based

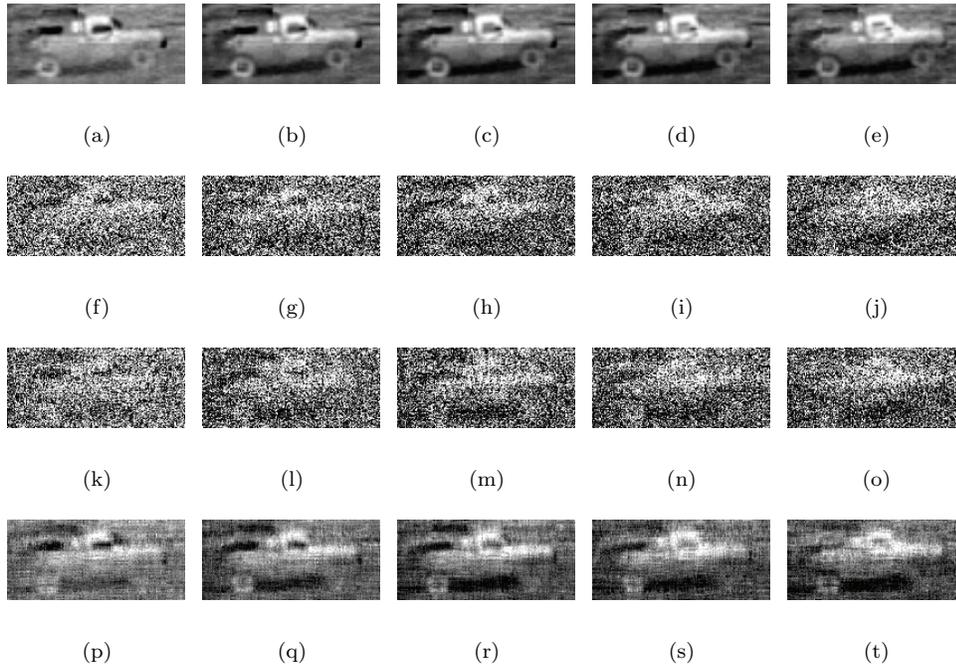


FIG. 4.7. Channels 30 to 34 of the processed multispectral images are presented: (a)–(e) Nonnoisy multispectral image. (f)–(j) Impaired multispectral image. (k)–(o) Results obtained with channel-by-channel SVD filtering. (p)–(t) Results obtained with rank-(30, 30, 30) approximation.

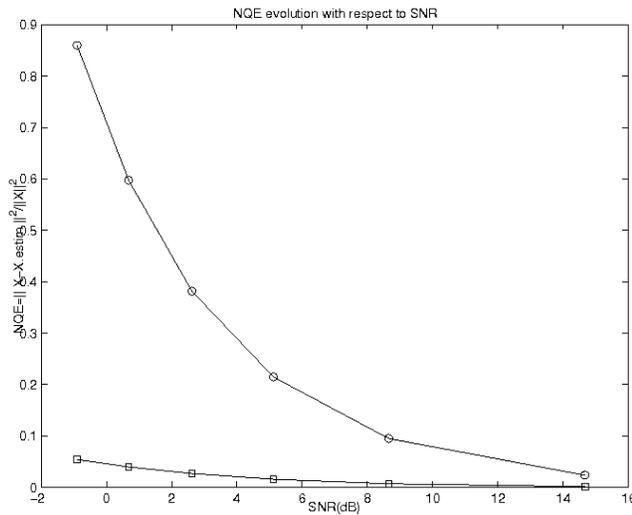


FIG. 4.8. NQE evolution with respect to SNR (from -1 to 15 dB): channel-by-channel SVD-based filtering of parameter 30 (-o-), and rank-(30, 30, 30) approximation (-□-).

filtering in regard to denoising. Moreover, the evolution of the NQE with respect to the SNR varying from -1 dB to 15 dB, represented in Figure 4.8, shows that the NQE obtained with $Wmm-(K_1, K_2, K_3)$ is lower than the NQE obtained with a previously existing method.

For this simulation the estimation quality, respect to the NQE criterion, is better for rank- (K_1, K_2, K_3) approximation, compared to channel-by-channel SVD-based filtering. Superiority of rank- (K_1, K_2, K_3) approximation compared to channel-by-channel SVD-based filtering is confirmed.

According to the simulations performed on a color image and on a multispectral image, it is possible to conclude that the more channels the image is composed of, the better the denoising. This can be explained by a better estimation of projectors on 1st-mode and 2nd-mode signal subspaces in the case of the multispectral image. Indeed, the number of spectral channels in a multispectral image is much larger than in a color image. Equivalently, I_3 is much larger than 3, so M_1 and M_2 are much larger than for a color image, and the estimation of matrices $\mathbf{C}^{(1),k}$ and $\mathbf{C}^{(2),k}$ presented in (3.4) are computed with more realization vectors.

4.4. Statistical performances. The goal of the following simulation is to test the robustness to noise of channel-by-channel SVD-based filtering of parameter K and of rank- (K_1, K_2, K_3) approximation, with respect to the NQE criterion. We process the Sailboat standard color image, impaired by an additive Gaussian noise, with SNR values varying from -0.7 dB to 15 dB; 100 trials are performed. For each trial one realization of additive Gaussian noise is simulated and added to the nonnoisy image. The mean and standard deviation are computed over the NQE values obtained each time the channel-by-channel SVD-based filtering and the rank- (K_1, K_2, K_3) approximation are run. The evolution of the mean NQE

$$(4.5) \quad m_{\text{NQE}} = \frac{1}{100} \sum_{i=1}^{100} \text{NQE}_i,$$

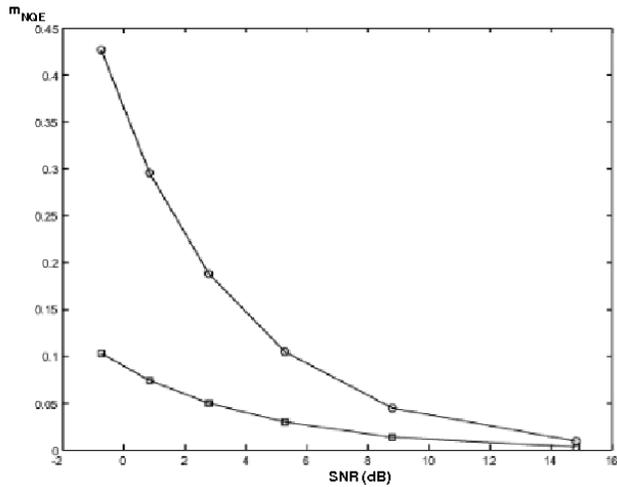
where index i refers to the i th noise realization, is represented in Figure 4.9(a) with respect to SNR. The evolution of the standard deviation of the NQE,

$$(4.6) \quad \text{std}_{\text{NQE}} = \sqrt{\frac{1}{100} \sum_{i=1}^{100} (\text{NQE}_i - m_{\text{NQE}})^2},$$

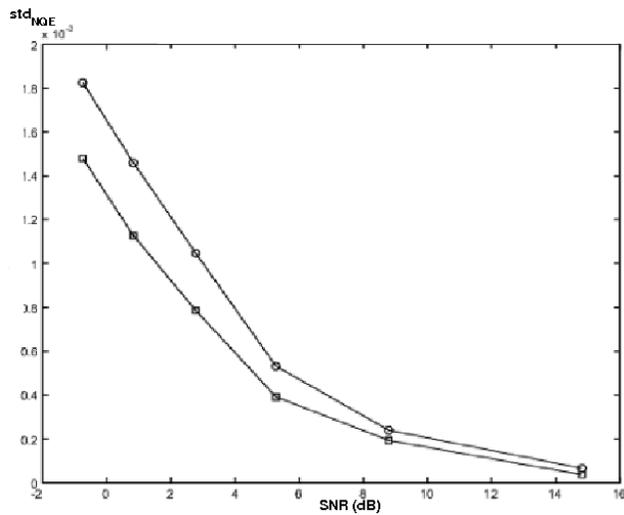
is represented in Figure 4.9(b), with respect to the SNR. Figure 4.9 shows that the mean and standard deviation values of the NQE obtained with rank- (K_1, K_2, K_3) approximation and computed over 100 noise realizations are both lower than the mean and the standard deviation values obtained with channel-by-channel SVD-based filtering. Thus, for these simulations, the rank- (K_1, K_2, K_3) approximation gives better results than channel-by-channel SVD-based filtering in regard to the robustness of tensor estimation and considering the NQE criterion.

4.5. Filtering of a multicomponent seismic type signal.

4.5.1. Filtering of a multicomponent seismic type signal impaired by an additive white Gaussian noise. In this simulation, a multicomponent seismic wave is received on a linear antenna composed of 10 sensors. The direction of propagation of the wave is assumed to be contained in a plane which is orthogonal to the antenna. The three components of the wave, represented as signal tensor \mathcal{X} , are called Component 1, Component 2, and Component 3 and are represented in Figures 4.10(a)–(c). In each seismic slice, the x-axis corresponds to the time sampling (200 or 100 time samples) and the y-axis corresponds to the spatial sensors (10 sensors). Each consecutive component presents a $\frac{\pi}{2}$ radian phase shift. The three components of noisy data tensor



(a)



(b)

FIG. 4.9. (○): results obtained with channel-by-channel SVD-based filtering of parameter 30; (□): results obtained with rank-(30,30,2) approximation. (a) Evolution of the mean NQE with respect to SNR (dB). (b) Evolution of the standard deviation of NQE with respect to SNR (dB).

\mathcal{R} are represented in Figures 4.10(d)–(f), where the additive noise is considered as white and Gaussian and for which the SNR = -10 dB. The classical Wiener filtering of parameter K (W_{cc-K}) of each component, with a signal subspace dimension fixed to $K = 8$, permits us to obtain the results presented in Figures 4.10(g)–(i). The multimode PCA-based filtering achieved by applying HOSVD-(8,8,3) to noisy data tensor permits to obtain the results presented in Figures 4.10(j)–(l). Finally, the results obtained with multiway Wiener filtering applied to the noisy data tensor are presented in Figures 4.10(m)–(o). The evolution of the NQE with respect to the SNR

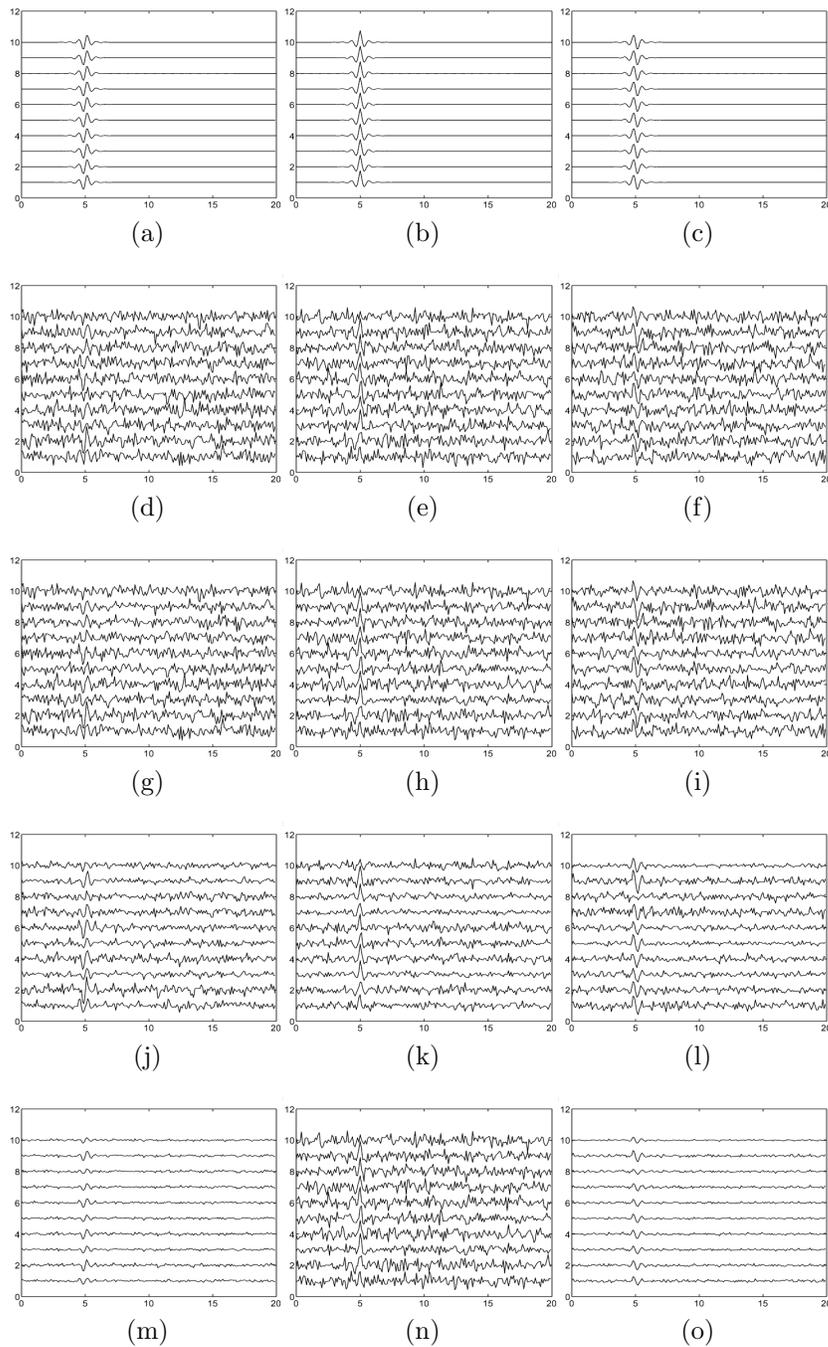
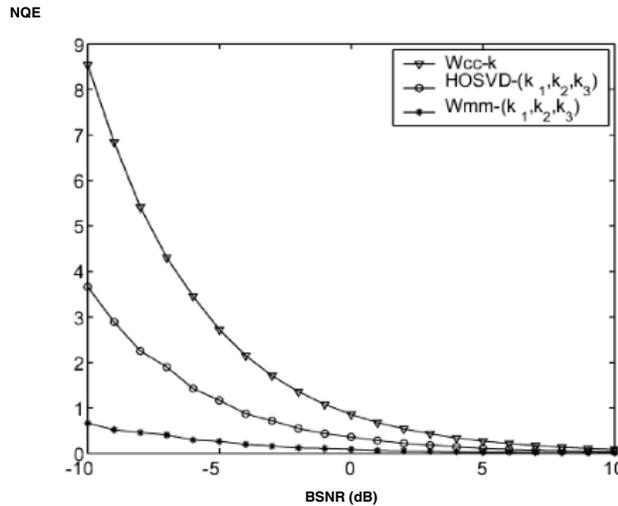
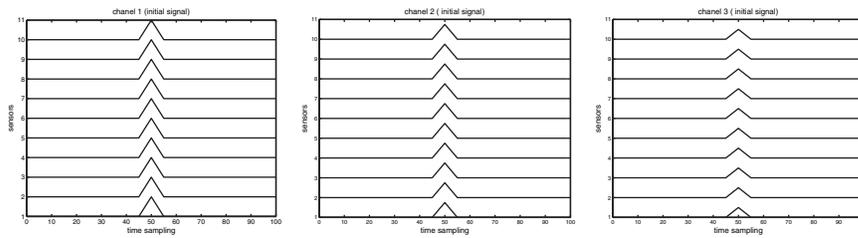


FIG. 4.10. *Nonnoisy, impaired, and processed seismic wave: the three polarization components.* (a)–(c) Components 1, 2, and 3 of the nonnoisy seismic wave. (d)–(f) Components 1, 2, and 3 of the seismic wave, impaired by an additive white Gaussian noise ($\text{SNR} = -10$ dB). (g)–(i) Wiener filtering applied component by component (W_{cc-K}), with rank $K = 8$. (j)–(l) $HOSVD-(K_1, K_2, K_3)$, with $(K_1, K_2, K_3) = (8, 8, 3)$. (m)–(o) multiway Wiener filtering ($W_{mm}-(K_1, K_2, K_3)$), with $(K_1, K_2, K_3) = (8, 8, 3)$.



(a)

FIG. 4.11. Evolution of the NQE with respect to the SNR (dB) for each tensor filtering method. (∇): Wiener filtering applied component by component ($Wcc-K$), with rank $K = 8$; (\circ): HOSVD- (K_1, K_2, K_3) , with $(K_1, K_2, K_3) = (8, 8, 3)$; (\bullet): multiway Wiener filtering ($Wmm-(K_1, K_2, K_3)$) with $(K_1, K_2, K_3) = (8, 8, 3)$.



(a)

(b)

(c)

FIG. 4.12. Multicomponent seismic signal. (a)–(c) Components 1 to 34 of the nonnoisy seismic wave.

(dB) is given in Figure 4.11. As well as in the case of color image filtering, in this simulation, the best quality, in terms of noise reduction, is given by multiway Wiener filtering since, for all considered SNR values, the NQE values given by this method are lower than the values given by both HOSVD- $(8, 8, 3)$ and $Wcc-8$.

4.5.2. Filtering of a multicomponent seismic type signal impaired by an additive correlated Gaussian noise. In this simulation, we consider a multicomponent seismic wave, impaired by a correlated Gaussian noise. The purpose here is to compare the performances of multiway filtering algorithms based on either second order moments or fourth order cumulants. Figures 4.12 and 4.13 show the efficiency, in terms of noise reduction, of rank- $\mathcal{C}(K_1, K_2, K_3)$ based filtering and rank- $\mathbf{C}_1(K_1, K_2, K_3)$ based filtering compared to rank- (K_1, K_2, K_3) approximation based on second order statistics, when seismic signals impaired by a correlated Gaussian noise are considered.

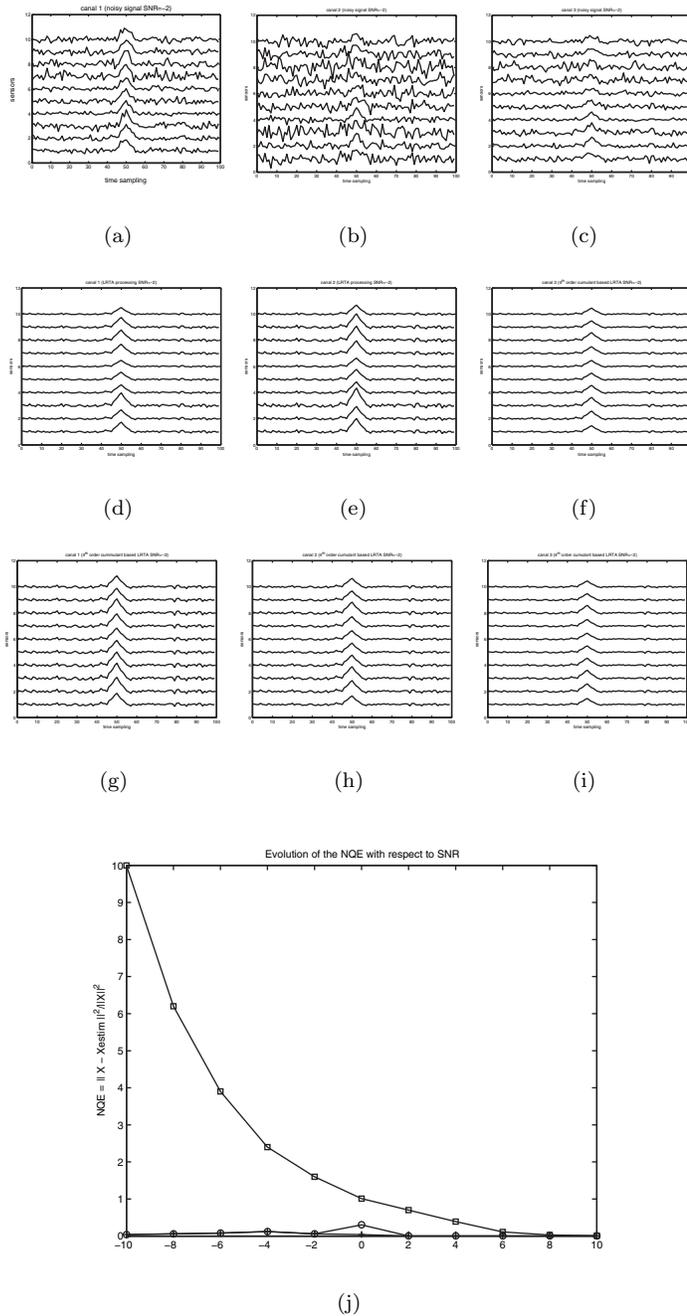


FIG. 4.13. Denoising of a multicomponent seismic wave impaired by an additive correlated Gaussian noise ($SNR = -2$ dB), using multiway filtering based on fourth order cumulants: comparison of rank- $C(8, 8, 3)$, and rank- $C_1(8, 8, 3)$. (a)–(c) Noised signal; components 1 to 3 impaired by a correlated Gaussian noise ($SNR = -2$ dB). (d)–(f) rank- $C(8, 8, 3)$ based filtering. (g)–(i) rank- $C_1(8, 8, 3)$ based filtering. (j) Evolution of NQE with respect to SNR (dB) for rank- $(8, 8, 3)$ approximation (\square), rank- $C_1(8, 8, 3)$ using fourth order cumulant slice matrix (\circ), and rank- $C(8, 8, 3)$ using fourth order cumulants ($+$).

5. Conclusion. In this paper, an overview on new mathematical methods dedicated to multicomponent data is presented. Multicomponent data are represented as tensors, that is, multiway arrays, and the tensor filtering methods that are presented rely on multilinear algebra. First we present how to perform channel-by-channel SVD-based filtering. Then we review three methods that take into account the relationships between each component of a processed tensor. The first method consists of an extension of the classical SVD-based filtering method. In the case of an additive white Gaussian noise, the signal tensor is estimated thanks to a multimode PCA achieved by applying a lower rank- (K_1, \dots, K_N) approximation to the noisy data tensor, or a lower rank- (K_1, \dots, K_N) truncation of its HOSVD. This method is implicitly based on second order statistics and relies on the orthogonality between n th-mode noise and signal subspaces. The second presented method consists of an improvement of the multimode PCA-based tensor filtering in the case of an additive correlated Gaussian noise. In this case, the covariance matrix involved in TUCKALS3 algorithm is replaced with the fourth order cumulant matrix of the related vectors. We reviewed a low computational load procedure involving the fourth order cumulant slice matrix instead of fourth order cumulants. This improved multimode PCA provides good performances compared to the multimode PCA method based on second order statistics, as was shown in the case of noise reduction in color images and multicomponent seismic waves.

Finally, the third reviewed method is a multiway version of the classical Wiener filtering. In extension to the one-dimensional case, the n th-mode Wiener filters are estimated by minimizing the mean squared error between the expected signal tensor and the estimated signal tensor obtained by applying the n th-mode Wiener filters to the noisy data tensor thanks to the n th-mode product operator. An alternating least squares algorithm has been presented to determine the optimal n th-mode Wiener filters. The performances of this multiway Wiener filtering and comparative results with multimode PCA have been presented in the case of additive white noise reduction in a color image and in a multicomponent seismic wave.

Appendix A. n th-mode Wiener filter analytical expression. The following computations are related to section 3.3. They rely on the definitions and properties of tensors and multilinear algebra that can be found in [11, 13, 14].

The mean squared error involved in multiway Wiener filtering is given by relation

$$(A.1) \quad e(\mathbf{H}^{(1)}, \dots, \mathbf{H}^{(N)}) = E \left[\|\mathcal{X}\|^2 \right] - 2E \left[\langle \mathcal{X}, \mathcal{R} \times_1 \mathbf{H}^{(1)} \times_2 \dots \times_N \mathbf{H}^{(N)} \rangle \right] + E \left[\|\mathcal{R} \times_1 \mathbf{H}^{(1)} \times_2 \dots \times_N \mathbf{H}^{(N)}\|^2 \right].$$

The Frobenius norm of a tensor is also equal to the norm of any of its n th-mode flattening matrices. In order to determine the expression of filter $\mathbf{H}^{(n)}$ associated with fixed filters $\mathbf{H}^{(m)}$, for all $m \neq n$, the n th-mode flattening of (A.1) is processed.

Let us define matrix $\mathbf{F}_{\mathbf{XR}}^{(n)}$ as

$$(A.2) \quad \mathbf{F}_{\mathbf{XR}}^{(n)} = \mathbf{X}_n \mathbf{T}^{(n)} \mathbf{R}_n^T$$

with

$$(A.3) \quad \mathbf{T}^{(n)} = \mathbf{H}^{(1)} \otimes \dots \otimes \mathbf{H}^{(n-1)} \otimes \mathbf{H}^{(n+1)} \otimes \dots \otimes \mathbf{H}^{(N)}.$$

Hence, for all $n = 1$ to N ,

$$(A.4) \quad \left\langle \mathcal{X}, \mathcal{R} \times_1 \mathbf{H}^{(1)} \times_2 \dots \times_N \mathbf{H}^{(N)} \right\rangle = \text{tr} \left(\mathbf{F}_{\mathbf{XR}}^{(n)} \mathbf{H}^{(n)T} \right).$$

Let us define matrix $\mathbf{G}_{\mathbf{RR}}^{(n)}$ as

$$(A.5) \quad \mathbf{G}_{\mathbf{RR}}^{(n)} = \mathbf{R}_n \mathbf{Q}^{(n)} \mathbf{R}_n^T$$

with

$$(A.6) \quad \mathbf{Q}^{(n)} = \mathbf{T}^{(n)T} \mathbf{T}^{(n)}$$

$$(A.6) \quad \mathbf{Q}^{(n)} = \mathbf{H}^{(1)T} \mathbf{H}^{(1)} \otimes \dots \otimes \mathbf{H}^{(n-1)T} \mathbf{H}^{(n-1)} \otimes \mathbf{H}^{(n+1)T} \mathbf{H}^{(n+1)} \otimes \dots \otimes \mathbf{H}^{(N)T} \mathbf{H}^{(N)}.$$

Hence for all $n = 1$ to N ,

$$(A.7) \quad \left\| \mathcal{R} \times_1 \mathbf{H}^{(1)} \times_2 \dots \times_N \mathbf{H}^{(N)} \right\|^2 = \text{tr} \left(\mathbf{H}^{(n)} \mathbf{G}_{\mathbf{RR}}^{(n)} \mathbf{H}^{(n)T} \right).$$

Minimization of mean squared error $e(\mathbf{H}^{(1)}, \dots, \mathbf{H}^{(N)})$. The expression of the n th-mode flattened mean squared error $e(\mathbf{H}^{(1)}, \dots, \mathbf{H}^{(N)})$ is the following:

$$(A.8) \quad e(\mathbf{H}^{(1)}, \dots, \mathbf{H}^{(N)}) = \text{E} \left[\|\mathbf{X}_n\|^2 \right] - 2\text{E} \left[\text{tr} \left(\mathbf{F}_{\mathbf{XR}}^{(n)} \mathbf{H}^{(n)T} \right) \right] + \text{E} \left[\text{tr} \left(\mathbf{H}^{(n)} \mathbf{G}_{\mathbf{RR}}^{(n)} \mathbf{H}^{(n)T} \right) \right].$$

Assuming that m -mode filters $\mathbf{H}^{(m)}$ are fixed for all $m \neq n$, mean squared error $e(\mathbf{H}^{(1)}, \dots, \mathbf{H}^{(N)})$ is minimal when its gradient with respect to n th-mode filter $\mathbf{H}^{(n)}$ is null,

$$(A.9) \quad \text{grad}(e) = \left[\frac{\partial e}{\partial \mathbf{H}^{(1)}}, \dots, \frac{\partial e}{\partial \mathbf{H}^{(N)}} \right]^T,$$

that is, when $\frac{\partial e}{\partial \mathbf{H}^{(n)}}$ are conjointly null for all $n = 1$ to N . Let us study $\frac{\partial e}{\partial \mathbf{H}^{(n)}}$ for a given n th-mode. The n th-mode filters $\mathbf{H}^{(m)}$ are supposed to be fixed for all $m \in \{1, \dots, N\} - \{n\}$. Then $\frac{\partial e}{\partial \mathbf{H}^{(n)}} = 0$ implies that

$$(A.10) \quad \text{E} \left[\frac{\partial}{\partial \mathbf{H}^{(n)}} \text{tr} \left(\mathbf{H}^{(n)} \mathbf{G}_{\mathbf{RR}}^{(n)} \mathbf{H}^{(n)T} \right) \right] = 2\text{E} \left[\frac{\partial}{\partial \mathbf{H}^{(n)}} \text{tr} \left(\mathbf{F}_{\mathbf{XR}}^{(n)} \mathbf{H}^{(n)T} \right) \right],$$

We compute then the derivatives on both sides in (A.10), taking into account the fact that $\mathbf{G}_{\mathbf{RR}}^{(n)}$ and $\mathbf{F}_{\mathbf{XR}}^{(n)}$ are independent from $\mathbf{H}^{(n)}$:

$$(A.11) \quad \frac{\partial}{\partial \mathbf{H}^{(n)}} \text{tr} \left(\mathbf{F}_{\mathbf{XR}}^{(n)} \mathbf{H}^{(n)T} \right) = \mathbf{F}_{\mathbf{XR}}^{(n)},$$

$$(A.12) \quad \frac{\partial}{\partial \mathbf{H}^{(n)}} \text{tr} \left(\mathbf{H}^{(n)} \mathbf{G}_{\mathbf{RR}}^{(n)} \mathbf{H}^{(n)T} \right) = 2\mathbf{H}^{(n)} \mathbf{G}_{\mathbf{RR}}^{(n)}.$$

Expression of $\mathbf{H}^{(n)}$, n th-mode Wiener filter. Replacing (A.11) and (A.12) into expression (A.10) leads to the expression of $\mathbf{H}^{(n)}$ n th-mode Wiener filter associated with fixed $\mathbf{H}^{(m)}$ $m \neq n$:

$$(A.13) \quad \mathbf{H}^{(n)} = \gamma_{\mathbf{XR}}^{(n)} \mathbf{\Gamma}_{\mathbf{RR}}^{(n)-1},$$

where

$$(A.14) \quad \gamma_{\mathbf{XR}}^{(n)} = \text{E} \left[\mathbf{F}_{\mathbf{XR}}^{(n)} \right]$$

is the $\mathbf{T}^{(n)}$ -weighted covariance matrix between the signal \mathbf{X}_n and the data \mathbf{R}_n and

$$(A.15) \quad \mathbf{\Gamma}_{\mathbf{RR}}^{(n)} = \mathbb{E} \left[\mathbf{G}_{\mathbf{RR}}^{(n)} \right]$$

is the $\mathbf{Q}^{(n)}$ -weighted correlation matrix of the data.

Appendix B. Assumptions and related expression of the n th-mode Wiener filter. The following computations are related to section 3.3. Let us consider matrices $\mathbf{T}^{(n)}$ and $\mathbf{Q}^{(n)}$ defined in (A.3) and (A.6). Their generic (i, j) -terms are denoted respectively by $T_{ij}^{(n)}$ and by $Q_{ij}^{(n)}$.

Weight matrix term independence. The terms of weight matrix $\mathbf{O}^{(n)} \in \mathbb{R}^{K_n \times M_n}$ are supposed mutually independent,

$$(B.1) \quad \mathbb{E} [o_{kl}o_{mn}] = \alpha_{kl}\delta_{km}\delta_{ln},$$

whatever k and $m \in \{1, \dots, K_n\}$, l and $n \in \{1, \dots, M_n\}$ and where α_{kl} is not null.

White and Gaussian noise condition. White and Gaussian noise condition applied to the n th-mode flattening \mathbf{N}_n can be expressed by

$$(B.2) \quad \mathbb{E} [n_{kl}n_{pq}] = \sigma_n^2\delta_{kp}\delta_{lq},$$

where $(k, p) \in \{1, \dots, K_n\}^2$, $(l, q) \in \{1, \dots, M_n\}^2$ and σ_n^2 is the n th-mode noise power.

Noise and signal independence. The condition on noise and signal independence can be expressed by

$$(B.3) \quad \mathbb{E} [x_{kl}n_{pq}] = 0$$

for all $(k, p) \in \{1, \dots, K_n\}^2$ and $(l, q) \in \{1, \dots, M_n\}^2$. Hence, $\mathbf{T}^{(n)}$ and $\mathbf{Q}^{(n)}$ -weighted (\mathbf{X}, \mathbf{N}) -covariance matrices are null:

$$(B.4) \quad \begin{aligned} \gamma_{\mathbf{XN}}^{(n)} &= \gamma_{\mathbf{NX}}^{(n)} = \mathbf{0}, \\ \mathbf{\Gamma}_{\mathbf{XN}}^{(n)} &= \mathbf{\Gamma}_{\mathbf{NX}}^{(n)} = \mathbf{0}. \end{aligned}$$

Indeed, their (i, j) -term is

$$(B.5) \quad \begin{aligned} \left(\gamma_{\mathbf{XN}}^{(n)} \right)_{ij} &= \sum_{k=1}^{M_n} \sum_{l=1}^{M_n} T_{kl}^{(n)} \mathbb{E} [x_{ik}n_{jl}], \\ \left(\mathbf{\Gamma}_{\mathbf{XN}}^{(n)} \right)_{ij} &= \sum_{k=1}^{M_n} \sum_{l=1}^{M_n} Q_{kl}^{(n)} \mathbb{E} [x_{ik}n_{jl}]. \end{aligned}$$

Expressions of weighted covariance matrices.

Covariance matrix $\gamma_{\mathbf{RR}}^{(n)}$. As $\mathbf{R}_n = \mathbf{X}_n + \mathbf{N}_n$, the expression of $\gamma_{\mathbf{RR}}^{(n)}$ reads

$$(B.6) \quad \gamma_{\mathbf{RR}}^{(n)} = \gamma_{\mathbf{XX}}^{(n)} + \gamma_{\mathbf{XN}}^{(n)} + \gamma_{\mathbf{NX}}^{(n)} + \gamma_{\mathbf{NN}}^{(n)}.$$

So according to (B.4), $\gamma_{\mathbf{RR}}^{(n)}$ weighted covariance matrix can be expressed by

$$(B.7) \quad \gamma_{\mathbf{RR}}^{(n)} = \gamma_{\mathbf{XX}}^{(n)} + \gamma_{\mathbf{NN}}^{(n)}.$$

Moreover,

$$(B.8) \quad \gamma_{\mathbf{XR}}^{(n)} = \gamma_{\mathbf{XX}}^{(n)} + \gamma_{\mathbf{XN}}^{(n)} = \gamma_{\mathbf{XX}}^{(n)}.$$

Covariance matrix $\Gamma_{\mathbf{RR}}^{(n)}$. Relations (B.6), (B.7), and (B.8) hold as well for $\Gamma_{\mathbf{RR}}^{(n)}$:

$$\Gamma_{\mathbf{RR}}^{(n)} = \Gamma_{\mathbf{XX}}^{(n)} + \Gamma_{\mathbf{XN}}^{(n)} + \Gamma_{\mathbf{NX}}^{(n)} + \Gamma_{\mathbf{NN}}^{(n)}$$

and

$$(B.9) \quad \Gamma_{\mathbf{RR}}^{(n)} = \Gamma_{\mathbf{XX}}^{(n)} + \Gamma_{\mathbf{NN}}^{(n)}.$$

Moreover,

$$(B.10) \quad \Gamma_{\mathbf{XR}}^{(n)} = \Gamma_{\mathbf{XX}}^{(n)} + \Gamma_{\mathbf{XN}}^{(n)} = \Gamma_{\mathbf{XX}}^{(n)}.$$

Expressions of $\Gamma_{\mathbf{NN}}^{(n)}$ and $\gamma_{\mathbf{NN}}^{(n)}$. According to (B.2), the (i, j) -term of $\Gamma_{\mathbf{NN}}^{(n)}$ is the following:

$$(B.11) \quad \left(\Gamma_{\mathbf{NN}}^{(n)}\right)_{ij} = \sum_{k=1}^{M_n} \sum_{l=1}^{M_n} Q_{kl}^{(n)} \mathbb{E}[n_{ik}n_{jl}] = \sigma_{\Gamma}^{(n)^2} \delta_{ij}$$

with

$$(B.12) \quad \sigma_{\Gamma}^{(n)^2} = \text{tr}(\mathbf{Q}^{(n)})\sigma_n^2.$$

Hence

$$(B.13) \quad \Gamma_{\mathbf{NN}}^{(n)} = \sigma_{\Gamma}^{(n)^2} \mathbf{I}_{I_n}.$$

The (i, j) -term of $\gamma_{\mathbf{NN}}^{(n)}$ can also be expressed by

$$\left(\gamma_{\mathbf{NN}}^{(n)}\right)_{ij} = \sum_{k=1}^{M_n} \sum_{l=1}^{M_n} T_{kl}^{(n)} \mathbb{E}[n_{ik}n_{jl}] = \sigma_{\gamma}^{(n)^2} \delta_{ij}$$

with

$$\sigma_{\gamma}^{(n)^2} = \text{tr}(\mathbf{T}^{(n)})\sigma_n^2.$$

Hence

$$(B.14) \quad \gamma_{\mathbf{NN}}^{(n)} = \sigma_{\gamma}^{(n)^2} \mathbf{I}_{I_n}.$$

Expressions of $\Gamma_{\mathbf{XX}}^{(n)}$ and $\gamma_{\mathbf{XX}}^{(n)}$. Considering the signal model (3.16),

$$(B.15) \quad \gamma_{\mathbf{XX}}^{(n)} = \mathbf{V}_s^{(n)} \gamma_{\mathbf{OO}}^{(n)} \mathbf{V}_s^{(n)T},$$

where

$$(B.16) \quad \gamma_{\mathbf{OO}}^{(n)} = \mathbb{E} \left[\mathbf{O}^{(n)} \mathbf{T}^{(n)} \mathbf{O}^{(n)T} \right].$$

According to (B.1), the generic term of $\gamma_{\mathbf{OO}}^{(n)}$ is

$$(B.17) \quad \left(\gamma_{\mathbf{OO}}^{(n)}\right)_{ij} = \sum_{k=1}^{M_n} \sum_{l=1}^{M_n} T_{kl}^{(n)} \mathbb{E}[n_{ik}n_{jl}] = \beta_i \delta_{ij},$$

where, for all $i = 1$ to K_n ,

$$(B.18) \quad \beta_i = \sum_{k=1}^{M_n} T_{kk}^{(n)} \alpha_{ik},$$

and where α_{ik} is defined in (B.1). So, $\gamma_{\mathbf{OO}}^{(n)}$ is a diagonal matrix:

$$(B.19) \quad \gamma_{\mathbf{OO}}^{(n)} = \begin{bmatrix} \beta_1 & 0 \\ & \ddots \\ 0 & \beta_{K_n} \end{bmatrix}.$$

The matrix $\mathbf{\Gamma}_{\mathbf{XX}}^{(n)}$ is also expressed as

$$(B.20) \quad \mathbf{\Gamma}_{\mathbf{XX}}^{(n)} = \mathbf{V}_s^{(n)} \mathbf{\Gamma}_{\mathbf{OO}}^{(n)} \mathbf{V}_s^{(n)T},$$

where $\mathbf{\Gamma}_{\mathbf{OO}}^{(n)}$ is the diagonal matrix

$$(B.21) \quad \mathbf{\Gamma}_{\mathbf{OO}}^{(n)} = \begin{bmatrix} \epsilon_1 & 0 \\ & \ddots \\ 0 & \epsilon_{K_n} \end{bmatrix},$$

and

$$(B.22) \quad \epsilon_i = \sum_{k=1}^{M_n} Q_{kk}^{(n)} \alpha_{ik},$$

where α_{ik} is defined in (B.1).

Final expression of $\mathbf{H}^{(n)}$, n th-mode Wiener filter. According to (B.8) and (B.15),

$$(B.23) \quad \gamma_{\mathbf{XR}}^{(n)} = \mathbf{V}_s^{(n)} \gamma_{\mathbf{OO}}^{(n)} \mathbf{V}_s^{(n)T}.$$

According to (B.9), (B.13), and (B.20),

$$\mathbf{\Gamma}_{\mathbf{RR}}^{(n)} = \mathbf{V}_s^{(n)} \mathbf{\Gamma}_{\mathbf{OO}}^{(n)} \mathbf{V}_s^{(n)T} + \sigma_\Gamma^{(n)2} \mathbf{I}_{I_n},$$

which can be expressed as

$$(B.24) \quad \mathbf{\Gamma}_{\mathbf{RR}}^{(n)} = \begin{bmatrix} \mathbf{V}_s^{(n)} & \mathbf{V}_b^{(n)} \end{bmatrix} \begin{bmatrix} \mathbf{\Gamma}_{\mathbf{OO}}^{(n)} + \sigma_\Gamma^{(n)2} \mathbf{I}_{K_n} & 0 \\ 0 & \sigma_\Gamma^{(n)2} \mathbf{I}_{I_n - K_n} \end{bmatrix} \begin{bmatrix} \mathbf{V}_s^{(n)T} \\ \mathbf{V}_b^{(n)T} \end{bmatrix}$$

with $\mathbf{V}_b^{(n)} \in \text{St}(I_n, I_n - K_n)$ the columnwise orthogonal matrix containing the noise subspace basis vectors. The assumption of noise and signal independence implies that the noise and signal subspaces are orthogonal:

$$(B.25) \quad \mathbf{V}_s^{(n)T} \mathbf{V}_b^{(n)} = \mathbf{0}.$$

Let us call

$$(B.26) \quad \mathbf{\Lambda}_s^{(n)} = \mathbf{\Gamma}_{\mathbf{OO}}^{(n)} + \sigma_\Gamma^{(n)2} \mathbf{I}_{K_n}$$

and

$$(B.27) \quad \Lambda_b^{(n)} = \sigma_{\Gamma}^{(n)2} \mathbf{I}_{I_n - K_n}.$$

Inserting the last expressions of $\gamma_{\mathbf{XR}}^{(n)}$ and $\Gamma_{\mathbf{RR}}^{(n)}$ (see (B.23) and (B.24)) into Wiener n th-mode filter expression (A.13) leads to

$$(B.28) \quad \mathbf{H}^{(n)} = \mathbf{V}_s^{(n)} \gamma_{\mathbf{OO}}^{(n)} \mathbf{V}_s^{(n)T} \begin{bmatrix} \mathbf{V}_s^{(n)} & \mathbf{V}_b^{(n)} \end{bmatrix} \begin{bmatrix} \Lambda_s^{(n)-1} & 0 \\ 0 & \Lambda_b^{(n)-1} \end{bmatrix} \begin{bmatrix} \mathbf{V}_s^{(n)T} \\ \mathbf{V}_b^{(n)T} \end{bmatrix},$$

which can be expressed as

$$(B.29) \quad \mathbf{H}^{(n)} = \begin{bmatrix} (\mathbf{V}_s^{(n)} \gamma_{\mathbf{OO}}^{(n)} \mathbf{V}_s^{(n)T} \mathbf{V}_s^{(n)}) & (\mathbf{V}_s^{(n)} \gamma_{\mathbf{OO}}^{(n)} \mathbf{V}_s^{(n)T} \mathbf{V}_b^{(n)}) \end{bmatrix} \begin{bmatrix} \Lambda_s^{(n)-1} \mathbf{V}_s^{(n)T} & 0 \\ 0 & \Lambda_b^{(n)-1} \mathbf{V}_b^{(n)T} \end{bmatrix}$$

Considering noise and signal orthogonality condition (B.25) and the fact that $\mathbf{V}_n^{(n)} \mathbf{V}_n^{(n)T} = \mathbf{I}_{K_n}$, the final Wiener n th-mode filter expression becomes

$$(B.30) \quad \mathbf{H}^{(n)} = \mathbf{V}_s^{(n)} \gamma_{\mathbf{OO}}^{(n)} \Lambda_{\Gamma_s}^{(n)-1} \mathbf{V}_s^{(n)T}.$$

Acknowledgments. We would like to thank the anonymous reviewers who contributed to the quality of this paper by providing helpful suggestions.

REFERENCES

- [1] K. ABED-MERAIM, H. MAÎTRE, AND P. DUHAMEL, *Blind multichannel image restoration using subspace based method*, in Proceedings of the IEEE International Conference on Acoustics Systems and Signal Processing, Hong Kong, China, 2003.
- [2] O. ALTER AND G. H. GOLUB, *Reconstructing the pathways of a cellular system from genome-scale signals by using matrix and tensor computations*, Proc. Natl. Acad. Sci. USA, 102 (2005), pp. 17559–17564.
- [3] H. C. ANDREWS AND C. L. PATTERSON, *Singular value decomposition and digital image processing*, IEEE Trans. Acoustics, Speech, and Signal Processing, 24 (1976), pp. 26–53.
- [4] H. C. ANDREWS AND C. L. PATTERSON, *Singular Value Decomposition (SVD) image coding*, IEEE Trans. Communications, 1976, pp. 425–432.
- [5] B. W. BADER AND T. G. KOLDA, *Algorithm 862: Matlab tensor classes for fast algorithm prototyping*, ACM Tran. Math. Software, 32 (2006).
- [6] E. BALA AND A. ERTÜZÜN, *A multivariate thresholding technique for image denoising using multiwavelets*, EURASIP J. ASP, 8 (2005), pp. 1205–1211.
- [7] S. BOURENNANE AND A. BENDJAMA, *Locating wide band acoustic sources using higher order statistics*, Appl. Acoustics, 63 (2001), pp. 235–251.
- [8] A. BENJAMA, S. BOURENNANE, AND M. FRIKEL, *Seismic wave separation based on higher order statistics*, in Proceedings of the IEEE International Conference on Digital Signal Processing and its Applications, Moscow, Russia, 1998.
- [9] R. BRO, *Multi-way Analysis in the Food Industry*, Ph.D. thesis, Royal Veterinary and Agricultural University, Denmark, 1998.
- [10] J. D. CAROLL AND J. J. CHANG, *Analysis of individual differences in multidimensional scaling via n -way generalization of Eckart–Young decomposition*, Psychometrika, 35 (1970), pp. 283–319.
- [11] L. DE LATHAUWER, *Signal Processing Based on Multilinear Algebra*, Ph.D. thesis, K. U. Leuven, E. E. Department (ESAT), Belgium, 1997.
- [12] L. DE LATHAUWER, B. DE MOOR, AND J. VANDEWALLE, *Dimensionality reduction in higher-order-only ICA*, Proc. IEEE Signal Processing Workshop on Higher Order Statistics, 7 (1997), pp. 316–320.
- [13] L. DE LATHAUWER, B. DE MOOR, AND J. VANDEWALLE, *A multilinear singular value decomposition*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1253–1278.

- [14] L. DE LATHAUWER, B. DE MOOR, AND J. VANDEWALLE, *On the best rank-(1) and rank- (r_1, \dots, r_N) approximation of higher-order tensors*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1324–1342.
- [15] Y. ECKART AND G. YOUNG, *The approximation of a matrix by another of lower rank*, Psychometrika, 1 (1936), pp. 211–218.
- [16] F. L. M. FREIRE AND T. J. ULRYCH, *Application of SVD to vertical seismic profiling*, Geophysics, 53 (1988), pp. 778–785.
- [17] F. GLANGEAUD AND J.-L. MARI, *Wave Separation*, Technip IFP, 1994.
- [18] G. H. GOLUB AND C.F. VAN LOAN, *Matrix Computations*, 3rd ed., John Hopkins University Press, Baltimore, 1996.
- [19] E. GONEN AND J. M. MENDEL, *Application of cumulants to array processing Part III: Blind beamforming for coherent signal*, IEEE Trans. Signal Process., 45 (1997), pp. 2252–64.
- [20] R. A. HARSHMAN AND M. E. LUNDY, *Research methods for multimode data analysis*, Praeger, New York, 1984, pp. 122–215.
- [21] M. HEMON AND D. MACE, *The use of Karhunen–Loeve transform in seismic data prospecting*, Geophys. Prospecting, 26 (1978), pp. 600–626.
- [22] J. HÅSTAD, *Tensor rank is np -complete*, J. Algorithms, 11 (1990), pp. 644–654.
- [23] K. HSU, *Wave separation and feature extraction of acoustic well-logging waveforms of triaxial recordings by singular value decomposition*, Geophysics, 55 (1990), pp. 176–184.
- [24] G. M. JACKSON, I. M. MASON, AND S. A. GREENHALGH, *Principal component transforms of triaxial recordings by singular value decomposition*, Geophysics, 56 (1991), pp. 176–184.
- [25] H. KIEERS, *Towards a standardized notation and terminology in multiway analysis*, J. Chemometrics, 14 (2000), pp. 105–122.
- [26] E. KOFIDIS AND P. A. REGALIA, *On the best rank-1 approximation of higher-order supersymmetric tensors*, SIAM J. Matrix Anal. Appl., 23 (2002), pp. 863–884.
- [27] T. G. KOLDA, *Orthogonal tensor decomposition*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 243–255.
- [28] P. M. KROONENBERG, *Three-mode Principal Component Analysis*, DSWO Press, Leiden, Netherlands, 1983.
- [29] P. M. KROONENBERG AND J. DE LEEUW, *Principal component analysis of three-mode data by means of alternating least squares algorithms*, Psychometrika, 45 (1980), pp. 69–97.
- [30] J. B. KRUSKAL, *Rank, decomposition, and uniqueness for 3-way and N -way arrays*, Multiway Data Analysis, Elsevier, Amsterdam, Netherlands, 1988.
- [31] J. L. LACOUME, P. O. AMBLARD, AND P. COMON, *Statistiques d'ordre supérieur pour le traitement du signal*, Masson, Paris, 1997.
- [32] X. LIU, *Ground roll suppression using the Karhunen–Loeve transform*, Geophysics, 64 (1991), pp. 564–566.
- [33] J. M. MENDEL, *Tutorial on higher order statistics (spectra) in signal processing and system theory: Theoretical results and some applications*, Proc. IEEE, 79 (1991), pp. 278–305.
- [34] D. MUTI, *Traitement du Signal Tensoriel. Application aux images en Couleurs et aux Signaux Sismiques*, Ph.D. thesis, Institut Fresnel, Université Paul Cézanne, Aix-Marseille III, Marseille, France, 2004.
- [35] D. MUTI AND S. BOURENNANE, *Multidimensional estimation based on a tensor decomposition*, in Proceedings of the IEEE Workshop on Statistical Signal Processing, St. Louis, 2003.
- [36] D. MUTI AND S. BOURENNANE, *Multidimensional signal processing using lower rank tensor approximation*, in Proceedings of the IEEE International Conference on Acoustics, Systems, and Signal Processing, Hong Kong, China, 2003.
- [37] D. MUTI AND S. BOURENNANE, *Traitement du signal par décomposition tensorielle*, in Proceedings of the GRETSI Symposium, Paris, France, 2003.
- [38] D. MUTI AND S. BOURENNANE, *Multidimensional filtering based on a tensor approach*, Signal Process. J., 85 (2005), pp. 2338–2353.
- [39] D. MUTI AND S. BOURENNANE, *Multiway filtering based on fourth order cumulants*, Appl. Signal Process., 7 (2005), pp. 1147–1159.
- [40] R. NEELAMANI, H. CHOI, AND R. BARANIUK, *Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems*, IEEE Trans. Signal Process., 52 (2004), pp. 418–433.
- [41] B. PORAT AND B. FRIEDLANDER, *Direction finding algorithms based on higher-order statistics*, IEEE Trans. Signal Process., 39 (1991), pp. 2016–2024.
- [42] C. PREZA, M. I. MILLER, L. J. THOMAS JR., AND J. G. MCNALLY, *Regularized method for reconstruction of three-dimensional microscopic objects from optical sections*, J. Opt. Soc. Amer. A, 9 (1992), pp. 219–228.
- [43] L. W. SHE AND B. ZHENG, *Multiwavelets based denoising of SAR images*, in Proceedings of the 5th International Conference on Signal Processing, Beijing, China, 8, 2000, pp. 321–324.

- [44] N. D. SIDIROPOULOS AND R. BRO, *On the uniqueness of multilinear decomposition of N -way arrays*, J. Chemometrics, 14 (2000), pp. 229–239.
- [45] N. D. SIDIROPOULOS, R. BRO, AND G. GIANNAKIS, *Parallel factor analysis in sensor array processing*, IEEE Trans. Signal Process., 48 (2000), pp. 2377–2388.
- [46] N. D. SIDIROPOULOS, G. GIANNAKIS, AND R. BRO, *Blind PARAFAC receivers for DS-CDMA systems*, IEEE Trans. Signal Process., 48 (2000), pp. 810–823.
- [47] A. SMILDE, R. BRO, AND P. GELADI, *Multi-way Analysis: Applications in the Chemical Sciences*, John Wiley, New York, 2004.
- [48] L. R. TUCKER, *Some mathematical notes on three-mode factor analysis*, Psychometrika, 31 (1966), pp. 279–311.
- [49] M. A. O. VASILESCU AND D. TERZOPOULOS, *Multilinear image analysis for facial recognition*, in Proceedings of the IEEE International Conference on Pattern Recognition (ICPR2002), Vol. 2, Quebec City, Canada, 2002.
- [50] M. A. O. VASILESCU AND D. TERZOPOULOS, *Multilinear independent components analysis*, in Proceedings of Learning 2004, Snowbird, UT, 2004.
- [51] M. A. O. VASILESCU AND D. TERZOPOULOS, *Multilinear independent components analysis*, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'05), 1 (2005), pp. 547–553.
- [52] H. WANG AND N. AHUJA, *Facial expression decomposition*, in Proceedings of the 9th IEEE International Conference on Computer Vision (ICCV2003), Vol. 2, Nice, France, 2003.
- [53] M. WAX AND T. KAILATH, *Detection of signals information theoretic criteria*, IEEE Trans. Acoustics Speech Signal Process., 33, 1985.
- [54] N. YUEN AND B. FRIEDLANDER, *DOA estimation in multipath: An approach using fourth order cumulant*, IEEE Trans. Signal Process., 45 (1997), pp. 1253–63.
- [55] N. YUEN AND B. FRIEDLANDER, *Asymptotic performance analysis of blind signal copy using fourth order cumulant*, Internat. Adaptative Control Signal Process., 48 (1996), pp. 239–65.
- [56] T. ZHANG AND G. H. GOLUB, *Rank-one approximation to high order tensor*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 534–550.