

FAST AND STABLE RECOVERY OF APPROXIMATELY LOW MULTILINEAR RANK TENSORS FROM MULTI-WAY COMPRESSIVE MEASUREMENTS

Cesar F. Caiafa*

Andrzej Cichocki

IAR - CCT La Plata, CONICET
CC5 (1894) V. Elisa, ARGENTINA

LABSP - Brain Science Institute, RIKEN
2-1 Hirosawa, Wako, 351-0198, JAPAN

ABSTRACT

We introduce a reconstruction formula that allows one to recover an N -order tensor $\underline{\mathbf{X}} \in \mathbb{R}^{I_1 \times \dots \times I_N}$ from a reduced set of multi-way compressive measurements by exploiting its low multilinear rank structure. It is proved that, in the matrix case ($N = 2$), the proposed reconstruction is stable in the sense that the approximation error is proportional to the one provided by the best low-rank approximation, i.e. $\|\mathbf{X} - \hat{\mathbf{X}}\|_2 \leq K \|\mathbf{X} - \mathbf{X}_0\|_2$, where K is a constant and \mathbf{X}_0 is the corresponding truncated SVD of \mathbf{X} . We also present simulation results indicating that the same stable behavior is observed with higher order tensors ($N > 2$). In addition, it is shown that, an interesting property of multi-way measurements allows us to build the reconstruction based on compressive linear measurements of fibers taken only in two selected modes, independently of the tensor order N . Simulation results using real-world 2D and 3D signals are presented illustrating our results and comparing the reconstructions against the best low multilinear rank approximations and the reconstructions obtained by using the Kronecker-CS approach.

Index Terms— Compressed Sensing (CS), Kronecker-CS, Low-rank tensors, Multi-way analysis, Tucker model.

1. INTRODUCTION

During the last years there has been an increased interest in *Compressed Sensing* (CS), whose aim is to reconstruct a signal based on a set of measurements that is much smaller than the original signal size. In classical CS, a signal $\mathbf{x} \in \mathbb{R}^n$ is reconstructed from a reduced set of m linear projections ($m < n$) $\mathbf{y} = \Phi \mathbf{x}$, where the sensing matrix $\Phi \in \mathbb{R}^{m \times n}$ is typically random or composed by few selected rows of the Fourier transform matrix. In order to make the CS problem solvable, it is necessary to impose constraints about the signal of interest, for example, it is assumed that the signal is

Also from Facultad de Ingeniería - UBA, Av. Paseo Colón 850. 4to. piso, Ala sur (1063), Buenos Aires, ARGENTINA.

Partially supported by grants from ANPCyT (Agencia Nacional de Promoción Científica y Tecnológica), PICT 2012 #1519; and CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas), PIP #114-201101-00021.

compressible by decomposing it in a known basis (*dictionary*), typically associated with a Wavelet transform. Many algorithms were developed to solve this problem involving iterative refinements of the solution by means of *Greedy algorithms* or by minimizing the ℓ_1 -norm of the solution (see [1] for an up to date summary of algorithms). These algorithms have found many applications in diverse fields such as in medical imaging, surveillance, machine learning, etc.

Recently, CS has been extended to problems involving the recovery of multidimensional datasets, such as matrices and tensors, by exploiting their associated low-rank structures and considering different models for how the measurements are taken. For example, in [2, 3] matrices are reconstructed from limited information or undersampled measurements by solving a convex optimization problem. More recently, similar ideas has been extended to tensor data allowing, for example, to estimate missing entries in tensors [4, 5], to reconstruct tensors from linear projections [6] and to perform tensor denoising [7].

In this paper we approach the problem of reconstructing an approximately low multilinear rank tensor and we assume that measurements are provided as a set of multilinear projections, i.e. multiplying each mode of the data tensor by a different sensing matrix. This model comes into scene naturally, for example, in the case of sensing 2D images by means of a separable operator as developed in [8, 9], i.e. by taking compressive measurements of columns and rows separately, imposing a Kronecker structure on the sensing operator. In [10], the Kronecker-CS model was proposed to deal with higher order tensors and applied to hyperspectral 3D images and video data. Recently, greedy algorithms specially designed to take advantage of the Kronecker structure and block sparsity of the representations were proposed in [11, 12] and applied to a variety of signal processing problems in [13]. Also, in [20], generalized tensor CS algorithms were developed using an ℓ_1 -norm minimization approach. More recently, in [14], the Kronecker sensing structure was used for tensor compression and a method involving a low-rank model fitting, followed by a per mode ℓ_0/ℓ_1 decompression, was proposed in order to recover a low-rank tensor based on the PARAFAC tensor decomposition model.

1.1. Tensor Notation and Definitions

Tensors (multi-way arrays) are denoted by underlined bold-face capital letters, e.g. $\underline{\mathbf{X}} \in \mathbb{R}^{I_1 \times \dots \times I_N}$ is an N th order tensor of real numbers. Matrices (2D arrays) are denoted by bold uppercase letters and vectors are denoted by bold-face lower-case letters, e.g. $\mathbf{X} \in \mathbb{R}^{I_1 \times I_2}$ and $\mathbf{x} \in \mathbb{R}^I$ are a matrix and a vector, respectively. The element (i_1, \dots, i_N) of a tensor is referred as $x_{i_1 \dots i_N}$. The Frobenius norm of a tensor is defined by $\|\underline{\mathbf{X}}\|_F = \sqrt{\sum_{i_1} \dots \sum_{i_N} x_{i_1 \dots i_N}^2}$. Given a tensor $\underline{\mathbf{X}} \in \mathbb{R}^{I_1 \times \dots \times I_N}$, its mode- n fibers are the vectors obtained by fixing all indices except i_n , which correspond to columns ($n = 1$), rows ($n = 2$), and so on. Mode- n unfolding of a tensor $\underline{\mathbf{X}} \in \mathbb{R}^{I_1 \times \dots \times I_N}$ yields a matrix $\mathbf{X}_{(n)} \in \mathbb{R}^{I_n \times \bar{I}_n}$ ($\bar{I}_n = \prod_{m \neq n} I_m$) whose columns are the corresponding mode- n fibers [15]. Given a multidimensional signal (tensor) $\underline{\mathbf{X}} \in \mathbb{R}^{I_1 \times \dots \times I_N}$ and a matrix $\Phi \in \mathbb{R}^{J \times I_n}$ the mode- n tensor by matrix product $\underline{\mathbf{Y}} = \underline{\mathbf{X}} \times_n \Phi \in \mathbb{R}^{I_1 \times \dots \times I_{n-1} \times J \times I_{n+1} \times \dots \times I_N}$ is defined by: $y_{i_1 \dots i_{n-1} j i_{n+1} \dots i_N} = \sum_{i_n=1}^{I_n} x_{i_1 \dots i_n \dots i_N} \phi_{j i_n}$, with $i_k = 1, 2, \dots, I_k$ ($k \neq n$) and $j = 1, 2, \dots, J$. It is noted that this corresponds to the product of matrix Φ by each one of the mode- n fibers of $\underline{\mathbf{X}}$ since $\mathbf{Y}_{(n)} = \Phi \mathbf{X}_{(n)}$.

The Tucker decomposition [16] provides a generalization of the low-rank approximation of matrices to the case of tensors, i.e. $\underline{\mathbf{X}} = \underline{\hat{\mathbf{X}}} + \underline{\mathbf{E}}$, where $\underline{\mathbf{E}}$ is an error tensor and the multilinear rank- (R_1, \dots, R_N) tensor approximation $\underline{\hat{\mathbf{X}}}$ (Tucker model) is defined as follows:

$$\underline{\hat{\mathbf{X}}} = \underline{\mathbf{G}} \times_1 \mathbf{A}_1 \times_2 \dots \times_N \mathbf{A}_N, \quad (1.1)$$

with *core tensor* $\underline{\mathbf{G}} \in \mathbb{R}^{R_1 \times \dots \times R_N}$ and *factor matrices* $\mathbf{A}_n \in \mathbb{R}^{I_n \times R_n}$ ($R_n \leq I_n$). A tensor $\underline{\mathbf{X}} \in \mathbb{R}^{I_1 \times \dots \times I_N}$ is said to have *multilinear rank*- (R_1, \dots, R_N) if such a decomposition is exact for a set of minimal values (R_1, \dots, R_N) , i.e. $\underline{\mathbf{X}} = \underline{\hat{\mathbf{X}}}$. We say that a tensor $\underline{\mathbf{G}} \in \mathbb{R}^{R_1 \times \dots \times R_N}$ is *full-rank* if all its unfolded matrices are full-rank matrices, i.e. $\text{rank}(\mathbf{G}_{(n)}) = R_n, \forall n$. A particularly interesting case of the Tucker model is when factor matrices $\mathbf{A}_n \in \mathbb{R}^{I_n \times R_n}$ are chosen as the truncated matrices of left singular vectors associated with the unfolding matrices $\mathbf{X}_{(n)}$. In this case, we obtain the so called truncated *Higher Order Singular Value Decomposition* (HOSVD) [16], which is denoted here as $\underline{\mathbf{X}}_0$. It is interesting to note that the HOSVD decomposition does not provide the best low multilinear rank approximation, however, we have that $\|\underline{\mathbf{X}} - \underline{\mathbf{X}}_0\|_F \leq \sqrt{N} \|\underline{\mathbf{X}} - \underline{\mathbf{X}}_{opt}\|_F$ [17]. It is noted that, in the matrix case, the SVD provides the best low rank approximation having orthogonal factors and a diagonal core matrix.

2. RECOVERY OF EXACT MULTILINEAR RANK- (R_1, \dots, R_N) TENSORS

We assume, for the moment, that we have available the following set of compressive multi-way measurements $\underline{\mathbf{Z}}_n \in$

$\mathbb{R}^{R_1 \times \dots \times R_{n-1} \times I_n \times R_{n+1} \times \dots \times R_N}$ ($n = 1, 2, \dots, N$):

$$\underline{\mathbf{Z}}_n = \underline{\mathbf{X}} \times_1 \Phi_1 \cdots \times_{n-1} \Phi_{n-1} \times_{n+1} \Phi_{n+1} \cdots \times_N \Phi_N, \quad (2.1)$$

where the sensing matrices $\Phi_n \in \mathbb{R}^{R_n \times I_n}$ ($R_n \ll I_n$) will be specified later. Note that eq. (2.1) indicates that the original tensor is multiplied by a set of different sensing matrices in all modes except in mode- n (see Fig. 1 (top)). The following theorem provides an explicit reconstruction formula, as illustrated in Fig. 1 (bottom), and states the condition under which the original tensor can be exactly recovered from this set of multi-way measurements (see the proof in Appendix 6).

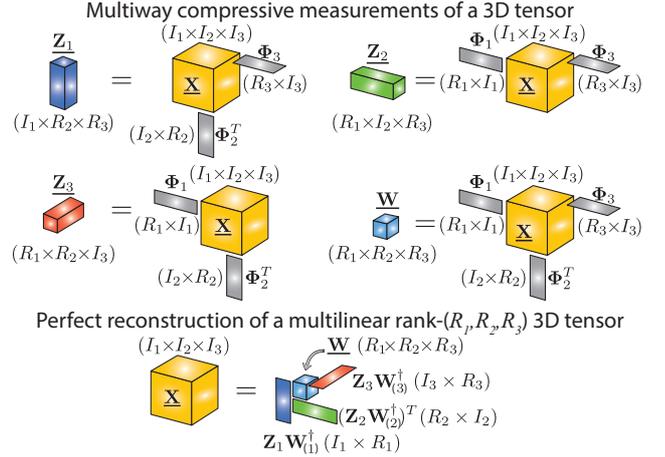


Fig. 1. Multi-way measurements and the reconstruction model.

Theorem 2.1. If tensor $\underline{\mathbf{X}} \in \mathbb{R}^{I_1 \times \dots \times I_N}$ has multilinear rank- (R_1, \dots, R_N) and matrices $\Phi_n \in \mathbb{R}^{R_n \times I_n}$ are such that the tensor $\underline{\mathbf{W}} = \underline{\mathbf{X}} \times_1 \Phi_1 \times_2 \dots \times_N \Phi_N \in \mathbb{R}^{R_1 \times \dots \times R_N}$ is full-rank, then the following reconstruction formula is exact, i.e. $\underline{\hat{\mathbf{X}}} = \underline{\mathbf{X}}$:

$$\underline{\hat{\mathbf{X}}} = \underline{\mathbf{W}} \times_1 \mathbf{Z}_1 \mathbf{W}_{(1)}^\dagger \cdots \times_N \mathbf{Z}_N \mathbf{W}_{(N)}^\dagger, \quad (2.2)$$

where “ \dagger ” stands for the Moore-Penrose pseudo-inverse of a matrix and $\mathbf{Z}_n \equiv (\underline{\mathbf{Z}}_n)_{(n)} \in \mathbb{R}^{I_n \times \bar{R}_n}$, with $\bar{R}_n = \prod_{m \neq n} R_m$.

In some applications, the available measurements are given as linear projections of mode- n fibers, e.g. in [8] rows and columns of an image are measured through random sensing matrices. It is interesting to note that all multi-way measurements defined in eq. (2.1) ($n = 1, 2, \dots, N$) can be computed from linear measurements taken only in two selected modes out of N . More explicitly, suppose we have at our disposal the linear measurements in modes m_1 and m_2 ($m_1 \neq m_2$) given by: $\mathbf{Y}_m = \Phi_m \mathbf{X}_{(m)}$, with $m = m_1, m_2$; then, it is easy to see that the mode- m unfolding matrix of every multi-way measurement $\underline{\mathbf{Z}}_n$, ($n \neq m$), can be written as follows:

$$(\underline{\mathbf{Z}}_n)_{(m)} = \mathbf{Y}_m (\Phi_N^T \otimes \dots \otimes \Phi_{m+1}^T \otimes \Phi_{m-1}^T \otimes \dots \otimes \Phi_{n+1}^T \otimes \mathbf{I} \otimes \Phi_{n-1}^T \otimes \dots \otimes \Phi_1^T).$$

For example, for a 3rd order tensor ($N = 3$), if we have available compressive measurements of columns ($\mathbf{Y}_1 = \Phi_1 \mathbf{X}_{(1)} \in \mathbb{R}^{R_1 \times I_2 I_3}$) and rows ($\mathbf{Y}_2 = \Phi_2 \mathbf{X}_{(2)} \in \mathbb{R}^{R_2 \times I_1 I_3}$), then the multi-way measurements can be computed as follows: $(\underline{\mathbf{Z}}_1)_{(2)} = \mathbf{Y}_2 (\Phi_3^T \otimes \mathbf{I}) \in \mathbb{R}^{R_2 \times I_1 R_3}$, $(\underline{\mathbf{Z}}_2)_{(1)} = \mathbf{Y}_1 (\Phi_3^T \otimes \mathbf{I}) \in \mathbb{R}^{R_1 \times I_2 R_3}$ and $(\underline{\mathbf{Z}}_3)_{(1)} = \mathbf{Y}_1 (\Phi_2^T \otimes \mathbf{I}) \in \mathbb{R}^{R_1 \times R_2 I_3}$.

3. STABLE RECOVERY OF APPROXIMATELY LOW MULTILINEAR RANK TENSORS

The reconstruction formula of eq. (2.2) gives an exact reconstruction when the original tensor has multilinear rank- (R_1, \dots, R_N) but, it becomes unstable when the tensor deviates from the low multilinear rank model because the pseudo-inverse $\|\mathbf{W}_n^\dagger\|_2 = 1/\sigma_{R_n}$ can be extremely large given that σ_{R_n} (smallest singular value of $\mathbf{W}_{(n)}$) can be arbitrarily small. In order to solve this problem, we propose a modified reconstruction formula by stabilizing the pseudo-inverse. Given the SVD decomposition of a generic matrix $\mathbf{W} = \mathbf{U}\Sigma\mathbf{V}^T$, we define the modified pseudo-inverse $\mathbf{W}^* = \mathbf{V}\Sigma^*\mathbf{U}^T$ where the diagonal matrix Σ^* is defined by

$$\sigma_i^* = \begin{cases} 1/\sigma_i & \text{if } \sigma_i > \tau \\ 0 & \text{otherwise} \end{cases} \quad (3.1)$$

It is noted that $\mathbf{W}^* \rightarrow \mathbf{W}^\dagger$ as $\tau \rightarrow 0$ and $\|\mathbf{W}^*\|_2 \leq 1/\tau$. Also, the following properties are easily verified: I) $\mathbf{W}\mathbf{W}^*\mathbf{W} = \mathbf{W} + \mathbf{H}$ with $\|\mathbf{H}\|_2 \leq \tau$; II) $\mathbf{W}^*\mathbf{W}\mathbf{W}^* = \mathbf{W}^*$ and III) $\|\mathbf{W}\mathbf{W}^*\|_2 = \|\mathbf{W}^*\mathbf{W}\|_2 = 1$. Thus, the proposed modified reconstruction is the following:

$$\hat{\underline{\mathbf{X}}} = \underline{\mathbf{W}} \times_1 \mathbf{Z}_1 \mathbf{W}_{(1)}^* \cdots \times_N \mathbf{Z}_N \mathbf{W}_{(N)}^*, \quad (3.2)$$

The following theorem states an upper bound of the approximation error based on the modified reconstruction formula of eq. (3.2) in the matrix case ($N = 2$).

Theorem 3.1. *Let matrix $\mathbf{X} \in \mathbb{R}^{I \times I}$ be approximated by a rank- R matrix $\mathbf{X}_0 \in \mathbb{R}^{I \times I}$, i.e. $\mathbf{X} = \mathbf{X}_0 + \mathbf{E}$ where $\|\mathbf{E}\|_2 \leq \epsilon$. If sensing matrices $\Phi_1, \Phi_2 \in \mathbb{R}^{R \times I}$ are such that $\mathbf{W} = \Phi_1 \mathbf{X} \Phi_2^T$ is non-singular and assuming $\tau > \sigma_R$, where σ_R is the smallest singular value of matrix \mathbf{W} , then the following reconstruction error upper bound holds true:*

$$\|\mathbf{X} - \hat{\mathbf{X}}\|_2 \leq a\tau + b\epsilon + c\frac{\epsilon^2}{\tau}, \quad \text{with} \quad (3.3)$$

where constants a, b and c are defined below.

Proof. Let $\mathbf{X}_0 = \mathbf{U}_1 \Gamma \mathbf{U}_2^T$ be the rank- R SVD, for convenience we introduce a change of bases for columns and rows by defining $\mathbf{A}_n = \mathbf{U}_n (\Phi_n \mathbf{U}_n)^{-1} \in \mathbb{R}^{I \times R}$, and we obtain

$$\mathbf{X}_0 = \mathbf{A}_1 \mathbf{G} \mathbf{A}_2^T \quad \text{with } \mathbf{G} = \mathbf{W} - \Phi_1 \mathbf{E} \Phi_2^T, \quad (3.4)$$

where $\mathbf{W} = \Phi_1 \mathbf{X} \Phi_2^T$. Thus, the mode-1 measurement matrix is $\mathbf{Z}_1 = \mathbf{X}_{(1)} \Phi_2^T = \mathbf{X}_0 \Phi_2^T + \mathbf{E} \Phi_2^T = \mathbf{A}_1 \mathbf{W}_{(1)} + \mathbf{F}_1$,

with $\mathbf{F}_1 = (\mathbf{I} - \mathbf{A}_1 \Phi_1) \mathbf{E}_{(1)} \Phi_2^T$ (where the fact that $\Phi_n \mathbf{A}_n = \mathbf{I}$ was used). Using a similar analysis in mode-2, we obtain that $\mathbf{Z}_2 = \mathbf{A}_2 \mathbf{W}_{(2)} + \mathbf{F}_2$, with $\mathbf{F}_2 = (\mathbf{I} - \mathbf{A}_2 \Phi_2) \mathbf{E}_{(2)}^T \Phi_1^T$.

Now, we observe that the reconstruction is given by

$$\hat{\mathbf{X}} = \mathbf{Z}_1 \mathbf{W}^* \mathbf{W} (\mathbf{Z}_2 \mathbf{W}^{*T})^T = \mathbf{Z}_1 \mathbf{W}^* \mathbf{Z}_2^T, \quad (3.5)$$

where the properties of the modified pseudo-inverse matrix were used. By inserting the expressions of \mathbf{Z}_1 and \mathbf{Z}_2 into eq. (3.5) we arrive at:

$$\hat{\mathbf{X}} = \mathbf{A}_1 \mathbf{W} \mathbf{W}^* \mathbf{W} \mathbf{A}_2^T + \mathbf{A}_1 \mathbf{W} \mathbf{W}^* \mathbf{F}_2^T + \mathbf{F}_1 \mathbf{W}^* \mathbf{W} \mathbf{A}_2^T + \mathbf{F}_1 \mathbf{W}^* \mathbf{F}_2^T. \quad (3.6)$$

Using the fact that $\mathbf{W} \mathbf{W}^* \mathbf{W} = \mathbf{W} + \mathbf{H}$ with $\|\mathbf{H}\|_2 \leq \tau$ and eq. (3.4), we obtain:

$$\hat{\mathbf{X}} - \mathbf{X} = -\mathbf{E} + \mathbf{A}_1 \Phi_1 \mathbf{E} \Phi_2^T \mathbf{A}_2^T + \mathbf{A}_1 \mathbf{H} \mathbf{A}_2^T + \mathbf{A}_1 \mathbf{W} \mathbf{W}^* \mathbf{F}_2^T + \mathbf{F}_1 \mathbf{W}^* \mathbf{W} \mathbf{A}_2^T + \mathbf{F}_1 \mathbf{W}^* \mathbf{F}_2^T. \quad (3.7)$$

In order to find a bound of $\|\mathbf{X} - \hat{\mathbf{X}}\|_2$ we need to find bounds for the norms of each one of the terms in the last equation. Thus, we finally arrive at eq. (3.3), where constants are identified as follows:

$$a = \|\mathbf{A}_1\| \|\mathbf{A}_2\|, \quad (3.8)$$

$$b = 1 + \|\mathbf{A}_1 \Phi_1\| \|\mathbf{A}_2 \Phi_2\| + \|\mathbf{A}_1\| (1 + \|\mathbf{A}_2 \Phi_2\|) \|\Phi_1\| + \|\mathbf{A}_2\| (1 + \|\mathbf{A}_1 \Phi_1\|) \|\Phi_2\|, \quad (3.9)$$

$$c = (1 + \|\mathbf{A}_1 \Phi_1\|) (1 + \|\mathbf{A}_2 \Phi_2\|) \|\Phi_1\| \|\Phi_2\|. \quad (3.10)$$

□

Corollary 3.1. *The optimal selection of the parameter τ , for fixed ϵ and R , is $\tau_{opt} = \epsilon \sqrt{\frac{\epsilon}{a}}$. In this case, the error bound of eq. (3.3) becomes:*

$$\|\mathbf{X} - \hat{\mathbf{X}}\|_2 \leq \epsilon(b + 2\sqrt{ac}) = \epsilon K. \quad (3.11)$$

Proof. The parameter τ should be chosen in order to minimize the right hand term in eq. (3.3). By taking its derivative and setting it to zero we obtain the optimal value τ_{opt} and, by inserting it in eq. (3.3), we obtain the desired optimal bound (3.11). □

4. SIMULATION RESULTS

We have applied the reconstruction formula of eq. (3.2) to different 2D and 3D data sets. Gaussian compressive linear measurements taken in modes 1 and 2. i.e. $\mathbf{Y}_1 = \Phi_1 \mathbf{X}_{(1)}$, $\mathbf{Y}_2 = \Phi_2 \mathbf{X}_{(2)}$ ($\Phi_n \in \mathbb{R}^{R \times I}$, $n = 1, 2$) were used to compute the corresponding multi-way measurements defined in eqs. (2.1). In order to compute the modified pseudo-inverse of matrices $\mathbf{W}_{(n)}$, the following heuristic value $\tau = \|\mathbf{W}_{(n)}\|_2 / R_n$ was chosen. It can be shown that the sampling ratio, i.e. the size of non-redundant available measurements divided the size of the

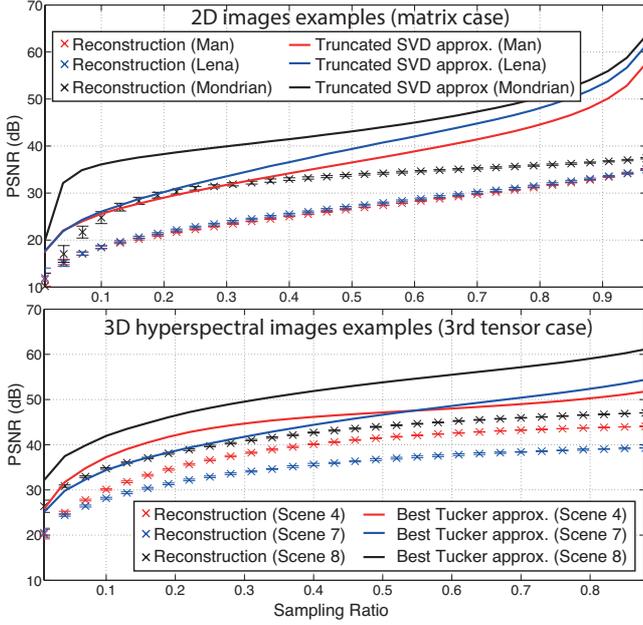


Fig. 2. Performance of reconstructions for 2D (top) and 3D (bottom) data sets. Mean value and standard deviation over 100 Monte Carlo simulations are shown. PSNRs associated with the best low rank (truncated SVD) and the best multilinear rank (obtained through the HOOI algorithm) approximations are also shown for reference.



Fig. 3. Reconstruction of a 2D data set example: Kronecker-CS solution based on measurements $\mathbf{W} = \Phi_1 \mathbf{X} \Phi_2^T$ assuming a Kronecker Wavelet dictionary (middle) and our Low-Rank approximation based on measurements $\mathbf{Y}_1 = \Phi_1 \mathbf{X}$ and $\mathbf{Y}_2 = \Phi_2 \mathbf{X}^T$ (right). In both methods $R = I/2$ was considered.

original signal is $2\frac{R}{T} - (\frac{R}{T})^2$. In Fig. 2 (top), the PSNR (Peak Signal to Noise Ratio) versus the sampling ratio is shown for for three different 2D images: “Man” (1024×1024), “Lena” (512×512) and “Mondrian” (512×512); and three 3D tensor datasets corresponding to hyperspectral images of natural scenes ($1024 \times 1024 \times 32$) [18] where $\Phi_3 = \mathbf{I}$. As a reference, we show the PSNR obtained by using the best low-rank approximation (2D case) and the best multilinear rank approximation (3D case). The latter was estimated by using the Higher Order Orthogonal Iteration (HOOI) algorithm implemented in [19]. It is interesting to note that, our method not only provide stable reconstructions but also very robust (very small standard deviation among realizations using different Gaussian matrices). In Fig. 3, we use $R = I/2$ and compare the reconstructions provided by our method (right)

against the Kronecker-CS solution (middle), based on measurements given by $\mathbf{W} = \Phi_1 \mathbf{X} \Phi_2^T$ and using a Basis Pursuit method implemented (SPGL1 algorithm) as proposed in [9, 10]. It is clear that our method obtains much better results with minimal computation effort. In this case, the sampling ratio of the Kronecker-CS method 25%, on the other hand, with our method we are able to use more effectively the available measurements corresponding to a sampling ratio of 75%, for the same value of R .

5. CONCLUSIONS

We have provided a new reconstruction formula that exploits the low multilinear rank structure of tensors assuming that linear projections of its mode- n fibers are at our disposal in, at least, two selected modes. Compared to the Kronecker-CS technique available for 2D images [8, 9] and tensors [10, 11, 12], the present method has the following remarkable advantages: I) It is very fast because it does not involve iterations which makes it suitable for large-scale problems; II) It is stable, in the sense that tensors which are well approximated by its associated low rank model are also well reconstructed by the proposed method, and III) It is robust, in the sense that the performance of the reconstruction does not depend on the actual Gaussian sensing matrices.

6. APPENDIX

Proof of Theorem 2.1. Let us consider the exact HOSVD decomposition $\underline{\mathbf{X}} = \underline{\mathbf{G}} \times_1 \mathbf{U}_1 \times_2 \cdots \times_N \mathbf{U}_N$, with core tensor $\underline{\mathbf{G}} \in \mathbb{R}^{R_1 \times \cdots \times R_N}$ and orthogonal factors $\mathbf{U}_n \in \mathbb{R}^{I_n \times R_n}$, which exists because tensor $\underline{\mathbf{X}}$ has multilinear rank- (R_1, \dots, R_N) . We consider, for convenience, a change of bases by assigning a new set of factors given by $\mathbf{A}_n = \mathbf{U}_n (\Phi_n \mathbf{U}_n)^{-1}$, thus, we have $\underline{\mathbf{X}} = \underline{\mathbf{G}} \times_1 \mathbf{A}_1 \times_2 \cdots \times_N \mathbf{A}_N$ where $\underline{\mathbf{G}} = \underline{\mathbf{X}} \times_1 \mathbf{A}_1^\dagger \times_2 \cdots \times_N \mathbf{A}_N^\dagger = \underline{\mathbf{W}}$. Note also that, the multi-way measurements are now simplified to $\underline{\mathbf{Z}}_n = \underline{\mathbf{W}} \times_n \mathbf{A}_n$ or, equivalently, $\mathbf{Z}_n = \mathbf{A}_n \mathbf{W}_{(n)}$. Taking into account that $\mathbf{W}_{(1)} \mathbf{W}_{(1)}^\dagger \mathbf{W}_{(1)} = \mathbf{W}_{(1)}$, the mode-1 version of eq. (2.2) is:

$$\hat{\mathbf{X}}_{(1)} = \mathbf{A}_1 \mathbf{W}_{(1)} \left(\mathbf{Z}_N \mathbf{W}_{(N)}^\dagger \otimes \cdots \otimes \mathbf{Z}_2 \mathbf{W}_{(2)}^\dagger \right)^T, \quad (6.1)$$

whose mode-2 unfolding matrix version becomes:

$$\hat{\mathbf{X}}_{(2)} = \mathbf{Z}_2 \mathbf{W}_{(2)}^\dagger \mathbf{W}_{(2)} \left(\mathbf{Z}_N \mathbf{W}_{(N)}^\dagger \otimes \cdots \otimes \mathbf{Z}_3 \mathbf{W}_{(3)}^\dagger \otimes \mathbf{A}_1 \right)^T.$$

Now, by substituting $\mathbf{Z}_2 = \mathbf{A}_2 \mathbf{W}_{(2)}$ in the previous equation and by repeating this process for the rest of modes $n = 3, 4, \dots, N$, we finally arrive at:

$$\hat{\mathbf{X}}_{(N)} = \mathbf{A}_N \mathbf{W}_{(N)} (\mathbf{A}_{N-1} \otimes \cdots \otimes \mathbf{A}_1)^T, \quad (6.2)$$

which proves that $\hat{\mathbf{X}} = \underline{\mathbf{X}}$.

7. REFERENCES

- [1] F Marvasti, A Amini, F Haddadi, M Soltanolkotabi, B Khalaj, A Aldroubi, S Sanei, and J Chambers, "A Unified Approach to Sparse Signal Processing," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, pp. 44, 2012.
- [2] EJ Candes and Y Plan, "Matrix Completion with Noise," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 925–936, 2010.
- [3] D Zachariah, M Sundin, M Jansson, and S Chatterjee, "Alternating Least-Squares for Low-Rank Matrix Reconstruction," *IEEE Signal Processing Letters*, vol. 19, no. 4, pp. 231–234, Apr. 2012.
- [4] E Acar, DM Dunlavy, TG Kolda, and M Mørup, "Scalable tensor factorizations for incomplete data," *Chemometrics and Intelligent Laboratory Systems*, vol. 106, no. 1, pp. 41–56, Mar. 2011.
- [5] J Liu, P Musialski, P Wonka, and J Ye, "Tensor Completion for Estimating Missing Values in Visual Data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 208–220, Jan. 2012.
- [6] H Rauhut and R Schneider, "Low-rank Tensor Recovery via Iterative Hard Thresholding," *Proc. SAMPTA 2013*, June 2013.
- [7] HF Alqadah and H Fan, "Low rank variational tensor recovery for multi-linear inverse problems," in *Signals, Systems and Computers (ASILOMAR), 2011 Conference Record of the Forty Fifth Asilomar Conference on*. 2011, pp. 1438–1442, IEEE.
- [8] R Robucci, JD Gray, LK Chiu, J Romberg, and P Hasler, "Compressive Sensing on a CMOS Separable-Transform Image Sensor," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1089–1101, May 2010.
- [9] Y Rivenson and A Stern, "Compressed Imaging with a Separable Sensing Operator," *Signal Processing Letters, IEEE*, vol. 16, no. 6, pp. 449–452, 2009.
- [10] MF Duarte and RG Baraniuk, "Kronecker Compressive Sensing," *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 494–504, Feb. 2012.
- [11] CF Caiafa and A Cichocki, "Block Sparse Representations of Tensors Using Kronecker Bases," *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pp. 2709–2712, 2012.
- [12] CF Caiafa and A Cichocki, "Computing Sparse Representations of Multidimensional Signals Using Kronecker Bases," *Neural Computation*, pp. 186–220, Dec. 2012.
- [13] CF Caiafa and A Cichocki, "Multidimensional compressed sensing and their applications," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2013 (in press).
- [14] ND Sidiropoulos and A Kyrillidis, "Multi-Way Compressed Sensing for Sparse Low-Rank Tensors," *IEEE Signal Processing Letters*, vol. 19, no. 11, pp. 757–760, Nov. 2012.
- [15] TG Kolda and BW Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.
- [16] L De Lathauwer, B De Moor, and J Vandewalle, "A Multilinear Singular Value Decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, pp. 1253–1278, 2000.
- [17] L Grasedyck, D Kressner, and C Tobler, "A Literature Survey of Low-rank Tensor Approximation Techniques," *GAMM-Mitteilungen*, vol. 36, no. 1, pp. 53–78, 2013.
- [18] DH Foster, SMC Nascimento, and K Amano, "Information limits on neural identification of colored surfaces in natural scenes," *Visual Neuroscience*, vol. 21, May 2004.
- [19] BW Bader, TG Kolda, "Matlab Tensor Toolbox version 2.2," available from www.sandia.gov/~tgkolda/TensorToolbox/
- [20] Q Li, D Schonfeld and S Friedland, "Generalized tensor compressive sensing," *Multimedia and Expo (ICME)*, 2013.