Tensor Versus Matrix Completion: A Comparison With Application to Spectral Data

Marco Signoretto, Raf Van de Plas, Bart De Moor, and Johan A. K. Suykens

***Abstract*—Tensor completion recently emerged as a generalization of matrix completion for higher order arrays. This problem formulation allows one to exploit the structure of data that intrinsically have multiple dimensions. In this work, we recall a convex formulation for minimum (multilinear) ranks completion of arrays of arbitrary order. Successively we focus on completion of partially observed spectral images; the latter can be naturally represented as third order tensors and typically exhibit intraband correlations. We compare different convex formulations and assess them through case studies.**

***Index Terms*—Hyperspectral imaging, image reconstruction, matrix completion, tensor completion.**

I. Introduction

MATRIX completion refers to the process of inferring missing entries from a partially specified table of numbers. The values could be missing for example because of problems in the acquisition process, or because the user manually identified what are believed to be outliers. The problem has been studied from an algebraic perspective [8] and recently witnessed a renewed interest, especially in relation to recommendation systems. In this domain, vendors want to infer the values of the missing entries on a product/customer grid of ratings. This task is ill-posed since there are infinitely many ways to fill in the missing values. It is then customary to exploit the common belief that users behavior is dictated by a limited number of common factors. This is translated into a low-rank assumption and makes the inference process feasible. In the recent literature the extension of this task to the case of arrays of arbitrary order, namely *tensors*, was studied [6], [10], [12]. Tensor representations naturally arise in many domains such as neuroscience (where one has, e.g., $\text{time} \times \text{frequency} \times \text{electrodes}$) and chemistry (e.g., $\text{samples} \times \text{emission} \times \text{excitation}$). In this letter, we focus on hyperspectral data that consist of collections of measurements sensed from contiguous spectral bands. Due to the specific nature of these data, the information content of the resulting third order array is expected to be redundant. It is then natural to ask whether treating the tensor as a unique object, hereby exploiting this property, can improve the estimation of missing entries with respect to alternative approaches in the literature. Although experimental results can be found elsewhere [6], [12], [10] the existing literature does not make an explicit comparison between different completion techniques. Our contribution in this research note is to discuss alternative approaches and to provide experimental evidence that partially fills this gap. In the next section, we briefly recall the necessary preliminaries on tensors. We then formalize the problem of completion for both matrices and higher order arrays in Section III. Successively we present case studies in Section IV and conclude by drawing our final remarks in Section V.

Manuscript received March 04, 2011; revised April 21, 2011; accepted April 26, 2011. Date of publication May 05, 2011; date of current version May 19, 2011. This work was supported by Research Council KUL: ProMeta, CoE EF/05/007 SymBioSys, GOA/10/09 MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC) en PFV/10/002 (OPTEC), IOF-SCORES4CHEM, several Ph.D./postdoc and fellow grants; Flemish Government: FWO: Ph.D./postdoc grants, G0321.06 (Tensors), G.0302.07 (SVM/Kernel), G.0558.08 (Robust MHE). IWT: Ph.D. Grants, SBO LeCoPro, O&O-Dsquare. Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO, Dynamical systems, control and optimization, 2007–2011). IBBT. EU: ERNSI.

M. Signoretto, B. De Moor, and J. A. K. Suykens are with the Department of Electrical Engineering, Katholieke Universiteit Leuven, B-3001 Leuven (Heverlee), Belgium (e-mail: marco.signoretto@esat.kuleuven.be; bart.demoor@esat.kuleuven.be; johan.suykens@esat.kuleuven.be).

R. Van de Plas is with the Vanderbilt University School of Medicine, Mass Spectrometry Research Center, Department of Biochemistry, Vanderbilt University, Nashville, TN 37232 USA (e-mail: raf.vandeplas@vanderbilt.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2011.2151856

II. Preliminaries on Tensors

We denote scalars by lower-case letters $(a, b, c, \ldots)$, vectors as capitals $(A, B, C, \ldots)$ and matrices as bold-face capitals $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \ldots)$. In particular, we denote by $\boldsymbol{I}_N$ the $N \times N$ identity matrix. We also use lower-case letters i, j in the meaning of indices and, with some abuse of notation, we will use I, J to denote the index upper bounds. Additionally, we write $\mathbb{N}_I$ to denote the set. We recall that Nth order tensors, which we denote by calligraphic letters $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots)$, are higher order generalizations of vectors (first order tensors) and matrices (second order tensors). We write $a_{i_1,\ldots,i_N}$ to denote $(\mathcal{A})_{i_1,\ldots,i_N}$. An *$n$-mode vector* of $\mathcal{A}$ is an element of $\mathbb{R}^{I_n}$ obtained from $\mathcal{A}$ by varying the index i_n and keeping the other indices fixed. The *n-rank* of $\mathcal{A}$, indicated by $\text{rank}_n(\mathcal{A})$, is the dimension of the space spanned by the n-mode vectors. Equivalently, the n-rank can be defined as the rank of the matrix that we define next.

Definition II.1 (n-Mode Matricization): Assume an Nth order tensor $\mathcal{A} \in \mathbb{R}^{I_1} \otimes \cdots \otimes \mathbb{R}^{I_N}$ and let $J = \prod_{p \in \mathbb{N}_N \setminus \{n\}} I_p$. The nth mode matrix unfolding, denoted as $\mathcal{A}_{\langle n \rangle}$, is then the matrix $\mathcal{A}_{\langle n \rangle} \in \mathbb{R}^{I_n} \otimes \mathbb{R}^{J}$, whose columns are the n-mode vectors.

Notice that the choice of the ordering of the columns of $\mathcal{A}_{\langle n \rangle}$ does not matter for practical purposes. It is enough that one sticks to the same rule to arrange the n-mode vectors as columns of the nth mode matrix unfolding.

III. Convex Completion of Matrices and Higher Order Tensors

Tensor completion amounts to recovering a multidimensional array $\mathcal{X} \in \mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \otimes \cdots \otimes \mathbb{R}^{I_N}$ from a sampling of its entries.

1070-9908/$26.00 © 2011 IEEE

More precisely, one has a set of measurements $Y \in \mathbb{R}^J$ that arise from the evaluation of $\mathcal{X}$ at a subset of $J < I_1 I_2 \cdots I_N$ indices

$$\mathfrak{O} = \left\{ \left(i_1^j, \ldots, i_N^j \right) \in \mathbb{N}_{I_1} \times \cdots \times \mathbb{N}_{I_N} : j \in \mathbb{N}_J \right\} \quad (1)$$

and want to infer the values of the remaining entries. Denote by $A_{\mathfrak{O}} : \mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \otimes \cdots \otimes \mathbb{R}^{I_N} \to \mathbb{R}^J$ the *sampling operator* defined by $(A_{\mathfrak{O}}(\mathcal{X}))_j := x_{i_1^j i_2^j \ldots i_N^j}$ for any $j \in \mathbb{N}_J$. The fact that there are infinitely many solutions to $A_{\mathfrak{O}}(\mathcal{X}) = Y$ shows that the completion is an ill-posed task.

A. Matrix Completion

When $N = 2$, namely for the case where $\mathcal{X}$ is actually a matrix $\boldsymbol{X}$, a natural approach to deal with such ill-posedness consists of taking the particular solution of $A_{\mathfrak{O}}(\boldsymbol{X}) = Y$ that has minimal rank:

$$\min\{\operatorname{rank}(\boldsymbol{X}) : A_{\mathfrak{O}}(\boldsymbol{X}) = Y, \boldsymbol{X} \in \mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2}\}. \quad (2)$$

A popular convex relaxation to this NP-hard problem arises from the definition of *nuclear norm* of $\boldsymbol{X}$:

$$\|\boldsymbol{X}\|_* := \left\{ \sum_{r \in \mathbb{N}_R} \sigma_r : \boldsymbol{X} = \sum_{r \in \mathbb{N}_R} \sigma_r U_r V_r^\top, \; \boldsymbol{U}^\top \boldsymbol{U} = I_R, \boldsymbol{V}^\top \boldsymbol{V} = I_R \right\}. \quad (3)$$

In fact, the nuclear norm can be shown to be the *convex envelope* of the rank function [4]. In turn, this makes the optimization problem $\min\{\|\boldsymbol{X}\|_* : A_{\mathfrak{O}}(\boldsymbol{X}) = Y, \boldsymbol{X} \in \mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2}\}$ the best possible convex relaxation of (2).

B. Generalization to Higher Order Tensors

In order to extend the approach to a tensor $\mathcal{X}$ of arbitrary order N, consider the function

$$g(\mathcal{X}) := \frac{1}{N} \sum_{n \in \mathbb{N}_N} \left\| \mathcal{X}_{\langle n \rangle} \right\|_*. \quad (4)$$

It can be shown that g is a well-defined norm on $\mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \otimes \cdots \otimes \mathbb{R}^{I_N}$ [12]. Hence, we write $\|\mathcal{X}\|_* := g(\mathcal{X})$ and, by extension, we call the latter *nuclear norm*. The concept of nuclear norm for tensors appeared for the first time in [10]. Its use as a penalty function to learn tensor-based models was successively thoroughly studied in [12]. The case of tensor completion arises as a special case of a more general problem class and can be dealt with by means of the optimization problem:

$$\min\{\|\mathcal{X}\|_* : A_{\mathfrak{O}}(\mathcal{X}) = Y, \mathcal{X} \in \mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \otimes \cdots \otimes \mathbb{R}^{I_N}\}. \quad (5)$$

It is easy to see that the approach is ultimately linked to the notion of n-ranks (Section II). In turn the latter are closely connected with the Tucker decomposition [16], also known as multilinear singular value decomposition (MLSVD) [3]. The interested reader is referred to [12] for an in-depth discussion. In general, it is worth remarking that since there is no unique way to generalize the concept or rank from matrices to higher order tensors, other meaningful approaches might be formulated. Finally, we remark that (5) is a convex problem for which different solution strategies can be devised [6], [10], [12], [15].

IV. The Case of Hyperspectral Data

Spectral data consist of collections of measurements sensed from contiguous spectral bands. We call *frame* the measurements corresponding to a single band that are organized within a matrix of size $I_1 \times I_2$. The complete dataset results from arranging the I_3 frames along the third dimension of a tensor $\mathcal{X}$ which, as a result, has dimensions $I_1 \times I_2 \times I_3$. Hyperspectral Remote Sensing (HRS) [2] is used for the detection and identification of minerals, terrestrial vegetation, and man-made materials. By contrast, Mass Spectral Imaging (MSI) studies the distribution in tissues of biochemical compounds such as metabolites, peptides, or proteins by their molecular masses [9], [13].

A. Completion and Decay of Spectra

A common approach when dealing with an ensemble of images amounts at vectorizing each image [17]. For the case of hyperspectral data, this corresponds to regarding the $I_1 \times I_2 \times I_3$ data tensor as the third $I_3 \times I_1 I_2$ matrix unfolding $\mathcal{X}_{\langle 3 \rangle}$. In practice, intensities measured at nearby wavelengths are normally highly correlated. It is then expected that each row of $\mathcal{X}_{\langle 3 \rangle}$ can be approximately restated as a linear combination of a small number of orthogonal vectors that play the role of *eigenimages*. This corresponds to a low rank assumption and ultimately motivates the completion of $\mathcal{X}_{\langle 3 \rangle}$ based upon the nuclear norm (3). For tensors, the idea of relying on the third mode unfolding to perform completion is found in [14]. However, we note that—especially when the number of spectral bands is limited—it might be not possible to interpolate the value of missing entries at a certain frequency based upon the intensity on the same locations in the other frequencies. In these cases, relying solely on spectral redundancy might not be ideal.

A possible workaround is based on the realization that the aforementioned approach discards the matrix representation of each frame. Different algorithms for images, such as [7], successfully rely on this additional structural information. For a given frame, the well-known Eckart–Young–Mirsky theorem states that the best low rank approximation in least squares sense is obtained via its truncated SVD decomposition. In particular, it is often the case that a good approximation is achieved by retaining a small number of singular values/vectors. Matrix completion performed independently on each frame exploits this prior belief only. However, since the entire data tensor is available, it makes sense to exploit this property as well as the redundancy along the spectral mode. It is reasonable to expect that the spectrum of the different matrix unfoldings has a fast decay (Fig. 1). Since minimizing the nuclear norm (4) is a proxy for the minimization of the multilinear ranks, tensor completion is seen to leverage this composite assumption.

B. Experimental Protocol

Next we assess these different approaches by experiments. For each trial, we draw uniformly at random an index set $\mathfrak{O}$ [see (1)] containing a prescribed fraction of the total number of entries of the data tensor. The estimation error incurred by each

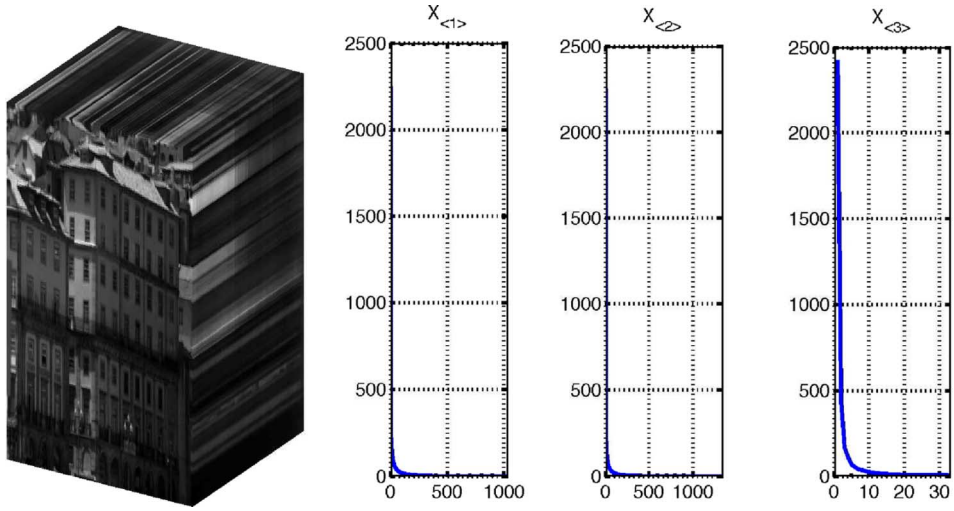


Fig. 1. Hyperspectral image represented as a $1017 \times 1340 \times 33$ tensor $\mathcal{X}$ along with the spectrum of the corresponding matrix unfoldings.

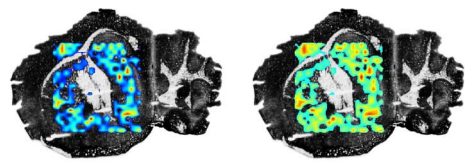
procedure is measured in terms of the *normalized root mean squared error* (NRMSE) function. Denote by $\mathfrak{O}^c$ the complement of the index set $\mathfrak{O}$ for a given experiment. If we now denote by J^c its cardinality and let $A_{\mathfrak{O}^c} : \mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \otimes \cdots \otimes \mathbb{R}^{I_N} \rightarrow \mathbb{R}^{J_c}$ be the corresponding sampling operator, we have

$$\mathrm{NRMSE}(\mathcal{X}, \hat{\mathcal{X}}) := \frac{\|A_{\mathfrak{O}^c}(\mathcal{X}) - A_{\mathfrak{O}^c}(\hat{\mathcal{X}})\|}{(\max(A_{\mathfrak{O}^c}(\mathcal{X})) - \min(A_{\mathfrak{O}^c}(\mathcal{X})))\sqrt{J^c}}. \tag{6}$$

In the following, we consider the cases where the approximation $\hat{\mathcal{X}}$ is obtained by *a*) tensor completion (indicated as *tensor* in the tables), *b*) combining the solutions of frame completion problems (indicated as *frame*), and *c*) third mode matrix completion (indicated as *mode-3*). In practice, to solve all the corresponding optimization problems we considered a MATLAB implementation of the *equality-constrained convex multilinear estimation* (E-CMLE) algorithm introduced in [12, Sect. 6.7 and Alg. 4], where we set $\beta = 1.1, \gamma_u = 2, \gamma_d = 1.1, c^{(0)} = 100, \mu = 1$ and $d^{(0)} = 1$. Notice that, since matrix completion is a special case of tensor completion, the same algorithm can be used to deal with the three cases (with certain simplification for the two matrix problems). Similar solution strategies are found in [6] and [10]. Note that, as E-CMLE, they are based on extending the *soft-thresholding operator* used for the second order case [1].

C. Mass Spectral Imaging

As a first case study, we considered MALDI mass spectral measurements that were performed on a sagittal section of the brain of a mouse at the University Hospital Leuven. Details about the experimental setting can be found in [18], [11]. Here we focus on a grid consisting of 22×22 measurement locations (pixels) situated at the center of the tissue section. For each location N equispaced mass-over-charge (m/z) bins within the range from 2500 to about 25 000 Dalton were considered to form the $22 \times 22 \times N$ MSI data tensor. Fig. 2 reports the intensity plots for two different values of m/z overlaid on the registered sectioned tissue. The (normalized) intensity of each pixel represents the number of molecules (or ions) collected for the specified m/z bin. In order to assess the three completion procedures, entries were artificially removed from the tensor data. The NRMSE results are reported in Table I together with running times on a standard desktop machine for increasing values of the fraction of sampled entries and two different values of N.



(a) (b)

Fig. 2. Intensity plots corresponding to two different frames. The square box of 22×22 measurement locations (pixels) is overlaid on the registered tissue section. (a) $\mathrm{m/z} = 2.80\mathrm{e}3$; (b) $\mathrm{m/z} = 3.03\mathrm{e}3$.

TABLE I
MSI CASE STUDY—NRMSE AND RUNNING TIMES

N=50 — fraction of entries

	0.1		0.3		0.5		0.7	
	NRMSE	time (s)	NRMSE	time (s)	NRMSE	time (s)	NRMSE	time (s)
frame	0.123	4.70e3	0.110	4.42e3	0.108	4.99e3	0.106	6.88e3
tensor	0.129	0.13e3	0.071	0.37e3	0.052	0.43e3	0.036	0.90e3
mode-3	**0.058**	0.01e3	**0.021**	0.01e3	**0.016**	0.09e3	**0.014**	0.02e3

N=4 — fraction of entries

	0.1		0.3		0.5		0.7	
	NRMSE	time (s)	NRMSE	time (s)	NRMSE	time (s)	NRMSE	time (s)
frame	0.172	0.02e3	0.147	0.02e3	0.149	0.02e3	0.148	0.03e3
tensor	**0.171**	0.05e3	**0.137**	0.07e3	**0.131**	0.07e3	**0.125**	0.07e3
mode-3	0.180	0.02e2	0.160	0.02e2	0.149	0.01e2	0.137	0.01e3

D. Hyperspectral Images of Natural Scenes

This example considers natural images that were previously used in [5]. In a first part, we focused on scene 6 (Braga) and 7 (Ribeira)[1]. Details on the hyperspectral system and processing are given at the referred webpage and in [5]. Each scene originally consisted of 33 frames obtained varying the peak-transmission wavelength from 400 nm to 720 nm. We down-sampled each frame independently to 203×268 for Ribeira and 204×267 for Braga. Successively, we performed completion for increasing fraction of missing entries. On the one hand, we

[1]The Matlab reflectance files for these scenes can be downloaded from http://personalpages.manchester.ac.uk/staff/david.foster/Hyperspectral_images_of_natural_scenes_04.html.

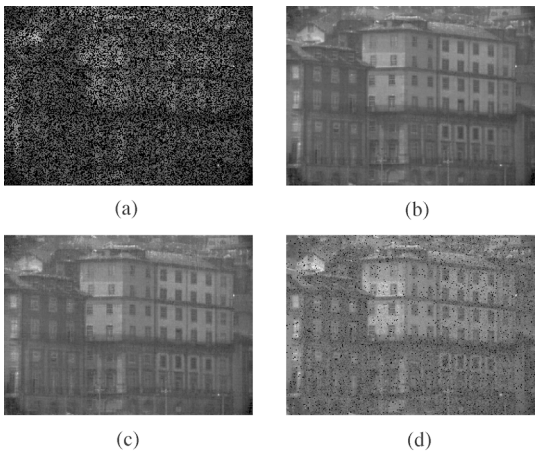


Fig. 3. Completion of the first frame in the ensemble given by the first 5 wavelengths of Ribeira. Approximately 50% of entries are missing. (a) Data—Ribeira @ 400 nm, (b) Tensor, (c) Frame, and (d) mode-3.

TABLE II
FIRST PART OF THE HRS CASE STUDY—NRMSE AND RUNNING TIMES

Ribeira (33 frames) — fraction of entries

	0.1		0.3		0.5		0.7	
	NRMSE	time (s)	NRMSE	time (s)	NRMSE	time (s)	NRMSE	time (s)
frame	0.092	3.78e3	0.061	3.72e3	0.0455	2.30e3	0.035	3.28e3
tensor	0.072	2.63e4	0.031	2.58e4	0.020	4.20e4	0.013	4.61e4
mode-3	**0.068**	0.27e3	**0.018**	0.31e3	**0.012**	0.325e3	**0.009**	0.42e3

Ribeira (5 frames) — fraction of entries

	0.1		0.3		0.5		0.7	
	NRMSE	time (s)	NRMSE	time (s)	NRMSE	time (s)	NRMSE	time (s)
frame	0.071	0.15e3	0.046	0.14e3	0.034	0.14e3	0.026	0.20e3
tensor	**0.067**	3.14e3	**0.034**	4.48e3	**0.023**	4.06e3	**0.016**	4.64e3
mode-3	0.191	0.02e3	0.119	0.02e3	0.070	0.02e3	0.036	0.02e3

Braga (33 frames) — fraction of entries

	0.1		0.3		0.5		0.7	
	NRMSE	time (s)	NRMSE	time (s)	NRMSE	time (s)	NRMSE	time (s)
frame	0.093	5.45e3	0.064	3.82e3	0.052	3.64e3	0.040	2.23e4
tensor	0.069	4.26e4	0.033	7.55e4	0.022	7.62e4	0.013	8.00e4
mode-3	**0.05**	0.35e3	**0.014**	0.41e3	**0.010**	0.45e3	**0.008**	0.493e3

Braga (5 frames) — fraction of entries

	0.1		0.3		0.5		0.7	
	NRMSE	time (s)	NRMSE	time (s)	NRMSE	time (s)	NRMSE	time (s)
frame	0.091	0.20e3	0.066	0.20e3	0.051	0.20e3	0.040	0.16e3
tensor	**0.083**	4.01e3	**0.048**	7.10e3	**0.033**	7.15e3	**0.023**	5.95e3
mode-3	0.187	0.02e3	0.114	0.02e3	0.062	0.03e3	0.031	0.03e3

TABLE III
MIXTURE OF IMAGES—NRMSE AND RUNNING TIMES

fraction of entries

	0.1		0.3		0.5		0.7	
	NRMSE	time (s)	NRMSE	time (s)	NRMSE	time (s)	NRMSE	time (s)
frame	**0.038**	0.16e3	**0.029**	0.11e3	**0.023**	0.18e3	**0.025**	0.17e3
tensor	0.046	2.86e3	0.031	2.71e3	0.024	2.67e3	0.026	2.64e3
mode-3	0.091	0.01e3	0.077	0.01e3	0.067	0.02e3	0.077	0.02e3

considered the data tensor consisting of the whole set of frames. On the other, we used as data tensor the ensemble consisting of the first five wavelengths only. This case is considered to simulate the situation where the information along the spectral mode is limited. With reference to this second case, Fig. 3 shows the plot of a data frame (in black the missing entries) and the corresponding reconstructed images. The performance of the different completion techniques is reported in Table II.

As a second experiment, we looked into the case where the spectral mode fails to satisfy the redundancy assumption. Specifically, we considered the whole set of eight spectral images available at the referred webpage. The first frame of each image was down-sampled to 102×134 and stacked along the third mode of a new data tensor. The results for this case are reported in Table III.

V. CONCLUSION

Contrary to alternative approaches, tensor completion exploits redundancy along all the modes of spectral data. Our experiments (Tables I and II) show that, when the number of spectral bands is limited, leveraging the structure of all the modes improves the imputation of missing entries. By contrast, if the number of bands is sufficiently high, the experiments show that working in the third mode only is more advantageous. Finally, when no correlation along the spectral mode is present relying solely on low ranks assumptions of the frames performs best (Table III).

ACKNOWLEDGMENT

The MSI data set was kindly provided by Prof. E. Waelkens, Department of Molecular Cell Biology, KULeuven.

REFERENCES

[1] J. F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, 2008.

[2] C. I. Chang, *Hyperspectral Data Exploitation: Theory and Applications*. Chichester, U.K.: Wiley-Blackwell, 2007.

[3] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Applicat.*, vol. 21, no. 4, pp. 1253–1278, 2000.

[4] M. Fazel, "Matrix Rank Minimization With Applications," Ph.D. Thesis, Elect. Eng. Dept., Stanford Univ., Stanford, CA, 2002.

[5] D. H. Foster, S. M. C. Nascimento, and K. Amano, "Information limits on neural identification of colored surfaces in natural scenes," *Visual Neurosci.*, vol. 21, no. 03, pp. 331–336, 2004.

[6] S. Gandy, B. Recht, and I. Yamada, "Tensor completion and low-n-rank tensor recovery via convex optimization," *Inv. Probl.*, vol. 27, no. 2, p. 19, 2011, doi:10.1088/0266-5611/27/2/025010.

[7] X. He, D. Cai, and P. Niyogi, "Tensor subspace analysis," *Adv. Neural Inf. Process. Syst.*, pp. 499–506, 2006.

[8] C. R. Johnson, "Matrix completion problems: A survey," in *Matrix Theory and Applications*. Providence, RI: AMS, 1990, vol. 40, pp. 171–198.

[9] A. J. Kaufman and S. Xiao, "High CO_2 levels in the proterozoic atmosphere estimated from analyses of individual microfossils," *Nature*, vol. 425, no. 6955, pp. 279–282, 2003.

[10] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," in *IEEE Int. Conf. Computer Vision (ICCV)*, Kyoto, Japan, 2009, pp. 2114–2121.

[11] F. Ojeda, M. Signoretto, R. Van de Plas, E. Waelkens, B. De Moor, and J. A. K. Suykens, "Semi-supervised learning of sparse linear models in mass spectral imaging," in *Pattern Recognition in Bioinformatics (PRIB)*, Nijgmegen, The Netherlands, 2010, pp. 325–334.

[12] M. Signoretto, L. De Lathauwer, and J. A. K. Suykens, Nuclear Norms for Tensors and Their Use for Convex Multilinear Estimation, Leuven, Belgium, Int. Rep. 10-186, ESAT-SISTA, K. U. Leuven, Lirias number: 270741, 2010.

[13] I. M. Taban, A. F. Altelaar, Y. E. M. van der Burgt, L. A. McDonnell, R. Heeren, J. Fuchser, and G. Baykut, "Imaging of peptides in the rat brain using MALDI-FTICR mass spectrometry," *J. Amer. Soc. Mass Spectrometry*, vol. 18, no. 1, pp. 145–151, 2007.

[14] R. Tomioka, K. Hayashi, and H. Kashima, "On the extension of trace norm to tensors," in *NIPS Workshop on Tensors, Kernels, and Machine Learning*, 2010.

[15] R. Tomioka, K. Hayashi, and H. Kashima, "Estimation of low-rank tensors via convex optimization," Arxiv Preprint arXiv:1010.0789 2011.

[16] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966.

[17] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cogn. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.

[18] R. Van de Plas, K. Pelckmans, B. De Moor, and E. Waelkens, "Spatial querying of imaging mass spectrometry data: A nonnegative least squares approach," in *Neural Information Processing Systems Workshop on Machine Learning in Computational Biology*, 2007.