QUANTIZED TENSOR ROBUST PRINCIPAL COMPONENT ANALYSIS

Anastasia Aidini [1,2]*, Grigorios Tsagkatakis* [1,2]*, Panagiotis Tsakalides* [1,2]

[1] Institute of Computer Science, Foundation for Research and Technology-Hellas (FORTH)
[2] Department of Computer Science, University of Crete
Heraklion, Crete, Greece, GR-70013

ABSTRACT

High-dimensional data structures, known as tensors, are fundamental in many applications, including multispectral imaging and color video processing. Compression of such huge amount of multidimensional data collected over time is of paramount importance, necessitating the process of quantization of measurements into discrete values. Furthermore, noise and issues related to the acquisition and transmission of signals frequently lead to unobserved, lost or corrupted measurements. In this paper, we introduce a tensor robust principal component analysis algorithm in order to recover a tensor with real-valued entries from a partly observed set of quantized and sparsely corrupted entries. We formulate the problem as a constrained maximum likelihood estimation of the sum of a low-rank tensor and a sparse tensor, through matricizations in each mode, in combination with a quantization and statistical measurement model. Experimental results on satellite derived land surface time-series demonstrate that directly operating with the quantized measurements, rather than treating them as real values, results in a low recovery error, while the proposed method is also capable of detecting temperature anomalies (e.g., forest fires).

***Index Terms*—** Tensors, Robust PCA, Quantization, Missing values, Image time-series

1. INTRODUCTION

The continuous and excessive growth of acquired observations introduces considerable challenges in terms of data storage and transfer. Thus, it is of paramount importance that multi-dimensional observations, encoded into tensor structures, are efficiently compressed. At the core of essentially all lossy compression algorithms is the process of mapping a range of values to a single value from a discrete set, known as quantization. Furthermore, noise and other issues in the transmission and acquisition of a signal, lead frequently to unobserved or lost measurements, while the observed ones are possibly corrupted and quantized.

The recovery process of a tensor from a set of sparsely corrupted measurements, known as Tensor Robust Principal Component Analysis (TRPCA), is relevant in many data processing applications, including background modeling from video [1–3], image denoising [1–5], and outlier detection [6–8]. The TRPCA problem has been solved with different sparsity patterns using the tensor singular value decomposition (t-SVD) [2–4, 9, 10]. However, the t-SVD based models do not fully exploit the low-rank structure of the data. In fact, the t-SVD decomposes a 3-dimensional tensor in first and second modes, but leaves the third mode to be addressed by circular convolution. To better exploit the low-rank structures in all 3 modes, an improvement of the tensor nuclear norm is presented in [1], tensor-based methods that use the decomposition of a tensor are described in [5], and approaches based on tensor unfoldings are proposed in [7, 11]. For the case of tensor completion from binary measurements, a maximum likelihood estimator using a nuclear norm constraint on the different matricizations of the underlying tensor is presented in [12]. In addition, a recovery model based on the tensor nuclear norm is described in [13], and a tensor-based method that uses the decomposition of the tensor is proposed in [14] to approximate low-rank tensors from binary measurements. Although different algorithms have been presented for the recovery of a tensor from a set of sparsely corrupted measurements, no prior work has been proposed for the recovery from multi-level quantized and possibly corrupted observations.

In this paper, we introduce a tensor robust principal component analysis algorithm in order to recover a real-valued tensor from partial quantized and sparsely corrupted measurements. We formulate the problem as a constrained maximum likelihood estimation of the sum of a low-rank tensor and a sparse tensor, through matricizations in each mode, in combination with a quantization and statistical measurement model. The recovered tensor is constructed as the weighted sum of the recovered unfolded matrices, using dynamic weights that boost the performance of our algorithm.

The key novelties of this work are: (i) a formal approach for the recovery of a tensor from a number of quantized and sparsely corrupted measurements; (ii) the investigation be-

This research work was supported by the Hellenic Foundation for Research and Innovation (HFRI) and the General Secretariat for Research and Technology (GSRT), under HFRI PhD Fellowship grant no. 1509 and HFRI Faculty grant no. 1725.

tween quantization and sampling as well as the existence of sparse outliers in high-order structured data; and (iii) the quantification of the performance of the method on time-series of satellite derived images of land surface temperature.

2. QUANTIZATION AND STATISTICAL MODEL

Let $\mathcal{M} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ be the N-th order tensor to be recovered that models a high dimensional signal, like an image time-series, whereby the order of the tensor is called the number of its dimensions, also known as ways or modes. We assume that the unknown tensor $\mathcal{M}$ is constructed as the sum of a low-rank tensor $\mathcal{Z}$ and a sparse tensor $\mathcal{S}$ of the same dimensions, i.e., $\mathcal{M} = \mathcal{Z} + \mathcal{S}$.

We aim at the recovery of the low-rank and the sparse components of $\mathcal{M}$ from partial quantized observations $\mathcal{Y} = \mathcal{P}_\Omega(Q(\mathcal{Z} + \mathcal{S}))$, where $\Omega \subseteq \{1, .., I_1\} \times \cdots \times \{1, .., I_N\}$ is the index set of observed entries and $\mathcal{P}_\Omega$ is a random sampling operator which keeps the entries in Ω and zeros out others. In addition, the function $\mathcal{Q} : \mathbb{R} \to \{1, .., K\}$ corresponds to a uniform quantizer that maps a real number to a set of K ordered labels,assuming that the set of quantization bin boundaries is known a priori. Then, our observations can be modeled as

$$\mathcal{Y}_{i_1..i_N} = Q(\mathcal{Z}_{i_1..i_N} + \mathcal{S}_{i_1..i_N} + \epsilon_{i_1..i_N}),\ (i_1, .., i_N) \in \Omega. \tag{1}$$

The quantities $\epsilon_{i_1..i_N}$ model the uncertainty on each measurement that follow either the logistic distribution with zero mean and unit scale, Logistic(0,1), or the standard normal distribution $\mathcal{N}(0, 1)$.

In terms of the likelihood of the observations $\mathcal{Y}_{i_1..i_N}$, the model in (1) can be written equivalently as

$$p(\mathcal{Y}_{i_1..i_N} | \mathcal{Z}_{i_1..i_N} + \mathcal{S}_{i_1..i_N}) = \Phi(\mathcal{U}_{i_1..i_N} - \mathcal{Z}_{i_1..i_N} - \mathcal{S}_{i_1..i_N}) - \Phi(\mathcal{L}_{i_1..i_N} - \mathcal{Z}_{i_1..i_N} - \mathcal{S}_{i_1..i_N}), \tag{2}$$

where the $I_1 \times \cdots \times I_N$ tensors $\mathcal{U}$ and $\mathcal{L}$ contain the upper and lower bin boundaries corresponding to the measurements $\mathcal{Y}_{i_1..i_N}$. Furthermore, the function $\Phi(x)$ corresponds to an inverse link function. For the logistic model (logistic noise), we use the inverse logit link function $\Phi_{\log}(x) = \frac{1}{1+e^{-x}}$, and for the probit model (standard normal noise), we use the inverse probit link function $\Phi_{\text{pro}}(x) = \int_{-\infty}^{x} \mathcal{N}(s \mid 0, 1)\, ds$. The proposed algorithm can be formulated for both noise models.

3. QUANTIZED TENSOR ROBUST PCA

In order to recover the real-valued missing entries of the tensor $\mathcal{M}$ from partial and possible corrupted quantized observations, we unfold the measurement tensor $\mathcal{Y} \in \mathbb{R}^{I_1 \times \ldots \times I_N}$ into N matrices and for each of them, we apply the following algorithm. Formally, the n-th of these matrices is called the mode-n matricization or unfolding of the tensor $\mathcal{Y}$, which is denoted as $\text{unfold}_n(\mathcal{Y}) = \mathbf{Y}_{(n)} \in \mathbb{R}^{I_n \times \prod_{j \neq n} I_j}$ and corresponds to a matrix with columns being the vectors obtained by fixing all indices of $\mathcal{Y}$ except the n-th index. Then, the recovered tensor is calculated as the weighted sum of the estimated tensors $\mathcal{M}_n$ of each unfolding, such that

$$\mathcal{M} \approx \sum_{n=1}^{N} a_n \cdot \mathcal{M}_n \tag{3}$$

with weights a_n that depend on the fitting error and satisfy $\sum_n a_n = 1$.

The estimated tensor in each unfolding is the folding of the sum of the recovered low-rank component $\mathbf{Z}_{(n)}$ and the sparse component $\mathbf{S}_{(n)}$, i.e., $\mathcal{M}_n = \text{fold}_n(\mathbf{Z}_{(n)} + \mathbf{S}_{(n)}), n = 1, .., N$. In particular, in order to recover the low-rank and the sparse components from quantized measurements, one seeks to minimize the negative log-likelihood of $\mathbf{Y}_{(n)_{j,k}},\ (j, k) \in \Omega_n$ (where Ω_n is the index set of observed entries of $\mathbf{M}_{(n)}$), given by (2), subject to a low-rank constraint on $\mathbf{Z}_{(n)}$ and a sparse constraint on $\mathbf{S}_{(n)}$, i.e., we seek to solve the constrained optimization problem:

$$\begin{aligned} \min_{\mathbf{Z}_{(n)}, \mathbf{S}_{(n)}} \quad & -\sum_{(j,k) \in \Omega_n} \log p(\mathbf{Y}_{(n)_{j,k}} \mid \mathbf{Z}_{(n)_{j,k}} + \mathbf{S}_{(n)_{j,k}}) \\ \text{subject to} \quad & \|\mathbf{Z}_{(n)}\|_* \leq \lambda \text{ and } \|\mathbf{S}_{(n)}\|_1 \leq \sigma \end{aligned} \tag{4}$$

The nuclear norm constraint $\|\mathbf{Z}_{(n)}\|_* \leq \lambda$ promotes low-rankness on $\mathbf{Z}_{(n)}$ [15] and the parameter $\lambda > 0$ is used to control its rank. In addition, the l_1-norm constraint $\|\mathbf{S}_{(n)}\|_1 \leq \sigma$ promotes sparsity among all entries of $\mathbf{S}_{(n)}$.

Since the gradients of the negative log-likelihood of the inverse logit and probit link functions are convex in each individual variable $\mathbf{Z}_{(n)}$ and $\mathbf{S}_{(n)}$ while holding the other fixed, the optimization problem in (4) can be solved efficiently. Starting with an initialization of the low-rank component $\mathbf{Z}_{(n)}$ as a random matrix with entries between the corresponding quantization bin boundaries, and the sparse component $\mathbf{S}_{(n)}$ with all zeros, the algorithm consists of an inner and an outer iteration. Both iterations are repeated until a maximum number of iterations is reached or the change in $\mathbf{Z}_{(n)} + \mathbf{S}_{(n)}$ between consecutive iterations is below a given threshold. Then, in each outer iteration $q = 1, 2, ...$, the algorithm performs two steps consecutively:

(i) We hold the sparse component fixed and optimize the low-rank component using an iterative procedure detailed below that forms the inner iteration $l = 1, 2, ...$

(ii) We hold the low-rank component fixed and optimize the sparse component, using a similar approach.

The problem of optimizing the low-rank component $\mathbf{Z}_{(n)}$ can be solved efficiently by performing a gradient step to reduce the objective function $F(\mathbf{Z}_{(n)}, \mathbf{S}_{(n)})$ followed by a projection step to make the solution satisfy the nuclear norm con-

straint. Specifically, the gradient step is given by

$$\hat{\mathbf{Z}}_{(n)}^{l+1} \leftarrow \mathbf{Z}_{(n)}^{l} - c \cdot \nabla F, \tag{5}$$

where c is the step-size. For simplicity, we use a constant step-size $c = \frac{1}{L}$, where L is the Lipschitz constant, which is given by $L_{\log} = \frac{1}{4}$ and $L_{\text{pro}} = 1$ for each model. The gradient of the objective function of each measurement, with respect to $\mathbf{Z}_{(n)}$, using the sparse component $\mathbf{S}_{(n)}$ estimated from the previous outer iteration, is given by

$$[\nabla F]_{jk} = \begin{cases} \frac{\Phi^{'}(\mathbf{L}_{(n)jk} - \mathbf{X}_{jk}) - \Phi^{'}(\mathbf{U}_{(n)jk} - \mathbf{X}_{jk})}{\Phi(\mathbf{U}_{(n)jk} - \mathbf{X}_{jk}) - \Phi(\mathbf{L}_{(n)jk} - \mathbf{X}_{jk})} & \text{if } (j,k) \in \Omega_n \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where $\mathbf{X} = \mathbf{Z}_{(n)} + \mathbf{S}_{(n)}$, and $\mathbf{L}_{(n)}$, $\mathbf{U}_{(n)}$ are the mode-n matricizations of $\mathcal{L}$ and $\mathcal{U}$. The derivative of the function $\Phi(x)$ can be calculated as $\Phi^{'}_{\log}(x) = \frac{1}{2+e^{-x}+e^{x}}$ and $\Phi^{'}_{\text{pro}}(x) = \mathcal{N}(x \mid 0, 1)$ for each model.

The projection step that aims to impose low-rankness on $\mathbf{Z}_{(n)}$ can be achieved by taking the singular value decomposition (SVD), $\tilde{\mathbf{U}}\tilde{\mathbf{S}}\tilde{\mathbf{V}}^{T}$, of the matrix $\hat{\mathbf{Z}}_{(n)}^{l+1}$ of the previous step, and by holding some of its singular values $\mathbf{s} = \text{diag}(\tilde{\mathbf{S}})$, depending on the parameter λ. Specifically, we keep about 90% of the information of the singular values in our experiments.

The problem of optimizing the sparse component $\mathbf{S}_{(n)}$ while holding the other variable fixed is solved using a similar approach. Specifically, at each outer iteration we perform the gradient step

$$\hat{\mathbf{S}}_{(n)}^{q+1} \leftarrow \mathbf{S}_{(n)}^{q} - c \cdot \nabla F, \tag{7}$$

where the gradient ∇F with respect to $\mathbf{S}_{(n)}$ is given by the same equation (6), using the low-rank component $\mathbf{Z}_{(n)}$ estimated from the inner iteration. Then, we perform the projection step which is given by $\mathbf{S}_{(n)}^{q+1} \leftarrow P_{\sigma}(\hat{\mathbf{S}}_{(n)}^{q+1})$, where $P_{\sigma}(x) = \text{sign}(x) \cdot \max(|x| - \sigma, 0)$ is the soft-thresholding operator on the vectorized version of $\hat{\mathbf{S}}_{(n)}^{q+1}$. The parameter σ can be estimated through experiments.

Finally, the weights $a_1, .., a_N$ in (3) can uniformly set to $\frac{1}{N}$. But, in some cases, the recovery in one unfolding maybe better than others. Therefore, instead of fixed weights, we use dynamic weights which depend on the fitting error

$$\text{fit}_n(\mathbf{Z}_{(n)} + \mathbf{S}_{(n)}) = \|\mathcal{P}_{\Omega}(\mathcal{Q}(\text{fold}_n(\mathbf{Z}_{(n)} + \mathbf{S}_{(n)}))) - \mathcal{Y})\|_F, \tag{8}$$

where $\|\mathcal{X}\|_F = \sqrt{< \mathcal{X}, \mathcal{X} >}$ is the Frobenius norm of $\mathcal{X}$.The smaller the fitting error is, the larger a_n should be. Specifically, we set

$$a_n = \frac{[\text{fit}_n(\mathbf{Z}_{(n)} + \mathbf{S}_{(n)})]^{-1}}{\sum_{i=1}^{N}[\text{fit}_i(\mathbf{Z}_{(i)} + \mathbf{S}_{(i)})]^{-1}}, \; n = 1, ..., N. \tag{9}$$

As demonstrated below, using dynamic weights a_n can improve the recovery quality of the recovered tensor.

Our model can be regarded as an extension of the quantized robust PCA for matrices, described in [16], to tensors.

4. EXPERIMENTAL RESULTS

The efficacy of the proposed algorithm is evaluated over a time-series of publicly available images of the land surface temperature, acquired by the MODIS satellite over the region of Brazil, providing valuable information about fires and thermal anomalies. Specifically, the time-series consists of 22 days of July and August of 2019, synthesizing a third-order tensor of size $64 \times 64 \times 22$, with two spatial variables and an additional temporal variable (days).

In our experiments, we quantize the images to $3, 4, 5$ and 6 bits, as the original images use 8 bits per pixel per band, we subsample their entries and we recover the whole time-series applying the proposed method. In addition, we used the logistic model, as the performance was the same for both noise models. To assess the recovery performance of our algorithm, we use the Peak-Signal-to-Noise-Ratio (PSNR) between the original and the estimated images of the time-series. A higher PSNR represents better the quality of the recovered image.

An important parameter that has to be examined is the number of nonzero elements in the sparse component of the recovered tensor. Figure 1 presents the recovery error for different values of the sparsity parameter σ and sampling percentages, using 4 bits of quantization. Have in mind that larger values of σ correspond to a bigger sparsity of the component $\mathcal{S}$ of the tensor. As we can observe from the results, more sparsity is required for a large number of observations, while less sparsity is needed for fewer available measurements. However, the performance is fairly similar for the various values of σ therefore, we chose the value $\sigma = 0.8$ in our experiments.

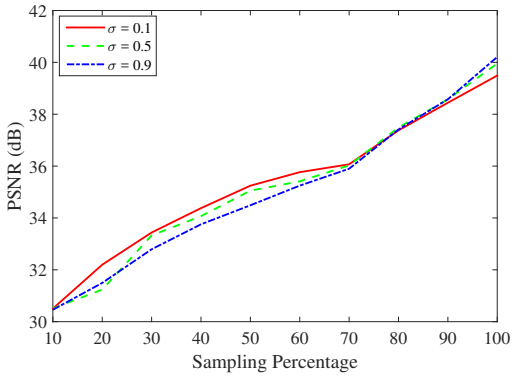


Fig. 1: Recovery error for different sparsity values and sampling percentages, using 4 bits of quantization.

Another important parameter that was examined was the number of bits of quantization. Specifically, we employed the proposed recovery algorithm for different numbers of bits of quantization and various sampling percentages, with the results reported in Figure 2. As it was expected, the more the bits of quantization, the better the quality of the recovered images. However, the improvement is more visible for a small

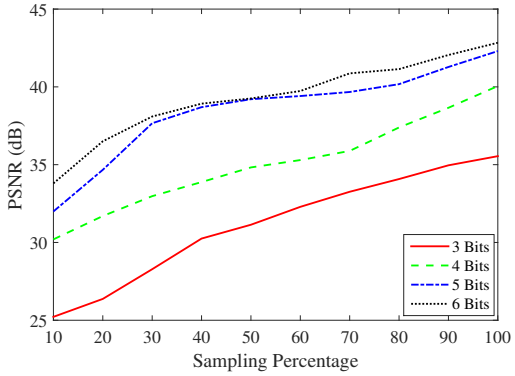


Fig. 2: Recovery error for different sampling percentages and bits of quantization.

Table 1: Recovery error for different sampling percentages on each mode matricization and on the weighted sum of them, using 3 bits of quantization.

PSNR	Sampling Percentage				
	10	30	50	70	100
Mode-1	24.43	26.62	28.43	33.15	35.32
Mode-2	23.38	25.96	31.33	33.17	35.35
Mode-3	25.92	28.52	30.46	32.35	34.89
Weighted sum	25.21	28.28	31.13	33.25	35.54

number of quantization bits.

An interesting investigation is the quality of the recovery for each unfolding of the tensor and as a sequence the impact of the dynamic weights on the recovery. Table 1 presents the recovery error for different sampling percentages on each mode matricization and on the weighted sum of them, using 3 bits of quantization. The results demonstrate that using bigger weight for the matricizations with smaller recovery error, can improve the quality of the recovered tensor, as the recovery differs for each unfolding. Note that all the results indicate that more measurements lead to better performance.

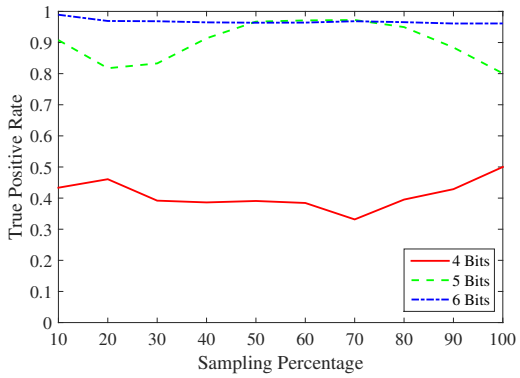
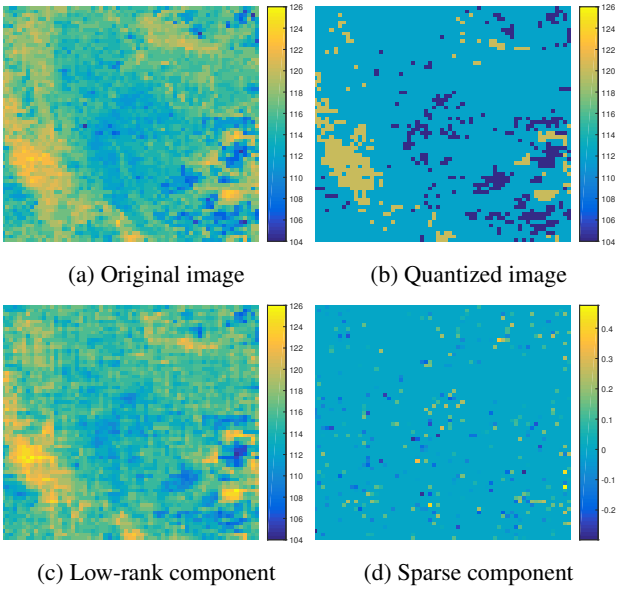


Fig. 3: True Positive Rate for different sampling percentages and bits of quantization measured only in the sparse component and the available observations.



(a) Original image (b) Quantized image

(c) Low-rank component (d) Sparse component

Fig. 4: The original and quantized to 5 bits images, as well the corresponding low-rank and sparse components.

A useful implication is that the sparse component of the recovered tensor can be considered as possible thermal anomalies in the time-series, indicating the potential locations of forest fire. In fact we selected this particular region and time period because of the extensive fires that affected a large portion of the Amazonian rain-forest [1]. Specifically, Figure 3 depicts the True Positive Rate for different sampling percentages and bits of quantization. Note that the True Positive Rate, defined as the number of true predicted positions over the true positions of the anomalies in the images, is measured only in the available observations in each case. The results indicate that our method can predict possible thermal anomalies in the images with a high probability for each number of measurements, using a large number of quantization bits.

Finally, in Fig. 4 we can see an image of the time-series and the corresponding quantized to 5 bits image, as well the recovered low-rank and sparse components, demonstrating the efficacy of our method.

5. CONCLUSION

The huge amount of multidimensional data collected over time can often be quantized and possibly corrupted due to noise and issues related to the acquisition and transmission of signals. To that end, we introduce a tensor robust principal component analysis algorithm that directly operates with the quantized measurements for the recovery of the real-valued tensor. Experimental results on satellite derived land surface time-series demonstrate the efficient performance of our algorithm, as well as the capability of the detection of possible anomalies in the data.

[1] IPAM Amazonia, 2019 `https://tinyurl.com/y5mpb7eh`

6. REFERENCES

[1] Yipeng Liu, Longxi Chen, and Ce Zhu, "Improved robust tensor principal component analysis via low-rank core matrix," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 6, pp. 1378–1389, 2018.

[2] Canyi Lu, Jiashi Feng, Wei Liu, Zhouchen Lin, Shuicheng Yan, et al., "Tensor robust principal component analysis with a new tensor nuclear norm," *IEEE transactions on pattern analysis and machine intelligence*, 2019.

[3] Shuting Cai, Qilun Luo, Ming Yang, Wen Li, and Mingqing Xiao, "Tensor robust principal component analysis via non-convex low rank approximation," *Applied Sciences*, vol. 9, no. 7, pp. 1411, 2019.

[4] Canyi Lu, Jiashi Feng, Yudong Chen, Wei Liu, Zhouchen Lin, and Shuicheng Yan, "Tensor robust principal component analysis: Exact recovery of corrupted low-rank tensors via convex optimization," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 5249–5257.

[5] Niannan Xue, George Papamakarios, Mehdi Bahri, Yannis Panagakis, and Stefanos Zafeiriou, "Robust low-rank tensor modelling using tucker and cp decomposition," in *2017 25th European Signal Processing Conference (EUSIPCO)*. IEEE, 2017, pp. 1185–1189.

[6] Hadi Fanaee-T and João Gama, "Tensor-based anomaly detection: An interdisciplinary survey," *Knowledge-Based Systems*, vol. 98, pp. 130–147, 2016.

[7] Jineng Ren, Xingguo Li, and Jarvis Haupt, "Robust pca via tensor outlier pursuit," in *2016 50th Asilomar Conference on Signals, Systems and Computers*. IEEE, 2016, pp. 1744–1749.

[8] Pan Zhou and Jiashi Feng, "Outlier-robust tensor pca," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2263–2271.

[9] Wolfgang Hackbusch, *Tensor spaces and numerical tensor calculus*, vol. 42, Springer Science & Business Media, 2012.

[10] Tao Li and Jinwen Ma, "Non-convex penalty for tensor completion and robust pca," *arXiv preprint arXiv:1904.10165*, 2019.

[11] Donald Goldfarb and Zhiwei Qin, "Robust low-rank tensor recovery: Models and algorithms," *SIAM Journal on Matrix Analysis and Applications*, vol. 35, no. 1, pp. 225–253, 2014.

[12] Anastasia Aidini, Grigorios Tsagkatakis, and Panagiotis Tsakalides, "1-bit tensor completion," *Electronic Imaging*, vol. 2018, no. 13, pp. 261–1, 2018.

[13] Baohua Li, Xiaoning Zhang, Xiaoli Li, and Huchuan Lu, "Tensor completion from one-bit observations," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 170–180, 2018.

[14] Navid Ghadermarzy, Yaniv Plan, and Ozgur Yilmaz, "Learning tensors from partial binary measurements," *IEEE Transactions on Signal Processing*, vol. 67, no. 1, pp. 29–40, 2018.

[15] Maryam Fazel, Haitham Hindi, Stephen P Boyd, et al., "A rank minimization heuristic with application to minimum order system approximation," in *Proceedings of the American control conference*. Citeseer, 2001, vol. 6, pp. 4734–4739.

[16] Andrew S Lan, Christoph Studer, and Richard G Baraniuk, "Matrix recovery from quantized and corrupted measurements," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 4973–4977.