This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LSP.2020.2990313, IEEE Signal Processing Letters

Matrix and Tensor Completion in Multiway Delay Embedded Space Using Tensor Train, with Application to Signal Reconstruction

Farnaz Sedighin, Andrzej Cichocki *Fellow, IEEE*, Tatsuya Yokota *Senior Member, IEEE* and Qiquan Shi

***Abstract*—In this paper, the problem of time series reconstruction in a multiway delay embedded space using Tensor Train decomposition is addressed. A new algorithm has been developed in which an incomplete signal is first transformed to a Hankel matrix and in the next step to a higher order tensor using extended Multiway Delay embedded Transform. Then, the resulting higher order tensor is completed using low rank Tensor Train decomposition. Comparing to previous Hankelization approaches, in the proposed approach, blocks of elements are used for Hankelization instead of individual elements, which results in producing a higher order tensor. Simulation results confirm the effectiveness and high performance of the proposed completion approach. Although in this paper we focus on single time series, our method can be straightforwardly extended to reconstruction of multivariate time series, color images and videos.**

***Index Terms*—Time series reconstruction, Tensor Train decomposition, Multiway delay embedded space, Rank incremental.**

Why Hankelization and tensorization?

I. Introduction

MULTIDIMENSIONAL data recorded by multi sensors has been widely used in various signal processing applications. Inefficiency of matrices for representing the multidimensional data, leads to move from matrices to tensors [1]. Tensors have been used in many applications such as time series and image completion (reconstruction) [2]–[5], source separation [6], neural network [7] and pattern recognition [8].

The acquired data is often incomplete or corrupted by noise or outliers due to different reasons. In such scenarios, using completion algorithms seems to be inevitable. Completion (reconstruction) is the problem of estimating missing or uncertain samples of a dataset using only its available elements [9]–[12].

Hankelization is a widely accepted approach in time series analysis and reconstruction [13]–[16]. Note that in a Hankel matrix all of the elements in each skew diagonal are the same. This is an important step in a time series reconstruction algorithm, called Singular Spectrum Analysis (SSA) [13], [17]–[19]. Tensor decomposition approaches such as Canonical Polyadic Decomposition (CPD) [2], Tucker decomposition [20], Hierarchical Tucker decomposition [21], have been also used for time series reconstruction.

Recently, in [20], [22], a new approach for image and time series completion has been proposed by applying multi dimensional Hankelization on an incomplete data. In this approach, the raw data has been first transformed to a higher order tensor using Multiway Delay embedded Transform (MDT). The transform is done by multiplying so called duplication matrices to the data, followed by a folding step, i.e., tensorization. Then a Tucker decomposition algorithm with rank incremental strategy has been applied for completion [20].

In this paper, a novel signal completion approach is developed based on Hankelization of block elements via multiway delay embedding. In the proposed approach, an incomplete time seres is transformed to a higher order tensor using a two stage procedure: first, time series is converted to a Hankel matrix and second, the Hankel matrix tensorized to a block Hankel tensor using its block elements (instead of individual elements). Such a Hankelization and tensorization allows us to exploit local correlation between different sub-windows of time series or patches of images. In addition, due to the Hankel structure, it is expected that the resulting higher order tensor has relatively low rank. Then, a tensor decomposition approach is needed for tensor completion. Comparing to tensor decomposition approaches such as CPD or Tucker decomposition, Tensor Train (TT) decomposition usually has a better ability to represent higher order tensors. In contrast to Tucker decomposition, TT decomposition does not suffer from curse of dimensionality and TT decomposition algorithms are more stable than standard CPD algorithms. Considering that the resulting tensor is quite a high order tensor, TT decomposition can be applied for efficient reconstruction. In TT decomposition, a tensor is decomposed to a series of third order interconnected core tensors as [12], [23]–[27]

$$\underline{\mathbf{X}} \cong \ll \underline{\mathbf{G}}^{(1)}, \underline{\mathbf{G}}^{(2)}, \dots, \underline{\mathbf{G}}^{(N)} \gg, \tag{1}$$

where $\underline{\mathbf{X}} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is the data tensor to be decomposed, $\underline{\mathbf{G}}^{(n)} \in \mathbb{R}^{R_n \times I_n \times R_{n+1}}$ is the n-th core tensor, R_n's are the ranks and $R_1 = R_{N+1} = 1$. To best of our knowledge, this is the first time of using TT decomposition in a multiway delay embedded space instead of signal (original) space.

In the proposed algorithm, similar to [20], the ranks of core tensors are not fixed but increased gradually during optimization process using a rank incremental strategy. In each iteration, the core tensors whose ranks should be increased are selected. Note that TT decomposition with rank incremental

This work was partially supported by the MES RF grant 14.756.31.0001.

F.Sedighin and A.Cichocki are with Skolkovo Institute of Science and Technology (SKOLTECH), 121205 Moscow, Russia (e-mail: F.Sedighin@skoltech.ru, A.Cichocki@skoltech.ru).

A.Cichocki ia also with the System Research Institute, Polish Academy of Science, 01-447 Warsaw, Poland, also with the college of Computer Science, Hangzhou Dianzu University, Hangzhou 310018, China. .

T.Yokota is with the Nagoya Institute of Technology, Nagoya, Japan (email: t.yokota@nitech.ac.jp) and with the RIKEN Center for Advanced Intelligence Project, Tokyo, Japan.

Q.Shi is with the Huawei Noah's Ark Lab, Hong Kong. (email: shiqiquan@huawei.com).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LSP.2020.2990313, IEEE Signal Processing Letters

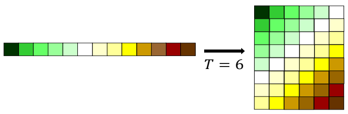

Fig. 1: Illustration of Hankelization of a time series for $T = 6$.

has been previously proposed in [28], but in a way that the ranks of all of the core tensors were increased simultaneously at each iteration, while in our approach only ranks of some automatically selected core tensors are increased.

Moreover, in our method, we have employed a new multi stage approach, i.e., the main algorithm is repeated several times for gradually decreased block sizes. In this approach, the algorithm starts with some block size, and then the resulting completed output is again processed by the algorithm with different block size (typically smaller than the previous one). Generally, our contributions can be summarized as follows

- Novel Hankelization of input data using blocks of elements (samples) instead of individual elements (samples).
- Using TT decomposition in a multiway delay embedded space instead of original signal space.
- Developing a new rank incremental strategy for gradually increasing the ranks of automatically selected core tensors in each iteration.
- Exploiting a novel multi stage strategy to gradually increase the performance by decreasing the block size in each stage.

The rest of this paper is organized as follows, the proposed Hankelization technique and the proposed algorithm are introduced in Sections II and III, respectively. Simulation results are presented in Section IV and in Section V we provide conclusion.

II. Proposed Hankelization Technique

In this paper, we have proposed a new Hankelization method to transform an original incomplete time series into a higher order tensor. The proposed approach includes two steps:

1- A time series (vector) is first transformed to a Hankel matrix using a specific window size of T. A sample for this transformation with $T = 6$ has been illustrated in Figure. 1.

2- After that, the proposed block Hankelization method is applied on the resulting Hankel matrix. Based on this approach, instead of Hankelizing individual elements of the resulting Hankel matrix, blocks of elements are Hankelized. Block Hankelization is achieved by multiplying a data matrix by block duplication matrices, followed by a tensorization step. A sample for a block duplication matrix is given in Fig 2. A block duplication matrix is composed of several identity matrices, where the size of each of them is $PT_k \times PT_k$, where $T_k, (k = 1, 2)$ is the window size and P is the block size. The amount of shift between two consecutive identity matrices is P. In the first step of the proposed block Hankelization method, two block duplication matrices are multiplied by a data matrix as $\mathbf{H}_e = \mathbf{S}_{H,1}\mathbf{H}\mathbf{S}_{H,2}^T$, where $\mathbf{H} \in \mathbb{R}^{I_1 \times I_2}$ is the data matrix and $\mathbf{S}_{H,1} \in \mathbb{R}^{PT_1(I_1/P-T_1+1)\times I_1}$ and $\mathbf{S}_{H,2}^T \in \mathbb{R}^{I_2 \times PT_2(I_2/P-T_2+1)}$ are two block duplication matrices and $\mathbf{H}_e \in \mathbb{R}^{PT_1(I_1/P-T_1+1)\times PT_2(I_2/P-T_2+1)}$ is the

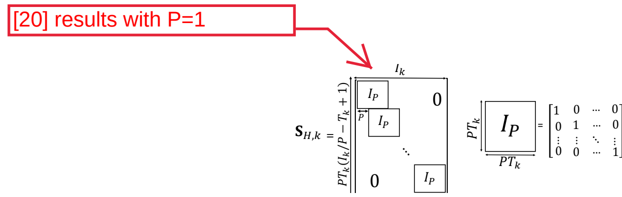

Fig. 2: The proposed block duplication matrix ($\mathbf{S}_{H,k}$), for transforming an original data matrix to a block Hankel matrix. This block Hankel matrix is then represented as a 6-th order tensor (see Fig. 3).

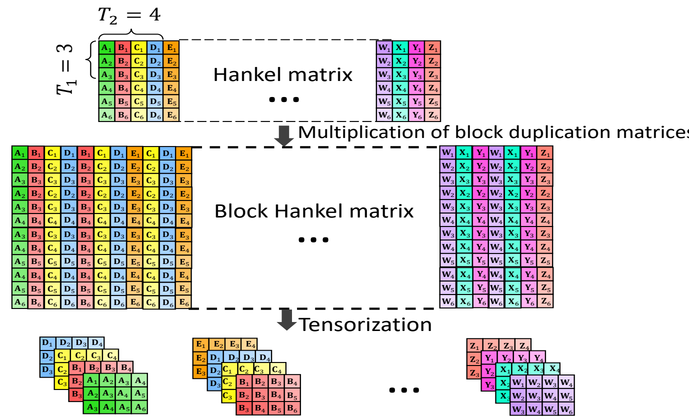

Fig. 3: The proposed block Hanklization method and tensorization to a 6-th order tensor.

resulting block Hankel matrix. Note that, when I_k/P is not integer, data matrix $\mathbf{H}$ will be zero padded in a way that I_k/P become integer. Then, a tensorization step is performed. So, $\mathbf{H}_e$ is reshaped to a 6-th order block Hankel tensor $\underline{\mathbf{H}}$ of size $P \times P \times T_1 \times D_1 \times T_2 \times D_2$, where $D_1 = (I_1/P - T_1 + 1)$ and $D_2 = (I_2/P - T_2 + 1)$. A sample result for block Hankelization method with $T_1 = 3$ and $T_2 = 4$ has been illustrated in Fig. 3. The top matrix in Fig. 3 is a Hankel matrix resulted from Hankelizing a vector (see Fig. 1), and each small color box with alphabet is a block matrix of size $P \times P$.

III. Proposed Algorithm

After multiway delay embedding the incomplete time series, the next step is to complete the 6-th order tensor using TT decomposition, by minimizing the following cost function:

$$J(\boldsymbol{\theta}) = \|\underline{\mathbf{\Omega}} \circledast (\underline{\mathbf{H}} - \widehat{\underline{\mathbf{H}}}(\boldsymbol{\theta}))\|_F^2, \tag{2}$$

where $\underline{\mathbf{H}}$ is the original incomplete tensor in a multiway delay embedded space, $\widehat{\underline{\mathbf{H}}}(\boldsymbol{\theta})$ is the estimated tensor in TT format with the core tensors $\boldsymbol{\theta} = (\underline{\mathbf{G}}^{(1)}, \underline{\mathbf{G}}^{(2)}, \ldots, \underline{\mathbf{G}}^{(N)})$ which can be written as $\widehat{\underline{\mathbf{H}}}(\boldsymbol{\theta}) = \ll \underline{\mathbf{G}}^{(1)}, \underline{\mathbf{G}}^{(2)}, \ldots, \underline{\mathbf{G}}^{(N)} \gg$, $\underline{\mathbf{\Omega}}$ is a tensor whose elements are 1 for the observed and 0 for the missing elements in the multiway delay embedded space and $\circledast$ is the element wise Hadamard product. By using a Majorization Minimization approach, minimization of (2) can be achieved by minimization of the following auxiliary cost function

$$\begin{aligned} J(\boldsymbol{\theta}|\boldsymbol{\theta}^k) = \|\underline{\mathbf{\Omega}} \circledast (\underline{\mathbf{H}} - \widehat{\underline{\mathbf{H}}}(\boldsymbol{\theta}))\|_F^2 + \\ \|(\underline{\mathbf{1}} - \underline{\mathbf{\Omega}}) \circledast (\widehat{\underline{\mathbf{H}}}(\boldsymbol{\theta}^k) - \widehat{\underline{\mathbf{H}}}(\boldsymbol{\theta}))\|_F^2, \end{aligned} \tag{3}$$

where $\underline{\mathbf{1}}$ is a 6-th order tensor whose all elements are equal to 1. Note that $J(\boldsymbol{\theta}|\boldsymbol{\theta}^k) \geq J(\boldsymbol{\theta})$ and $J(\boldsymbol{\theta}^k|\boldsymbol{\theta}^k) = J(\boldsymbol{\theta}^k)$, so

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LSP.2020.2990313, IEEE Signal Processing Letters

This is how imputation is done.

minimization of $J(\boldsymbol{\theta}|\boldsymbol{\theta}^k)$ with respect to $\boldsymbol{\theta}$ also results in a decrease in $J(\boldsymbol{\theta})$. The cost function (3) can be re-written as

$$J(\boldsymbol{\theta}|\boldsymbol{\theta}^k) = \|\underline{\tilde{\mathbf{H}}} - \underline{\widehat{\mathbf{H}}}(\boldsymbol{\theta})\|_F^2, \tag{4}$$

where $\boxed{\underline{\tilde{\mathbf{H}}} = \underline{\boldsymbol{\Omega}} \circledast \underline{\mathbf{H}} + (\underline{\mathbf{1}} - \underline{\boldsymbol{\Omega}}) \circledast \underline{\widehat{\mathbf{H}}}(\boldsymbol{\theta}^k)}$. For brevity in the notations, we remove $\boldsymbol{\theta}$ in the rest of equations. Note that the elements of the two terms of (3) are not overlapped, so sum of their Frobenius norms is equivalent to the Frobenius norm of their summation. Equation (4) implies that the minimization of (2), and consequently the completion approach, is equivalent to several iterations of TT estimation of $\underline{\tilde{\mathbf{H}}}$ while in each iteration the elements in which $\underline{\boldsymbol{\Omega}} = 1$ are replaced with the observed ones. The cost function (4) has been minimized by using Alternating Least Square (ALS) method.

For ALS estimation of (4), the function "`tt_als`" has been used [29]. In the first iteration, the core tensors of $\underline{\widehat{\mathbf{H}}}(\boldsymbol{\theta})$ are initialized randomly. In the next iterations, the estimated core tensors of the previous iteration used as the initial values for the current iteration and the new elements of the core tensors whose ranks are increased are initialized randomly.

As mentioned in Introduction, the ranks of the core tensors are increased gradually in each iteration based on an approach which will be discussed in the following subsection.

A. Rank Incremental Strategy

An important problem in solving the optimization problem (4) is how to determine the ranks of core tensors, i.e., R_n's. In this paper we have developed an algorithm that gradually increases the TT ranks. The proposed approach is based on gradual increasing the ranks of the most significant core tensor. So, the core tensor whose omission results in the largest approximation error is determined and its ranks are increased.

The estimated tensor, $\underline{\widehat{\mathbf{H}}}$, can be matricized as $\widehat{\mathbf{H}}_{(n)} = \mathbf{G}_{(2)}^{(n)}(\mathbf{G}_{(1)}^{>n} \otimes \mathbf{G}_{(n)}^{<n})$ for $n = 1, 2, \cdots, N$, where $\underline{\mathbf{G}}^{>n} = \ll \underline{\mathbf{G}}^{(n+1)}, \underline{\mathbf{G}}^{(n+2)}, \cdots, \underline{\mathbf{G}}^{(N)} \gg \in \mathbb{R}^{R_{n+1} \times I_{n+1} \times \cdots \times I_N}$ and $\underline{\mathbf{G}}^{<n} = \ll \underline{\mathbf{G}}^{(1)}, \underline{\mathbf{G}}^{(2)}, \cdots, \underline{\mathbf{G}}^{(n-1)} \gg \in \mathbb{R}^{I_1 \times \cdots \times I_{n-1} \times R_n}$, and $\underline{\mathbf{G}}^{>N} = \underline{\mathbf{G}}^{<1} = 1$ [12]. Recall that $\mathbf{H}_{(n)}$ is a matrix resulted from mode n-th matricization of tensor $\underline{\mathbf{H}}$ and $\otimes$ is the Kronecker product. The error resulted from the omission of the n-th core tensor can be expressed as

$$e(n) = \|\mathbf{G}_{(2)}^{(n)\dagger}(\boldsymbol{\Omega}_{(n)} \circledast (\mathbf{H}_{(n)} - \widehat{\mathbf{H}}_{(n)}))\|_F^2, \quad n = 1, 2 \cdots, N \tag{5}$$

where "$\dagger$" denotes the Moore-Penrose inverse. After computing $e(n)$ for each n, core tensor $\underline{\mathbf{G}}^{(n)}$ corresponding to the largest error $(e(n))$ is selected and its ranks are increased. Since the core tensors, except the first and the last ones, have two ranks, both of the ranks of the selected core tensor (and corresponding ranks of its neighbors) are increased (for the first and the last ones only one rank is increased). This procedure can be performed for increasing simultaneously the ranks of more than one core tensors in each iteration, i.e., by selecting several core tensors generating the largest errors. The ranks can be increased by any predetermined value. Note that, the ranks of core tensors are also limited as $R_n \leq R_{n-1} \times I_{n-1}$ and $R_n \leq I_n \times R_{n+1}$. So, the rank of a selected core tensor $\underline{\mathbf{G}}^{(n_s)}$, after one iteration of rank incremental, is equal to

$$R_{n_s}^{new} \Leftarrow \min(R_{n_s} + \text{step}, R_{n_s-1} \times I_{n_s-1}, I_{n_s} \times R_{n_s+1}), \tag{6}$$

where "step" is the value of rank increasing and R_{n_s} and $R_{n_s}^{new}$ are the previous and new ranks of $\underline{\mathbf{G}}^{(n_s)}$, respectively.

The ALS estimation and the rank incremental steps repeated until the difference of the error, i.e., $\|\boldsymbol{\Omega} \circledast (\mathbf{H} - \widehat{\mathbf{H}})\|_F^2$ between two consecutive iterations becomes less than a threshold or the number of iterations reaches its maximum. Note that $\widehat{\mathbf{H}}$ is the completed matrix in the signal space and $\boldsymbol{\Omega}$ is the matrix with 1 for the observed and 0 for the missing elements.

After estimating the completed tensor in the multiway delay embedded space, the signal should be re-transformed to the signal space using inverse of Hankelization or deHankelization [20]. In this paper, deHankelization has been done by averaging the blocks corresponding to a specific block in the original tensor, i.e., the blocks with the same alphabet. So, firstly, the frontal slices with the same color (see Fig. 3) are averaged and then the diagonal blocks of the resulting matrices are averaged to produce the final blocks. The series of final blocks are then reformatted to the final estimated Hankel matrix and consequently the reconstructed time series.

Considering that the cost function is nonconvex and in order to improve the performance, we have additionally developed a novel multi stage strategy. The optimization is done in several stages, starting from large block sizes and move to smaller ones. In this approach, the result of the first stage is considered as a good initialization for the next stage and so on. So the algorithm starts with some block sizes and then the result is used again as input of the algorithm with smaller block size. This can help us to increase the quality of the algorithm comparing to the situation when the algorithm performs only one stage. Our simulation results show that if the block size is increased, instead of being decreased, the final solution will be somewhat blurred and does not have enough quality.

In the noiseless cases, the observed elements are remained unchanged, so the final estimated matrix is computed as $\widehat{\mathbf{H}} = \boldsymbol{\Omega} \circledast \mathbf{H} + (\mathbf{1} - \boldsymbol{\Omega}) \circledast \widehat{\mathbf{H}}$, but in noisy cases, all of the elements are estimated by the proposed algorithm. Finally, a smoothing procedure, by replacing each of the estimated elements by the average of its four neighbors, is applied.

IV. Simulation Results

The performance of the proposed algorithm is compared with the state of the art MDT [20] and SSA [13] algorithms. Simulations will be illustrated here of reconstruction of real life speech signal and also synthetic time series defined as

$$\begin{aligned} h(n) = {} & n^2/L + 2n\sin(n\pi/k_1) + n\cos(n\pi/k_2) - n + \\ & 4\text{round}(200\cos(2n^3/k_3)))/L, \quad n = 1, 2, \cdots, L \end{aligned} \tag{7}$$

where L is the signal length, $k_1 = 2000$, $k_2 = 30$, $k_3 = 2 \times 10^7$ and "round" changes a non-integer value to the nearest integer value. In the first simulation, several incomplete and noisy time series have been reconstructed using the three algorithms. For the proposed algorithm, a two stage strategy has been used with $P = 2$ for the first stage and $P = 1$ for the second stage. Ranks of two selected core tensors were increased by one in each iteration and we have selected $T_1 = T_2 = 2$. The results have been shown in Fig. 4. In this figure, the first and the second rows correspond to the artificial time series and speech

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LSP.2020.2990313, IEEE Signal Processing Letters

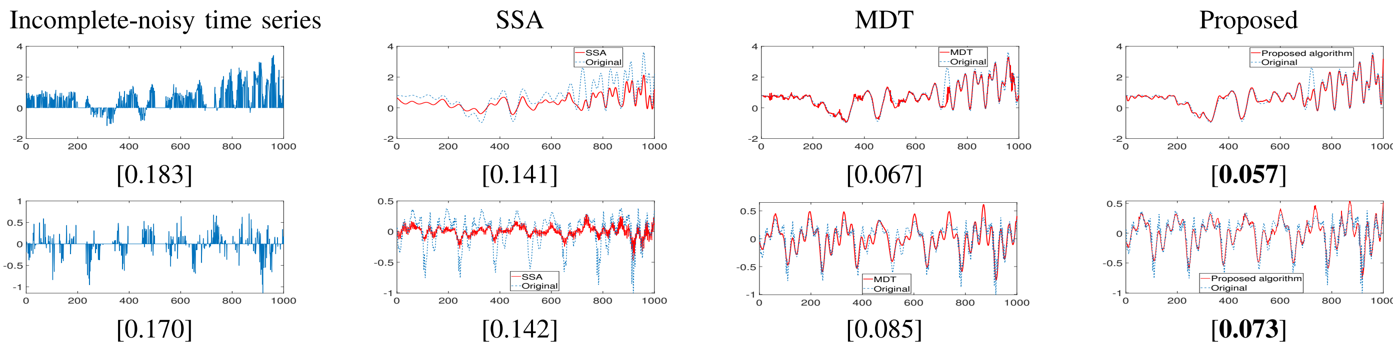

Fig. 4: Reconstructed time series from incomplete and noisy time series. First row corresponds to the reconstruction of artificial time series (7) while the second row illustrates the reconstruction of speech signal.

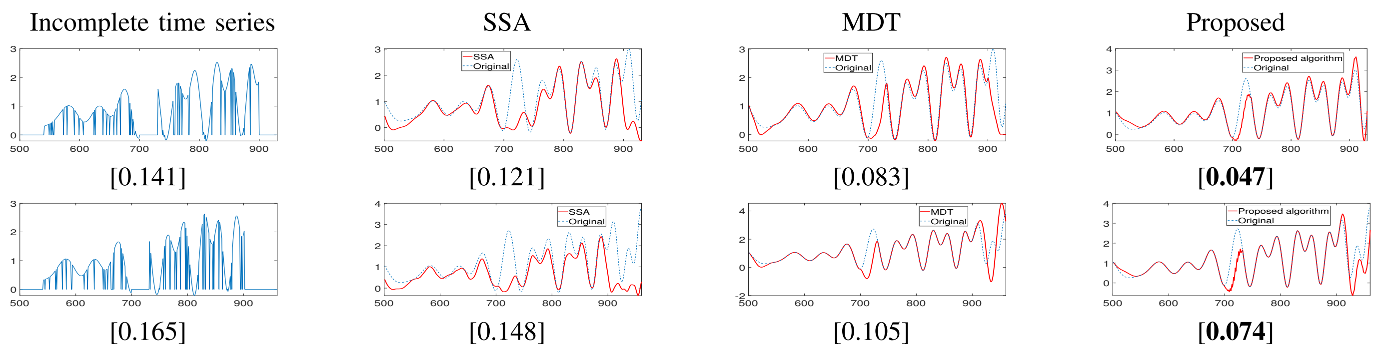

Fig. 5: Reconstructed time series when several last elements of the original time series have been also missed. In the first row the last 30 and in the second row the last 60 elements of the time series have been missed.

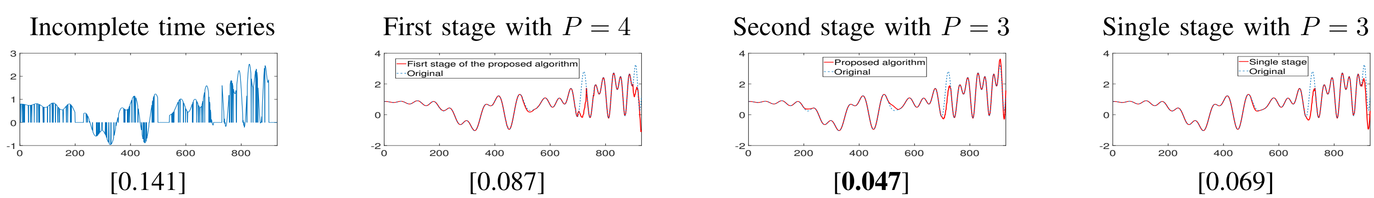

Fig. 6: Comparing the performance of our multi stage strategy with a single stage approach.

signal, respectively. Hierarchical Hankelization (with $P = 1$) has been applied for using MDT, with $\tau_1 = \tau_2 = 2$. So in this case we applied Tucker decomposition with hierarchical Hnakelization. The window size (T) is equal to 100 for the first and 300 for the second row. The Normalized Root Mean Square Error (NRMSE) for the incomplete and the reconstructed time series have been also shown beneath each figure in brackets. The SNR's of the noisy time series were 11.66 and 7.77 dB's respectively. The missing rate for the artificial time series is 53% and for the speech signal is 74%. The reconstructed time series and also and NRMSE's confirmed the better performance of the proposed approach comparing to the performance of other two approaches.

In the next simulation, the algorithms were also compared when 30 and 60 last samples of the time series have been also missed. The incomplete time series were noise free (see Fig. 5). For better illustration of the performance, only last 500 elements are shown. The results confirmed that the proposed method allows us to perform not only completion, but it also has ability of forecasting of the future samples.

Finally, our novel multi stage strategy has been investigated. The results are illustrated in Fig. 6. In this figure, the results of two stages of the algorithm for $P = 4$ and $P = 3$ for the first and second stages, are shown. The results were compared with single stage with $P = 3$. Intensive computer simulations confirmed that using several stages can improve the quality of the reconstructed signal comparing to the single stage strategy.

V. Conclusion

A new approach for time series reconstruction has been developed based on transforming the incomplete data to a multiway delay embedded space and using TT decomposition to perform reconstruction. Block Hankelization approach has been introduced by transforming the incomplete data to a relative high order tensor. Then low rank TT decomposition has been applied for completion. The ranks of the core tensors were determined automatically by gradually increasing them using a proposed rank incremental strategy. The computer simulations confirmed high performance of the proposed algorithm.

References

[1] A. Cichocki, D. Mandic, L. D. Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. Phan, "Tensor Decompositions for Signal Processing Applications: From two-way to multiway component analysis," *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 145–163, 2015.

[2] S. Trickett, L. Burroughs, and A. Milton, "Interpolation Using Hankel Tensor Completion," in *SEG Technical Program Expanded Abstracts 2013*. Society of Exploration Geophysicists, 2013, pp. 3634–3638.

[3] G. Ely, S. Aeron, N. Hao, and M. E. Kilmer, "5D Seismic Data Completion and Denoising Using a Novel Class of Tensor Decompositions," *Geophysics*, vol. 80, no. 4, pp. V83–V95, 2015.

[4] S. Kouchaki, S. Sanei, E. L. Arbon, and D.-J. Dijk, "Tensor based singular spectrum analysis for automatic scoring of sleep eeg," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 1, pp. 1–9, 2014.

[5] Y. Liu, Z. Long, and C. Zhu, "Image completion using low tensor tree rank and total variation minimization," *IEEE Transactions on Multimedia*, vol. 21, no. 2, pp. 338–350, 2018.

[6] D. Nion, K. N. Mokios, N. D. Sidiropoulos, and A. Potamianos, "Batch and Adaptive PARAFAC-Based Blind Separation of Convolutive Speech Mixtures," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 6, pp. 1193–1207, 2010.

[7] W. Wang, Y. Sun, B. Eriksson, W. Wang, and V. Aggarwal, "Wide Compression: Tensor Ring Nets," *learning*, vol. 14, no. 15, pp. 13–31, 2018.

[8] A. Cichocki, "Era of Big Data Processing: A New Approach via Tensor Networks and Tensor Decompositions," *arXiv preprint arXiv:1403.2048*, 2014.

[9] M. Filipović and A. Jukić, "Tucker Factorization with Missing Data with Application to Low-n-rank Tensor Completion," *Multidimensional Systems and Signal Processing*, vol. 26, no. 3, pp. 677–692, Jul 2015.

[10] E. Acar, D. M. Dunlavy, T. G. Kolda, and M. Mørup, "Scalable Tensor Factorizations for Incomplete Data," *Chemometrics and Intelligent Laboratory Systems*, vol. 106, no. 1, pp. 41 – 56, 2011, multiway and Multiset Data Analysis.

[11] 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, 2013.

[12] L. Yuan, Q. Zhao, L. Gui, and J. Cao, "High-order Tensor Completion via Gradient-Based Optimization Under Tensor Train Format," *Signal Processing: Image Communication*, vol. 73, pp. 53–61, 2019.

[13] H. Hassani and D. Thomakos, "A Review on Singular Spectrum Analysis for Economic and Financial Time Series," *Statistics and its Interface*, vol. 3, no. 3, pp. 377–397, 2010.

[14] S. L. Brunton, B. W. Brunton, J. L. Proctor, E. Kaiser, and J. N. Kutz, "Chaos as an intermittently forced linear system," *Nature communications*, vol. 8, no. 1, p. 19, 2017.

[15] Q. Shi, J. Yin, J. Cai, A. Cichocki, T. Yokota, L. Chen, M. Yuan, and J. Zeng, "Block Hankel Tensor ARIMA for Multiple Short Time Series Forecasting," *In Proceedings of AAAI 2020*.

[16] A. Agarwal, M. J. Amjad, D. Shah, and D. Shen, "Model Agnostic Time Series Analysis via Matrix Estimation," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 2, no. 3, p. 40, 2018.

[17] H. Hassani, A. Webster, E. S. Silva, and S. Heravi, "Forecasting us tourist arrivals using optimal singular spectrum analysis," *Tourism Management*, vol. 46, pp. 322–335, 2015.

[18] H. Hassani, M. Kalantari, and M. Yarmohammadi, "An Improved SSA Forecasting Result Based on a Filtered Recurrent Forecasting Algorithm," *Comptes Rendus Mathematique*, vol. 355, no. 9, pp. 1026–1036, 2017.

[19] H. Briceño, C. M. Rocco, and E. Zio, "Singular Spectrum Analysis for Forecasting of Electric Load Demand," *CHEMICAL ENGINEERING*, vol. 33, 2013.

[20] T. Yokota, B. Erem, S. Guler, S. K. Warfield, and H. Hontani, "Missing Slice Recovery for Tensors Using a Low-Rank Model in Embedded Space," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 8251–8259.

[21] C. Da Silva and F. J. Herrmann, "Low-rank Promoting Transformations and Tensor Interpolation-Applications to Seismic Data Denoising," in *76th EAGE Conference and Exhibition 2014*, 2014.

[22] T. Yokota and H. Hontani, "Tensor Completion with Shift-invariant Cosine Bases," in *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. IEEE, 2018, pp. 1325–1333.

[23] A. Cichocki, N. Lee, I. Oseledets, A.-H. Phan, Q. Zhao, D. P. Mandic *et al.*, "Tensor networks for dimensionality reduction and large-scale optimization: Part 1 low-rank tensor decompositions," *Foundations and Trends® in Machine Learning*, vol. 9, no. 4-5, pp. 249–429, 2016.

[24] A. Cichocki, A.-H. Phan, Q. Zhao, N. Lee, I. Oseledets, M. Sugiyama, D. P. Mandic *et al.*, "Tensor Networks for Dimensionality Reduction and Large-Scale Optimization: Part 2 applications and future perspectives," *Foundations and Trends® in Machine Learning*, vol. 9, no. 6, pp. 431–673, 2017.

[25] I. V. Oseledets, "Tensor-Train Decomposition," *SIAM Journal on Scientific Computing*, vol. 33, no. 5, pp. 2295–2317, 2011.

[26] C.-Y. Ko, K. Batselier, W. Yu, and N. Wong, "Fast and accurate tensor completion with tensor trains: A system identification approach," *arXiv preprint arXiv:1804.06128*, 2018.

[27] J. A. Bengua, H. N. Phien, H. D. Tuan, and M. N. Do, "Efficient tensor completion for color image and video recovery: Low-rank tensor train," *IEEE Transactions on Image Processing*, vol. 26, no. 5, pp. 2466–2479, 2017.

[28] L. Grasedyck, M. Kluge, and S. Krämer, "Variants of Alternating Least Squares Tensor Completion in the Tensor Train Format," *SIAM Journal on Scientific Computing*, vol. 37, no. 5, pp. A2424–A2450, 2015.

[29] I. Oseledets, "TT-Toolbox, version 2.2," http://spring.inm.ras.ru/osel, 2009.