

DISCOVERY OF HIDDEN STRUCTURES IN MICROSEISMIC DATA USING
TENSOR DECOMPOSITIONS AND MULTIWAY COMPONENT ANALYSIS

A Thesis

by

MAXIM ANDREEVICH YATSENKO

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Eduardo Gildin
Co-Chair of Committee,	Richard L. Gibson
Committee Member,	Akhil Datta-Gupta
Head of Department,	Jeff Spath

May 2019

Major Subject: Petroleum Engineering

Copyright 2019 Maxim Andreevich Yatsenko

ABSTRACT

Microseismic data is used by companies to analyze and check numerous processes, including horizontal well performance, directional drilling and, most recently, hydraulic fracturing. The microseismic data analysis approach is important and there is a lot more to discover within microseismic technology with an application of the correct data analysis approach. For example, different visualization methods could potentially contain hidden structures, not visible by traditional methods.

This work proposes a new methodology to access those hidden structures. In particular, machine learning tools, such as Tensor Decomposition (TD) and Multiway Component Analysis (MWCA), were utilized to gain more information from a previously existing pool of microseismic data. The extracted hidden structures can be used to learn more about source location, from which the information about fracture propagation could be inferred within the reservoir. This potentially gives a fast and cost-effective technique to analyze hydraulic fracturing processes.

The work further illustrates applicability to a real microseismic study of the noise reduction and model reduction methods, based on the same machine learning techniques. A special case of TD, Higher-Order Singular Value Decomposition (HOSVD) is used to decompose the data, while MWCA is used to show the relationship between the decomposed structure and hidden structures within the dataset. Finally, possible steps to improve the technology are outlined. Since the applications of MWCA and TD are still emerging, future enhancements to this methodology are expected.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Gildin, my co-chair Dr. Gibson, and all the professors I have been working with during the past two years, for their support and inspiration for my research, as well as their constant guidance throughout my supplemented coursework.

I would also like to thank my friends, colleagues and the department faculty for making my time at Texas A&M University an unforgettable and lifechanging experience.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of the primary advisor Professor Eduardo Gildin of the Department of Petroleum Engineering, co-chair Professor Richard Gibson of the Department of Geology & Geophysics, and Professor Akhil Datta-Gupta of the Department of Petroleum Engineering.

The microseismic data analyzed in Section 3 was provided by Professor Richard Gibson. The analysis in Section 3 regarding HOSVD application to microseismic data was conducted in collaboration with Milan Brankovic from the Department of Geology & Geophysics. All other work conducted for the thesis was completed by the author independently.

Parts of the conclusions of Section 3 are intended to be published as an SPE-195522-MS paper, "A Novel Approach to Discovery of Hidden Structures in Microseismic Data Using Machine Learning Techniques" (Control # 19EURO-P-592-SPE) and presented during the SPE Europec featured at 81st EAGE Annual Conference & Exhibition to be held 3-6 June 2019 in London, United Kingdom.

Funding Sources

Graduate study was supported by a fellowship from Texas A&M University, as well as by Texas A&M Energy Institute and The Fulbright U.S. Student Program. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Fulbright Association or Energy Institute.

NOMENCLATURE

BSS	Blind Signal Separation
CPD	Canonical Polyadic Decomposition
HOSVD	Higher-Order Singular Value Decomposition
MLSVD	Multilinear Singular Value Decomposition
MWCA	Multiway Component Analysis
NGL	Natural Gas Liquid
POD	Proper Orthogonal Decomposition
PVD	Population Value Decomposition
QTN	Quantized Tensor Networks
SVD	Singular Value Decomposition
SCA	Simultaneous Component Analysis
TD	Tensor Decomposition

TABLE OF CONTENTS

	Page
ABSTRACT	ii
ACKNOWLEDGEMENTS	iii
CONTRIBUTORS AND FUNDING SOURCES.....	iv
NOMENCLATURE.....	v
TABLE OF CONTENTS	vi
LIST OF FIGURES.....	viii
LIST OF TABLES	x
1. INTRODUCTION.....	1
1.1. Preface.....	1
1.2. Summary of Literature Review	3
1.3. Thesis Objective and Scope	5
1.4. Thesis Organization.....	5
2. PROBLEM BACKGROUND.....	7
2.1. Introduction to Microseismic Monitoring	7
2.2. The Problem and Mathematical Framework / Seismic Inversion	9
2.2.1. Seismic Inversion	9
2.2.2. Mathematical Framework.....	9
2.3. Machine Learning Through HOSVD – Introduction to HOSVD / MWCA	16
2.3.1. Machine Learning Techniques	16
2.3.2. Singular Value Decomposition (SVD).....	18
2.3.3. Tucker Decompositions.....	18
2.3.4. Higher-Order Singular Value Decomposition (HOSVD)	20
2.3.5. Multiway Component Analysis and Principal Component Analysis	21
2.4. Hidden Structures.....	24
3. PROPOSED METHODOLOGY AND NUMERICAL EXAMPLES.....	26
3.1. Tensorization of Microseismic Data	26

3.2. Denoising and Data Compression	28
3.2.1. The Importance of Model and Noise Reduction in Microseismic Data	28
3.2.2. Description of the Methodology	29
3.2.3. Results and Possible Future Research	36
3.3. Hidden Structures	37
3.3.1. Fractures Research.....	38
3.3.2. Other Ideas in Fracture Research	45
3.3.3. Research in Four Dimensions and Variation in Time	46
3.4. Overview of the Proposed Methodology	47
3.5. Main Findings and Future Work	48
4. CONCLUSIONS	51
REFERENCES	52

LIST OF FIGURES

	Page
Figure 1.1 U.S. Crude and Natural Gas Liquids (NGLs) production. From [29], distributed for educational and other non-commercial purposes.....	1
Figure 2.1 Graphical description of a possible microseismic event.....	7
Figure 2.2. General structure of microseismic data.....	8
Figure 2.3 Fourth-order tensor and its tensor diagram representation. From [12], distributed with author’s permission.	10
Figure 2.4 Fifth-order tensor and its tensor diagram representation. From [12], distributed with author’s permission.	10
Figure 2.5 Sixth-order tensor and its tensor diagram representation. From [12], distributed with author’s permission.	11
Figure 2.6 Graphical tensor diagrams representations of higher-order tensors. From [12], distributed with author’s permission.	11
Figure 2.7 Mode-N product example. From [14], distributed under a Creative Commons license (Attribution-Noncommercial)	12
Figure 2.8 Tensor diagram of general n-mode unfolding. Modified from [12].	14
Figure 2.9. An example of a mixed signal. From [32], distributed under the terms of the Creative Commons Attribution 4.0 International License.	17
Figure 2.10 Tensor diagram representation of a general Tucker Decomposition. Modified from [12]	19
Figure 2.11 Tensor diagram representation of a Higher-Order Singular Value Decomposition. Modified from [12].....	20
Figure 2.12. Principal axes of a linear dataset. Reproduced using the code from [33], distributed under The MIT License.	22
Figure 2.13. Projection of a dataset to single principal component. Reproduced using the code from [33], distributed under The MIT License.	23
Figure 2.14. Original picture of various faces. The dataset is from [34]	23
Figure 2.15. An example of eigenfaces. The dataset is from [34]	23

Figure 3.1. Tensor representation of a three-dimensional microseismic data tensor	27
Figure 3.2. Graphical visualization of a three-dimensional microseismic data tensor.....	27
Figure 3.3. Tensor representation of a four-dimensional microseismic data tensor	28
Figure 3.4. Truncation of a Singular Value Decomposition	30
Figure 3.5. The graph of singular values of a microseismic data matrix at location 22 ..	31
Figure 3.6. Effect of an artificial noise addition to the general singular values representation of a dataset.....	32
Figure 3.7. Singular values of a decomposed three-dimensional microseismic tensor....	33
Figure 3.8. The comparison of a 12-singular-values dataset (left) and a 96-singular- values dataset (right).....	34
Figure 3.9. Microseismic data compression results	36
Figure 3.10. Example of a far and near source position.....	38
Figure 3.11. Graphical definition of polarization.....	39
Figure 3.12. An example of synthetic data tensor.....	40
Figure 3.13. Comparison between logarithms of first (left) and second (right) functions for near and far sources.....	41
Figure 3.14. Representation of Decay as a function of position in x-, and y- direction...	42
Figure 3.15. Representation of Decay as a function of vertical position and distance between the source and receivers.....	43
Figure 3.16. Tensor representation of HOSVD performed on a four-dimensional microseismic data tensor.....	46
Figure 3.17. Fourth factor matrix and hidden structures related to time.....	47
Figure 3.18 Graphical description of the proposed methodology.....	48

LIST OF TABLES

	Page
Table 2.1 Tensor Notation Summary Table.	16

1. INTRODUCTION

1.1. Preface

Unconventional reservoirs, such as shales and tight gas, continue to play an ever-increasing role in the world with the volume of crude oil production from these reservoirs steadily increasing year after year, in a nearly exponential fashion. This can be seen from Figure 1.1, where the US conventional crude curve is on a declining trend, whilst the US tight crude grows considerably and is expected to grow in similar manner in the future.

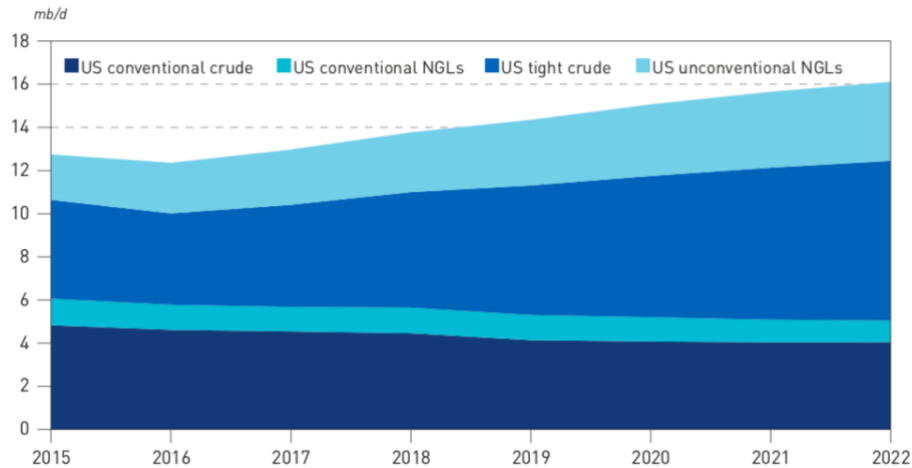


Figure 1.1 U.S. Crude and Natural Gas Liquids (NGLs) production. From [29], distributed for educational and other non-commercial purposes.

The growth of unconventional reservoirs production is mainly connected to the progress made with hydraulic fracturing, which is an important part of tight crude oil extraction. Despite the recent advances with technology, however, there are still issues that should be addressed. Currently, the great concern surrounds the environmental effect of hydraulic fracturing operations [24, 37]. The environmental effect is mostly connected

to the ineffectiveness of the oil industry to guarantee well integrity and control over fracture propagation. As a consequence, the potential contamination and poisoning of local water sources is a great concern in society [37].

The key to address this concern is to ensure well integrity and high standards of safety and responsibility, while performing fluid injection operations. That can be achieved by imposing a set of rigorous operating principles within a company, but more importantly, by improving the technologies such as microseismic data, and improving understanding of their connection to hydraulic fracturing. Ideally, there should be an efficient and fast quality control of fracture propagation during fracturing operations that does not rely on conventional time-consuming simulations.

Microseismic data can be connected to hydraulic fracturing processes. It was recently found that natural fractures during hydraulic operations can play a minor role in generating the microseismicity or guiding fracture growth, hence indicating that the microseismic shearing is associated with hydraulic fractures [36]. To this end, microseismic technology is arguably the most effective way to understand underground processes during hydraulic operations [1, 5] and to get insights about rock structure and geometry [2, 25], as well as stratigraphic deformations. The paper from Schlumberger [4] emphasizes this idea by providing case studies from Texas that further illustrate the value microseismic data has given to geoscientists and petroleum engineers dealing with stimulation operations.

This thesis concentrates on the application of recent advances in machine learning, more specifically tensor decompositions, to analyze high-dimensional microseismic data.

It is shown that with a decomposed representation of microseismic data, hidden structures, that were not visible by traditional visualization methods, could be extracted. These structures, for example, include the data matrix that corresponds to a variation of parameters with time. This helps track and predict changes in source location, which could potentially be tied to fracture propagation. Additionally, the structure of the decomposition can hint to a seismic wave source mechanism, which can imply some information about fractures and potentially help with their visualization and characterization. Furthermore, the application of tensor decompositions is an efficient model reduction technique, and it helped to achieve 73% data compression, which simplifies real-time research during hydraulic operations. These ideas are explored further in Section 3 of the paper.

1.2. Summary of Literature Review

The research in the area of microseismic data connection to hydraulic fracturing is currently very active [6], with recent papers proposing ways to analyze fractures [27] or estimate rock permeability [28]. However, despite the research done with microseismic data technology, there is a vast room for improvement both in terms of the physics behind microseismic imaging, and in terms of the current data science approach to microseismic data processing [26].

It is possible to use tools from modern data science and machine learning to further advance research in this area. This work is focused on a specific type of machine learning techniques called Blind Signal Separation (BSS). Several machine learning techniques within BSS, such as tensor decompositions and multiway component analysis, can be a

promising tool to obtain more information from microseismic data while improving safety and efficiency of hydraulic fracturing.

Tensor decompositions, and Tucker decompositions (TD) in particular, have been successfully applied to various fields. Examples of TD applications to other fields include epidemiology, where HOSVD (a type of TD) was utilized as a basis for a real-time event detection software, particularly for early detection of disease outbreaks [9]. Additionally, the HOSVD technique has been used for image and face recognition [10, 38] as it is possible to extract certain parts of a picture and recombine it in a specific way. This helps to create a compilation of various faces which can potentially be used for suspect sketching. Also, there are cases of using HOSVD in neuroscience to track changes in brain and help detect problems at an early stage [11]. In petroleum engineering the HOSVD has been used as a permeability reparametrization tool tied to history matching processes [30, 31].

Despite the successful application of tensor decompositions in other areas, however, an insufficient amount of research is done to apply the HOSVD and MWCA techniques to microseismic data analysis.

The fMRI and EEG recordings used in [11] are visually similar to general microseismic data structures. Considering that the decomposed data analysis approach used in [11] helped to see hidden structures related to changes in brain, the HOSVD technique has a big potential with hidden structures extraction for microseismic data, given their structural similarities.

This is the main reason behind the choice of HOSVD and MWCA as a data analysis tool for microseismic data and the main reason why the application of the HOSVD technique to microseismic data is relevant and important.

1.3. Thesis Objective and Scope

The main objective of this research is, hence, to apply HOSVD and MWCA to microseismic data and propose a workflow to extract more information from the decomposed structure alone, which leads to more insights about seismic signal source, which ultimately can infer information about fractures without using conventional time-consuming simulations.

Other goals include investigation of a possibility to simultaneously compress and denoise multidimensional microseismic data, as well as to outline steps for further research connected to the technology. Given how recent the technology is, there is vast opportunity for possible improvements of both HOSVD and MWCA, which could help to discover even more hidden information from microseismic data. Hence, tensor decomposition can be an important addition to a set of hydraulic fracturing quality assessment tools.

1.4. Thesis Organization

The thesis proposal is organized as follows. In Section 2, a brief introduction to the important concepts related to the proposed workflow is given. The concepts include microseismic monitoring, seismic inversion, overview of machine learning techniques, SVD, HOSVD and MWCA. Additionally, in Section 2 a brief mathematical background for tensors and tensor products is presented.

Section 3 offers an exact workflow, its practical application to a numerical example, results and main findings. The section consists of three different subsections representing each step of the proposed methodology. Subsection 3.4 includes a complete overview of the methodology.

Conclusions are summarized in Section 4. Parts of this work are based on the technical paper SPE-195522-MS that is expected to be published and presented during the SPE Europec featured at 81st EAGE Annual Conference & Exhibition.

2. PROBLEM BACKGROUND

This section gives a brief introduction to important concepts underlying the theory of the proposed workflow.

2.1. Introduction to Microseismic Monitoring

Microseismic monitoring is a technique that records and locates microseismic events, thus generating microseismic data. Those events are generated by a seismic wave, which, in turn, could be caused by high rates and high pressures of hydraulic fracturing [3]. An example of a physical set up of a microseismic experiment is depicted in Figure 2.1, where treatment well represents a well upon which the fracturing is performed, while monitoring well is used to detect seismic wave with a set of receivers.

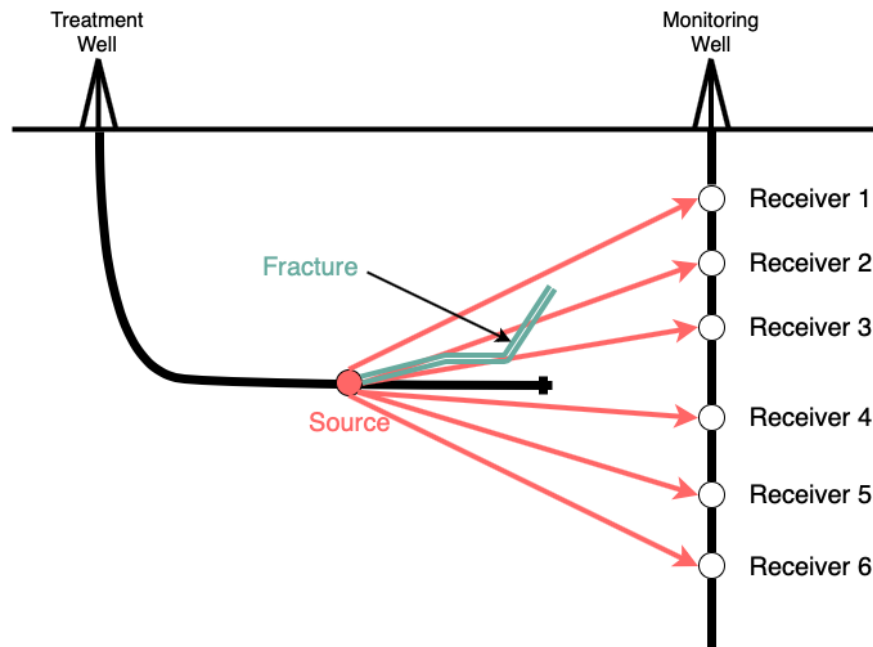


Figure 2.1 Graphical description of a possible microseismic event.

The set of sensors, or receivers, depicted on the right side of the Figure 2.1, are used to record the time and intensity of seismic wave arrivals. Each sensor measures displacement caused by a seismic wave in X- Y- and Z-directions.

There are two types of waves of interest: P-wave, or Primary wave, and S-wave, or Secondary wave [39]. The primary wave is a longitudinal wave with motion along the line of propagation in the case of isotropic and homogeneous solids. It is generally faster and, therefore, it is the first one to appear on a seismogram. The secondary wave is a transverse wave where the motion is perpendicular to the direction of propagation. It follows the Primary wave.

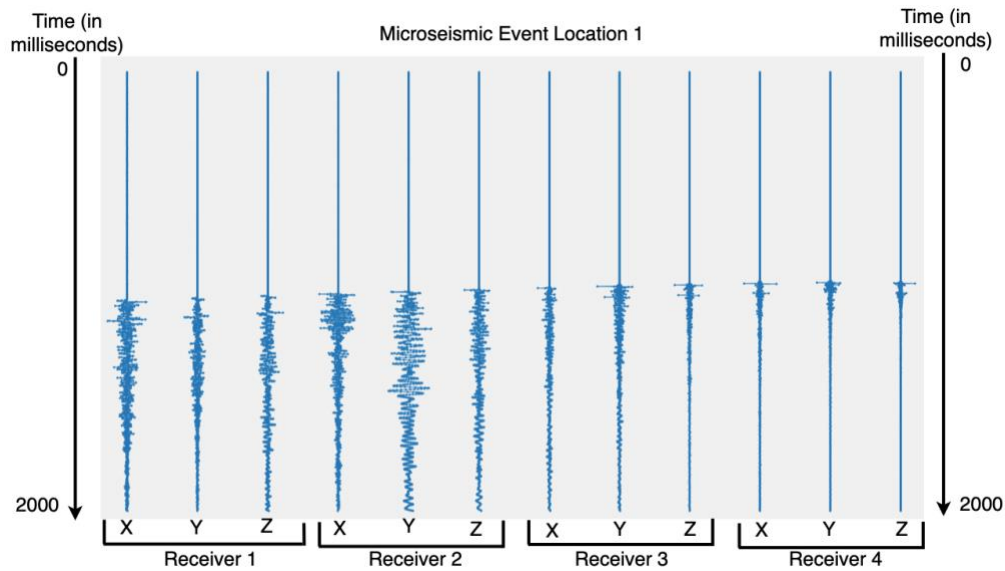


Figure 2.2. General structure of microseismic data.

An example of recorded data for one location is shown in Figure 2.2, where x-axis represents variation of receivers and y-axis represents change in time, in milliseconds.

2.2. The Problem and Mathematical Framework / Seismic Inversion

2.2.1. Seismic Inversion

With a correct analysis of microseismic data recorded by sensors it is possible to get information about rock structure, faults and fractures. The process of transforming recorded seismic data into a meaningful description of a reservoir is called seismic inversion [40]. Therefore, the problem, as stated in subsection 1.3, is connected to seismic inversion. In order to proceed with a solution to the stated problem, it is important to give a brief background of the appropriate mathematical framework.

2.2.2. Mathematical Framework

Subsection 2.2.2.1 starts with an introduction to general tensor notation used in the workflow.

In this work, Tensor Diagrams were used, as defined in [12]. Subsection 2.2.2.2 is intended to give a brief introduction to the notation, since there are no general rules for tensor diagrams representation in the literature.

The subsequent sections (2.2.2.3-2.2.2.8) provide clarification to some of the concepts of tensor algebra, such as unfolding, as it could be performed in several ways. Subsection 2.2.2.8 is followed by subsection 2.2.2.9 which contains Table 2.1 that summarizes the main mathematical concepts.

2.2.2.1. Tensor Notation

In this work, vectors are described with underlined letters (a), while matrices are denoted with double-underlined letters (B). Tensors are denoted with triple-underlined letters (G). Tensors with four and more modes (also called ways, or dimensions) are also

triple-underlined. It is important to emphasize that the number of lines does not depend on a mode of a tensor.

Additionally, a matrix could be denoted as such: $\underline{\underline{B}}^{(n)}$. This is a representation of component matrices, or factor matrices. This should not be confused with $\underline{\underline{X}}_{(n)}$, which stands for a mode-n unfolding of a tensor $\underline{\underline{X}}$.

2.2.2.2. Tensor Diagrams

Tensor diagrams are useful for visualization of complex tensors. There is no general standard for tensor diagrams representation, therefore in this work the definitions presented in [12] were used. They are shown in Figures 2.3-2.6.

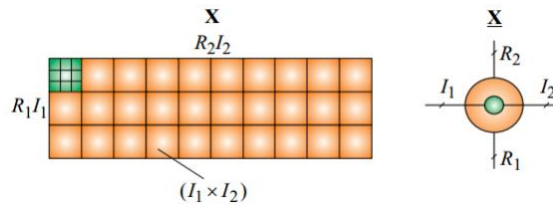


Figure 2.3 Fourth-order tensor and its tensor diagram representation. From [12], distributed with author's permission.

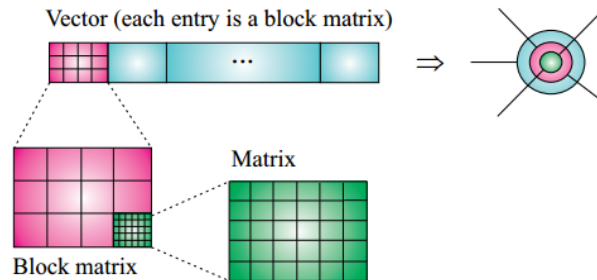


Figure 2.4 Fifth-order tensor and its tensor diagram representation. From [12], distributed with author's permission.

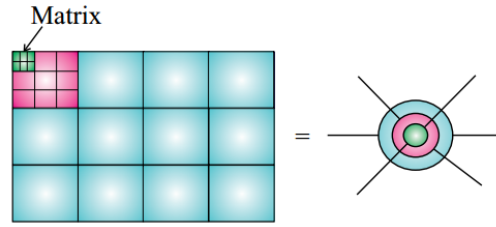


Figure 2.5 Sixth-order tensor and its tensor diagram representation. From [12], distributed with author's permission.

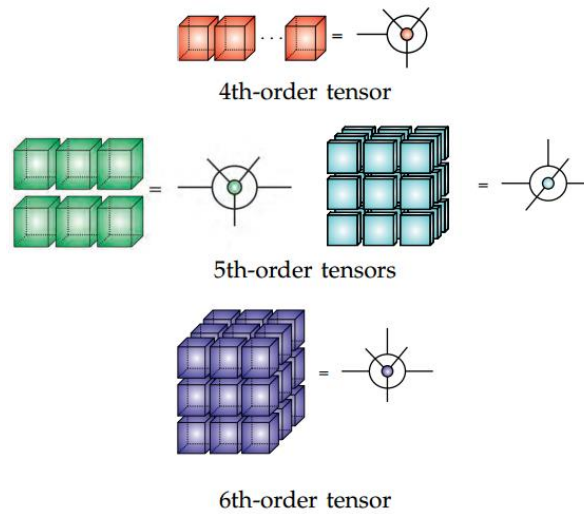


Figure 2.6 Graphical tensor diagrams representations of higher-order tensors. From [12], distributed with author's permission.

2.2.2.3. Mode-N Product

The mode- n product of a general tensor $\underline{\underline{A}} \in R^{I_1 \times \dots \times I_N}$ and a general vector $\underline{b} \in R^{I_n}$ is defined as a tensor $\underline{\underline{C}} = \underline{\underline{A}} \times_n \underline{b}$, where $\underline{\underline{C}} \in R^{I_1 \times \dots \times I_{n-1} \times I_{n+1} \times \dots \times I_N}$.

For example, 3-dimensional tensor mode-1 multiplication with a vector will reduce the rank of a tensor to 2 (Fig. 2.7).

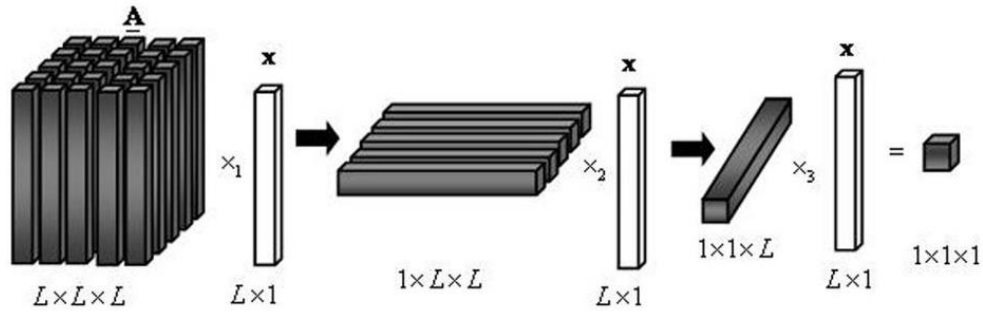


Figure 2.7 Mode-N product example. From [14], distributed under a Creative Commons license (Attribution-Noncommercial)

A mode-1 or mode-2 multiplication of a 3-dimensional tensor with a matrix instead of a vector is based on the similar principles. However, one mode of the matrix will add up to the tensor and, therefore, the rank of the tensor will be the same as prior to the mode-n product.

2.2.2.4. Outer Product

The outer product of a tensor $\underline{A} \in R^{I_1 \times \dots \times I_N}$ and a tensor $\underline{B} \in R^{J_1 \times \dots \times J_M}$, as defined in Eq. 2.1, is a tensor \underline{C} that has $(N+M)$ th order.

$$\underline{C} = \underline{A} \circ \underline{B} \quad (2.1)$$

In other words, the result of an outer product contains all modes of all matrices (or vectors, or tensors) that the product contained. Outer product of two vectors will produce a matrix.

2.2.2.5. Kronecker Product

The Kronecker product of a tensor $\underline{A} \in R^{I_1 \times \dots \times I_N}$ and a tensor $\underline{B} \in R^{J_1 \times \dots \times J_N}$, is defined in Eq. 2.2, where the tensor \underline{C} has an $(N \cdot N)$ th order.

$$\underline{\underline{C}} = \underline{\underline{A}} \otimes \underline{\underline{B}} \quad (2.2)$$

In other words, the traditional matrix multiplication contrasts the Kronecker Product, since instead of actual multiplication of 2x2 matrix $\underline{\underline{A}}$ and 2x2 matrix $\underline{\underline{B}}$, each element of $\underline{\underline{A}}$ is multiplied by the whole matrix $\underline{\underline{B}}$. As a result, a product of two 2x2 matrices will result in one 4x4 matrix, as it is shown in Eq. 2.3.

$$\underline{\underline{A}} \otimes \underline{\underline{B}} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \otimes \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a \cdot \begin{bmatrix} e & f \\ g & h \end{bmatrix} & b \cdot \begin{bmatrix} e & f \\ g & h \end{bmatrix} \\ c \cdot \begin{bmatrix} e & f \\ g & h \end{bmatrix} & d \cdot \begin{bmatrix} e & f \\ g & h \end{bmatrix} \end{bmatrix} \quad (2.3)$$

2.2.2.6. Khatri-Rao Product

The Khatri-Rao product is the “matching column-wise” Kronecker product [15]. The Khatri-Rao product of a matrix $\underline{\underline{A}} \in R^{I \times J}$ and a matrix $\underline{\underline{B}} \in R^{K \times J}$ is defined in equation 2.4, where the matrix $\underline{\underline{C}} \in R^{IK \times J}$.

$$\underline{\underline{C}} = \underline{\underline{A}} \odot \underline{\underline{B}} \quad (2.4)$$

To simplify the definition above, the Khatri-Rao product can be considered an attempt to do a Kronecker product column-wise. An example is given in Eq. 2.5, where the columns of $\underline{\underline{A}}$ are taken to perform a Kronecker product with columns from $\underline{\underline{B}}$.

$$\underline{\underline{A}} \odot \underline{\underline{B}} = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \odot \begin{bmatrix} g & h & i \\ j & k & l \end{bmatrix} = \begin{bmatrix} ag & bh & ci \\ aj & bk & cl \\ dg & eh & fi \\ dj & ek & fl \end{bmatrix} \quad (2.5)$$

2.2.2.7. Hadamard Product

The Hadamard product of a matrix $\underline{\underline{A}} \in R^{I \times J}$ and a matrix $\underline{\underline{B}} \in R^{I \times J}$ is defined in Eq. 2.6, where the matrix $\underline{\underline{C}} \in R^{I \times J}$.

$$\underline{\underline{C}} = \underline{\underline{A}} \circledast \underline{\underline{B}} \quad (2.6)$$

The Hadamard product, as defined in Eq. 2.6, is simply an elementwise multiplication of matrices $\underline{\underline{A}}$ and $\underline{\underline{B}}$, which could be extended to tensors without changes to the approach. Because of the definition, it is important to have two matrices, or tensors, of the same size. An example of Hadamard product is given in Eq. 2.7.

$$\underline{\underline{A}} \circledast \underline{\underline{B}} = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \circledast \begin{bmatrix} g & h & i \\ j & k & l \end{bmatrix} = \begin{bmatrix} ag & bh & ci \\ dj & ek & fl \end{bmatrix} \quad (2.7)$$

2.2.2.8. Mode-N Unfolding

For a tensor $\underline{\underline{\underline{A}}} \in R^{I_1 \times I_2 \times \dots \times I_N}$ mode-n unfolding (also called matricization, or flattening) is defined as $\underline{\underline{\underline{A}}}_{(n)} \in R^{I_n \times (I_1 \times I_2 \times \dots \times I_{n-1} \times I_{n+1} \times \dots \times I_N)}$. Therefore, the result of a mode-n unfolding of a tensor $\underline{\underline{\underline{A}}}$ is a matrix. The process of unfolding is presented in Figure 2.8 with tensor diagrams, as well as clarified with an example, presented in equations 2.10-2.14. The background to the Tensor Diagrams notation is detailed further in subsection 2.2.2.2.

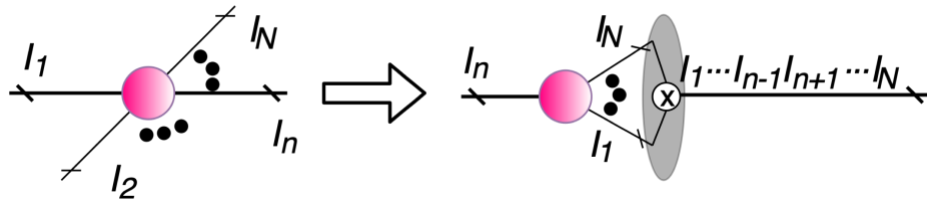


Figure 2.8 Tensor diagram of general n-mode unfolding. Modified from [12].

It should be noted that the definition of unfolding used in this work is based on the definition in [15]. However, there are different ways to define unfolding. For example, Python TensorLy (high level API for tensor methods in Python) is based on a different definition of unfolding, for the reasons of improved computational power in comparison to standard methods [17].

An example of three-modes (3x4x2) tensor $\underline{\underline{X}}$ suitable for unfolding is given in equations 2.10-2.11, where each of the equation is a representation of the tensor's slice.

Possible unfoldings of the tensor $\underline{\underline{X}}$ are given in equations 2.12-2.14.

$$\underline{\underline{X}}(:, :, 1) = \begin{bmatrix} 0 & 2 & 4 & 6 \\ 8 & 10 & 12 & 14 \\ 16 & 18 & 20 & 22 \end{bmatrix} \quad (2.10)$$

$$\underline{\underline{X}}(:, :, 2) = \begin{bmatrix} 1 & 3 & 5 & 7 \\ 9 & 11 & 13 & 15 \\ 17 & 19 & 21 & 23 \end{bmatrix} \quad (2.11)$$

Mode-1 unfolding:

$$\underline{\underline{X}}_{(1)} = \begin{bmatrix} 0 & 2 & 4 & 6 & 1 & 3 & 5 & 7 \\ 8 & 10 & 12 & 14 & 9 & 11 & 13 & 15 \\ 16 & 18 & 20 & 22 & 17 & 19 & 21 & 23 \end{bmatrix} \quad (2.12)$$

Mode-2 unfolding:

$$\underline{\underline{X}}_{(2)} = \begin{bmatrix} 0 & 8 & 16 & 1 & 9 & 17 \\ 2 & 10 & 18 & 3 & 11 & 19 \\ 4 & 12 & 20 & 5 & 13 & 21 \\ 6 & 14 & 22 & 7 & 15 & 23 \end{bmatrix} \quad (2.13)$$

Mode-3 unfolding:

$$\underline{\underline{X}}_{(3)} = \begin{bmatrix} 0 & 8 & 16 & 2 & 10 & 18 & 4 & 12 & 20 & 6 & 14 & 22 \\ 1 & 9 & 17 & 3 & 11 & 19 & 5 & 13 & 21 & 7 & 15 & 23 \end{bmatrix} \quad (2.14)$$

2.2.2.9. Notation Summary Table

Table 2.1 is a quick summary of all the concepts described in subsections 2.2.2.3-2.2.2.8.

Table 2.1 Tensor Notation Summary Table

Notation	Meaning
$\underline{\underline{X}}_{(n)}$	Mode-n unfolding of $\underline{\underline{X}}$
$\underline{\underline{C}} = \underline{\underline{A}} \times_n \underline{\underline{b}}$	Mode-n product
$\underline{\underline{C}} = \underline{\underline{a}} \circ \underline{\underline{b}}$	Outer product
$\underline{\underline{C}} = \underline{\underline{A}} \otimes \underline{\underline{B}}$	Kronecker product
$\underline{\underline{C}} = \underline{\underline{A}} \odot \underline{\underline{B}}$	Khatri-Rao product
$\underline{\underline{C}} = \underline{\underline{A}} \circledast \underline{\underline{B}}$	Hadamard product
$\underline{\underline{X}} = [[\underline{\underline{A}}; \underline{\underline{B}}^{(1)}, \dots, \underline{\underline{B}}^{(n)}]]$	The same as: $\underline{\underline{X}} = \underline{\underline{A}} \times_1 \underline{\underline{B}}^{(1)} \times_2 \underline{\underline{B}}^{(2)} \dots \times_n \underline{\underline{B}}^{(n)}$

2.3. Machine Learning Through HOSVD – Introduction to HOSVD / MWCA

2.3.1. Machine Learning Techniques

Machine learning, which is a part of a data science toolkit, is an emerging and powerful tool that gives a framework to work with microseismic data. The specific type of machine learning techniques that are intended to implement to the workflow are generally referred to as Blind Signal Separation (BSS), also known as Blind Source Separation [8].

BSS is a machine learning technique of original data recovery from mixed signals generated by several sources [18]. An example of a mixed signals source is depicted in Fig. 2.9 [32], where a transparent picture of a standing person is printed on top of another picture. The idea is to use BSS techniques to separate them and restore two original pictures. Another practical application of BSS is noise and speech recognition in recording devices [18].

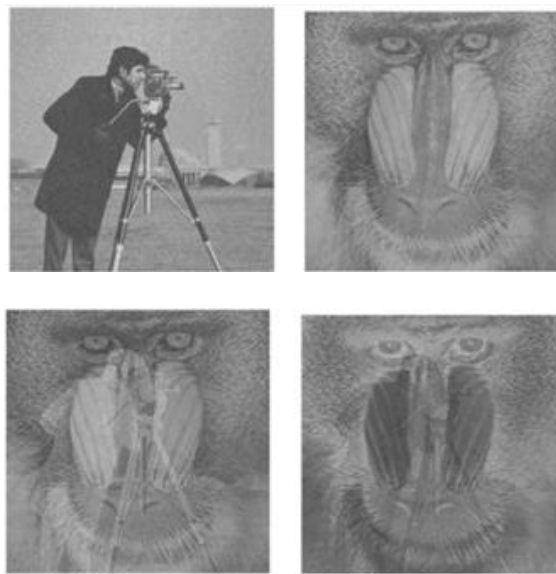


Figure 2.9. An example of a mixed signal. From [32], distributed under the terms of the Creative Commons Attribution 4.0 International License.

There are various tools within BSS, including tensor decompositions, Singular Value Decomposition (SVD) and component analysis. Moreover, there are different ways to analyze components, however, the emphasis of this research is on MWCA and Principal Component Analysis (PCA). PCA is sometimes referred as Proper Orthogonal Decomposition (POD) in petroleum industry and history matching [19, 23].

2.3.2. Singular Value Decomposition (SVD)

The SVD of a matrix $\underline{\underline{A}}$ is defined in equation (2.15), where $\underline{\underline{U}}$ is a matrix with left singular vectors, $\underline{\underline{V}}$ is a matrix with right singular vectors and $\underline{\underline{D}}$ is a diagonal matrix with singular values [16].

$$\underline{\underline{A}} = \underline{\underline{U}} \underline{\underline{D}} \underline{\underline{V}}^T \quad (2.15)$$

The most important properties to note here are as follows:

- $\underline{\underline{U}}$ and $\underline{\underline{V}}$ are orthogonal matrices, therefore, $\underline{\underline{U}}^T = \underline{\underline{U}}^{-1}$ and $\underline{\underline{V}}^T = \underline{\underline{V}}^{-1}$.
- $\underline{\underline{D}}$ is a diagonal matrix with a diagonal structured in a decreasing order.

2.3.3. Tucker Decompositions

With extensive usage of tensor data structures, tensor decomposition techniques rapidly emerged. The tensor decompositions can be considered to be the next important step in data analysis [12] and this work shows the applicability of the methods to current microseismic data research. This work is concentrated on Tucker Decomposition. The general structure of Tucker decomposition is shown in Figure 2.10 in the form of Tensor Diagram notation with core tensor $\underline{\underline{\underline{S}}} \in R^{R_1 \times R_2 \times \dots \times R_5}$ and five factor matrices $\underline{\underline{\underline{B}}}^{(n)} \in R^{I_n \times R_n}$. As seen from the diagram, the number of factor matrices depends on the order of the core tensor $\underline{\underline{\underline{S}}}$. Mathematically, Tucker Decomposition could be written as in Eq. 2.16, where $\underline{\underline{\underline{T}}} \in R^{I_1 \times I_2 \times \dots \times I_5}$ is the original tensor before decomposition.

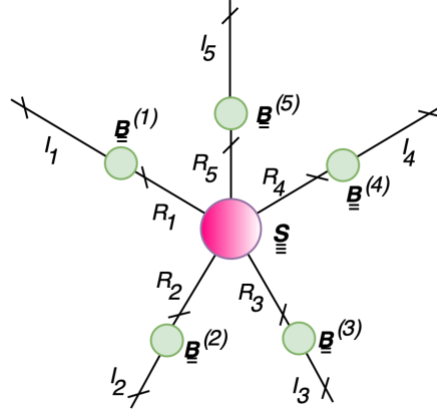


Figure 2.10 Tensor diagram representation of a general Tucker Decomposition.
Modified from [12]

$$\underline{\underline{T}} \cong \underline{\underline{S}} \times_1 \underline{\underline{B}}^{(1)} \times_2 \underline{\underline{B}}^{(2)} \times_3 \dots \times_5 \underline{\underline{B}}^{(5)} = [[\underline{\underline{S}}; \underline{\underline{B}}^{(1)}, \underline{\underline{B}}^{(2)}, \dots, \underline{\underline{B}}^{(5)}]] \quad (2.16)$$


Additionally, the Tucker decomposition can be expressed as a Kronecker product as in Eq. 2.17, with numbering of factor matrices from 1 to N. It is generally nonunique, although constraints may be imposed and the rule will not hold true [12].

$$\underline{\underline{T}}_{(n)} \cong \underline{\underline{B}}^{(n)} \underline{\underline{S}}_{(n)} \left(\underline{\underline{B}}^{(1)} \otimes \dots \otimes \underline{\underline{B}}^{(n-1)} \otimes \underline{\underline{B}}^{(n+1)} \otimes \dots \otimes \underline{\underline{B}}^{(N)} \right)^T \quad (2.17)$$

Some selected factor matrices can be identity matrices. If this holds true, the decomposition is called Tucker-(K,N), where N is the order of the core tensor and K is the number of non-identity factor matrices. For example, Tucker-(2,3) has 3-2=1 identity matrix. This particular case with only one factor matrix is often called Tucker-2 Decomposition [12]. Tucker-2 decomposition narrows to Eq. 2.18 [22].

$$\underline{\underline{T}} \cong \underline{\underline{S}} \times_1 \underline{\underline{B}}^{(1)} \times_2 \underline{\underline{B}}^{(2)} = [[\underline{\underline{G}}; \underline{\underline{B}}^{(1)}, \underline{\underline{B}}^{(2)}]] \quad (2.18)$$

2.3.4. Higher-Order Singular Value Decomposition (HOSVD)

Higher-Order Singular Value Decomposition (HOSVD), also known as Multilinear Singular Value Decomposition (MLSVD), is a special form of a Tucker decomposition where all factor matrices $\underline{\underline{B}}^{(n)}$ are orthogonal and the core tensor $\underline{\underline{S}} \in R^{R_1 \times R_2 \times R_3 \times R_4}$ is orthogonal, as shown in Figure 2.11 and Eq. 2.19. The symbol  represents orthogonality. It is not to confuse the symbol with a type of motion on a fracture representation common in geophysics.

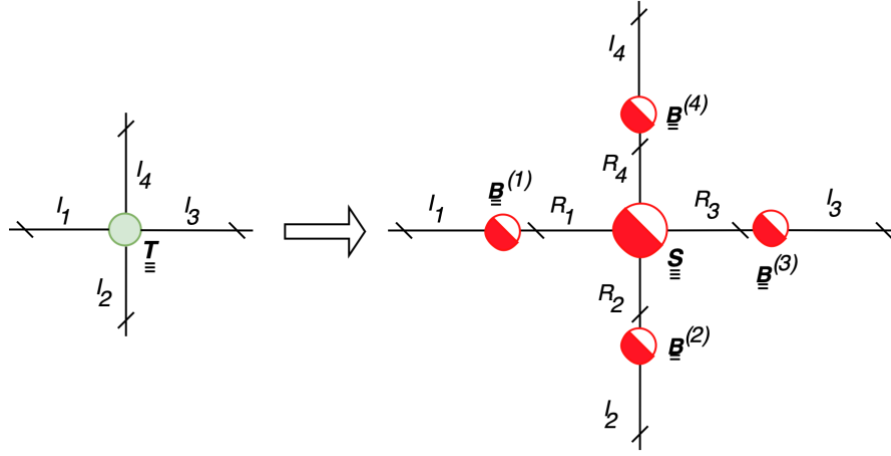


Figure 2.11 Tensor diagram representation of a Higher-Order Singular Value Decomposition. Modified from [12]

$$\underline{\underline{T}} \cong \underline{\underline{S}} \times_1 \underline{\underline{B}}^{(1)} \times_2 \underline{\underline{B}}^{(2)} \times_3 \underline{\underline{B}}^{(3)} \times_4 \underline{\underline{B}}^{(4)} = [[\underline{\underline{S}}; \underline{\underline{B}}^{(1)}, \underline{\underline{B}}^{(2)}, \underline{\underline{B}}^{(3)}, \underline{\underline{B}}^{(4)}]] \quad (2.19)$$

Orthogonality of the core tensor implies that each slice of the core tensor is orthogonal with its every other slice, as shown in Eq. 2.20. Additionally, the Frobenius norm of the first slice is the biggest. The Frobenius norm decreases whilst marching along slices. In that sense, the structure reminds the SVD, since the values of the diagonal matrix

$\underline{\underline{D}}$ in the SVD are likewise in a decreasing order (see subsection 2.3.2 for details). The Frobenius norm of a general matrix $\underline{\underline{A}}$ is defined in Eq. 2.21. The $\langle \dots \rangle$ notation, used in Eq. 2.20, represents an inner product of all the objects enclosed in brackets.

$$\langle \underline{\underline{S}}_{:,k,:}, \underline{\underline{S}}_{:,l,:} \rangle = 0, \quad \text{for } k \neq l \quad (2.20)$$

$$\|\underline{\underline{A}}\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} \quad (2.21)$$

The factor matrices could be computed while performing the SVD with mode- n unfoldings of $\underline{\underline{T}}$. That way, the factor matrices $\underline{\underline{B}}^{(n)}$ could be found which can later be used to approximate the core tensor $\underline{\underline{S}}$ by consecutive mode- n multiplication of the original tensor $\underline{\underline{T}}$ with all transposed factor matrices, as shown in equation 2.22 [12].

$$\underline{\underline{S}} = \underline{\underline{T}} \times_1 \underline{\underline{B}}^{(1)T} \times_2 \underline{\underline{B}}^{(2)T} \times_3 \dots \times_n \underline{\underline{B}}^{(n)T} = [[\underline{\underline{T}}; \underline{\underline{B}}^{(1)T}, \underline{\underline{B}}^{(2)T}, \dots, \underline{\underline{B}}^{(n)T}]] \quad (2.22)$$

Another method to compute the core tensor is to use a modified version of Grassman Manifolds [20], where mode- n unfoldings of $\underline{\underline{T}}$ are transposed before applying the SVD and then to use the matrix $\underline{\underline{V}}$ instead (the notation is as defined in subsection 2.3.2). More information about the method can be found in [20].

2.3.5. Multiway Component Analysis and Principal Component Analysis

The Principal Component Analysis (PCA) is a type of group analysis and an important technology to identify links between data. With PCA, hidden structures can be

found, while analyzing correlation and variability of the components across data structures [12, 21].

Another term for principal components is singular vectors or eigenvectors. The values, by which the principal components are scaled, are determined by eigenvalues, or singular values. Singular vectors can help to see principal axes, which could be used to describe the dataset [33]. For example, a distorted one-dimensional dataset can be described with only one eigenvector, as shown in Figure 2.12, while the second eigenvector, and a smaller vector, helps to describe the dataset variation completely.

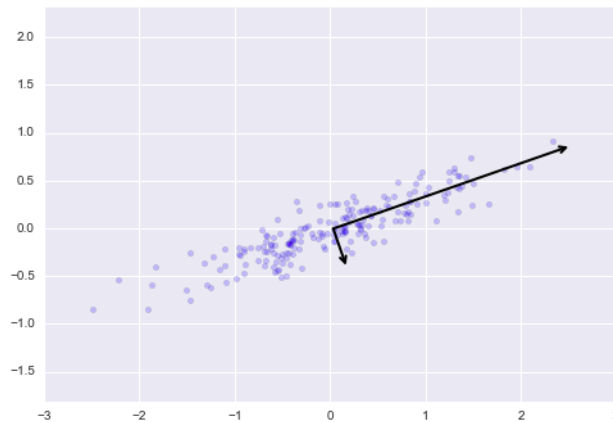


Figure 2.12. Principal axes of a linear dataset. Reproduced using the code from [33], distributed under The MIT License.

The PCA helps make a projection of the complete dataset to single principal component by ignoring smaller singular vectors, as shown in Figure 2.13.

Similar to the example shown in Fig. 2.13, the PCA could be performed with a general image. For example, principal components can extract most significant characteristics of an image. This technique can serve as a basis for image detection and can help to create so called eigenfaces, e.g. visualization of the images associated with the

first several principal components [33]. The original image is shown in Fig. 2.14, while the eigenfaces created from these images are shown in Fig. 2.15.

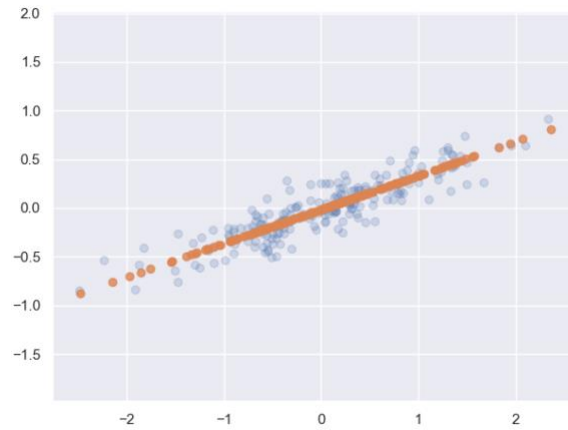


Figure 2.13. Projection of a dataset to single principal component. Reproduced using the code from [33], distributed under The MIT License.



Figure 2.14. Original picture of various faces. The dataset is from [34]



Figure 2.15. An example of eigenfaces. The dataset is from [34]

Principal Component Analysis has a connection to the SVD, described in subsection 2.3.2, since the diagonal matrix \underline{D} consists of singular values examined with PCA, while two other matrices after decomposition represent a set of singular vectors.

The Multiway Component Analysis is a natural extension of a general component analysis to high-dimensional structures. The application of MWCA assumes that dimensions of the high-dimensional data are naturally split into different modalities [21]. Therefore, by including additional modes, extra connections across the multi-block data structures can be obtained, which helps to determine even more hidden structures across the data, comparing to a standard two-way component analysis. This is the main idea of application of tensorization and MWCA to microseismic data, which will be explored further in Section 3.

2.4. Hidden Structures

The term hidden structures without additional clarification is unclear. A particular meaning behind the term in this work can be split into two main categories.

The first category contains structures related to information about a source of a seismic wave. Tensor decompositions and constraints were used to get an idea about source location and source mechanism from the decomposed structure alone. The work is concentrated around the first category of hidden structures, and the methodology is discussed in details in Section 3.

The second category contains structures related to variation of seismic parameters in time. Future research with four-dimensional tensors could give an idea about trends in source location which could imply information about trends of fracture propagation.

However, the goal of this work is only to introduce the need for research in the area of second category of hidden structures, and more research is needed to propose and generalize the methodology. To summarize, the second category of hidden structures are introduced in this work, however, they are out of the scope of this research.

3. PROPOSED METHODOLOGY AND NUMERICAL EXAMPLES

This section contains the proposed workflow of application of MWCA and TD to microseismic data. Each subsection of Section 3 represents a different step of the methodology. The overview of the methodology is presented in subsection 3.4. Finally, the main findings of the research are discussed in subsection 3.5.

3.1. Tensorization of Microseismic Data

The microseismic data used in this work has a general structure depicted in Figure 2.2 of Section 2.

From Fig. 2.2, sample interval of the seismic data used in the research is one millisecond. However, in the case represented in Figure 2.2, there are only four receivers with three directional dependences (X, Y, and Z). In the actual data used in the research there were 12 receivers per location (with three directional dependences, which gives 36 different wavelets per location). This data is stored in a computer as a set of float numbers in the form of a matrix. Therefore, it gives one matrix per location with size 2000x36.

As it was mentioned in subsection 2.3.5, the reason behind working with tensors of a third- or fourth-order, rather than matrices, is that tensors will generally provide more information about the connection of various data structures between each other. Moreover, as it will be proved further in subsection 3.2, higher-order tensor structure will give more efficient model reduction.

In this work 24 different locations of receiver strings have been used. Given that there is one matrix per location, all the matrices were combined together in a third-order

tensor form with size 2000x36x24. The general structure of a microseismic data tensor is presented in Figure 3.1 in a tensor notation form, where mode I_1 is time in milliseconds, mode I_2 represents variation of receivers (x-axis of Fig. 2.2), and mode I_3 represents variation of locations. A three-dimensional visual representation of the tensor is also presented in Figure 3.2.

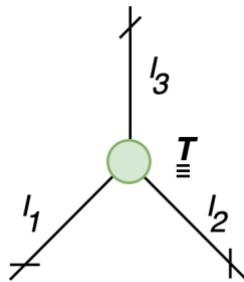


Figure 3.1. Tensor representation of a three-dimensional microseismic data tensor

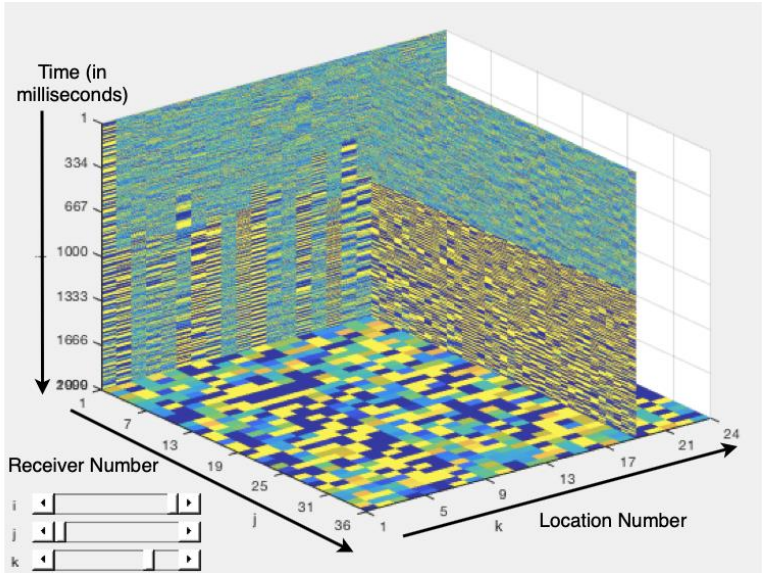


Figure 3.2. Graphical visualization of a three-dimensional microseismic data tensor

In case a 4D seismic data set is available, it is possible to work with microseismic data in a fourth-order tensor form as an alternative. On the whole, 4D seismic research represents the data tensor presented in Fig 3.1 and Fig. 3.2 that is available at different times. Therefore, it is possible to represent all the data in the form of a vector consisted of third-order tensors, which by itself is a four-dimensional structure. The four-dimensional tensor is presented in Fig. 3.3, where the additional fourth mode I_4 represents variation of seismic research in time.

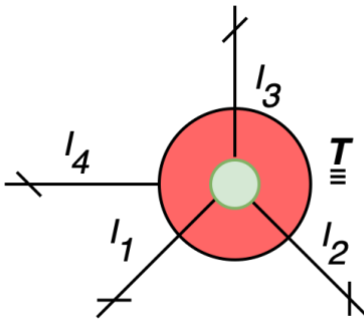


Figure 3.3. Tensor representation of a four-dimensional microseismic data tensor

It is important to mention that some locations within the data tensor in Figure 3.2 had very clear traces of signal, whilst others had no signal at all, or a mixture between signal and pure noise. This variety of data structures helped generalize the denoising and compression techniques and observe that they have vast applicability.

3.2. Denoising and Data Compression

3.2.1. The Importance of Model and Noise Reduction in Microseismic Data

Generally, microseismic data can contain big data structures. In the case provided, there were 24 locations with 12 receivers, implying that only one second of research

contained approximately 1,000,000 data points. Assuming that each float number is stored with 7 decimal digits precision (32 bits per float), one hour of research will contain 14.4 Gigabytes of data structures. This prevents usage of cloud platform solutions and real-time operations, since, considering average uploading capacity of 5 mbps, 1 hour of microseismic research is equivalent to 6.4 hours of data transition. On top of that, the usage of promising recent DAS technology can even further increase the number of variables [7].

Moreover, the microseismic data can contain a significant amount of distortion. To account for that, it is important to either use sensors with a better precision, which could be expensive, or use a denoising technique. Therefore, it is important to analyze the possibility to use the proposed methodology as a noise reduction technique.

3.2.2. Description of the Methodology

The connection between Singular Value Decomposition and Principal Component Analysis was mentioned earlier in subsection 2.3.5. In essence, the matrices after Singular Value Decomposition represent a combination of principal components, or singular vectors, while the diagonal matrix $\underline{\underline{D}}$, as defined in subsection 2.3.2, is a combination of singular values. The singular values are structured in a decreasing order which draws a conclusion that some singular vectors play a bigger role in overall picture than the others. This idea can be traced back to the Fig. 2.12, where one of the principal axes has a much bigger influence than the other ones.

The general structure of the SVD of any matrix $\underline{\underline{A}}$ is shown in Fig. 3.4.

In Fig. 3.4, the columns of $\underline{\underline{U}}$ and rows of $\underline{\underline{V}}^T$ represent singular vectors, or principal components. Their significance is represented by the diagonal matrix $\underline{\underline{D}}$ containing singular values related to the singular vectors. It is useful to think of the singular values as of scale factors.

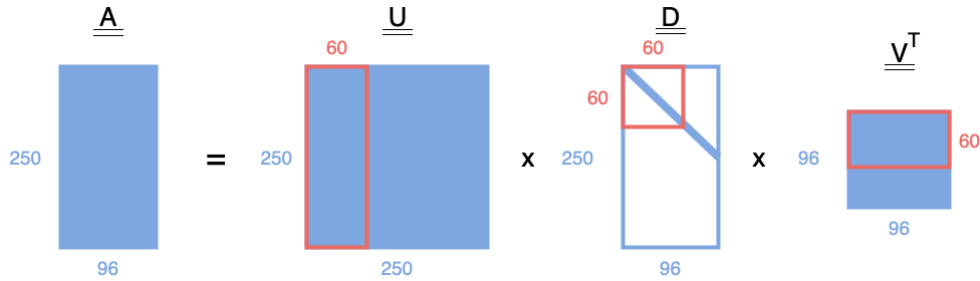


Figure 3.4. Truncation of a Singular Value Decomposition

The plot of singular values after the SVD of microseismic data, similar to the one shown in Fig. 2.2, is presented in Fig. 3.5. It can be seen that a lot of data points within the diagonal matrix $\underline{\underline{D}}$ have values close to zero, therefore, the significance of a large quantity of singular vectors is small. The drop of significance could be spotted more easily whilst using a semi-log graph of normalized singular values, presented on the right side of the Figure 3.5. In Fig. 3.5, the significant drop could be seen at around the 55th singular value.

The idea behind model reduction is to reduce insignificant singular values from the diagonal matrix $\underline{\underline{D}}$, simultaneously reducing the singular vectors corresponding to them. The fact that there is a big number of small singular values can be explained by the general repetitiveness of microseismic data structure. Since the influence of the small singular values to the complete picture is insignificant, the truncated version of singular

values decomposition will give a visually indistinguishable approximation of original microseismic data. The truncation process is described in Fig. 3.4. The threshold for truncation is recommended to be found using normalized semi-log plot of singular values, as shown in Fig. 3.5.

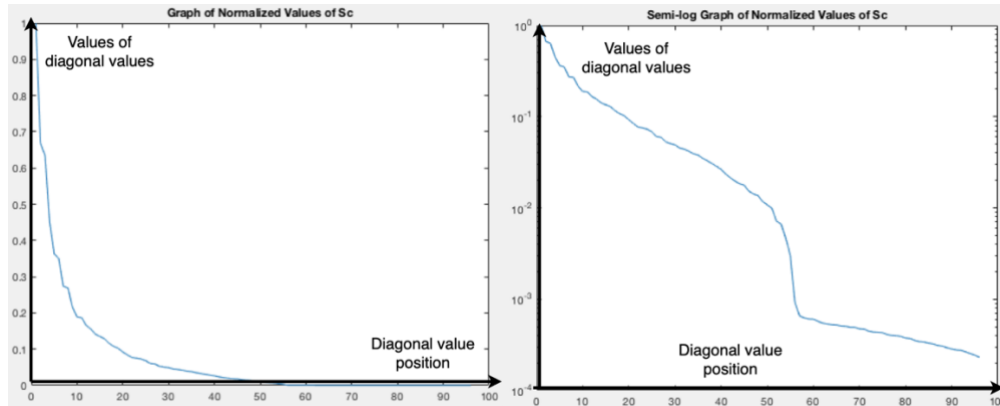


Figure 3.5. The graph of singular values of a microseismic data matrix at location 22

It was found that small singular values represent smaller seismic wave amplitudes and, generally, they are related to noise in a dataset. This can be proved by an artificial addition of noise to the microseismic data. The singular values representation of a microseismic dataset before the experiment is presented in the top part of the Fig. 3.6, while the representation of singular values after addition of artificial noise is presented in the bottom part of the Fig. 3.6. It is clear that the addition of noise affects only the part with the smaller singular values. Therefore, the truncation of the decomposed microseismic data serves both as a model reduction technique and a noise reduction technique.

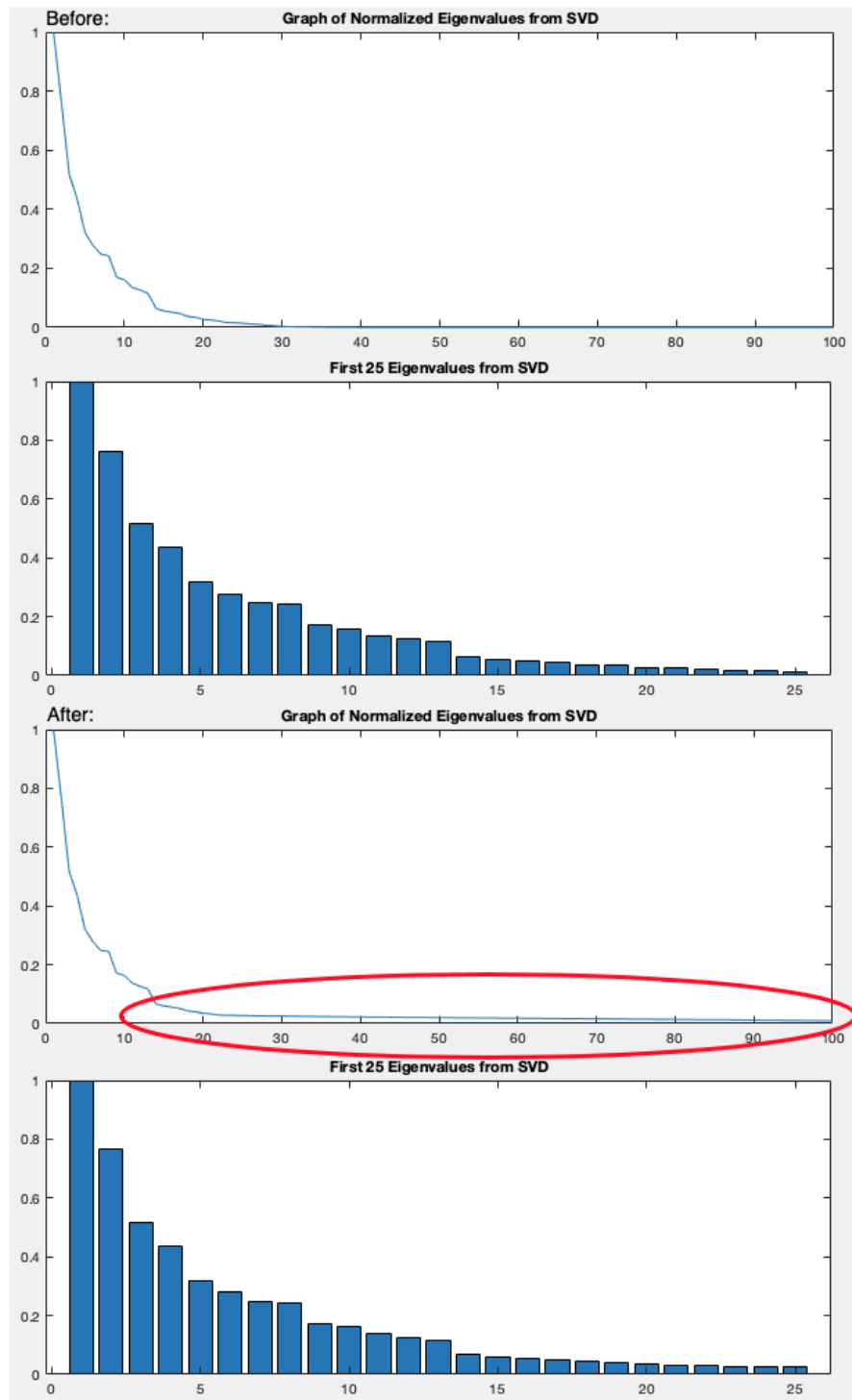


Figure 3.6. Effect of an artificial noise addition to the general singular values representation of a dataset

The number of singular values in the diagonal \underline{D} , and as a consequence of singular vectors, is related to original size of the matrix \underline{A} , as seen from the Figure 3.4. The number of singular values corresponds to the smallest number of variables within matrix dimensions. For example, in the case described in Fig. 3.4 and Fig. 3.5, the number of principal components is 96, its dimension with the smallest number of variables.

The original microseismic data matrix, however, had dimensions 2000x36, which included X-, Y-, and Z-components of 12 different receivers. By selecting only X-component of the dataset, the matrix is of the size 2000x12, meaning that the SVD of the matrix will give only 12 principal components. Fewer number of principal components reduce model reduction capabilities, as can be seen from Figure 3.8, where it is emphasized that 12 principal components are not enough to perform any model reduction and denoising of the dataset. To account for this issue, it is possible to synthetically increase the number of variables among the modes of the matrix by reshaping it from a 2000x12 to a 250x96 matrix.

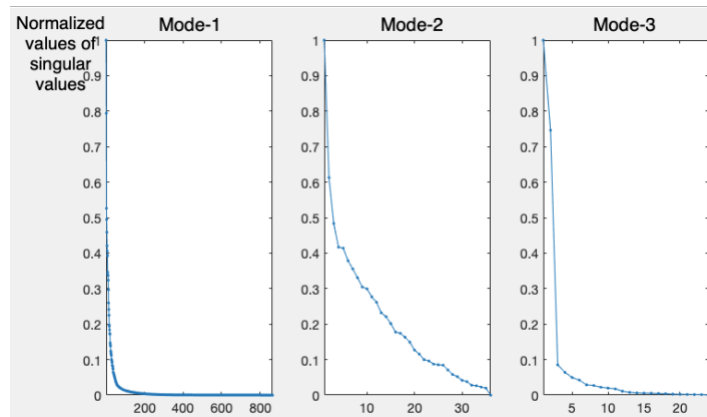


Figure 3.7. Singular values of a decomposed three-dimensional microseismic tensor

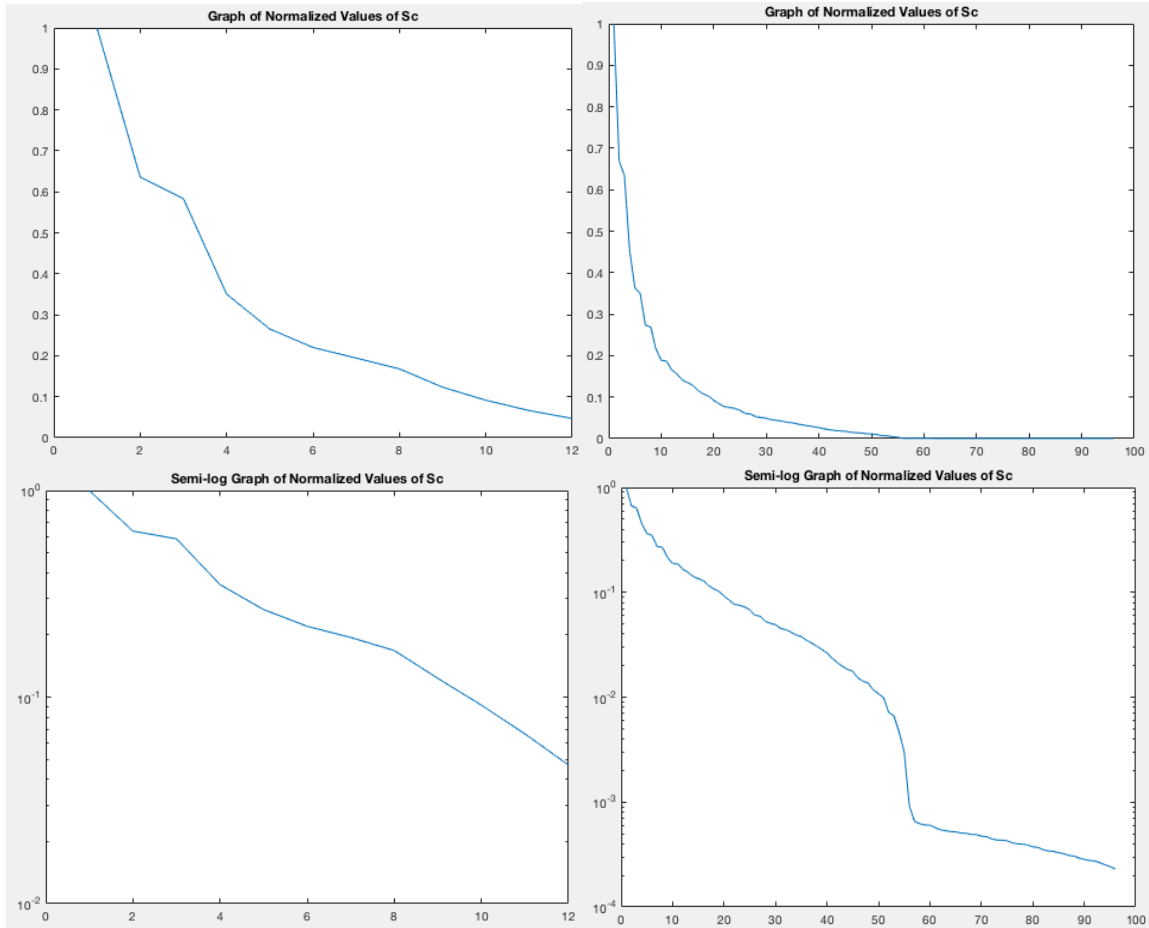


Figure 3.8. The comparison of a 12-singular-values dataset (left) and a 96-singular-values dataset (right)

It should be noted, that the reshaping of the original microseismic data matrix leads to a distorted picture of seismic signals that requires more singular vectors to represent it. However, there is an important tradeoff, since with the increased complexity and number of required singular vectors, there is also an increase of singular values available after decomposition to represent the original matrix.

The idea to increase the number of variables described above can be further extended to the increase of dimensionality. It can be shown that the increase in dimensionality of the original data tensor helps reduce more variables and helps denoise the dataset more effectively.

The HOSVD of the data tensor described in Fig. 3.1 and Fig. 3.2 will yield three factor matrices due to the three-dimensional nature of the original tensor (see more in subsections 2.3.3 and 2.3.4). Each of the matrices in the HOSVD, similarly to the SVD, represents a set of singular vectors, or principal components, or eigenvectors, while the core tensor $\underline{\underline{S}}$ represents a combination of singular values, or eigenvalues, similarly to the SVD.

Additionally, the Frobenius norm of the core tensor is decreasing from slice to slice in a similar manner to decreased order of singular values in SVD, as mentioned previously in subsection 2.3.4. Therefore, it is possible to apply a similar technique to HOSVD. However, the benefit of the HOSVD is the three sets of singular values, since each factor matrix contains their own singular vectors related to a specific mode, as Figure 3.7 shows.

3.2.3. Results and Possible Future Research

The result of the model reduction can be shown using a comparison of number of variables needed to represent the original tensor of microseismic data. The number of variables is directly proportional to size of a disk storage required to store the data, hence, it is a good representation of model reduction.

The original version of a third-order tensor presented in Figure 3.2 is of the size $2000 \times 36 \times 24$ and consists of 1,728,000 variables. In contrast, the decomposed HOSVD version of the tensor truncated at 179 at the first mode, 36 at the second mode, and 17 at the third mode, as given in Fig. 3.7, consists of only 469,252 variables. Figure 3.9 summarizes this information. Overall, HOSVD technique helped to reduce 73% of original microseismic data.

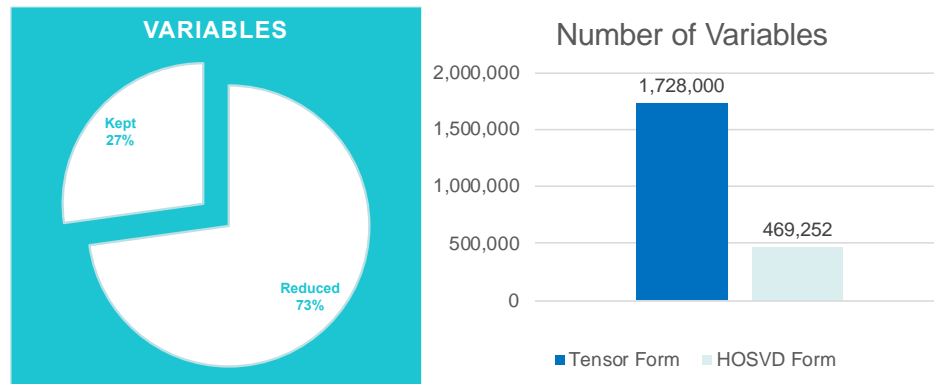


Figure 3.9. Microseismic data compression results

Similarly to the synthetic increase of variables in each mode performed earlier for an SVD, the dimensionality of the original data tensor could be synthetically increased with Quantized Tensor Networks-HOSVD approach (QTN-HOSVD) [12].

The idea behind QTN is to divide the original tensor into as many modes as possible by creating artificial modes, which could potentially reduce the model to the point of supercompression [12, 35]. However, more research is needed to test the technology of tensor networks and its application to microseismic data.

Moreover, instead of working with all tensor of microseismic data, it could be possible to pick certain valuable fibers from the very beginning. This technique is called CUR, or Skeleton Cross-Approximation [13]. However, it could be challenging to apply this technique to high dimensional microseismic tensors, since it is mostly applied to two-dimensional tensors in literature.

3.3. Hidden Structures

The beauty of decompositions is in the possibility of imposing additional constraints, otherwise a group of matrices after a decomposition could have no physical meaning [12]. For this purpose, it is interesting to test MWCA and see if the technology can be applied to complex microseismic data structures.

MWCA is a direct extension of two-way component analysis to higher-dimensional tensors. The main idea is to apply constraints to HOSVD and get hidden information out of the structure of a core tensor and factor matrices. Since microseismic data is often used in connection to hydraulic fracturing, this work is concentrated on a source of seismic wave prediction during hydraulic fracturing operations. Moreover, an attempt was made to infer an idea about a source mechanism from a structure of the decomposed tensor alone.

It should be noted that the idea of artificial increase of dimensionality, used in subsection 3.2.3, and subsequent mixing different modalities within the same dimension should not be carried onto MWCA technique, since the distorted representation can misrepresent connections between different modalities. This could lead to a false idea about hidden structures. Therefore, dimensions should be split into different modalities while applying HOSVD for hidden structures extraction, and no artificial increase of dimensionality should be performed.

3.3.1. Fractures Research

MWCA can be used to draw a connection between the decomposed data and a source position, which could be tied to fracture analysis.

A simple example of how an event position can affect the HOSVD results can be displayed by comparing two P-wave sources: near-case source that is close to receivers and far-case, as shown in Figure 3.10.

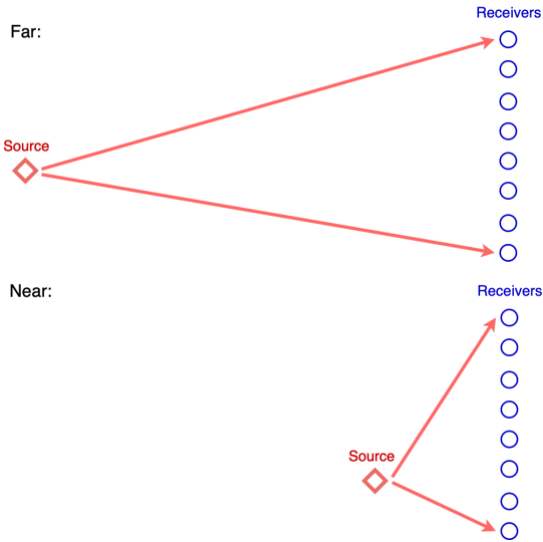


Figure 3.10. Example of a far and near source position

Polarization specifies the geometrical orientation of the wave and in the case of primary wave it propagates parallel to the direction of polarization vector. The angular deviation of the P-wave polarization from horizontal plane is called azimuth. It is shown in Figure 3.11.

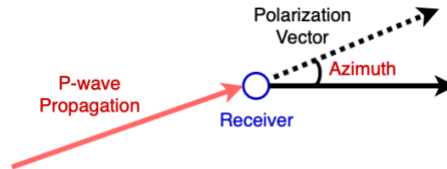


Figure 3.11. Graphical definition of polarization

For the far-case event, the azimuth of polarization of the P-wave does not vary significantly from receiver to receiver, in contrast to the near-case event. Therefore, most of the data can be described with only one vector component that is parallel to the horizontal plane. However, for the near event case, two vectors are needed to properly describe the data. As a result, a variation between the influences of singular vectors is expected while comparing far-case events and near-case events.

To test the correlation between the recorded waveforms and the HOSVD structure, synthetic data has been created to simulate positions of sources at various x-, y-, and z-coordinates. The origin is defined at the position of the lowest receiver. The height of the source is related to the z-coordinate, while the distance between the source and the receiver is related to the x-coordinate. The angle of the source relative to the receivers is mainly affected by the y-coordinate. As a consequence, the frequencies and amplitudes of the P-waves and S-waves differ between events according to their specific location.

An example of the synthetic data tensor for one source position is shown in Figure 3.12, where X, Y, and Z-components of each receiver were separated within the third mode. The tensor modes in Figure 3.12, hence, are time in milliseconds, receiver variation and X, Y, Z-component variation.

For each data tensor three distinct functions have been created, corresponding to three factor matrices from the tensor decomposition, as it was previously shown in Figure 2.11. The elements of the first function indicate the singular vectors, or columns, of the first factorization matrix. Each element of the function represents the "influence" of the corresponding singular vector.

The influence is defined as Frobenius norm of the singular values taken from the core tensor and related to the specific column. An example of the element i of the first, second, and third functions are shown in Eq. 3.1 – 3.3, respectively, where x , y , and z are indices of singular values related to the column i .

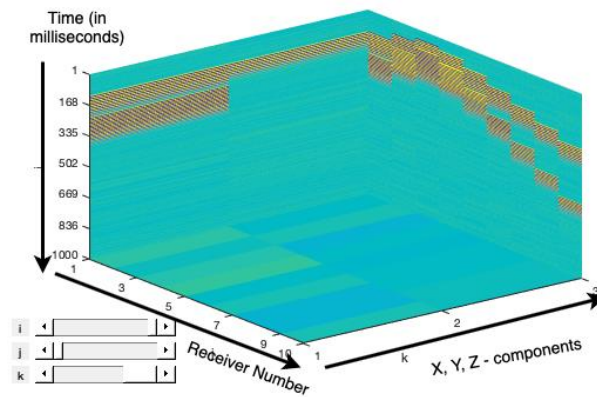


Figure 3.12. An example of synthetic data tensor

$$f_1(i) = \sqrt{\sum_y \sum_z \sigma_{iyz}^2} \quad (3.1)$$

$$f_2(i) = \sqrt{\sum_x \sum_z \sigma_{xiz}^2} \quad (3.2)$$

$$f_3(i) = \sqrt{\sum_x \sum_y \sigma_{xyi}^2} \quad (3.3)$$

Figure 3.13 shows how the natural logarithms of the functions vary for the far-case and near-case. It is easier to spot variation between "influences" of each element with the logarithms of the functions.

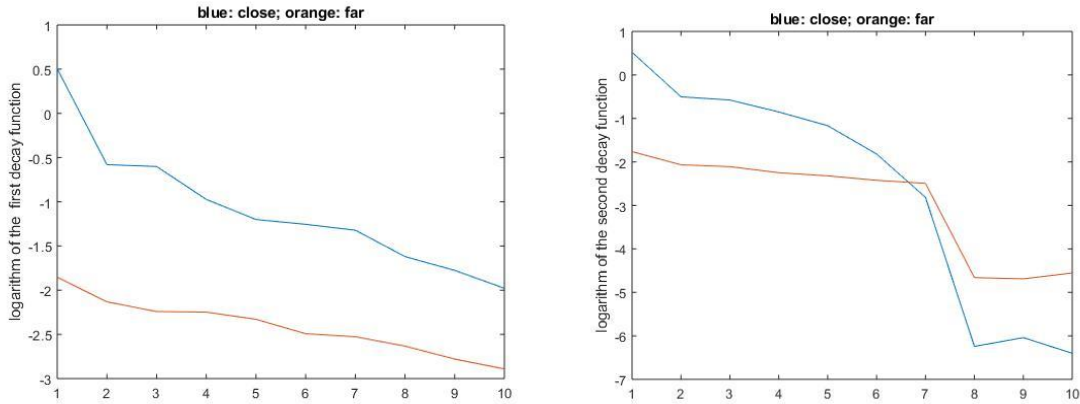


Figure 3.13. Comparison between logarithms of first (left) and second (right) functions for near and far sources

Each factor matrix, and consequently each function described by Eq. 3.1-3.3, is related to a dimension of original microseismic data tensor. Therefore, it might be expected that the first two functions are related to each other, since they represent the time variation mode and the receiver variation mode. The third function, however, has a different nature and it is worth concentrating on it first.

The behavior of the functions presented in Figure 3.13 can be described by one decay value. The decay of the function f (Eq. 3.1-3.3) is defined as in Eq. 3.4, where n is the last element of the function. To simplify, from Fig. 3.13, the decay is based on the three first elements and the three last elements of the function.

$$Decay(f) = \log \left(\sqrt[3]{\frac{f(1) \times f(2) \times f(3)}{f(n-2) \times f(n-1) \times f(n)}} \right) \quad (3.4)$$

By varying y -coordinate and distance between the source and receivers, it is possible to create a map of decay as a function of both parameters. The results are presented in Figure 3.14, where the decay of the third function is presented as a function of x - and y -coordinates.

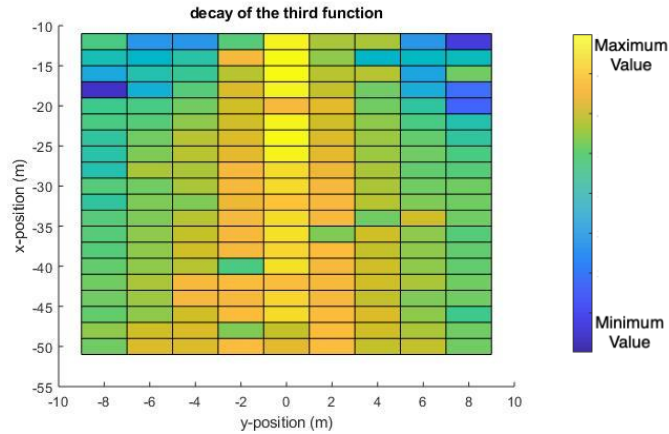


Figure 3.14. Representation of Decay as a function of position in x -, and y - direction

From Figure 3.14, it can be noted that there is a clear relationship between the decay function, as defined in Eq. 3.4, and the location of the source. Moreover, relative angle between the source orientation and receivers string, or the y -position of the source, seems to affect the value of the decay, which can be explained by a change in S-wave

polarization with a change in position of the source. The main idea is that the decay function, as defined in Eq. 3.4, can be used as a way to learn more about an S-wave and even determine its polarization from the structure of decomposition alone. The polarization of the S-wave, in turn, can provide information about the source mechanism, which can be tied to a fracture orientation and potentially help with their visualization.

As mentioned before, the first and second functions, based on Eq. 3.1 and Eq. 3.2, are related to each other. This can additionally be proved by looking at the representation of their decay function while varying position of the source. The source position is varied along the z-direction, or vertical position in Fig. 3.15, and the x-direction, or distance from receivers in Fig 3.15. Since the origin of the coordinate system is at the position of the lowest receiver, the lowest receiver is at 0 meters of the vertical position and 0 meters of the distance from receivers according to the Fig. 3.15. While visualizing the decay of the functions separately, no distinct correlation can be found, however, the decay of the summation of the first two functions, as plotted on the right side of the Fig. 3.15, clearly shows the correlation between the significance of the first few singular vectors and the distance from the receivers and the source.

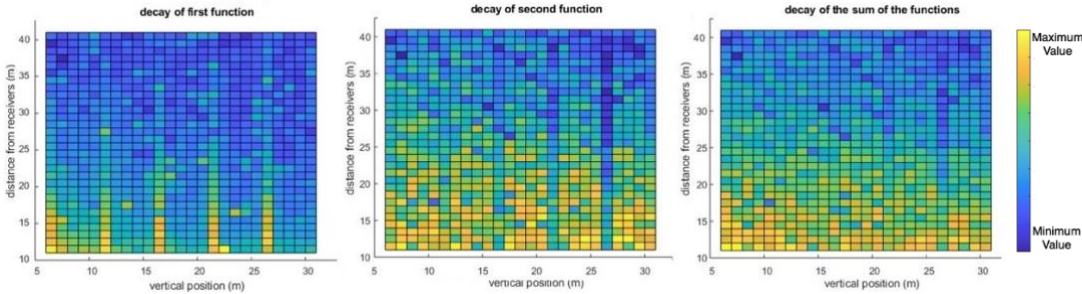


Figure 3.15. Representation of Decay as a function of vertical position and distance between the source and receivers

The decay in Figure 3.15 is, in essence, a discrete function relative to sources, whose position is defined with x-, y-, and z-coordinates. A curve-fitting has been applied to express this discrete function as a continuous 5-term polynomial function of the structure shown in Eq. 3.5, where c_0 , c_1 , c_2 , c_3 , and c_4 are coefficients of the continuous function $D(x, y, z)$.

$$D(x, y, z) = c_0 + c_1 \times (x^2 + y^2)^{0.15} + c_2 \times y^2 \times (x^2 + y^2)^{0.15} + c_3 \times y^2 \times z + c_4 \times z \times (x^2 + y^2)^{0.15} \quad (3.5)$$

The polynomial fitting is implemented using Greedy algorithm based on the modified error function approach, where the influence of large error between calculated decay and actual value is reduced. Since it was developed in collaboration with Milan Brankovic, the information regarding the exact implementation of the error function can be found in SPE-195522-MS paper, "A Novel Approach to Discovery of Hidden Structures in Microseismic Data Using Machine Learning Techniques" (Control # 19EURO-P-592-SPE), as mentioned in the Contributors and Funding Sources section.

The polynomial fitting gives a somewhat convolutional neural networks approach, in the sense that it represents a function that maps the relation between the decomposed structure of the original tensor and the location of the seismic wave source. The approach has been tested on a large number of synthetic events with distances that ranged from 10 meters to over 50 meters. As a result, for the majority of events, the location of the source has been determined with an average of 80% accuracy. Moreover, it is important to remember that the third function is found to be related to the S-wave polarization. Using this information, it could be possible to obtain a source mechanism and, hence, understand

the orientation of a fracture, however further research is needed to develop a generalized methodology.

To summarize, only using the results of HOSVD, it was possible to approximately estimate the location of the source. Therefore, the application of HOSVD and MWCA techniques could potentially provide quick insights on fracture propagation as well as additional information during hydraulic fracturing operations, if the connection between the source location and fracture is established.

3.3.2. Other Ideas in Fracture Research

In addition to decay functions, it is interesting to analyze factorization matrices to extract the polarization of the P-wave from the receivers. The definition of polarization is shown in Figure 3.11 and it helps to determine the angle at which the source of the wave is expected. By using both the polarization and source location estimation based on the decay, more accurate results are expected.

The methodology to determine the polarization from the decomposed structure of the data tensor alone is developed by Milan Brankovic and is described in SPE-195522-MS paper, "A Novel Approach to Discovery of Hidden Structures in Microseismic Data Using Machine Learning Techniques" (Control # 19EURO-P-592-SPE), as mentioned in the Contributors and Funding Sources section.

Additionally, it could be possible to implement a different approach to MWCA. For example, Canonical Polyadic Decomposition (CPD) could be performed instead of HOSVD. CPD has a diagonal structure of a core tensor [15], therefore it could be easier

to look at relationships between the core tensor and factor matrices. This, in turn, could reveal more insights to hidden structures of a microseismic data tensor.

3.3.3. Research in Four Dimensions and Variation in Time

The third order tensor, presented in Figure 3.1, Figure 3.2, or Figure 3.12, can be further extended to a fourth order tensor to account for variation in time, as described in subsection 3.1 and Figure 3.3. As a reminder, the microseismic research could be done at different times at the same locations, such as Time = 1 hour or Time = 2 hours. If this is the case, then there is a set of three-dimensional tensors available for research. It is possible to create a vector of the three-dimensional structures, which is a fourth-order tensor.

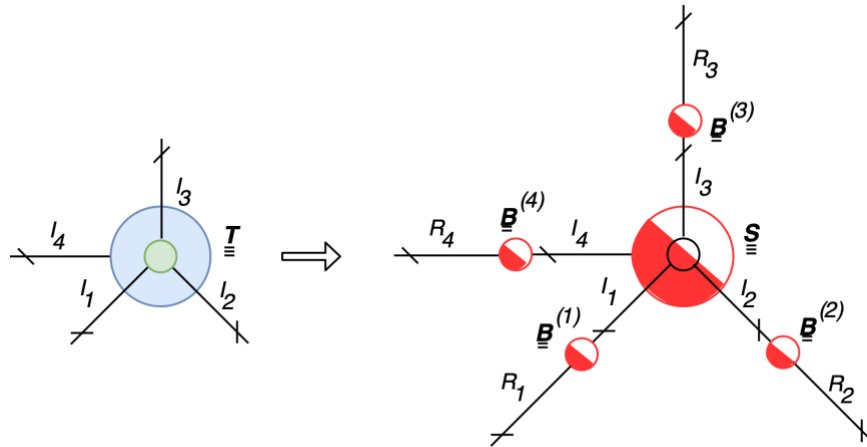


Figure 3.16. Tensor representation of HOSVD performed on a four-dimensional microseismic data tensor

The Higher-Order Singular Value decomposition of the four-dimensional data structure $\underline{\underline{T}} \in R^{I_1 \times I_2 \times I_3 \times I_4}$ will result in a fourth-order core tensor $\underline{\underline{S}} \in R^{R_1 \times R_2 \times R_3 \times R_4}$ and

four factor matrices related to it, as shown in Figure 3.16, Figure 3.17 and Eq. 3.6, where the additional fourth matrix is $\underline{\underline{B}}^{(4)} \in R^{I_4 \times R_4}$.

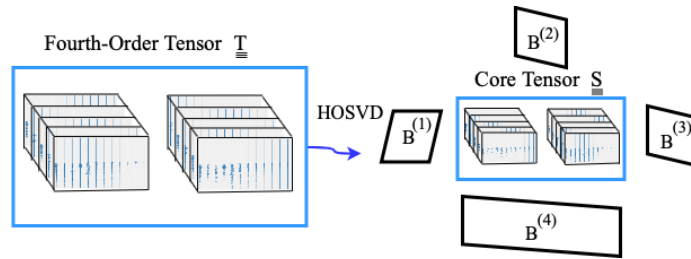


Figure 3.17. Fourth factor matrix and hidden structures related to time

The matrix $\underline{\underline{B}}^{(4)}$ is, then, directly related to variation of data in time, the fourth dimension of the structure, and can give valuable insights on variation of seismic response.

That can potentially serve as a basis for prediction of source location in time, which can be a valuable tool to infer data about fracture propagation during hydraulic fracturing, however this is not the scope of the paper and more research should be done in order to create a systematic methodology.

3.4. Overview of the Proposed Methodology

The proposed methodology concept map is presented in Figure 3.18.

Starting at the top left corner of Figure 3.18, the tensorization of available microseismic data should be performed. With the high-dimensional microseismic data structure, it is possible to perform four-dimensional HOSVD and MWCA analysis. The analysis is related to the variation of seismic wave source in time and could potentially serve as a basis for prediction of fracture propagation trends. In case the variation in time is not available, HOSVD and MWCA could be performed with a three-dimensional data structure. The denoising and model reduction algorithm could be applied, and source

location and source mechanism analysis could be performed based on the structure of decomposition.

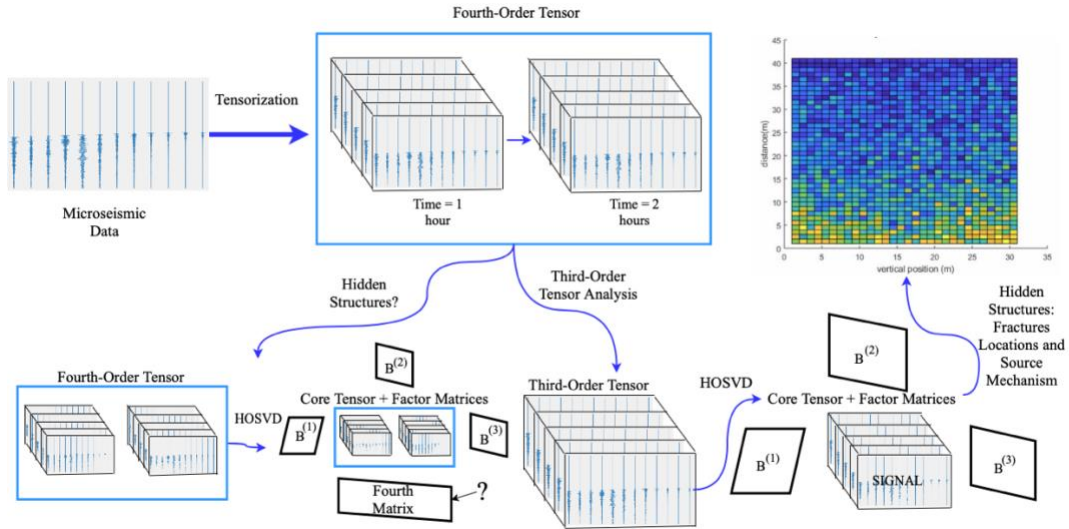


Figure 3.18 Graphical description of the proposed methodology.

3.5. Main Findings and Future Work

Among the main findings are the following:

- By implementation of a better data analysis approach to microseismic data, additional information, or hidden structures, could be revealed to give insights about variation of seismic parameters with time, which could potentially serve as a basis for fracture propagation prediction.
- Quick seismic inversion was performed to get information about source location and source mechanism by using only the structure of a decomposed version of a microseismic data tensor.

- Tensor decompositions were used to significantly reduce the data in the 70% range and further supercompression could be achieved by implementing a synthetic increase of dimensionality.
- The denoising algorithm, based on singular values analysis, helped work with noisier data cases, which could save time and money for petroleum companies.
- It was found that model reduction and denoising technique based on HOSVD works better with a cubic structure, since it gives more eigenvalues after decomposition. However, mixing different modalities within the same dimension could lead to a distorted idea about hidden structures. Therefore, dimensions should be split into different modalities while applying HOSVD for hidden structures extraction.

The technologies of HOSVD and MWCA are recent and still emerging. Therefore, there are many challenging problems to be addressed in future research:

- The possibility of implementation of Quantized Tensor Networks could be addressed to further compress high-dimensional microseismic data structures. Additionally, a combination between HOSVD and QTT could be developed, which is called QTT-HOSVD decomposition, or, more generally, QTT-Tucker decomposition.
- In contrast to working with all tensor of microseismic data, it is interesting to look at picking on certain valuable fibers. The technique is called CUR, or Skeleton Cross-Approximation. However, not a lot of research has been

done to extend the technique from second-order tensors to higher dimensional tensors.

- Finally, a different approach to MWCA could be implemented. In particular, Canonical Polyadic Decomposition could be used instead of HOSVD, which might provide even more hidden structures and insights from microseismic data.

4. CONCLUSIONS

An alternative method for identifying structures by finding hidden information within microseismic data has been proposed. With machine learning techniques, precisely tensor decompositions and component analysis, quick seismic inversion has been performed to get information about source locations and source mechanisms by using only the structure of a decomposed version of a microseismic data tensor. Moreover, the extracted information could potentially give insights about fracture propagation. It is a cost-effective and fast technique, as opposed to conventional simulation methods.

Furthermore, this methodology yielded 70% model reduction whilst using the special type of Tucker Decomposition, HOSVD, due to its specific orthogonal structure.

The proposed workflow is general and highly applicable, given the variation of data used for the research.

Finally, since tensor decompositions and multiway component analysis are still very young and emerging technologies, significant improvements to the technologies are expected. That implies even further benefits, making the application of the machine learning techniques to microseismic data a very promising field for future research.

REFERENCES

- [1] Wang, Hua, Meng Li, & Xuefeng Shang. "Current developments on micro-seismic data processing." *Journal of Natural Gas Science and Engineering* 32: 521-537, 2016.
- [2] Langenbruch, Cornelius, & Serge A. Shapiro. "Quantitative analysis of rock stress heterogeneity: Implications for the seismogenesis of fluid-injection-induced seismicity." *Geophysics* 80.6: WC73-WC88, 2015.
- [3] Sleefe, G. E., N. R. Warpinski, and B. P. Engler. "The use of broadband microseisms for hydraulic fracture mapping". No. SAND-93-0091C; CONF-931088-1. Sandia National Labs., Albuquerque, NM (United States), 1993.
- [4] Le Calvez, Joël, et al. "Hydraulic fracturing insights from microseismic monitoring." *Oilfield Review* 28.2, 2016.
- [5] Zhang, Xiaopu, et al. "An Efficient Neural-Network-Based Microseismic Monitoring Platform for Hydraulic Fracture on an Edge Computing Architecture." *Sensors* 18.6: 1828, 2018.
- [6] Rassenfoss, Stephen. "Using microseismic to see where fracturing is happening." *Journal of Petroleum Technology*, 16 October 2018.
- [7] Daley T. M., Freifeld B. M., Ajo-Franklin J., Dou S., Pevzner R., Shulakova V., Kashikar S., et al. "Field testing of fiber-optic distributed acoustic sensing (DAS) for subsurface seismic monitoring." *The Leading Edge* 32, 2013.

- [8] Cichocki, Andrzej, & Shun-ichi Amari. "Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications, Volume 1." John Wiley & Sons, 2002.
- [9] Fanaee-T, Hadi, & João Gama. "Eigenevent: An algorithm for event detection from complex data streams in syndromic surveillance." *Intelligent Data Analysis* 19.3: 597-616, 2015.
- [10] Vasilescu, M. Alex O., & Demetri Terzopoulos. "Multilinear analysis of image ensembles: Tensorfaces." *European Conference on Computer Vision*. Springer, Berlin, Heidelberg, 2002.
- [11] De Vos, Maarten. "Decomposition methods with applications in neuroscience." PHD Thesis, Katholieke Universiteit Leuven, 2009.
- [12] Cichocki, Andrzej. "Era of big data processing: A new approach via tensor networks and tensor decompositions." *arXiv preprint:1403.2048*, 2014.
- [13] Mahoney, M. W., Maggioni, M., & Drineas, P. "Tensor-CUR decompositions for tensor-based data". *SIAM Journal on Matrix Analysis and Applications*, 30(3), 957-987, 2008.
- [14] Geng, Xiurui, et al. "A high-order statistical tensor based algorithm for anomaly detection in hyperspectral imagery." *Scientific reports* 4: 6869, 2014.
- [15] Kolda, Tamara G., & Brett W. Bader. "Tensor decompositions and applications." *SIAM review* 51.3: 455-500, 2009.
- [16] Golub G. H., & Van Loan, C. F. "Matrix Computations". Johns Hopkins University Press: fourth edition, 2013.

- [17] Kossaifi, Jean, Yannis Panagakis, & Maja Pantic. "Tensorly: Tensor learning in python." arXiv preprint arXiv:1610.09555, 2016.
- [18] Hyvarinen, Aapo, Juha Karhunen, & Erkki Oja. "Independent Component Analysis." John Wiley & Sons, 2001.
- [19] Hinze, Michael, & Stefan Volkwein. "Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: Error estimates and suboptimal control." Dimension reduction of large-scale systems. Springer, Berlin, Heidelberg, 2005.
- [20] Lui, Yui Man, J. Ross Beveridge, & Michael Kirby. "Action classification on product manifolds." Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. IEEE, 2010.
- [21] Beckmann, Christian F., et al. "Group comparison of resting-state fMRI data using multi-subject ICA and dual regression." Neuroimage 47, 2009.
- [22] Phan, Anh Huy, & Andrzej Cichocki. "Tensor decompositions for feature extraction and classification of high dimensional datasets." Nonlinear theory and its applications, IEICE 1.1, 2010.
- [23] Ghasemi, Mohamadreza, & Eduardo Gildin. "Localized Model Reduction in Porous Media Flow." IFAC-PapersOnLine 48.6, 2015.
- [24] Leimkuhler, Joe, & Greg Leveille. "Unconventional Resources." Society of Petroleum Engineers. doi:10.2118/0112-026-TWA, 2012.
- [25] Juhong, Th Li. "Microseismic data analysis in unconventional oil and gas exploration." SPG/SEG 2016 International Geophysical Conference, Beijing,

- China, 20-22 April 2016. Society of Exploration Geophysicists and Society of Petroleum Geophysicists, 2016.
- [26] Drew, Julian, & Joelle Calvez. "Complexities in the Quick Analysis And Interpretation of Microseismic Data." Society of Exploration Geophysicists, 2007.
- [27] Tafti, Tayeb A., & Fred Aminzadeh. "Characterizing Fracture Network in Shale Reservoir Using Microseismic Data." Society of Petroleum Engineers, doi:10.2118/153814-MS, 2012.
- [28] Turuntae, S. B., E. V. Zenchenko, & E. I. Eremeeva. "Permeability Evaluation From Microseismic Data (laboratory Study)." Society of Exploration Geophysicists, 2010.
- [29] Organization of the Petroleum Exporting Countries. OPEC World Oil Outlook, October 2017. Available from: <http://www.opec.org>
- [30] Afra, S., Gildin, E., & Tarrahi, M. "Heterogeneous reservoir characterization using efficient parameterization through higher order SVD (HOSVD)". American Control Conference (ACC), IEEE, 2014.
- [31] Afra, S., & Gildin, E. "Tensor based geology preserving reservoir parametrization with Higher Order Singular Value Decomposition (HOSVD)". Computers & Geosciences Journal, 2016.
- [32] Abbass, M. Y., & Kim, H. J. "Blind image separation using pyramid technique". Image Video Processing Journal, Springer International Publishing, 2018.
- [33] VanderPlas, J. "Python Data Science Handbook". O'Reilly Media, 2016.

- [34] Pedregosa, Fabian, et al. "Scikit-learn: Machine Learning in Python", JMLR Volume 12, 2011.
- [35] Khoromskij, B. N. "O(dlogN)-quantics approximation of N-d tensors in high-dimensional numerical modeling". Constructive Approximation Volume 34, Springer, 2011.
- [36] Alumbaugh D., Bevc D. "High-resolution microseismic-source locations and moment-tensor solutions from the Permian Basin". Society of Exploration Geophysicists, 2018.
- [37] Rahm D. "Regulating hydraulic fracturing in shale gas plays: The case of Texas". Energy Policy Volume 39, Elsevier, 2011.
- [38] Tan H., Zhang Y. "Expression-independent face recognition based on higher-order singular value decomposition". IEEE International Conference on Machine Learning and Cybernetics, 2008.
- [39] Warpinski N. "Microseismic monitoring: Inside and out". Society of Petroleum Engineers, SPE 118537, 2009.
- [40] Wang Y. "Seismic Inversion: Theory and Applications". John Wiley & Sons, 2016.