University of Maribor

Faculty of Electrical Engineering and Computer Science

Aleš Holobar

BLIND DECOMPOSITION OF CONVOLUTIVE MIXTURES OF CLOSE-TO-ORTHOGONAL PULSE SOURCES APPLIED TO SURFACE ELECTROMYOGRAMS

SLEPA DEKOMPOZICIJA KONVOLUTIVNIH MEŠANIC SKORAJ ORTOGONALNIH IMPULZNIH IZVOROV, UPORABLJENA PRI POVRŠINSKIH ELEKTROMIOGRAMIH

PhD Thesis

Maribor, October 2004

Author: Aleš Holobar

Title: Blind decomposition of convolutive mixtures of close-to-orthogonal pulse sources applied to surface electromyograms

UDC: 621.391:61(043.3)

Keywords: digital signal processing, blind source separation, multi-channel blind deconvolution, pulse sources, time-frequency analysis, higher-order statistics, w-slices, Newton-Gauss optimization, inverse correlation matrix based method, MIMO system identification, surface electromyographic signals, surface EMG decomposition, motor unit, innervation pulse train, motor unit action potential

Copyright: System software laboratory, Faculty of Electrical Engineering and Computer Science, University of Maribor, Slovenia

University of Maribor

Faculty of Electrical Engineering and Computer Science

Aleš Holobar

Blind decomposition of convolutive mixtures of close-to-orthogonal pulse sources applied to surface electromyograms

Committee members:

Prof. dr. Igor Tičar, chairman

Prof. dr. Damjan Zazula, supervisor

Prof. dr. Roberto Merletti, supervisor

Prof. dr. Tadej Bajd, member

Prof. dr. Gorazd Lešnjak, member

Maribor, 26th of October 2004

BLIND DECOMPOSITION OF CONVOLUTIVE MIXTURES OF CLOSE-TO-ORTHOGONAL PULSE SOURCES APPLIED TO SURFACE ELECTROMYOGRAMS

UDC: 621.391:61(043.3)

Keywords:

digital signal processing, blind source separation, multi-channel blind deconvolution, pulse sources, time-frequency analysis, higher-order statistics, w-slices, Newton-Gauss optimization, inverse correlation matrix based method, MIMO system identification, surface electromyographic signals, surface EMG decomposition, motor unit, innervation pulse train, motor unit action potential

Abstract:

In this doctoral dissertation blind decomposition of convolutive mixtures of close-to-orthogonal pulse source signals is addressed. Three novel decomposition approaches based on the time-frequency analysis, higher-order statistics and algebraic approach are developed. Furthermore, necessary conditions for the reconstruction of complete pulse sequences and their impulse responses are investigated and a thorough and detailed study of the factors influencing its performance, such as noise and non-orthogonality of sources, are carried out. Although derived in the case of more measurements than sources (overdetermined system), the algebraic approach is extended to slightly underdetermined systems (with more sources than measurements). In contrast with other decomposition techniques, the proposed approaches work well also in the case of not completely orthogonal source signals. Finally, all the proposed solutions are applied to the surface electromyographic (EMG) signals.

The thesis begins with an overview of existing methods and techniques for blind source separation. The methods for both multiplicative and convolutive cases are critically assessed and mutually compared. Next, a brief introduction to the physiology of the human muscles is given. In order to provide the basis for evaluation of the decomposition results on the real surface EMG signals, the properties of the motor unit (MU) innervation pulse trains, and generation of the motor unit action potentials (MUAPs) at the end-plate are explained. The factors influencing the shape and amplitudes of MUAPs, detected at the skin surface, are also identified. Afterwards, the decomposition methods for both intra-muscular and surface EMG signals are critically evaluated and the influence of the superimposed MUAPs studied. Finally, the assumed data model of surface EMG signals is introduced and its main limitations and assumption clarified. Surface EMG signals are modelled as a multi-channel, linear, shift-invariant multiple-input-multiple-output (MIMO) system.

In the second part of this dissertation, three different approaches to blind source separation of the convolutive mixtures of general pulse source signals are derived. Firstly, the over-determined case is assumed (the case with more measurements than sources) and two novel approaches introduced. The first one utilizes the Wigner-Ville time-frequency distributions and enables the reconstruction of both the pulse source signals and the corresponding MIMO system responses. The second approach enables automatic reconstruction of the MIMO system responses and is based on higher-order cumulants. Next, the decomposition is extended to the underdetermined case (the case with more sources than measurements) and a completely novel approach to blind deconvolution of pulse source signals, so called inverse correlation matrix based method, is derived.

The introduced approaches are tested on both synthetic and real surface EMG signals. Firstly, the impacts of the number of active MUs, their firing frequencies, depth in the muscle tissue, etc., as well as the influence of noise are evaluated on the synthetic signals. The results prove the superiority of the inverse correlation based method. Applying it to the over-determined case at a high signal to noise ratio (SNR), almost all simulated MU innervation pulse trains are completely reconstructed. In the under-determined case with number of sources exceeding the number of measurements by factor 1.4, approximately a half of the simulated MUs are completely identified. The performance also drops with the SNR. At SNR of 0 dB, approximately 30 % of the MUs identified at SNR = 20 dB are reconstructed. In all cases, the decomposed innervation pulse trains exhibit a perfect match with the reference synthetic source signals. The other two methods are significantly less efficient.

Finally, all three decomposition approaches are applied to the real surface EMG signals, recorded during an isometric 5 % and 10 % contractions of the dominant biceps brachii muscle of 9 healthy young male subjects. Again, the inverse correlation based method proves to be superior. Altogether, 30 and 56 MUs' innervation pulse trains are completely reconstructed from the 5 % and 10 % muscle contraction measurements, respectively. The reconstructed MU firing patterns are compared against various physiologically induced limitations and prove to be in agreement with expectations and careful visual analysis.

SLEPA DEKOMPOZICIJA KONVOLUTIVNIH MEŠANIC SKORAJ ORTOGONALNIH IMPULZNIH IZVOROV, UPORABLJENA PRI POVRŠINSKIH ELEKTROMIOGRAMIH

UDK: 621.391:61(043.3)

Ključne besede:

digitalna obdelava signalov, slepa ločitev izvorov, večkanalna slepa dekonvolucija, impulzni izvor, časovno-frekvenčna analiza, statistike višjih redov, w-rezine, Newton-Gaussova optimizacija, metoda z inverzom korelacijske matrike, identifikacija sistemov MIMO, površinski elektromiogram, dekompozicija površinskih elektromiogramov, motorična enota, inervacijski vlaki impulzov, akcijski potencial motorične enote

Povzetek:

V doktorski disertaciji obravnavamo postopke za slepo ločevanje konvolutivnih mešanic impulznih izvornih signalov in vpeljemo tri nove metode, ki omogočajo popolno rekonstrukcijo izvornih signalov oz. odzivov posameznih mešalnih kanalov (filtrov) v sistemu MIMO (*multiple-input-multiple-output*). Pri izpeljavi najprej privzamemo, da je meritev več kot izvorov, preučimo pa tudi možnosti dekompozicije, ko je meritev manj kot izvorov. Ovrednotimo še vpliv šuma, neortogonalnosti izvornih signalov in razmerja med številom meritev in številom izvorov.

Raziskavo pričnemo z opredelitvijo ciljev slepe ločitve signalov in pregledom obstoječih metod. Osredotočimo se predvsem na ločitev konvolutivnih mešanic izvornih signalov, opišemo pa tudi najpogosteje uporabljene metode za ločitev multiplikativnih mešanic. Opisane metode kritično ovrednotimo in jih medsebojno primerjamo. V nadaljevanju na kratko opišemo anatomijo in fiziološke omejitve električnih signalov v človeških mišicah, poseben poudarek pa namenimo opisu nastanka, potovanja in ponora akcijskih potencialov v mišičnem tkivu. Razpravo nadaljujemo s pregledom obstoječih metod za dekompozicijo signalov EMG. Predstavljene metode kritično ovrednotimo in navedemo njihove prednosti in omejitve. Nazadnje definiramo pogoje, pod katerimi lahko večkanalne meritve površinskega EMG modeliramo kot sistem MIMO s končnimi, vzročnimi in časovno nespremenljivimi sistemskimi odzivi.

V nadaljevanju izpeljemo tri splošne postopke za slepo ločitev skoraj ortogonalnih impulznih izvornih signalov. Najprej obravnavamo konvolutivne mešanice z več meritvami kot izvori in opišemo dva nova postopka za njihovo ločitev. Prvi temelji na časovno-frekvenčni analizi, drugi pa na statistikah višjega reda. Dekompozicijo nato razširimo na konvolutivne mešanice z več izvori kot meritvami. V okviru algebrajskih

pristopov in statistik drugega reda razvijemo popolnoma nov koncept ločevanja konvolutivne mešanice, t. i. metodo z inverzom korelacijske matrike.

Uspešnost izpeljanih postopkov preverimo na umetnih in realnih površinskih signalih EMG. Pri podajanju rezultatov vpeljemo ustrezne kvantitativne mere za vrednotenje opisanih postopkov in preučimo vpliv šuma in razmerja med številom meritev in številom izvorov. Rezultati na umetnih signalih potrjujejo superiornost metode z inverzom korelacijske matrike (metode IC). Če je meritev več kot izvorov, uspe metoda IC pri visokem razmerju signal-šum (SNR) popolnoma rekonstruirati skoraj vse inervacijske vlake impulzov simuliranih ME. V poddeterminiranem sistemu, kjer število izvorov za faktor 1,4 presega število meritev, uspe metoda IC popolnoma rekonstruirati približno polovico vseh simuliranih ME. Učinkovitost dekompozicije pada tudi z razmerjem signal-šum. V povprečju uspe metoda IC pri SNR = 0 dB rekonstruirati približno 30 % tistih ME, ki so bile rekonstruirane pri SNR = 20 dB. Rekonstruirani inervacijski vlaki impulzov pa se tudi pri nizkih razmerjih signal-šum skoraj popolnoma ujemajo z referenčnimi umetnimi vlaki. Preostali dve metodi sta manj učinkoviti.

Metoda IC se izkaže kot najbolj učinkovita tudi pri realnih površinskih signalih, posnetih med izometrično konstantno 5- in 10-odstotno skrčitvijo dominantne mišice biceps brachii pri devetih mladih, zdravih poskusnih osebah moškega spola. Skupno metoda IC popolnoma rekonstruira inervacijske vlake 30 ME pri 5-odstotni mišični skrčitvi, oz. 56 ME pri 10-odstotni mišični skrčitvi. Rekonstruirani inervacijski vlaki impulzov izpolnjujejo vse fiziološke omejitve.

Acknowledgements

Foremost, I would like to thank my mentor Prof. Damjan Zazula for his patience, support, guidance and encouragement over the last 8 years and especially throughout the course of my PhD work. My sincere gratitude is also extended to my comentor Prof. Roberto Merletti for all his guidance, support and valuable discussions on my research topic. I would especially like to thank Dr. Dario Farina, Marco Gazzoni, Marco Pozzo and all the members of the LISiN laboratory at Politecnico di Torino, Italy, for their support in using their SEMG simulator and in the design and implementation of the experimental protocol. I would also like to thank my committee members Prof. Gorazd Lešnjak and Prof. Tadej Bajd for their valuable suggestions and expertise.

I am sincerely grateful to all of my colleagues from the System Software Laboratory at Faculty of Electrical Engineering and Computer Science, Maribor, who have given me excellent support and friendship for all these years. Special thanks go to Matjaž Divjak for all his suggestions and the corrections he has contributed to this dissertation. I would also like to thank Prof. Zvonko Fazarinc. His unique and gifted attitude to many aspects of life, has always been, and will always be, a great source of inspiration for me.

Last, but not in any measure, the least, I must thank my parents and my sister for their strong support and inspiration. I can never thank enough my beloved Katja, mostly for her moral support, patience and understanding.

This work was supported by the Slovenian Ministry of Education, Science, and Sport (Contract No. S2-796-010/21301/2000 and Programme funding P2-0041), and by the European Shared Cost Project *NEW – Neuromuscular assessment in the Elderly Worker* (Contract No. QLK6-2000-00139).

Thank you all.

Aleš Holobar

Table of contents

List of symbols…….…….…….... iv

List of abbreviations.…….... vii

1. Introduction……..……..……... 1

1.1 Explanations of frequently used synonyms…...……..…….…..... 5

2. Blind Source Separation…….…….…….. 9

2.1 Separation of linear multiplicative mixtures…..……..……...……... 10

2.1.1 Separation by maximization of non-Gaussianity……...……..……..….. 11

2.1.2 Separation based on maximum likelihood……...…….…… 11

2.1.3 Separation of time signals…….…….... 12

2.1.4 Identification of underdetermined MIMO systems....…… 13

2.2 Separation of linear convolutive mixtures…….....…….…..……..... 14

2.2.1 Maximum likelihood based approaches..... 14

2.2.2 Separation in time-frequency plane…….……..... 15

2.2.3 Reformulation of convolutive mixture to multiplicative one…… 15

3. Electromyogram..... 17

3.1 Physiology of human muscles..... 18

3.1.1 Motor unit action potentials……..... 18

3.2 EMG decomposition……..... 20

3.3 Assumed data model of surface EMG signals..... 22

3.3.1. Extension of the convolutive MIMO vector form…..... 25

3.3.2. Ambiguities of blind source separation..... 27

4. Decomposition of convolutive mixtures of sources – more measurements than sources..... **29**

4.1 Method based on time-frequency distributions..... 30

4.1.1 Discrete time-frequency distributions……..... 31

4.1.2 Separation of general sources..... 32

4.1.3 Blind separation of pulse sources..... 33
4.1.4 Decomposition of close-to-orthogonal sources..... 34
4.1.5 Noise influence..... 37
4.2 W-slices and higher-order cumulants……..... 38
4.2.1 Higher-order statistics..... 39
4.2.2 Coarse estimation of system responses: w-slices..... 42
4.2.3 Nonlinear optimisation of coarse estimates..... 45
4.2.4 Influence of non-Gaussian noise and nonorthogonal sources..... 48
Appendix 4A: Derivatives of cumulants 49
5. Inverse correlation based method..... 51
5.1 Decomposition of orthogonal sources……..... 52
5.2 Decomposition of non-orthogonal sources..... 54
5.2.1 Separation of the superimposed sources 55
5.3 Noise influence..... 59
5.3.1 Noise in the space of sources..... 60
5.4 Decomposition of under-determined systems……..... 63
Appendix 5A: Inverse of diagonally dominant matrix..... 66
Appendix 5B: Probability of overlapped pulses..... 69
Appendix 5C: Average number of pulses in the product of the superimposed pulse trains of sources..... 74
Appendix 5D: Properties of orthogonal projection matrix..... 78
6. Surface EMG decomposition results..... 83
6.1 Results on synthetic surface EMG signals..... 84
6.1.1 Method based on time-frequency distributions…..... 85
6.1.2 Method based on higher-order statistics..... 90
6.1.3 Inverse correlation based method……..... 95
6.2 Results on real surface EMG signals…..... 104
6.2.1 Inverse correlation based method……..... 107
6.2.2 Method based on time-frequency distributions..... 119
6.2.3 Method based on higher-order statistics..... 123

7. Discussion ……. 127

7.1 Comparison of the results on synthetic surface EMG signals ….…… 127

7.2 Comparison of the results on the real surface EMG signals .....…….... 131

7.3 Computational complexity……..... 133

8. Conclusions.…… 135

References..... 139

Biography..... 139

List of symbols

$\mathbf{A}^{\#}$	Moore-Penrose pseudoinverse of matrix $\mathbf{A}$,
$\mathbf{B}$	whitening matrix
$\mathbf{C}_{ijk}$	cumulant matrix calculated from *i*-th, *j*-th and *k*-th measurement
$\mathbf{C}_m(\tau_1, \tau_2)$	third-order cumulant matrix of *m*-th measurement (*m* denotes the fixed measurement)
$\mathbf{D}_{\overline{\mathbf{ss}}}(n, f)$	spatial time-frequency distribution matrix of extended sources
$\mathbf{D}_{\overline{\mathbf{xx}}}(n, f)$	spatial time-frequency distribution matrix of extended measurements
$\delta(\tau)$	Dirac impulse
F	total number of pulses in a pulse source
$\boldsymbol{\varphi}(m,l)$	kernel of a time-frequency distribution
G_{n_i}	set of indices denoting the sources that are simultaneously active at given time moment n_i
g_i	cardinal number of set G_{n_i}
γ_{3,s_j}	skewness (central value in the third-order cumulant matrix) of source s_j
$\mathbf{H}$	mixing matrix
$h_{ij}(l)$	(i,j)-th system response of assumed MIMO system
$Ind(n)$	activity index
K	extension factor (the number of delayed repetitions added to each measurement)
$\kappa = \lVert\mathbf{H}\rVert\lVert\mathbf{H}^{-1}\rVert$	condition number of matrix $\mathbf{H}$
L	maximal length of system responses (in samples)
M	number of original measurements

N	number of original sources
p	probability that the pulses of two arbitrary sources overlap
R	number of subsets of sources (throughout the thesis we assume the sources can be classified into R subsets, each comprising Q sources, and that the sources from the same subset cannot overlap with each other.
$r_{i,j}^{inv}$	(i,j)-th element of matrix $\mathbf{R}_{\bar{\mathbf{s}}}^{-1}$
$\mathbf{R}_{\mathbf{s}}$	correlation matrix of sources
$\mathbf{R}_{\bar{\mathbf{x}}}$	correlation matrix of extended measurements
$\mathbf{R}_{\bar{\mathbf{s}}}$	correlation matrix of extended sources
Q	number of sources constituting each subset of sources (see explanation for R)
$\mathbf{s}(n)$	vector of pulse sources
$\bar{\mathbf{s}}(n)$	extended vector of pulse sources
$\mathbf{S}_{\mathrm{a}}$	matrix comprising all anti-casual parts of all auto-cumulants and all cross-cumulants that are based on two different measurements
$\mathbf{S}_{\mathrm{c}}$	matrix comprising all casual parts of all auto-cumulants and all cross-cumulants that are based on two different measurements
$s_j(n)$	j-th pulse source
σ^2	noise variance
$\mathbf{T}$	orthogonal projection matrix
$v_{n_i,G_{n_i}}(n)$	superimposition of sources which are simultaneously active at given time moment n_i
$\mathbf{W}$	matrix of cumulant weights
$\boldsymbol{\omega}(n)$	noise vector
$\omega_i(n)$	noise detected in i-th measurement

$\mathbf{x}(n)$	vector of all noise-free measurements
$\overline{\mathbf{x}}(n)$	extended vector of noise-free measurements
$x_i(n)$	i-th noise-free measurement
$y_i(n)$	i-th noisy measurement
$\mathbf{y}(n)$	vector of noisy measurements
$\overline{\mathbf{y}}(n)$	extended vector of noisy measurements

List of abbreviations

AP	Action potential
BSS	Blind source separation
EMG	Electromyogram
FIR	Finite impulse response
HOS	Higher-order statistics
i.i.d	Independent, identically distributed
IC	Inverse correlation
IPI	Inter-pulse interval
MIMO	Multiple-input-multiple-output
ML	Maximum likelihood
MU	Motor unit
MUAP	Motor unit action potential
MVC	Maximum voluntary contraction
STFD	Spatial time-frequency distribution
TF	Time-frequency

1.

Introduction

Digital signal processing has become one of the most important tools in industry processes, research activities and many fields of science. It is also one of the fundamental building stones of the modern society. More and more of everyday activities rely on prompt and precise information, which is always mediated by the signals. Signal processing can be found everywhere, form the wide area of telecommunications, economy, seismology, radar and sonar applications, exploration of space, medicine and biology, down to the microcosm and quantum mechanics. The signals transmit the information about the world we live in and crucially influence our knowledge and understanding of the external world.

Although the signals are often thought to be invented by the humans, they are quite common in the nature itself. Throughout the evolution various living beings learned how to process different signals from the external world. The most obvious examples of such a “natural” signal processing are the human senses of sight and hearing. They shape our everyday life. Yet, we are only beginning to understand the complicated processes in our brains behind them. Less obvious, but not less important, are the electrical signals in human muscles. They control the movement and give us the feedback about the muscle activation. Moreover, we can easily track the relative position of our arms, even blindfolded. This means the human brains are quite capable of calculating the inverse kinematics of our extremities, given just the electrical signals from the muscles. Considering the subject investigated in this dissertation probably the most interesting feature of our brains is their ability to separate the sounds of different sources out of their compound mixtures. This phenomenon can easily be observed at the parties. Even though there are several simultaneously speaking persons, we can easily focus on one of them and ignore the sounds from the others. Without even realizing we are performing one of the very difficult signal processing tasks, so called blind source separation.

From the scientific point of view the signal processing is a wide research field. It includes noise and artefact removal, frequency analysis and synthesis, pattern recognition, coding and decoding, speech and image analysis, etc. Accordingly, many different theories and tools ranging from plain filtering, time-frequency and time-scale analysis, second- and higher-order statistics (HOS), to cepstral analysis and neural networks were developed in the past. One of the many tools which are becoming increasingly important is so called Blind Source Separation (BSS). Its goal is to reconstruct the original sources out of their

multiplicative or convolutive mixtures. Although being a relatively young scientific field, it has been successfully applied to many different areas.

In this dissertation the separation of linear convolutive mixtures is addressed. Although there were many attempts in the past, no general method applicable to arbitrary convolutive mixture of optional signals exists. The approaches developed so far focus on specific mixtures and exploit the information offered by the maximum likelihood estimators [15, 27, 81, 84, 104, 126, 133], second- and higher-order statistic [33, 111, 117, 123, 124, 134, 166, 167], and time-frequency and time-scale analysis [5, 84, 101, 106, 109, 121]. Supposing mutually independent sources, they usually enable only the reconstruction of impulse responses, while the sources are generally reconstructed only up to an unknown filtering effect [17, 84, 106]. In our work we limit the discussion to the linear convolutive mixtures of pulse sources. The latter can be modelled as time-domain sequences of pulses, i.e. the time signals with their energy concentrated in the separated time moments – pulses. Except a common assumption on statistically independent sources, no information about the pulse sources or the coefficients of the mixing filters is supposed. However, we assume several concurrent measurements of the same source signals are given, i.e. so called multiple-input-multiple-output (MIMO) system.

The assumed data model corresponds to biomedical and most of the telecommunication [75, 56, 58, 114, 148, 168, 169] signals and has been exploited by many approaches [9, 18, 39, 49, 51, 60, 89, 94, 102, 103, 129, 144, 145, 171]. However, they all suffer from at least two major drawbacks. Firstly, the majority of decomposition techniques suppose the number of measurements greater than the number of sources. From the practical point of view, this is a highly critical assumption as there are many physical and physiological reasons speaking against it [56, 113]. Secondly, as mentioned above, the source signals are only partially reconstructed. To be exact, the state-of-the-art decomposition techniques work well only when the number of sources is small and their contributions (impulse responses) do not overlap significantly.

In this dissertation, three novel approaches for blind decomposition of convolutive mixtures of pulse source signals are introduced. The possible routes towards a perfect reconstruction of pulse source signals and their impulse responses are examined, and the influence of noise and non-orthogonal (statistically dependent) source signals are evaluated. The identification of slightly underdetermined MIMO systems (systems in which the number of sources slightly exceeds the number of measurements) is also studied.

The novel decomposition approaches are tested on surface electromyographic (SEMG) signals [19, 113, 146]. The latter comprise a large number of active sources with relatively short impulse responses. Namely, during voluntary contractions of skeletal muscles small functional groups of muscle fibres, called motor units (MUs), contract synchronously producing a measurable electrical signal called motor-unit action potential (MUAP). When recorded by an array of electrodes at low muscle contraction level, the MUs contributing to the SEMG signals can be treated as close-to-orthogonal pulse source signals whose cross-correlations do not exceed approx. 5-10 % of their maximal possible

values [56, 114, 139, 158]. The SEMG signals can, therefore, serve as a perfect example of convolutive mixtures of pulse source signals investigated in this dissertation. Finally, the application of developed techniques to the SEMG signals increases the value of our work, as a major step towards a complete reconstruction of constituent MUAP trains is made (see Section 3.2).

To summarise, two different scientific fields; blind source separation and surface EMG decomposition are studied. Although all decomposition approaches are based on general pulse signals we frequently use the surface EMG signals to exemplify and illustrate the efficiency of applied solutions, or to simply demonstrate the complexity of convolutive mixture under investigation.

The main assumptions and expected limitations are summarised in the following thesis:

In the case of close-to-orthogonal pulse source signals and more measurements than sources, a complete blind reconstruction of convolutive mixing matrix and source pulse trains out of their convolutive mixture is possible. When the number of sources exceeds the number of measurements only a limited number of source signals can be completely reconstructed. Processing the surface EMG signal, the impulse responses correspond to the motor-unit action potentials, while the pulse source signals stand for the motor-unit innervation pulse trains.

Enlightening the decomposition assumptions, possibilities and limitations at least three groups of hypotheses can be introduced. They separately address the expectations in the case of non-orthogonal source signals, in the case of more sources than measurements and in the presence of noise:

1. *In the case of more measurements than sources, a complete reconstruction of pulse source signals and their impulse responses out of their convolutive mixtures is possible. When the number of sources exceeds the number of measurements only the projections of source signals into the subspace of measurements can be observed. As a result, the impulse responses in the convolutive mixing matrix interfere and only a limited number of source signals can be completely reconstructed.*

2. *Supposing mutually independent source signals, the influence of convolutive mixing matrix can be eliminated by the second- and higher-order statistics. When processing non-orthogonal source signals, the impulse responses in statistical moments and higher-order cumulants interfere what introduces additional errors to the decomposition process. As a result, the second- and higher-order statistics can be used to separate only orthogonal, or at least close-to-orthogonal, source signals whose pulses do not overlap significantly. Similar limitations can be encountered when using time-frequency analysis.*

3. *In noisy environments, only a limited number of source pulse sequences can be completely reconstructed. Their exact number depends on the signal-to-noise ratio and on the power of source signals. The latter (in the case of pulse signals) corresponds to the number of pulses and, consequently, to the source firing rate. The sources with higher firing rates are, therefore, expected to be more successfully recognised than the sources with lower firing rates. Among all th*e *proposed approaches the methods based on HOS are expected to be superior in noise rejection.*

This dissertation is organized in 8 chapters. First, in Chapter 2, overview of the state-of-the-art in the blind source separation is outlined. Many solutions from the multiplicative cases can be straightforwardly generalized to the convolutive ones. Hence, the different approaches to the decomposition of both multiplicative and convolutive mixtures are critically evaluated and mutually compared.

In Chapter 3, the main principles of the anatomy and physiology of human muscles are briefly summarized. The descriptions given should only serve as a basis for the evaluation of the results on real SEMG (Chapter 6). Hence, the main focus is on the generation of action potentials in muscle fibres, MU firing frequencies, and the filtering effect of the subcutaneous tissue. Next, an overview of the state-of-the- art EMG decomposition techniques is given, along with their main pros and cons. Finally, the conditions under which the surface EMG signals can be modelled as linear, time-invariant MIMO systems are enlightened.

In Chapters 4 and 5, three novel approaches to the decomposition of convolutive mixtures of close-to-orthogonal pulse source signals are derived. First, in Chapter 4 the identification of overdetermined MIMO systems (systems with more measurements than sources) is studied and two novel decomposition methods introduced. The first one is based on time-frequency analysis, while the second one utilises the higher-order statistics. In Chapter 5, the decomposition is extended to slightly underdetermined MIMO systems. Based on algebraic approaches and second-order statistics, a novel concept of blind source separation, the so called inverse correlation based method, is introduced.

In Chapter 6, the introduced decomposition methods are tested on both synthetic and real surface EMG signals. The influence of noise and non-orthogonal source signals are evaluated and the impact of different MU parameters studied, while in Chapter 7 the decomposition results are mutually compared. We conclude our work in Chapter 8 where the principal hypotheses are evaluated and the limitations of the introduced approaches are discussed.

1.1 Explanations of frequently used synonyms

We have explained that two different areas (blind source separation and surface EMG decomposition) will be studied in this dissertation. Although we will try to strictly separate them, the basic terminology of both areas will be used. For example, when describing the EMG signals we will be talking about the innervation pulse trains and action potentials. On the other hand, deriving novel decomposition approaches we will use more general terms such as pulse source signals and system responses. While the innervation pulse trains can certainly be modelled as pulse signals, we will prefer to use more specific names whenever possible. General notations will only be used in Chapters 4 and 5 in order to stress the generality of the derived decomposition methods.

In the sequel, the most common terms from both areas (blind source separation and surface EMG decomposition) are clarified. They should not be understood as definitions. They are just explanations of the most frequently used synonyms.

Pulse source: a source, whose signals can occupy only binary values 0 and a, where a denotes an arbitrary nonzero real number. Samples with value a are referred to as pulses. In the sequel, the term "pulse source" will frequently be used as a synonym for "source pulse signal". Although this can induce some uncertainties, it is quite common convention in the field of blind source separation.

Inter-pulse interval (IPI): time distance between two successive pulses in a source pulse train.

Superimposition of signals: the element-wise sum of signals. In this dissertation, two different superimpositions – the superimposition of pulse sources and the superimposition of motor unit action potentials – will be studied.

Close-to-orthogonal: the phrase "close-to-orthogonal" will be used to denote the sources whose cross-correlations do not exceed 10 % of their maximal possible values. In the cases of pulse sources, this implies only up to 10 % of pulses of two arbitrary source pulse trains can overlap.

System response: the contribution to a single measurement caused by a single firing of a pulse source. It is commonly referred to also as a system impulse response and is defined in more detail in Section 3.3. In the sequel, the motor unit action potentials will be modelled as special forms of system responses.

Motor unit (MU): a small functional group of muscle fibres which are innervated by the same alpha motor neuron and contract synchronously (Section 3.1). Throughout this dissertation, MU will be understood as an origin of the surface EMG signals, with its innervation pulse train as a close-to-orthogonal pulse source signal, and its action potential as a system response.

Innervation pulse train: the time sequence of action potentials in alpha motor neuron, which are transferred through neuromuscular junction to all the muscle fibres of the same MU. As they trigger the responses of MUs, actions potentials in motor neurons are commonly referred to as pulses. Throughout our work, the innervation pulse trains will be modelled as close-to-orthogonal pulse sources. From the blind source separation viewpoint, this is a very coarse approximation because the pulses in neurons are at least a few ms long and occupy more than two different values (when sampled with high frequency). Nevertheless, only the MUs' responses (MUAPs) to innervation pulse trains are measured by the surface electrodes, while the shape of neuron potentials do not impact the shapes of MUAPs. Therefore, action potentials in a motor neuron can be conceptually replaced by true binary pulses. The innervation pulse trains and MUAPs are described more in detail in Sections 3.1 and 3.2.

Action potential (AP): a rapid change in the transmembrane potential in a neuron or muscle fibre triggered by electric, chemical or mechanical excitation of the fibre membrane. It is usually several ms long and propagates along the neuron or muscle fibre from its origin towards the end of the fibre.

Motor unit action potential (MUAP): a superimposition of filtered action potentials of all muscle fibres comprising the corresponding MU as detected in a single EMG measurement. The subcutaneous tissue between a particular muscle fibre and a measuring electrode acts as a low-pass filter and attenuates the single fibre AP. The MUAP shape depends on many physiological parameters (the MU depth in muscle tissue, the number of MU's fibres, their conduction velocities, etc.), on the properties of the detection system (the type of spatial filter, inter-electrode distance, the shape of electrodes, etc.), as well as on the relative position of single fibres with respect to the detection point (the spatial filter position). The shape and the amplitude of MUAPs also vary as a function of the distance from the innervation zone. As a result, the MUAPs of different MUs, as well as MUAPs of the same MU measured at different positions in space, will differ significantly. In this dissertation, the MUAPs will be commonly modelled as special forms of system responses.

Electromyogram (EMG): a time record of superimposed MUAP sequences from all MUs which are active in detection volume of measuring electrodes.

Spatial filter: a configuration of pick-up electrodes used to measure a single EMG measurement. There are many possible electrode configurations (spatial filters). The simplest one is a unipolar electrode (Fig. 1.1 a). Single differential (SD) spatial filter (Fig. 1.1 b) consists of two electrodes while the measurement is formed by subtracting the signals measured on both electrodes. Double differential (DD) filter is depicted in Fig. 1.1 d and comprises three electrodes. The EMG measurement is formed by, first, adding together the signals from two side electrodes and, afterwards, subtracting twice the signal from the central electrode. Spatial filter can also be classified according to their orientation with respect to the muscle fibres. Longitudinal spatial filters (Fig. 1.1 b and d) have their electrodes aligned with the muscle fibres, while in the case of transversal

filters (Fig. 1.1 c and e) the electrodes are placed transversal to the orientation of muscle fibres.

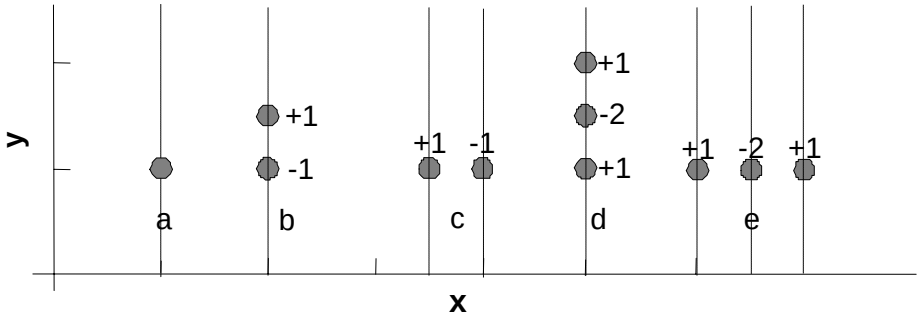


Figure 1.1: Spatial filters commonly used for surface EMG recoding: a – unipolar electrode, b – longitudinal single differential filter, c – transversal single differential filter, d – longitudinal double differential filter and e – transversal double differential filter. Electrodes are depicted by grey circles, while the corresponding numbers denote the weights with which the signals from electrodes are multiplied before they are summed up to yield the final EMG measurement. The muscle fibres are assumed aligned with the y axis.

2.

Blind source separation

Blind source separation (BSS) [4, 24, 78, 79, 84, 88, 97] is becoming an increasingly important tool in the filed of signal processing. For the last decade it has been successfully applied to the areas of radar [32, 42, 79] and sonar signal analysis [78, 106], speech recognition [84, 106, 122, 81], telecommunications [5, 45], separation of seismic signals [151, 79], image processing [35, 162, 174] and, lately, to the analysis of biomedical data [149, 8, 87, 86, 105, 156, 175, 173]. Its goal is to reconstruct the original sources given just their multiplicative or convolutive mixtures. The mixing process is usually denoted by the so called mixing matrix, which is, of course, unknown. In the case of multiplicative mixtures, the mixing matrix consists of scalar factors (weights). In the case of convolutive mixtures, the mixing matrix comprises impulse responses of different channels (convolutive filters) [79, 106]. Usually several measurements of the same sources are available, which results in MIMO data model [33, 100, 163, 164].

This chapter provides a short overview of the state of the art in the blind separation of linear mixtures. The main focus is on convolutive mixtures, although the methods separating the multiplicative mixtures are also described. There are at least two reasons for that. Firstly, many methods for separation of convolutive mixtures prove to be straightforward generalizations of the solutions developed for the multiplicative cases. Secondly, every convolutive mixture can always be reformulated into the multiplicative one, opening a possibility of solving it by the multiplicative BSS techniques.

The decomposition techniques described in this chapter are classified with respect to their basic decomposition principle. In the case of multiplicative mixtures (Section 2.1), the methods based on the maximization of non-Gaussianity, maximum likelihood estimators, and time structure of sources, are discussed. Although the main focus is on the separation of overdetermined MIMO systems, the methods separating the underdetermined MIMO systems are also described. Section 2.2 outlines possible routes towards separation of linear convolutive mixtures. The methods based on maximum likelihood, Fourier transform, and reformulation of the convolutive mixture to the multiplicative one are briefly summarized. Major assumptions, along with pros and cons of each group of methods, are also presented.

2.1 Separation of linear multiplicative mixtures

In the case of linear multiplicative mixtures the mixing channels are assumed to be memoryless, i.e. the mixtures are supposed to be instantaneous, depending only on the current values of the sources, while no delayed versions of the same source can be mixed together. In the assumed data model the sources are first multiplied by the unknown factors (weights) and, afterwards, summed up to form the observed measurement:

$$x_i(n) = \sum_j h_{ij} s_j(n), \tag{2.1}$$

where $x_i(n)$ denotes the *i*-th measurements, h_{ij} is unknown mixing factor, and $s_j(n)$ stands for the *j*-th source. We commonly suppose the factors h_{ij} constant in time. Another common assumption is that the number of observed measurements exceeds the number of sources.

Using the matrix notation, (2.1) can be rewritten as

$$\mathbf{x}(n) = \mathbf{H}\mathbf{s}(n), \tag{2.2}$$

where $\mathbf{x}(n) = [x_1(n),....,x_M(n)]^T$ and $\mathbf{s}(n) = [s_1(n),....,s_N(n)]^T$ denote the vectors of measurements and sources, respectively, while **H** stands for the unknown $M \times N$ mixing matrix:

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1N} \\ h_{21} & h_{22} & \dots & h_{2N} \\ \vdots & \vdots & & \vdots \\ h_{M1} & h_{M2} & \dots & h_{MN} \end{bmatrix}. \tag{2.3}$$

There are many approaches to the blind separation of multiplicative mixtures. The decomposition techniques are based on mutual information and Kullback-Leibner divergence [2, 3, 130, 10, 80, 28], on second- and higher-order statistics [5, 13, 30, 41, 65, 66, 67, 73, 84, 108, 110, 123, 124], on time-frequency [6, 14, 17, 21, 37, 73, 84, 120] and time-scale analysis [84, 9, 148, 171], on algebraic [4, 11, 29, 38, 68, 84, 112, 173] and geometric approaches [28, 84, 82, 118, 135] and, finally, on the maximum likelihood estimators [10, 15, 126, 133]. The majority of them supposes independent and identically distributed (i.i.d.) sources (the source samples are modelled as independent, equally distributed random variables) [4, 22, 23, 24, 38, 78, 152]. It has been shown [61, 79, 84] that the separation of i.i.d sources is possible only if at most one source has Gaussian distribution. The other BSS methods consider the case of temporally correlated sources [13, 65, 66, 97, 153]. Identifiability in such an approach is granted even when the signals are normally distributed. The third possible route to multiplicative BSS is to exploit the inequalities in the source sample distributions, i.e. non-stationarities of sources [14, 17, 21, 37, 73, 120].

In the sequel, the most frequently used groups of methods for separation of linear multiplicative mixtures of sources are coarsely outlined. Their detailed overview can be found in [4, 24, 84, 106].

2.1.1 Separation by maximization of Nongaussianity

Vast majority of the multiplicative BSS methods are based on the central limit theorem, which states that the distribution of a sum of i.i.d. random variables tends toward a Gaussian distribution. In other words, supposing the samples of sources i.i.d. non-Gaussian random variables, the central limit theorem guarantees that their linear mixture is more Gaussian than any of the original samples. Moreover, the mixture becomes least Gaussian exactly when it equals to the sample of one of the sources. Hence, in the separation process we look for such a separation matrix which maximizes the non-Gaussianity of the measurements. Although the exact distributions of the measurements are usually unknown, we can always estimate them out of the signal samples.

The non-Gaussianity of the measurements can also be estimated indirectly. One of the most frequently used measures utilizes higher-order cumulants [111, 166, 167]. They are known to be zero for all the Gaussian random processes. This fact is effectively utilized by the separation methods which typically use the gradient based optimization in order to numerically maximize the values of third- or higher-order cumulants of measurements. The second possible measure of non-Gaussianity is negentropy [84]. It is defined as a difference between the entropy of the observed random variable and the entropy of the Gaussian random variable with the same mean and variance. Being nonnegative and zero if and only if the observed random variable is Gaussian, it is very natural measure for the optimisation.

The main drawback of the methods that use cumulant-based or negentropy-based measures is that they do not distinguish among the sources. In the optimisation process the global maximum is found, which implies only one source can be successfully reconstructed. To reconstruct several sources, the separation algorithm must be run several times using different initialization points. It is easy to see that this is not a reliable solution. More consistent methods combine the results of the described approaches with different orthogonalization algorithms which utilize the mutual orthogonality of the whitened vectors of measurements [82, 83]. With their help all the sources can be reconstructed.

2.1.2 Separation based on maximum likelihood

A second, very popular class of multiplicative BSS methods comprises the maximum likelihood (ML) estimators. Given the probability density of the sources and their mixing matrix, the likelihood function of the mixture can be derived from the well-known result of density of a linear transformation [84, 126]. Usually, due to its algebraic simplicity, the logarithm of the likelihood is numerically optimized. This makes no difference, since the maximum of the logarithm is obtained at the same point as the maximum of the likelihood. ML-based methods simply try to find the mixing matrix (and the source probability densities) which gives the highest probability for the observations. The numerical maximization of the likelihood function is usually performed by different gradient-based algorithms [15, 48, 126, 133].

The log-likelihood function of the mixture is, in general, function of both the elements of the mixing matrix and the probability densities of the sources. The former are considered the parameters of the assumed data model and can easily be optimized. The optimization of the source probability densities is much more difficult problem, because, in general, their estimation cannot be reduced to the estimation of a finite parameter set. There are two possible ways to avoid the nonparametric density estimation [15, 27, 84, 126, 133]:

1. Sometimes the densities of the sources (or at least their family) might be known in advance. In this case this prior information can be used in the optimization process, making the log-likelihood function dependent of the mixing matrix only.

2. The unknown densities can also be approximated by a family of densities that depend on a finite and small set of parameters. Usually, very simple and general parameterizations of the densities yield satisfactory results.

The class of ML-based separation methods comprises many different solutions. Among them, the methods based on the infomax principle [10, 27], entropy and Kullback-Leibner divergence [28, 80, 82], and mutual information [131, 72, 84] are the most popular. Their main drawback is their computational complexity.

2.1.3 Separation of time signals

The BSS methods considered so far model the source samples as i.i.d non-Gaussian random variables. Measuring non-Gaussianity of the mixtures, they can only be used for separation of mixtures with at most one Gaussian source. In the case of two or more Gaussian sources, the described approaches fail [13, 14, 28, 84] and we are forced to seek for an alternative solution. Often the observed sources are time signals. As such they contain much more structure than the simple set of strictly independent random variables. This structure can then be used to improve the separation, or even to make it feasible. Actually, exploiting their time structures also Gaussian sources can be separated.

In the past, two large families of the time-structure based methods evolved. The first one exploits the time dependencies among the samples in each source. In other words, the samples of each source are modelled as dependent, identically distributed random variables. The large majority of the representatives of this family employs so called joint (simultaneous) diagonalization of several correlation matrices at different time lags [13, 16, 28, 31, 66].

The second family of the time-structure based methods takes advantage of the nonstationarity of the sources. Here, the source samples are modelled as independent random variables with time-dependent (nonstationary) distributions. Most of the representatives of this family rely on time-frequency analysis and on the fact that the mixing matrix is invariant to the Fourier transform [14, 21, 37, 73, 132].

Due to their simplicity and implicit capability of separating the Gaussian sources, the time-structure based methods gained a lot of interest over the past few years. Many efficient solutions were proposed for both multiplicative and convolutive mixtures. One of many possible routes towards the separation of convolutive mixtures is described in Section 4.1.

2.1.4 Identification of underdetermined system

So far we supposed the linear multiplicative mixtures have more measurements than sources. The described separating methods aimed to estimate the mixing matrix, while the sources were reconstructed by simply applying the inverse of the mixing matrix to the measurements. This is no longer the case when the mixtures with more sources than measurements are processed. In such a case, the mixing matrix cannot be inverted any more and the exact values of the sources cannot be reconstructed, even if we know the mixing matrix exactly. Hence, the separation of the underdetermined systems is a difficult problem and can be compared to the mathematically well understood problem of solving an underdetermined system of equations. Although infinitely many possible solutions exist, we can usually gain satisfactory results by limiting the properties of the sources [69, 81, 84, 104].

Separating the underdetermined system we face two large problems. First, we have to estimate the mixing matrix, and second, we have to estimate the unknown source signals. Supposing the mixing matrix known, the sources can be estimated by applying its pseudo-inverse to the measurements [68, 92]. Although simple, this is not the optimal solution. Much more consistent approach is based on the maximum likelihood estimator [27, 84, 126, 133]. The problem with the ML estimator is that it cannot be expressed in analytic form and must be optimized numerically, which highly increases its computational complexity.

In general, the mixing matrix is unknown and must be estimated from the measurements. Again, the most popular approaches are based on the ML estimators. Among them, so called Monte Carlo methods based on the Bayesian estimation and numerical integration, are becoming increasingly important. The most popular methods of this family utilize so called Gibbs sampling to estimate the missing probability densities [125, 59]. Their drawback is that they are very computationally demanding. Much more attractive are the so called orthogonalization methods [84, 96] which are based on a rather unrealistic assumption of quasi-orthogonal columns of the mixing matrix.

2.2 Separation of linear convolutive mixtures

There are at least two reasons why the multiplicative mixtures are relatively rare in practice. First, measuring the observed signals with the array of sensors, the finite propagation speed of the signals induces different time delays of the same source in different measurements. Second, due to the multi-path propagation in the observed medium each measurement contains different time-delayed versions of the same source. This calls for the convolutive mixture model, i.e. the model, where the contribution of each source to each measurement is modelled by a convolution. Taking MIMO system into consideration, the *i*-th measurement can be expressed as

$$x_i(n) = \sum_{j=1}^{N} \sum_{l=0}^{L-1} h_{ij}(l) s_j(n-l), \quad i=1,...,M, \tag{2.4}$$

where $s_j(n)$ and $h_{ij}(l)$ denote the unknown source and finite convolution filter, respectively. The goal behind the separation of convolutive mixtures is to reconstruct the convolutive filter $h_{ij}(l)$ and the sources $s_j(n)$, given just the measurements $x_i(n)$. As in the multiplicative case, the samples of the sources are assumed to be i.i.d. random variables, with unknown probability densities.

The separation of convolutive mixtures appears in the literature under many different names: multichannel blind deconvolution, blind identification of MIMO systems, independent component analysis (ICA) with convolutive mixtures, multichannel blind identification, etc. The approaches developed so far are based on ML estimators [16, 27, 84, 100, 133], time-frequency analysis [17, 98, 101, 106, 109, 119, 121, 128, 157] and reformulation of the convolutive mixtures to the multiplicative ones [17, 70, 75 154]. Supposing mutually independent sources, they usually enable only coarse estimations of the convolutive filters, while the sources are generally reconstructed only up to the unknown filtering effect [17, 44, 84, 106, 119]. This section provides a quick overview of the most popular decomposition techniques.

2.2.1 Maximum likelihood based approaches

As already explained in the previous section, the ML based methods are of the most popular ones. Being based on the probability density of the sources, they are inherently capable of capturing the prior knowledge about the sources. Moreover, they often provide the superior performance and can easily be extended to the underdetermined case. Their main drawback is their computational complexity.

It should not come as a surprise that the ML estimators can also be extended to the convolutive case. Actually, the proposed solutions are more or less direct generalizations of the solutions developed for the multiplicative cases [17, 98, 101, 106, 109, 121, 119, 128, 140, 157]. Being based on numerical optimisation of a large number of unknown parameters, they inherently suffer from a high computational complexity and slow

convergence. Although numerous attempts to speed them up were made in the past, the effective solutions were proposed only for the simplest convolutive model possible, the so called single-input-single-output (SISO) system.

2.2.2 Separation in time-frequency plane

The largest, and probably the most effective class of convolutive BSS methods utilises the Fourier transform. It is based on well known fact that the convolution in the time domain becomes the product between the Fourier transforms in the frequency domain. In other words, using the time-frequency transformation, the convolutive mixture can be transformed into a set of multiplicative mixtures in the time-frequency domain [17, 101, 106, 109, 121, 119, 157, 128, 98].

The main drawback of the time-frequency based methods is that we have to estimate the mixing matrix for every possible frequency bin. This is not a trivial problem, due to the implicit indeterminacy of the multiplicative BSS problems (Section 3.3.2). Namely, no matter how successful the decomposition, when compared to the original sources, the reconstructed sources can always appear randomly permutated and even multiplied by the unknown scaling factor. Reconstructing the sources in the frequency domain, the permutation and the scale of the sources is generally different for every frequency bin. To be able to reconstruct the source in the time domain, we must first find the way how to find the frequency components that belong to the same source. This problem has been known as a permutation problem and has been intensively studied in the past [119, 157, 128]. The most effective solutions utilize the spectral continuity of the sources [141 127], information about the direction of arrival (DOA) [85, 128] and various combinations of the abovementioned solutions [84, 157]. Nevertheless, the permutation problem remains, especially when decomposing large convolutive MIMO systems. As a result, the time-frequency methods have only been applied to relatively small overdetermined MIMO systems, with no more than 10 simultaneously active sources.

2.2.3 Reformulation of convolutive mixture into multiplicative one

By adding the delayed repetitions of each source to the vector of sources $\mathbf{s}(n) = [s_1(n),....,s_N(n)]^T$, every convolutive mixture can be transformed into the multiplicative one (see Section 3.3) and solved with one of many possible solutions mentioned in Section 2.1. Although the number of sources is highly increased, this is a very attractive solution as there is no permutation problem. The main drawback of such a solution lies in the statistical properties of the newly formed sources which (at least in general case) are not statistically independent anymore. This proves to have a crucial impact on the separation performance. Although many attempts to overcome this setback were made, the up-to-date methods enable only reconstruction of sources up to the unknown filtering effect [11, 17, 75].

3.

Electromyogram

Electrical activity of muscles has been under the scientific investigations for centuries. Already in the 18th century, Galvani discovered a surprising property of muscles, which contract when elicited by the electrical stimulation. He also showed the muscle responses can be measured in the form of electrical potentials. In 1929, Adrian and Bronk used concentric needle electrode to measure the electrical potentials inside the live skeletal muscle. Studies that followed revealed the secret behind the electrical signals in the muscles. It was soon realized how the muscle fibres are organized into the basic building blocks of muscles, so called motor units (MUs) and numerous studies of their electrical properties emerged.

For a long time, the EMG recordings have been based on the expensive and invasive needle electrodes. It was only in the second half of the previous century when the investigations of so called surface EMG (SEMG), i.e. the EMG measured by the electrodes placed upon the skin surface, emerged (Basmajian 1974, Deluca 1984, Merletti and LoConte 1995). Due to their implicit non-invasive nature, surface electrodes have numerous advantages [76, 99, 134, 166, 167, 170, 161]. There is significantly less discomfort, what allows for unlimited repetition of the long-term electromuscular monitoring in exactly the same place. Furthermore, recording of surface EMG is inexpensive and gives global information about muscle activity [19, 58, 113, 146]. Nevertheless, the SEMG hasn't made its way to the clinical practice yet. This is mainly due to the poor morphological information caused by the filtering effects of the subcutaneous tissue. Moreover, there is usually a large number of active MUs in the detection volume of a surface electrode, causing the detected signals to be highly complex. Many attempts to enhance the single MU information in SEMG were made in the past, nevertheless, the decomposition of the SEMG signals remains a very delicate process.

This chapter provides a short overview of main principles behind the muscle activities. The main focus is on the generation and extinction phenomena of the action potentials in muscle fibres, and on the filtering effect of the surrounding tissue. These descriptions will serve as a basis for evaluation of decomposition results in Chapter 6. In Section 3.2, the state of the art in the decomposition of both needle and surface EMG is briefly outlined, while in Section 3.3 the mathematical model of SEMG which will be used in this dissertation is formulated.

3.1 Physiology of human muscles

Human muscles consist of 10 to 150 mm long and 5 to 90 μm thin muscle fibres, which are in the tendon regions attached to the bones. Their orientation varies substantially between the muscles. Although the muscle fibres can be aligned with the direction of the tendons, they more often form a non-zero angle with them. Beside the muscle fibres, the muscle tissue comprises various other structures: nerves, blood vessels, connecting tissue, etc. There is also a layer of subcutaneous fat separating the muscle from the skin.

Each muscle fibre is innervated by a single motor neuron which transmits the control commands from the central nervous system in a form of the firing pulse trains. Normally several muscle fibres are innervated by the same neuron, forming a basic functional unit of the muscles, so called motor unit (MU). The number of fibres in each MU varies considerably between the muscles. The smallest MU can be found in the eye muscles and comprise less than 10 fibres. The largest MU belong to the large leg muscles and combine more than 2000 fibres [137]. The size of MU varies considerably also within the same muscle. It is important to notice that the fibres belonging to the same MU are not adjacent to each other, but rather spread in space within the radius of approx. 2 to 15 mm.

Considering the time needed to achieve their peak tension the muscle fibres can be classified into two disjunctive classes: slow or type I, and fast or type II. The fibre type is determined by the central nervous innervation. Consequently, all the fibres from the same MU are of the same type, while all muscles comprise all types of MUs. The area of neuromuscular junctions between the muscle fibres and motor neuron is called innervation zone. Although the fibre is always innervated approximately in the middle between both tendons, the location of the innervation zone can vary considerably with respect to the muscle belly (the fibres are usually not parallel to each other).

Every pulse in motor neuron triggers a contraction of MU's fibres. During a voluntary contraction the whole sequences of pulses, so called innervation pulse trains, are transmitted by the motor neuron. The number of pulses per second, i.e. MU firing rate, depends on the type of muscle fibres and on the muscle contraction level. At low contraction levels the small and slow MUs of type I are first activated. Their firing rate varies between 6 and 10 Hz. When the contraction force is increased, larger MUs of type I, as well as the fast MUs of type II are activated. Their firing rates hardly exceed 20-30 Hz, under the normal conditions. By increasing the firing rate, the contractions triggered by each individual pulse in motor neuron are merged together to form constant and smooth muscle contraction.

3.1.1 Motor unit action potentials

Every pulse in the motor neuron triggers a brief change in the transmembrane potential of the muscle fibre, so called action potential (AP). This is mainly caused by the change of the membrane's permeability to sodium (Na^+) ions, which are located in the space

between the muscle fibres. The Na^+ ions enter the fibre and cause the depolarization of the transmembrane potential which changes its sign from the negative (approx. -70 mV) to the positive one (approx. +20 mV). Afterwards, the membrane's permeability to sodium ions decreases, while the permeability to potassium ions K^+ increases. Outflow of the K^+ ions causes the repolarisation of the transmembrane potential. Finally, using the Na^+/K^+ pumps the Na^+ and K^+ ions are actively transported out of and into the muscle fibre, respectively [47].

A fibre (and therefore a motor unit) cannot be activated again until the AP due to a previous activation is almost completed. This is well known as a refractory period. In addition, during sustained muscle contractions, single MU fire at most 20-30 pulses per second. In rare cases there may be so called "doublets", that is two firings of the same MU separated by only 10-15 ms.

The AP in the muscle fibre is initiated at the neuromuscular junction and propagates in both directions towards the tendon regions where it is gradually suppressed. The AP amplitude (approx. 100 mV), duration (3 to 5 ms) and its propagation/conduction velocity (3 to 5 m/s) depend both on the fibre diameter and on its type. The single fibre AP can be measured also in the close vicinity of the muscle fibre. The passive electric media surrounding the muscle fibres acts as a low-pass filter which decreases the AP amplitude and increases its duration. Moreover, the contributions from different muscle fibbers can be thought summed together [56, 60, 89]. As a result, measuring the electrical potential at an arbitrary position in space, the AP from different fibres will appear superimposed in both time and space. Again, due to the filtering effect of the surrounding tissue, the contributions from the fibres that are close to the measuring site will be stronger [138]. Usually, the detected AP from all the fibres belonging to the same MU are thought to be summed together forming so call motor unit action potential (MUAP). The superimposition of the single fibre APs can now be replaced by the superimposition of MUAPs. It is very important to realize this is nothing else but a very practical convention, which highly reduces the number of the components in the detected signal, while still enabling a study of the MU's properties.

It is also important to realize the shape of MUAP can vary enormously with the position of the detection point. MUAPs from the same MU, detected at different positions in space, will differ substantially from each other. Not only will they have different time delays (due to the finite speed of the MUAP propagation along the fibres), the filtering effect of the surrounding tissue, as well as the differences in the single fibre conduction velocities, will cause them to differ substantially in shape, length and amplitude. Even bigger differences can be expected between the MUAPs of different MUs. As already explained, the MUs differ in both the number and the type of muscle fibres. Fast MUs will have shorter and higher AP than the slow ones.

Although being ambiguous and implicitly dependent on the many hidden factors, the term MUAP is well established in the field of biomedical signal processing, as well as in the filed of neurophysiology.

3.2 EMG decomposition

Decomposition of the EMG signals to constituent MUAP trains has proved to be a very important clinical issue. Reconstructed sequences of innervation pulses provide the basis for research studies and clinical examinations of MU control properties, recruitment strategies, inter-pulse interval (IPI) variability, short term MU synchronization, and myoelectrical manifestation of fatigue, to name just a few areas of investigation. The existing computer-aided EMG decomposition methods have been mainly focused on the intra-muscular EMG signals [1, 18, 20, 39, 49, 51, 60, 64, 89, 94, 102, 103, 129, 143, 144, 145]. Being based on pattern recognition [43] and clustering in the time domain, on spatial filters, and on time-frequency and time-scale analysis, they typically exploit the differences in morphology of MUAPs and comprise two steps. Firstly, individual MUAP is identified from the raw inter-muscular EMG or from their time-frequency/time-scale transforms. Secondly, the identified MUAP is classified and assigned to the best fit class of previous decomposed MUAP shapes. The most frequently used measures include the correlation and amplitude, length and shape of MUAPs. In the case of multichannel EMG, the decomposition results from different EMG recordings can be mutually compared, which significantly improves the decomposition efficiency. Unfortunately, most of the methods fail when a large number of MUAPs become superimposed, the complete decomposition is possible only if the number of simultaneously active MUs is relatively small. On the other hand, even at low contraction levels, sufficiently large number of MUs contracts asynchronously to produce highly complex interference patterns.

Most recent pattern recognition decomposition techniques make use of the solutions from graph theory and artificial intelligence to cope with the complex superimposed MUAPs patterns [18, 95, 159]. Their main goal is to reduce the set of possible MUAP combinations in order to speed up the decomposition of the detected pattern. They build upon different priors about the MU firing patterns to predict the most probable solution. Artificial intelligence based methods have another nice property. They constantly reassess their own results and learn from the mistakes. As a result, they are capable of adapting to the time-varying properties of the detected signals. Their results are quite promising and the most successful approaches have already been transferred to the everyday clinical practice [95, 159]. Unfortunately, their use is restricted to the needle EMG, while in the case of surface EMG signals the decomposition fails.

In the case of needle EMG, the contributions of MUs are relatively strong (with the amplitude of approx. 100 mV) and short (about 10 ms). On the contrary, in the case of surface EMG, the filtering effect of subcutaneous tissue attenuates the morphological differences of MUAPs, decreases the MUAP amplitude (to approx. 100 μV), while the MUAP duration is highly increased (up to 25 ms). Moreover, in the case of needle electrodes we can selectively observe the action potentials of only a few active motor units, or even of a single muscle fibre. In the SEMG case, on the other hand, we deal with several tens of active MUs and the measurable contributions from other muscles not being under the clinical investigation, what is often referred to as muscle cross-talk [57]. Consequently, surface EMG comprises highly complex patterns [165] and the approaches

proposed for the decomposition of needle EMG fail. Nevertheless, the non-invasive nature of SEMG and its wide potential to offer global information about the muscle activity opened numerous application areas. The information extracted from the SEMG signals is being exploited in several different clinical studies mainly concerned with the timing of the muscle activation [54, 116], the EMG amplitude modulation [36], changes in the frequency content of the EMG signals [7, 55], and the conduction velocity estimation [53].

Many attempts to enhance the single MU information and to suppress the cross-talk in SEMG were made in the past. As already mentioned, the first approaches, which were based on the methods for separation of needle EMG, soon proved to be ineffective. Meanwhile, the SEMG acquisition became a rather developed and matured measuring technique, providing all the required reliability, robustness, and repeatability. Flexible high-density 2D arrays of electrodes have been developed, opening the possibilities of recording up to a few tens or even hundreds of the SEMG measurements simultaneously. Furthermore, with the development of the computer power more sophisticated and computationally complex separation techniques, capable of processing and combining the information from such a large multichannel recordings, became feasible. Among the most popular ones are certainly the methods based on neural networks [91, 161].

The second class of the surface EMG decomposition approaches comprises the time-frequency and time-scale based methods. The former became a popular tool for extraction of information out of the dynamic EMG, i.e. the EMG signals recorded during the dynamic muscle contractions. In such a case, the recorded signals can be modelled as nonstationary and the time-frequency analysis seems to be quite a natural tool for them. Taking the large number of active MUs, their time-varying firing rates, and their complex activation and deactivations patterns into account [50], also the signals recorded during the isometric contractions can be modelled as nonstationary. Nevertheless, the time-frequency based methods proved to have limited success when it comes to the separation of surface EMG signals.

On the other hand, the methods based on time-scale analysis try to find the mother wavelet which would perfectly resemble the properties of the one of the detected MUAPs [9, 148]. The construction of such a perfect wavelet is highly critical. First, it depends on the shape of detected MUAPs, and, secondly, it is highly time-consuming. Moreover, even if we find a perfect mother wavelet, it is highly unlikely that it will be able to significantly suppress the contributions from other MUs. Consequently, the wavelet-based decomposition gained little prosperity.

Only recently some more sophisticated decomposition methods have begun to emerge. Exploiting the advanced blind source separation techniques (Section 2) they provide new insights of clinical interest as they, for example, make their way towards the removal of artefacts [8, 105, 150, 156, 160], cross-talk suppression [52] and the identification of the EMG constituent MUAPs [4, 62, 172]. Nevertheless, the decomposition of the SEMG signals to their constituent MUAP trains, even at low muscle contractions, remains a very delicate process. Although the analysis of individual MUs and their firing patterns is very

important in the clinical electromyography [34, 40, 50, 53, 63, 90, 113, 144, 142, 147, 158], no approach developed so far enables their complete reconstruction based on SEMG signals only.

3.3 Assumed data model of surface EMG signals

EMG recordings can be modelled as a multi-channel linear, time-invariant system, as long as they have been taken in stable, controllable, and stationary measurement session. This implies an isometric muscle contraction and avoiding of the appearance of fatigue-induced changes. Dealing with sampled multi-channel EMG recordings, the discrete, shift-invariant MIMO modelling is most feasible. Each source (channel input) in such MIMO system is considered a MU innervation pulse train triggering the muscle, while the system responses (channels) correspond to the MUAPs as captured by a spatial filter (pick-up electrodes). The individual EMG measurements represent the model outputs.

The shapes of system responses in the assumed MIMO system are of no importance. Hence, any property of MU (e.g. its depth in the muscle tissue, the number of fibres and their conduction velocities, etc.), as well as the properties of the detection system (the spatial filter used, inter-electrode distance, etc.) influencing the shapes of MUAPs can be modelled in the system responses. As explained in Chapter 1, we will often interchange the terms “MUAP” and “system response”. Although in this thesis both expressions will denote the same thing, we will prefer to use the term “system response” whenever we talk about the BSS, and the term “MUAP” whenever we talk about the EMG.

Let now $h_{ij}(l)$ denote the *(i,j)*-th system response, i.e. the MUAP of the *j*-th MU as detected in the *i*-th measurements. Further suppose all the system responses (MUAPs) of length *L* samples (if the system responses differ in length they can always be padded by 0 to reach the same length). Then the contribution of the *j*-th source (MU) to the *i*-th measurements can be modelled by the following convolution:

$$\sum_{l=0}^{L-1} h_{ij}(l) s_j(n-l) \tag{3.1}$$

where $s_j(n)$ stands for the *j*-th source. Taking into account the linearity of the modelled MIMO system, the *i*-th measurement can be written as a sum of contributions from all the sources (MUs):

$$x_i(n) = \sum_{j=1}^{N} \sum_{l=0}^{L-1} h_{ij}(l) s_j(n-l), \quad \text{i=1,...,M} \tag{3.2}$$

where, assuming multiple inputs and outputs, vector notations follow: $\mathbf{x}(n) = [x_1(n),....,x_M(n)]^T$ for the transposed vector of M surface EMG measurements, and $\mathbf{s}(n) = [s_1(n),....,s_N(n)]^T$ for the vector of N pulse sources.

When the influence of noise is considered we get:

$$\mathbf{y}(n) = \mathbf{x}(n) + \boldsymbol{\omega}(n) \tag{3.3}$$

where $\mathbf{y}(n) = [y_1(n),....,y_M(n)]^T$ denotes the transposed vector of M noisy measurements, and $\boldsymbol{\omega}(n) = [\omega_1(n),....,\omega_M(n)]^T$ stands for the noise vector. The assumed data model is depicted in Fig. 3.1.

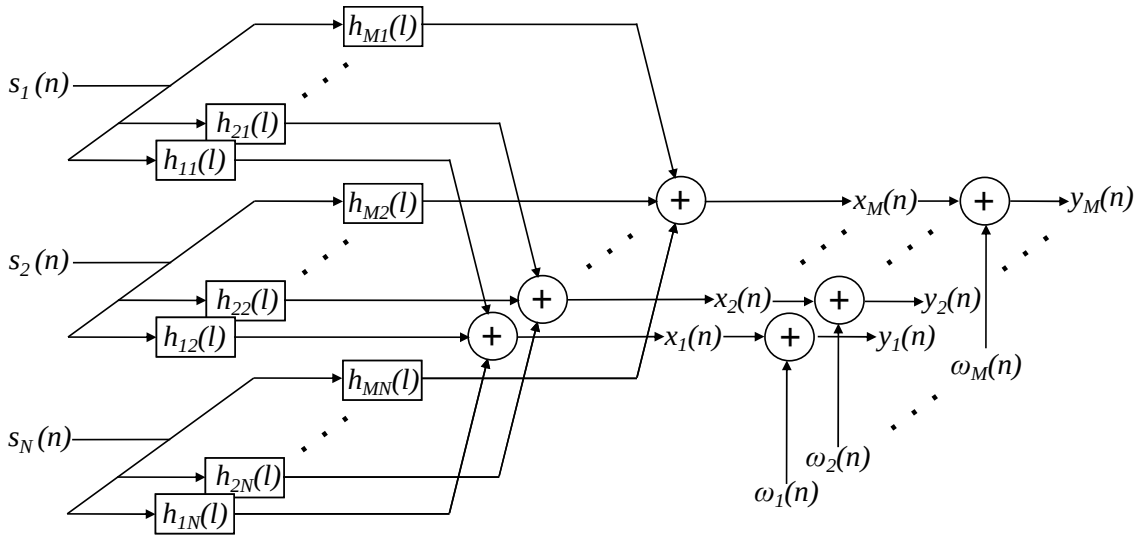


Figure 3.1: Assumed convolutive MIMO system with sources $s_j(n)$, system responses $h_{ij}(l)$, measurements $x_i(n)$, noise $\omega_i(n)$, and noisy measurements $y_i(n)$. The operation of convolution is depicted by rectangles, addition with circles.

We have explained the shape and the length of the system responses depend on numerous parameters (detection system, the depth of the MU in the muscle tissue, the number of muscle fibres constituting the corresponding MU, etc.). But the physiological properties of the muscles guarantee the system responses are limited in their amplitude and length:

$$\sum_{l=0}^{L-1} \left| h_{ij}(l) \right| < \infty, \quad \text{i=1,...,M, j= 1,...,N} \tag{3.4}$$

which implies the stability of the modelled MIMO system. We can further assume the assumed MIMO system is casual:

$$h_{ij}(l) = 0\,, \quad l<0, \quad i=1,...,M,\ j=1,\ldots,N. \tag{3.5}$$

Throughout this thesis, we will often suppose the system inputs $s_j(n)$ mutually independent, with their cross correlations equal to zero:

$$\mathbf{R_s} = \lim_{T\to\infty} \frac{1}{T} \sum_{n=1}^{T} \mathbf{s}(n)\mathbf{s}^H(n) = diag[r_{1,1}, r_{2,2},..., r_{N,N}] \tag{3.6}$$

where $\mathbf{R_s}$ denotes the correlation matrix of sources, $\mathbf{s}^H(n)$ conjugate transpose of the vector of sources **s***(n)*, *diag[]* stands for the diagonal matrix with its parameters on the diagonal, and $r_{j,j} = \lim_{T\to\infty} \frac{1}{T} \sum_{n=1}^{T} s_j(n) s_j^*(n)$ for the autocorrelation function of source $s_j(n)$ with $s_j^*(n)$ as a complex conjugate value of $s_j(n)$. Dealing with the biomedical signals the sources are not strictly orthogonal, while their cross correlations hardly exceed the 10 to 15 % of their maximal possible value. Where needed, the non-ortogonality of sources will be carefully outlined.

The additive noise $\boldsymbol{\omega}(n) = [\omega_1(n),....,\omega_M(n)]^T$ is commonly modelled as stationary, temporally and spatially white zero-mean Gaussian random process, independent from the sources:

$$E[\boldsymbol{\omega}(n+\tau)\boldsymbol{\omega}^H(n)] = \sigma^2\delta(\tau)\mathbf{I} \tag{3.7}$$

where *E[:]* stands for mathematical expectation, $\delta(\tau)$ for the Dirac impulse (delta function), σ^2 for the noise variance, and **I** denotes the identity matrix.

3.3.1. Extension of the convolutive MIMO vector form

Convolutive relationship described in (3.2) can always be expressed in the matrix form:

$$\mathbf{x}(n) = \mathbf{H}\bar{\mathbf{s}}(n) \tag{3.8}$$

where **H** stands for the so called mixing matrix of size $M \times NL$ which contains the unit sample responses $h_{ij}(l)$:

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_{11} & \cdots & \mathbf{h}_{1N} \\ \vdots & \ddots & \vdots \\ \mathbf{h}_{M1} & \cdots & \mathbf{h}_{MN} \end{bmatrix} \tag{3.9}$$

where

$$\mathbf{h}_{ij} = \left[h_{ij}(0), h_{ij}(1), h_{ij}(2), \ldots, h_{ij}(L-1)\right] \tag{3.10}$$

denotes the $1 \times L$ vector of *(i,j)*-th systems response. The extended vector of sources $\bar{\mathbf{s}}(n)$ takes the following form:

$$\bar{\mathbf{s}}(n) = [s_1(n),....,s_1(n-L+1),....,s_N(n),....,s_N(n-L+1)]^T . \tag{3.11}$$

Supposing more measurements than sources, a positive integer *K* can be found which satisfies:

$$KM > N(L+K-1) . \tag{3.12}$$

Extend the vector of noisy measurements $\mathbf{y}(n)$ by *K-1* delayed repetitions of each measurement:

$$\bar{\mathbf{y}}(n) = [y_1(n), y_1(n-1),....,y_1(n-K+1),....,y_M(n),....,y_M(n-K+1)]^T . \tag{3.13}$$

Extending the noise vector $\boldsymbol{\omega}(n)$ in the same manner, (3.8) evolutes to

$$\bar{\mathbf{y}}(n) = \mathbf{H}\bar{\mathbf{s}}(n) + \bar{\boldsymbol{\omega}}(n) . \tag{3.14}$$

The mixing matrix **H** is now of size $KM \times N(L+K-1)$ and still comprises the system responses $h_{ij}(l)$:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{11} & \cdots & \mathbf{H}_{1N} \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{M1} & \cdots & \mathbf{H}_{MN} \end{bmatrix} \tag{3.15}$$

with $K \times (L+K-1)$ matrix

$$\mathbf{H}_{ij} = \begin{bmatrix} h_{ij}(0) & h_{ij}(1) & \cdots & \cdots & h_{ij}(L-1) & 0 & \cdots & 0 \\ 0 & h_{ij}(0) & \cdots & \cdots & h_{ij}(L-2) & h_{ij}(L-1) & \cdots & 0 \\ \vdots & & \ddots & & & & \ddots & \vdots \\ 0 & \cdots & 0 & h_{ij}(0) & \cdots & & \cdots & h_{ij}(L-1) \end{bmatrix}. \tag{3.16}$$

The extended vector of sources $\bar{\mathbf{s}}(n)$ now yields:

$$\bar{\mathbf{s}}(n) = [s_1(n),....,s_1(n-L-K+1),....,s_N(n),....,s_N(n-L-K+1)]^T . \tag{3.17}$$

In (3.8) and (3.14) we intentionally used the same notation for the two different mixing matrices. This introduced some level of ambiguity in our notations. However, in order to keep the equations simple and clear the mixing matrix will always be denoted only by **H**. When needed, the factor *K* will be defined explicitly.

The extension of Eq. (3.8) to the over-determined system of Eq. (3.14) is not an innovation. In fact, it is a common approach when dealing with the convolutive mixtures of sources and has been extensively used in the past [46]. For the clarity reasons we briefly explained only its main ideas.

3.3.2. Ambiguities of blind source separation

The goal in blind source separation is to identify the mixing matrix **H** and to reconstruct the sources $\bar{\mathbf{s}}(n)$ given only the vector of measurements $\bar{\mathbf{x}}(n)$. However the blind separation hides two implicit ambiguities [8, 9, 46].

BSS Ambiguity 1: It is impossible to determine the order of the sources, because with both the vector of sources $\bar{\mathbf{s}}(n)$ and the mixing matrix **H** unknown any permutation of sources (and columns in **H**) is a solution of (3.14) [46].

BSS Ambiguity 2: Without any further information it is impossible to determine the variance (energies) of sources, because any scalar multiplier can be exchanged between an arbitrary source and the corresponding column of the mixing matrix **H**:

$$\bar{\mathbf{x}}(n) = \mathbf{H}\bar{\mathbf{s}}(n) = \sum_{i=1}^{K(N+L)} \frac{\mathbf{h}_i}{\alpha_i} \alpha_i \bar{s}_i(n) \tag{3.18}$$

where α_i denotes an arbitrary scalar multiplier and $\mathbf{h}_i$ stand for the *i*-th column of mixing matrix **H**.

Although different scalar factors α_i change the values of both the extended sources $\bar{s}_i(n)$ and the corresponding columns **H** we won't change their denotations. We are going to use the denotations **H** and $\bar{\mathbf{s}}(n)$, even though in some sections the values of the scalar factors α_i will be exactly defined. If not explained differently, throughout the dissertation both ambiguities will be neglected.

4.

Decomposition of convolutive mixtures of sources – more measurements than sources

In previous chapters, we introduced the main principles of blind source separation and gave a brief overview and evaluation of the state of the art in the decomposition of convolutive mixtures. We further described the properties of human muscles, explained the process of generation, conduction and extinction of motor unit action potentials, and discussed the different surface EMG recording techniques (spatial filters). We identified the main problems in surface EMG decomposition and briefly evaluated the existing decomposition techniques. Finally, we focused on the multi-channel surface EMG and, using the terminology of the communication theory, pointed out the conditions under which the surface EMG can be modelled as a linear, shift-invariant MIMO system.

In this chapter (and the next), we one will derive three novel decomposition approaches. We will first examine the overdetermined case with more measurements than sources and describe two different ways towards our final goal – the decomposition of convolutive mixtures of close-to-orthogonal pulse sources.

The first approach, derived in Section 4.1, is based on the time-frequency analysis and utilises the solutions introduced by A. Belouchrani and M. Amin [14]. We briefly discuss their work and show how their ideas can be extended to the case of pulse sources in order to estimate both the mixing matrix $\mathbf{H}$ and the sources $\mathbf{s}(n)$ up to the unknown scalar factor. The second decomposition algorithm is based on higher-order statistics and is derived in Section 4.2. We demonstrate how higher-order cumulants can be applied to a MIMO system to get a very robust estimation of the system responses $h_{ij}(l)$. The third approach, the so called inverse correlation based method, is treated in the subsequent chapter, where we discuss the underdetermined case (the case with more sources than measurements). For each novel approach the influence of noise and non-orthogonal sources is also analysed.

It is of high importance to realize that the novel approaches assume general pulse sources and do not depend on the type of the mixing matrix. They only suppose the mixing matrix stationary and the average inter-pulse distance greater than the length of the system responses. The latter is certainly true in the case of EMG signals, where the MU firing frequency is usually well below 30 Hz, while the average length of MUAPs can be

estimated to 20-25 ms. The latter guarantees all three approaches can be applied to surface EMG signals, as demonstrated in Chapter 6.

Let us now assume the number of sources does not exceed the number of measurements and derive the first decomposition approach, the so called method based on time-frequency distributions.

4.1 Method based on time-frequency distributions

Time-frequency (TF) analysis is one of the most frequently used tools in the filed of signal processing. It also plays an important role in blind source separation, because it enables us to separate sources with Gaussian distribution. Although it is most frequently used for separation of multiplicative mixtures [14, 17, 73, 120], it is becoming more and more important also when it comes to the separation of convolutive mixtures [4, 17, 33, 84].

Very efficient approach to separation of multiplicative mixtures of nonstationary sources was introduced by A. Belouchrani and M. Amin [14]. Their method exploits the differences in energy locations of sources in TF domain and is based on the joint diagonalization of spatial matrices which are constructed from the time-frequency distributions of all possible pairs of measurements. Similar approach was also suggested for the convolutive cases [17, 84]. By extending the vector of measurements with the delayed repetitions of each measurement (Subsection 3.3.1) every finite convolutive mixture of sources can be rewritten in the matrix form. But, as described in the sequel, this introduces some novel difficulties which prevent the exact reconstruction of convoluted sources.

The section is organized in five subsections. In Subsection 4.1.1, we briefly define the time-frequency distributions. Subsection 4.1.2 summarizes the main ideas of the approach that was introduced by A. Belouchrani and M. Amin, along with its main pros and cons. The extension of the approach to the case of pulse sources is given in Subsections 4.1.3 and 4.1.4. We study the influences of non-orthogonal sources and introduce some measures and procedures for overcoming their disturbing influences. We conclude our discussion in Subsection 4.1.5 where the influence of noise is carefully outlined, along with some possible countermeasures.

4.1.1 Discrete time-frequency distributions

The discrete TF distribution of Cohen class for signal $x_i(n)$ is defined as [37]:

$$\mathbf{D}_{x_i x_i}(n, f) = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \boldsymbol{\varphi}(m,l) x_i(n+m+l) x_i^*(n+m-l) e^{-4\pi \cdot f \cdot l} \tag{4.1}$$

where n and f denote time and frequency indices, respectively, x_i^* is complex conjugated value of x_i, and $\boldsymbol{\varphi}(m,l)$ stands for the kernel that characterises the TF distribution. Cross TF distribution of signals $x_i(n)$ and $x_j(n)$ is defined as [37]:

$$\mathbf{D}_{x_i x_j}(n, f) = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \boldsymbol{\varphi}(m,l) x_i(n+m+l) x_j^*(n+m-l) e^{-4\pi \cdot f \cdot l} . \tag{4.2}$$

Using definitions in (4.1) and (4.2) we construct the so called spatial TF distribution (STFD) matrices of extended measurements $\bar{\mathbf{x}}(n)$ [14]:

$$\mathbf{D}_{\bar{\mathbf{x}}\bar{\mathbf{x}}}(n, f) = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \boldsymbol{\Theta}(m,l) \bar{\mathbf{x}}(n+m+l) \bar{\mathbf{x}}^H(n+m-l) e^{-4\pi \cdot f \cdot l} \tag{4.3}$$

where ^{H} denotes matrix conjugate transpose. $\mathbf{D}_{\bar{\mathbf{x}}\bar{\mathbf{x}}}(n, f)$ is a four-dimensional data structure whose *(i,j)*-th element equals the discrete (cross-) time-frequency distribution of measurements $\bar{x}_i(n)$ and $\bar{x}_j(n)$:

$$\left[\mathbf{D}_{\bar{\mathbf{x}}\bar{\mathbf{x}}}(n, f)\right]_{i,j} = \mathbf{D}_{\bar{x}_i \bar{x}_j}(n, f) \quad \textit{for} \quad i,j=1,...,N. \tag{4.4}$$

Neglecting the noise and taking into account the linearity of the assumed MIMO system we can write:

$$\mathbf{D}_{\bar{\mathbf{x}}\bar{\mathbf{x}}}(n, f) = \mathbf{H}\mathbf{D}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}(n, f)\mathbf{H}^T \tag{4.5}$$

where, for every fixed index pair (n, f), $\mathbf{D}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}(n, f)$ denotes the STFD matrix of extended sources.

In the sequel, the off-diagonal elements of $\mathbf{D}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}(n, f)$ matrices (values of cross-TF distributions) will be referred to as crossterms, while the diagonal elements of $\mathbf{D}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}(n, f)$ will be called autoterms. As a result, matrices $\mathbf{D}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}(n, f)$ will be diagonal if and only if all their crossterms will be equal to zero [73].

4.1.2 Separation of general sources

The blind separation approach introduced by A. Belouchrani and M. Amin [14] is based on assumption of more measurements than sources and comprises two steps. In the first step the extended vector of measurements $\bar{\mathbf{x}}(n)$, defined in subsection 3.3.1, is whitened by the $KM \times N(K+L-1)$ matrix **B**, which satisfies

$$\mathbf{B}\left(\lim_{T\to\infty}\frac{1}{T}\sum_{n=1}^{T}\bar{\mathbf{x}}(n)\bar{\mathbf{x}}^{*}(n)\right)\mathbf{B}^{T}=\mathbf{B}\mathbf{R}_{\bar{\mathbf{x}}}\mathbf{B}^{T}=\mathbf{B}\mathbf{H}\mathbf{R}_{\bar{\mathbf{s}}}\mathbf{H}^{T}\mathbf{B}^{T}=\mathbf{I} \tag{4.6}$$

where $\mathbf{R}_{\bar{\mathbf{s}}}$ stands for the correlation matrix of extended sources and $\mathbf{R}_{\bar{\mathbf{x}}}$ for the correlation matrix of extended measurements. Denoting the matrix square root of $\mathbf{R}_{\bar{\mathbf{s}}}$ by $\mathbf{R}_{\bar{\mathbf{s}}}^{\frac{1}{2}}$ it is obvious from (4.6) that $\mathbf{U}=\mathbf{B}\mathbf{H}\mathbf{R}_{\bar{\mathbf{s}}}^{\frac{1}{2}}$ is a unitary matrix of size $N(K+L-1)\times N(K+L-1)$. The whitening step has, therefore, reduced the mixing matrix **H** of size $(KM \times N(K+L-1))$ to the unitary matrix **U** of size $N(K+L-1)\times N(K+L-1)$. The **B** matrix can be constructed as a matrix square root of the inverse of the $\mathbf{R}_{\bar{\mathbf{x}}}$ matrix [13, 14].

For the whitened vector of extended sources $\bar{\mathbf{z}}(n)$ the linearity of the modelled MIMO system still holds:

$$\bar{\mathbf{z}}(n)=\mathbf{B}\bar{\mathbf{x}}(n)=\mathbf{B}\mathbf{H}\bar{\mathbf{s}}(n)=\mathbf{U}\mathbf{R}_{\bar{\mathbf{s}}}^{-\frac{1}{2}}\bar{\mathbf{s}}(n). \tag{4.7}$$

Multiplying the $\mathbf{D}_{\bar{\mathbf{x}}\bar{\mathbf{x}}}(n,f)$ matrices by **B** yields the whitened STFD matrices of extended measurements [14, 17]:

$$\mathbf{D}_{\bar{\mathbf{z}}\bar{\mathbf{z}}}(n,f)=\mathbf{B}\mathbf{D}_{\bar{\mathbf{x}}\bar{\mathbf{x}}}(n,f)\mathbf{B}^{T}=\mathbf{U}\mathbf{R}_{\bar{\mathbf{s}}}^{-\frac{1}{2}}\mathbf{D}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}(n,f)\mathbf{R}_{\bar{\mathbf{s}}}^{-\frac{1}{2}}\mathbf{U}^{T}. \tag{4.8}$$

Even when the sources $\mathbf{s}(n)$ are orthogonal, the extended sources $\bar{\mathbf{s}}(n)$ do not need to be and the matrices $\mathbf{R}_{\bar{\mathbf{s}}}$ and $\mathbf{R}_{\bar{\mathbf{s}}}^{-\frac{1}{2}}\mathbf{D}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}(n,f)\mathbf{R}_{\bar{\mathbf{s}}}^{-\frac{1}{2}}$ are in general block diagonal. According to (4.8) any whitened STFD matrix $\mathbf{D}_{\bar{\mathbf{z}}\bar{\mathbf{z}}}(n,f)$ is block diagonal in the basis of the columns of **U** and the missing unitary matrix **U** can be retrieved by block joint diagonalization of set of $\mathbf{D}_{\bar{\mathbf{z}}\bar{\mathbf{z}}}(n,f)$ matrices in different TF points *(n,f)* [12, 17]. The algorithm of joint block diagonalization is a generalization of well-known Jacobi's technique for diagonalization of Hermitian matrices and is described in more detail in [12].

Knowing the matrices **B** and **U**, the mixing matrix **H** can be estimated as [13, 14]:

$$\mathbf{H}=\mathbf{B}^{\#}\mathbf{U}, \tag{4.9}$$

where $^{\#}$ denotes Moore-Penrose pseudoinverse. However, due to its inherent ambiguity the joint block diagonalization technique introduces an unknown block diagonal unitary matrix **D** [12, 17]. Therefore the original sources can be reconstructed only up to the unknown filtering effect [17]:

$$\bar{\mathbf{s}}_e(n) = \mathbf{U}^H \mathbf{B}\bar{\mathbf{x}}(n) = \mathbf{D}\mathbf{R}_{\bar{\mathbf{s}}}^{-\frac{1}{2}}\bar{\mathbf{s}}(n), \tag{4.10}$$

where matrices **D** and $\mathbf{R}_{\bar{\mathbf{s}}}^{-\frac{1}{2}}$ are unknown.

4.1.3 Blind separation of pulse sources

A. Belouchrani and M. Amin considered general orthogonal sources with nonorthogonal delayed repetitions of each source. On the other hand, the energy of pulse sources is concentrated around the unknown time moments (pulses). Furthermore, dealing with the biomedical signals, such as the EMG, the upper limit of the MU firing frequency (Section 3.1.1) guarantees also the extended sources are orthogonal as long as the maximal delay in measurements does not exceed the minimal inter-pulse interval. Considering the BSS ambiguity 2 the correlation matrix of extended sources can now be written as:

$$\mathbf{R}_{\bar{\mathbf{s}}} = \lim_{T\to\infty} \frac{1}{T} \sum_{n=1}^{T} \bar{\mathbf{s}}(n)\bar{\mathbf{s}}^H(n) = \mathbf{I}. \tag{4.11}$$

This eliminates the unknown matrices $\mathbf{R}_{\bar{\mathbf{s}}}$ and $\mathbf{R}_{\bar{\mathbf{s}}}^{-\frac{1}{2}}$ in (4.7), (4.8) and (4.10). The only unknown left is the block diagonal matrix **D** in (4.10), which is introduced by the joint block diagonalization [12, 17]. Block diagonalization can now be replaced by joint diagonalization [25, 31], what would enable reconstruction of sources up to a scalar factor. However, to be able to use the joint diagonalization we must ensure the STFD matrices $\mathbf{D}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}(n, f)$ which enter the reconstruction of unknown matrix **U** are also diagonal.

Generally speaking, the $\mathbf{D}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}(n, f)$ matrices are, even in the case of pulse sources, block diagonal. The reason is hidden in the kernel $\boldsymbol{\varphi}(m,l)$ of TF distributions, which is used to average the sources in time and, hence, spread the information around pulses. Excluding the kernel $\boldsymbol{\varphi}(m,l)$, (4.1) yields the so called Wigner-Ville TF distribution

$$\mathbf{D}_{\bar{x}_i\bar{x}_j}(n, f) = \mathbf{W}\mathbf{V}_{\bar{x}_i\bar{x}_j}(n, f) = \sum_{l=-\infty}^{\infty} \bar{x}_i(n+l)\bar{x}_j^{\,*}(n)e^{-4\pi\cdot f\cdot l}. \tag{4.12}$$

But new problems arise. The Wigner-Ville spectra suffer high sensitivity to the crossterms (values of cross TF distributions), which, again, make the source STFD

matrices block diagonal. Namely, the cross Wigner-Wille distribution of the *i*-th and *j*-th sources in an arbitrary time point n_k is a summation of all the pulses from the *i*-th and *j*-th sources, which satisfy the following relation:

$$n_k = \tfrac{1}{2}(n_{i,p} + n_{j,q}) \tag{4.13}$$

where $p, q \in (-\infty, \infty)$ and $n_{i,p}$ ($n_{j,q}$) denotes the time moment in which the *p*-th (*q*-th) pulse of the *i*-th (*j*-th) source appeared. In other words, the cross TF distribution $\mathbf{WV}_{\bar{s}_i \bar{s}_j}(n, f)$ differs from zero in every time moment which lies between the arbitrary pulses of the i-th and j-th source.
We can reduce the number of pulses contributing to $\mathbf{WV}_{\bar{s}_i \bar{s}_j}(n, f)$ by shrinking the calculation of the Wigner-Wille distribution in (4.12) to the finite interval [-*a*,*a*]:

$$\mathbf{D}_{\bar{x}_i \bar{x}_j}(n, f) = \mathbf{PWV}_{\bar{x}_i \bar{x}_j}(n, f) = \sum_{l=-a}^{a} \bar{x}_i(n+l)\bar{x}_j^{\,*}(n) e^{-4\pi \cdot f \cdot l} \tag{4.14}$$

where *a* denotes an arbitrarily large positive integer number. Setting the limit *a* to zero, all the crossterms are left out and the $\mathbf{D}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}(n, f)$ matrices begin to show their diagonal structure. The TF distribution in (4.12) now yields:

$$\mathbf{D}_{\bar{x}_i \bar{x}_j}(n, f) = \mathbf{PWV}_{\bar{x}_i \bar{x}_j}(n) = \bar{x}_i(n)\bar{x}_j^{\,*}(n). \tag{4.15}$$

Having ensured the diagonality of the $\mathbf{D}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}(n, f)$ matrices, we can use the joint diagonalization [13, 14, 25, 26, 31] of a set of $\mathbf{PWV}_{\bar{\mathbf{x}}\bar{\mathbf{x}}}(n)$ matrices to reconstruct the missing unitary matrix **U** and the sources up to a scalar factor:

$$\bar{\mathbf{s}}(n) = \mathbf{U}^H \mathbf{B}\bar{\mathbf{x}}(n) = \mathbf{A}\bar{\mathbf{s}}(n) \tag{4.16}$$

where **A** is unknown diagonal matrix.

4.1.4 Decomposition of close-to-orthogonal sources

In the previous subsection we supposed the sources strictly orthogonal. This implies several restrictions that are hard to meet in reality. Processing the EMG signals, for example, the innervation pulse trains will be independent only at very low levels of muscle contractions. When the contraction level increases the innervation trains become more and more correlated (up to 10 % or 15 %) and their pulses begin to overlap. Moreover, by extending the vector of measurements (Subsection 3.3.1) we have virtually increased the number of sources from *NL* to *N(L+K-1)*, which further decreases the assumed orthogonality of sources.

Let $G_{n_0} = \{j_{1,n_0}, ..., j_{p,n_0}\}$ denote a set of indices of sources which all trigger in a given time moment n_0 (they all share the overlapped pulse at n_0). $\mathbf{D}_{\overline{\mathbf{ss}}}(n_0, f)$ matrix will be far from diagonal as it will have p^2 elements different from zero (all autoterms and all crossterms at the positions that are contained in G_{n_0}). A large number of such matrices will cause the join diagonalization to reconstruct a wrong unitary matrix **U**. To be exact, using the eigendecomposition $\mathbf{D}_{\overline{\mathbf{ss}}}(n_0, f) = \mathbf{U}_\Lambda \mathbf{\Lambda} \mathbf{U}_\Lambda^T$ we can write

$$\mathbf{D}_{\overline{\mathbf{zz}}}(n_0, f) = \mathbf{U}\mathbf{D}_{\overline{\mathbf{ss}}}(n_0, f)\mathbf{U}^T = \mathbf{U}\mathbf{U}_\Lambda \mathbf{\Lambda} \mathbf{U}_\Lambda^T \mathbf{U}^T \tag{4.17}$$

which proves that the diagonalization of $\mathbf{D}_{\overline{\mathbf{zz}}}(n_0, f)$ produces the unitary matrix $\mathbf{U}\mathbf{U}_\Lambda$.

There is yet another condition that has to be met when using the joint diagonalization. To guarantee the uniqueness of the reconstructed unitary matrix **U,** the set of STFD matrices entering the joint diagonalization must contain the contributions of all active sources, including their delayed repetitions [26, 73]. But, to guarantee their diagonality in the column-space of matrix **U**, the matrices entering the joint diagonalization must contain the contribution of a single source only. This implies that we must diagonalize at least $N(L+K-1)$ STFD matrices $\mathbf{D}_{\overline{\mathbf{zz}}}(n, f)$ (one for each source).

In other words, to guarantee the reconstruction of original mixing matrix (and consequently the reconstruction of original sources) we must be able not only to distinguish between the diagonal and nondiagonal $\mathbf{D}_{\overline{\mathbf{ss}}}(n, f)$ matrices, but also to distinguish the contributions of different sources. With unitary matrix **U** unknown we can only rely on the information contained in the $\mathbf{D}_{\overline{\mathbf{zz}}}(n, f)$ matrices.

Let $\mathbf{D}^i_{\overline{\mathbf{ss}}}(n, f)$ denote the diagonal STFD matrix of sources with a single nonzero autoterm d_{ii} at the i-th diagonal position (i.e. with contribution from the i-th extended source). Bearing in mind that **U** is unitary we can derive the following relation:

$$\mathbf{D}^i_{\overline{\mathbf{zz}}}(n, f)\mathbf{D}^j_{\overline{\mathbf{zz}}}(n, f) = \mathbf{U}\mathbf{D}^i_{\overline{\mathbf{ss}}}(n, f)\mathbf{D}^j_{\overline{\mathbf{ss}}}(n, f)\mathbf{U}^T = 0, \; if \;\; i \neq j. \tag{4.18}$$

Denote by $\mathbf{D}^{ij}_{\overline{\mathbf{ss}}}(n, f)$ the STFD matrix of sources with contributions of the i-th and j-th extended source. $\mathbf{D}^{ij}_{\overline{\mathbf{ss}}}(n, f)$ will have four nonzero elements d_{ii}, d_{ij}, d_{ji} and d_{jj} at the positions (i,i), (i,j), (j,i) and (j,j), respectively. We derive

$$\mathbf{D}^{ij}_{\overline{\mathbf{zz}}}(n, f)\mathbf{D}^i_{\overline{\mathbf{zz}}}(n, f) = \mathbf{U}\mathbf{D}^{ij}_{\overline{\mathbf{ss}}}(n, f)\mathbf{D}^i_{\overline{\mathbf{ss}}}(n, f)\mathbf{U}^T \neq 0 \tag{4.19}$$

and

$$\mathbf{D}^{ij}_{\overline{\mathbf{zz}}}(n, f)\mathbf{D}^j_{\overline{\mathbf{zz}}}(n, f) = \mathbf{U}\mathbf{D}^{ij}_{\overline{\mathbf{ss}}}(n, f)\mathbf{D}^j_{\overline{\mathbf{ss}}}(n, f)\mathbf{U}^T \neq 0. \tag{4.20}$$

According to (4.18), (4.19) and (4.20) the STFD matrices of sources (and, hence, also the STFD matrices of whitened measurements) with a contribution from a single source form the orthogonal basis of $N(L+K-1)$-dimensional space.

Define the following measure:

$$\left\langle \mathbf{M}_1, \mathbf{M}_2 \right\rangle = \frac{trace(\mathbf{M}_1 \mathbf{M}_2)}{trace(\mathbf{M}_2)} \tag{4.21}$$

where *trace(***M***)* denotes the sum of diagonal elements of matrix **M** (the matrix trace). We can then easily check the following relations:

$$\left\langle \mathbf{D}^{ij}{}_{\overline{\mathbf{zz}}}(n,f), \mathbf{D}^{i}{}_{\overline{\mathbf{zz}}}(n,f) \right\rangle = \frac{trace(\mathbf{D}^{ij}{}_{\overline{\mathbf{zz}}}(n,f)\mathbf{D}^{i}{}_{\overline{\mathbf{zz}}}(n,f))}{trace(\mathbf{D}^{i}{}_{\overline{\mathbf{zz}}}(n,f))} = d_{ii} \tag{4.22}$$

and

$$\left\langle \mathbf{D}^{ij}{}_{\overline{\mathbf{zz}}}(n,f), \mathbf{D}^{j}{}_{\overline{\mathbf{zz}}}(n,f) \right\rangle = \frac{trace(\mathbf{D}^{ij}{}_{\overline{\mathbf{zz}}}(n,f)\mathbf{D}^{j}{}_{\overline{\mathbf{zz}}}(n,f))}{trace(\mathbf{D}^{j}{}_{\overline{\mathbf{zz}}}(n,f))} = d_{jj} \tag{4.23}$$

where d_{ii} and d_{jj} stand for the (i,i)-th and (j,j)-th element of $\mathbf{D}^{ij}_{\overline{\mathbf{ss}}}(n,f)$, respectively.

The quest for $\mathbf{D}^{i}{}_{\overline{\mathbf{zz}}}(n,f)$ matrices with contributions from single, but different sources has now reduced to much easier problem of finding the orthogonal basis of the $N(L+K-1)$-dimensional space. This is a very important result as the criterion function (4.21) enables us to control the $\mathbf{D}_{\overline{\mathbf{zz}}}(n,f)$ matrices entering the joint diagonalization, and, consequently, guarantees the uniqueness of matrix **U**. We must still ensure the set of the $\mathbf{D}_{\overline{\mathbf{zz}}}(n,f)$ matrices entering the joint diagonalization contains the matrices with contributions from all the sources.

Suppose now we fail to find the matrices with a single nonzero autoterm of the *i*-th source. We are than forced to include the non-optimal $\mathbf{D}_{\overline{\mathbf{zz}}}(n,f)$ matrices with contributions of the *i*-th and several other sources. In order to minimize the devastating influence of crossterms [73], we have to ensure the number of sources contributing to the selected $\mathbf{D}_{\overline{\mathbf{zz}}}(n,f)$ matrices is as small as possible. We can still rely on criterion function (4.21) and on the fact that the trace of a matrix equals to the sum of its diagonal elements:

$$trace(\mathbf{D}^{ij}{}_{\overline{\mathbf{zz}}}(n,f)) = d_{ii} + d_{jj}\,. \tag{4.24}$$

The selection of the optimal $\mathbf{D}_{\overline{\mathbf{zz}}}(n,f)$ matrices using the criterion (4.21) is described in more detail in [71].

4.1.5 Noise influence

Minimizing the interval *[a,a]* on which the Wigner-Ville distribution is calculated we annulled the time averaging of measurements and, hence, lost the averaging of the zero-mean noise. As a result, we reduced the robustness of the presented approach and exposed it to the influence of noise [13, 14, 37, 84]. The time averaging can be partially replaced by averaging of the $\mathbf{D}_{\overline{\mathbf{zz}}}(n, f)$ matrices with contributions from a single (not necessarily the same) source:

$$\frac{1}{T}\sum_{j=1}^{T}\left[\mathbf{D}^{j}_{\overline{\mathbf{zz}}}(n_j, f)+\mathbf{D}_{\overline{\boldsymbol{\omega}}\overline{\boldsymbol{\omega}}}(n, f)\right]=\mathbf{U}\left[\frac{1}{T}\sum_{j=1}^{T}\mathbf{D}^{j}_{\overline{\mathbf{ss}}}(n_j, f)\right]\mathbf{U}^{T}+\frac{1}{T}\sum_{j=1}^{T}\mathbf{D}_{\overline{\boldsymbol{\omega}}\overline{\boldsymbol{\omega}}}(n_j, f) \qquad (4.25)$$

where

$$\mathbf{D}_{\overline{\boldsymbol{\omega}}\overline{\boldsymbol{\omega}}}(n, f)=\mathbf{B}\overline{\boldsymbol{\omega}}(n)\overline{\boldsymbol{\omega}}(n)^{T}\mathbf{B}^{T} \qquad (4.26)$$

denotes the influence of noise $\overline{\boldsymbol{\omega}}(n)$. Recall the noise is modelled as Gaussian, white and zero-mean (Subsection 3.4):

$$\overline{\boldsymbol{\omega}}(n)\propto N(\mathbf{0},\sigma^{2}\mathbf{I}) \qquad (4.27)$$

where $\mathbf{0}$ stands for $N(L+K-1)\times 1$ vector with all elements equal to zero, while the sign $\propto$ should be read as »distributed according to«. Then the noise influence in (4.25) can be expressed as:

$$\lim_{T\to\infty}\frac{1}{T}\sum_{n=1}^{T}\mathbf{D}_{\overline{\boldsymbol{\omega}}\overline{\boldsymbol{\omega}}}(q, f)=\sigma^{2}\mathbf{B}\mathbf{B}^{T}. \qquad (4.28)$$

The noise variance σ^{2} can be estimated from the correlation matrix $\mathbf{R}_{\overline{\mathbf{x}}}$ (the exact procedure is explained in Section 5.2), while the whitening matrix $\mathbf{B}$ is known. This allows us to simply subtract the influence of noise from (4.25).

The influence of noise can further be reduced by averaging the reconstructed sources

$$\overline{\mathbf{s}}(n)=\mathbf{U}^{H}\mathbf{B}\overline{\mathbf{x}}(n). \qquad (4.29)$$

As stated in (3.17), the presented decomposition approach will reconstruct *(L+K-1)* delayed repetitions of each source, which can be aligned in time and averaged afterwards.

Nevertheless, the noise will also influence the criterion (4.21), which guarantees the uniqueness of the reconstructed matrix $\mathbf{U}$. We can, therefore, expect a drop of performance in the environments with strong noise and heavily overlapped sources.

4.2 W-slices and higher-order cumulants

The theory of higher-order statistics (HOS) has been well established, also in the field of MIMO systems [33, 41, 111, 117, 123, 124, 134, 166, 167]. It has been shown that the system's responses can be deconvolved whenever the system's input can be modelled as an independent identically distributed (i.i.d.), zero-mean, random white noise [111, 166, 167]. In addition to the system identification, HOS are frequently applied to the field of radar and sonar signals, image processing, seismology, and also speech analysis. Although suffering from a very high computational complexity, HOS-based methods have many nice properties. They can be successfully applied when the number of sources highly exceeds the number of measurements, while the superimpositions of impulse responses do not hinder a thorough decomposition. Moreover, by being insensitive to all Gaussian random processes the cumulants of order higher than two are also inherently resistant to the Gaussian noise.

In the last decade the use of HOS has expanded to the field of blind source separation [4, 107, 111, 166, 167]. Supposing the sources independent non-Gaussian random processes, the central limit theorem assures the distribution of their sum converges towards the Gaussian distribution. Because all higher-order cumulants of the Gaussian process equal zero, the HOS can be used as a measure of the nongaussianity of the mixed sources and, hence, serve as a basic criterion in source decoupling. This idea has been used in numerous blind source separation approaches. Their comprehensive overview can be found in [4, 84].

In this chapter, we will study one of many possible routes towards the HOS-based MIMO system identification, the so called w-slices. The method is based on Barlett-Brillinger-Rosenblatt equation [111] and enables reconstruction of the system's responses. The fact that the pulse sources cannot be modelled as i.i.d. white random noise does not corrupt the cumulants, as long as they are computed at the lags that do not exceed the minimum inter-pulse distance. As already explained, dealing with biomedical signals, the minimum inter-pulse distance exceeds the length of the modelled system responses, i.e. the length of MUAPs (Section 3.1.1).

In the subsequent subsections, we first briefly outline the definition of higher-order cumulants. We explain the theory behind the w-slices and extend it to the field of the MIMO systems. We also discuss its application to the convolutive mixtures of pulse sources and demonstrate its capability to resist the Gaussian noise. We further show that w-slices can be used to compute only a coarse estimation of system responses, and describe the Newton-Gauss method [61] for their optimisation. We conclude the chapter with the discussion about the influences of the noise and of nonorthogonal sources.

4.2.1 Higher-order statistics

Given a set of random variables $X_1,....,X_M$, their joint cumulant of order r is defined by

$$cum(X_1^{k_1},...,X_M^{k_M}) = (-j)^r \left. \frac{\partial^r \ln \Phi(\omega_1,...,\omega_M)}{\partial^{k_1}\omega_1 ... \partial^{k_M}\omega_M} \right|_{\omega_1=\cdots=\omega_M=0} \tag{4.30}$$

where $k_1 + \cdots + k_M = r$ and $\Phi(\omega_1,...,\omega_M)$ denotes the joint characteristic function of the variables $X_1,....,X_M$:

$$\Phi(\omega_1,...,\omega_M) = E[e^{j(\omega_1 X_1 + \omega_2 X_2 + \cdots \omega_M X_M)}]. \tag{4.31}$$

For a stationary discrete time random process $X(n)$ the cumulant of order r is defined as a cumulant of the random variables $X(n), X(n+\tau_1),...., X(n+\tau_{r-1})$:

$$C_{r,X}(\tau_1,\tau_2,...,\tau_{r-1}) = cum(X(n), X(n+\tau_1),..., X(n-\tau_{r-1})). \tag{4.32}$$

The cumulant definition in (4.30) is highly theoretical. In order to calculate the exact values of cumulants one should know the distributions of the observed random variables. In practice, we often rely on the fact that the *r*-th order cumulant can be expressed with the moments of the *r*-th and lower orders. Supposing $X(n)$ zero-mean the third order cumulant, for example, can be written as:

$$\begin{aligned} C_{3,X}(\tau_1,\tau_2) &= cum(X(n), X(n+\tau_1), X(n+\tau_2)) = \\ &= \lim_{T\to\infty} \frac{1}{T} \sum_{n=1}^{T} X(n)X(n+\tau_1)X(n+\tau_2). \end{aligned} \tag{4.33}$$

Suppose now the inputs (sources) in our MIMO system i.i.d., white and zero-mean random processes. According to the Barlett-Brillinger-Rosenblatt equation [111], the cumulant of the measurements $\mathbf{x}(n)$ can be written as:

$$\begin{aligned} C_{x_{i_1},...,x_{i_r}}(\tau_1,...,\tau_{r-1}) &= cum(x_{i_1}(n), x_{i_2}(n+\tau_1),..., x_{i_M}(n-\tau_{r-1})) = \\ &= \sum_{j=1}^{N} \gamma_{r,s_j} \cdot \left(\sum_{l=0}^{L-1} h_{i_1 j}(l) h_{i_2 j}(l+\tau_1) \cdots h_{i_M j}(l+\tau_{r-1}) \right), \end{aligned} \tag{4.34}$$

where γ_{r,s_j} denotes the central value in the *r*-th order cumulant matrix of source s_j:

$$\gamma_{r,s_j} = C_{s_j,...,s_j}(0,....,0). \tag{4.35}$$

Let $\mathbf{H}(l)$ denote the matrix which comprises all system's responses:

$$\mathbf{H}(l)=\begin{bmatrix} h_{11}(l) & h_{12}(l) & \cdots & h_{1N}(l) \\ h_{21}(l) & h_{22}(l) & \cdots & h_{2N}(l) \\ \vdots & \vdots & \ddots & \vdots \\ h_{M1}(l) & h_{M2}(l) & \cdots & h_{MN}(l) \end{bmatrix}. \tag{4.36}$$

According to (3.2) we can write

$$\mathbf{x}(n)=\sum_{l=0}^{L}\mathbf{H}(l)\mathbf{s}(n-l)+\boldsymbol{\omega}(n). \tag{4.37}$$

Define now the *m*-th matrix of the third order cumulants of measurements, where *m* *(*$1\le m\le M$ *)* corresponds to the fixed measurement:

$$\mathbf{C}_m(\tau_1,\tau_2)=\begin{bmatrix} C_{x_1,x_1,x_m}(\tau_1,\tau_2) & C_{x_1,x_2,x_m}(\tau_1,\tau_2) & \cdots & C_{x_1,x_M,x_m}(\tau_1,\tau_2) \\ C_{x_2,x_1,x_m}(\tau_1,\tau_2) & C_{x_2,x_2,x_m}(\tau_1,\tau_2) & \cdots & C_{x_2,x_M,x_m}(\tau_1,\tau_2) \\ \vdots & \vdots & \ddots & \vdots \\ C_{x_M,x_1,x_m}(\tau_1,\tau_2) & C_{x_M,x_2,x_m}(\tau_1,\tau_2) & \cdots & C_{x_M,x_M,x_m}(\tau_1,\tau_2) \end{bmatrix}. \tag{4.38}$$

Using (4.34) and (4.38), the $\mathbf{C}_m(\tau_1,\tau_2)$ matrix can be rewritten as [166]

$$\mathbf{C}_m(\tau_1,\tau_2)=\sum_{l=0}^{L-1}\mathbf{H}(l+\tau_1)\left[\sum_{i=1}^{N}\boldsymbol{\Gamma}_i h_{mi}(l+\tau_2)\right]\mathbf{H}^T(l), \tag{4.39}$$

where $\boldsymbol{\Gamma}_i$, i=1,...,N, denotes the cumulant matrix of sources with a single nonzero element γ_{3,s_i} at diagonal position (*i,i*):

$$\boldsymbol{\Gamma}_i=\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & \ddots & \cdots & 0 \\ \vdots & \vdots & \gamma_{3,s_i} & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \tag{4.40}$$

$\mathbf{C}_m(\tau_1,\tau_2)$ is a four-dimensional structure which contains third-order auto-cumulants of measurement $x_m(n)$ and all cross-cumulants with this measurement. According to (4.38) and (4.39) the cumulants can be calculated in two different ways. They can be estimated out of the system outputs (4.38), or determined from the model (4.39). If the two matrix equations are made equal at fixed lags τ_1 in τ_2, a system of non-linear equations appears, whose solution leads to unknown system responses $h_{ij}(l)$ and input skewness γ_{3,s_i}. If the system responses are of length *L*, the total number of unknowns equals

$$NML + N \tag{4.41}$$

where M stands for the number of measurements, N for the number of sources and L for the length of system responses. The first term in (4.41) stands for the unknown samples in responses $h_{ij}(l)$, and the second for the unknown skewnesses γ_{3,s_i} (the inputs are supposed temporally and spatially uncorrelated) .

On the other hand, we need at least that many non-linear equations as we have unknowns. The number of possible equations depends on different values obtainable in $\mathbf{C}_m(\tau_1, \tau_2)$. In general, the third-order auto-cumulant matrix $\mathbf{C}_{x_i,x_i,x_i}(\tau_1, \tau_2)$ generates $\frac{L(L+1)}{2}$ different values, the cross-cumulant matrix based on 2 different measurements $\mathbf{C}_{x_i,x_i,x_j}(\tau_1, \tau_2)$ contributes $\frac{L(3L-1)}{2}$, while the cross-cumulant matrix based on 3 different measurements contributes $3L(L-1)+1$ different values. When dealing with a MIMO system with N inputs and M outputs, the total number of different third-order cumulant matrices in $\mathbf{C}_m(\tau_1, \tau_2)$ is as follows: 1 auto-cumulant matrix, 2(M-1) cross-cumulant matrices based on 2 measurements, and $\frac{(M-2)(M-1)}{2}$ cross-cumulant matrices based on 3 measurements. Thus, the total number of different cumulant values in $\mathbf{C}_m(\tau_1, \tau_2)$ yields:

$$\begin{aligned} &\frac{L(L+1)}{2} + 2(M-1)\frac{L(3L-1)}{2} + \frac{(M-2)(M-1)}{2}(3L(L-1)+1) = \\ &= [3M(M-1)+1]\frac{L^2}{2} + [M(7-3M)-3]L + [M^2 - 3M + 2]. \end{aligned} \tag{4.42}$$

The system of nonlinear equations can be resolved and thoroughly decomposed whenever (4.42) computes higher than (4.41). In spite of this theoretical result, all the errors accompanying the cumulant estimation hinder successful closed-form decomposition in practice. It has been shown, however, that a solution may be found through non-linear optimisation [61]. The corresponding search for the global optimum needs an initialisation point which begins the optimisation in the vicinity of the global solution. In other words, to find the system channel responses, i.e., the MUAPs in our case, a cumulant-based optimisation must begin with as good approximation of these responses as possible.

4.2.2 Coarse estimation of system responses: w-slices

Among various approaches having been proposed for MIMO system identification, we want to rely upon those with as little a priori information needed as possible. One of them is so called w-slice method [155]. Beside the minimum number of system channels to be modelled, the only additional information needed in this approach is about the values of the starting samples of system-channel responses, as we will see in the sequel.

MIMO w-slice method is based on a presumption that such a matrix of weights

$$\mathbf{W}(\tau_2)=\begin{bmatrix} w_{11}(\tau_2) & \cdots & w_{1M}(\tau_2) \\ \vdots & \ddots & \vdots \\ w_{M1}(\tau_2) & \cdots & w_{MM}(\tau_2) \end{bmatrix} \tag{4.43}$$

exists that the following relationship can be derived:

$$\mathbf{C}_m(\tau_1,\tau_2)\mathbf{W}(\tau_2)=\sum_{l=0}^{L-1}\mathbf{H}(l+\tau_1)\mathbf{G}(l)\,, \tag{4.44}$$

where

$$\mathbf{G}(l)=\left[\sum_{i=1}^{N}\mathbf{\Gamma}_i h_{mi}(l+\tau_2)\right]\mathbf{H}^T(l)\mathbf{W}(\tau_2)=\begin{bmatrix} \delta(l) & 0 & \cdots & 0 \\ 0 & \delta(l) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \delta(l) \end{bmatrix}, \tag{4.45}$$

and $\delta(l)$ denotes the Dirac impulse. From (4.44) it follows

$$\mathbf{C}_m(\tau_1,\tau_2)\mathbf{W}(\tau_2)=\sum_{l=0}^{L-1}\mathbf{H}(l+\tau_1)\mathbf{G}(l)=\begin{bmatrix} h_{11}(\tau_1) & h_{12}(\tau_1) & \cdots & h_{1N}(\tau_1) \\ h_{21}(\tau_1) & h_{22}(\tau_1) & \cdots & h_{2N}(\tau_1) \\ \vdots & \vdots & \ddots & \vdots \\ h_{M1}(\tau_1) & h_{M2}(\tau_1) & \cdots & h_{MN}(\tau_1) \end{bmatrix}. \tag{4.46}$$

Hence, the causal part ($\tau_1 \geq 0$) of matrix $\mathbf{C}_m(\tau_1,\tau_2)$ can be used to identify the system's responses.

Recall the system responses are causal (Subsection 3.4):

$$h_{ij}(l)=0\,,\ l<0,\quad i=1,...,M,\ j=1,\ldots,N. \tag{4.47}$$

Therefore, for the anti-causal part ($\tau_1 \le 0$) of the matrix $\mathbf{C}_m(\tau_1,\tau_2)$ Eq. (4.46) yields:

$$\mathbf{C}_m(0,\tau_2)\mathbf{W}(\tau_2) = \sum_{l=0}^{L-1}\mathbf{H}(l+\tau_1)\mathbf{G}(l) = \begin{bmatrix} h_{11}(0) & h_{12}(0) & \cdots & h_{1N}(0) \\ h_{21}(0) & h_{22}(0) & \cdots & h_{2N}(0) \\ \vdots & \vdots & \ddots & \vdots \\ h_{M1}(0) & h_{M2}(0) & \cdots & h_{MN}(0) \end{bmatrix} \tag{4.48}$$

when $\tau_1 = 0$, and

$$\mathbf{C}_m(\tau_1 < 0,\tau_2)\mathbf{W}(\tau_2) = \sum_{l=0}^{L-1}\mathbf{H}(l+\tau_1)\mathbf{G}(l) = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \tag{4.49}$$

when $\tau_1 < 0$. Let the indices (τ_1,τ_2) extend over the length of the system responses $(-L+1 \le \tau_i \le L-1)$. Further let

$$\mathbf{C}^a_{ijk} = \begin{bmatrix} C_{x_i,x_j,x_k}(-L+1,-L+1) & \cdots & C_{x_i,x_j,x_k}(-L+1,L-1) \\ \vdots & \ddots & \vdots \\ C_{x_i,x_j,x_k}(0,-L+1) & \cdots & C_{x_i,x_j,x_k}(0,L-1) \end{bmatrix} \tag{4.50}$$

denote the anti-causal, and

$$\mathbf{C}^c_{ijk} = \begin{bmatrix} C_{x_i,x_j,x_k}(0,-L+1) & \cdots & C_{x_i,x_j,x_k}(0,L-1) \\ \vdots & \ddots & \vdots \\ C_{x_i,x_j,x_k}(L-1,-L+1) & \cdots & C_{x_i,x_j,x_k}(L-1,L-1) \end{bmatrix} \tag{4.51}$$

the causal part of entire cumulant matrix $C_{x_i,x_j,x_k}(\tau_1,\tau_2)$. According to (4.48) and (4.49) we can write:

$$\begin{bmatrix} \mathbf{C}^a_{11m} & \cdots & \mathbf{C}^a_{M1m} \\ \vdots & \ddots & \vdots \\ \mathbf{C}^a_{1Mm} & \cdots & \mathbf{C}^a_{MMm} \end{bmatrix} \cdot \mathbf{W} = \begin{bmatrix} \mathbf{h}^a_{11} & \cdots & \mathbf{h}^a_{1N} \\ \vdots & \ddots & \vdots \\ \mathbf{h}^a_{1M} & \cdots & \mathbf{h}^a_{MN} \end{bmatrix} \tag{4.52}$$

where $\mathbf{h}^a_{ij} = [h_{ij}(0), h_{ij}(-1),..., h_{ij}(-L+1)]^T = [h_{ij}(0),0,...,0]^T$ denotes the anti-causal part of the system's response $h_{ij}(l)$, while

$$\mathbf{W} = \begin{bmatrix} \mathbf{w}_{11} & \cdots & \mathbf{w}_{1M} \\ \vdots & \ddots & \vdots \\ \mathbf{w}_{M1} & \cdots & \mathbf{w}_{MM} \end{bmatrix}, \tag{4.53}$$

with

$$\mathbf{w}_{ij} = [w_{ij}(-L+1), w_{ij}(-L+2), ..., w_{ij}(L-1)], \tag{4.54}$$

stands for the entire matrix of weights. In the same way, the causal part of (4.46) can be rewritten as:

$$\begin{bmatrix} \mathbf{C}^c_{11m} & \cdots & \mathbf{C}^c_{M1m} \\ \vdots & \ddots & \vdots \\ \mathbf{C}^c_{1Mm} & \cdots & \mathbf{C}^c_{MMm} \end{bmatrix} \cdot \mathbf{W} = \begin{bmatrix} \mathbf{h}^c_{11} & \cdots & \mathbf{h}^c_{1N} \\ \vdots & \ddots & \vdots \\ \mathbf{h}^c_{1M} & \cdots & \mathbf{h}^c_{MN} \end{bmatrix} \tag{4.55}$$

where $\mathbf{h}^c_{ij} = [h_{ij}(L-1), ..., h_{ij}(0)]^T$ stands for the causal part of the system response $h_{ij}(l)$.

The results in (4.52) and (4.55) are independent of the fixed measurement m. Hence, the different rows of the left-most matrix in (4.52) and (4.55) can be constructed out of the cumulants with the different fixed measurements m. Combine now the anti-causal and causal parts of all auto-cumulants and all cross-cumulants based on two different measurements into the matrices $\mathbf{S}_a$ and $\mathbf{S}_c$, respectively:

$$\mathbf{S}_a = \begin{bmatrix} \mathbf{C}^a_{111} & \mathbf{C}^a_{211} & \cdots & \mathbf{C}^a_{M11} \\ \mathbf{C}^a_{122} & \mathbf{C}^a_{222} & \cdots & \mathbf{C}^a_{M22} \\ \vdots & \vdots & & \vdots \\ \mathbf{C}^a_{1MM} & \mathbf{C}^a_{2MM} & \cdots & \mathbf{C}^a_{MMM} \end{bmatrix}, \tag{4.56}$$

$$\mathbf{S}_c = \begin{bmatrix} \mathbf{C}^c_{111} & \mathbf{C}^c_{211} & \cdots & \mathbf{C}^c_{M11} \\ \mathbf{C}^c_{122} & \mathbf{C}^c_{222} & \cdots & \mathbf{C}^c_{M22} \\ \vdots & \vdots & & \vdots \\ \mathbf{C}^c_{1MM} & \mathbf{C}^c_{2MM} & \cdots & \mathbf{C}^c_{MMM} \end{bmatrix} \tag{4.57}$$

According to (4.52) and (4.55), the matrix of weights **W** can be computed as:

$$\mathbf{W} = \mathbf{S}_a^{\#} \begin{bmatrix} \mathbf{E} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{E} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{E} \end{bmatrix} = \mathbf{S}_a^{\#} \mathbf{D} \tag{4.58}$$

where $\mathbf{E}=[1,0,...,0]$ and $\mathbf{0}=[0,...,0]$ stand for the vectors of length L and $^{\#}$ denotes the Moore-Penrose matrix pseudoinverse. The right-hand side of (4.58) guarantees the uniqueness of the weights **W**, however, it forces the channel responses $h_{ij}(l)$ to begin by 1 if $i=j$, and by 0 otherwise.

Finally, the system responses can be calculated as:

$$\begin{bmatrix} \mathbf{h}_{11}^{c} & \mathbf{h}_{12}^{c} & \cdots & \mathbf{h}_{1N}^{c} \\ \mathbf{h}_{21}^{c} & \mathbf{h}_{22}^{c} & \cdots & \mathbf{h}_{2N}^{c} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{h}_{M1}^{c} & \mathbf{h}_{M1}^{c} & \cdots & \mathbf{h}_{MN}^{c} \end{bmatrix} = \mathbf{S}_{c}\mathbf{W} \tag{4.59}$$

where $\mathbf{h}_{ij}^{c}=[h_{ij}(L-1),...,h_{ij}(0)]^{T}$ stands for the causal part of the system response $h_{ij}(l)$.

Matrices $\mathbf{S}_c$ and $\mathbf{S}_a$ comprise only auto-cumulants and cross-cumulants based on the pairs of different measurements. As already explained, the different rows of matrices $\mathbf{S}_c$ and $\mathbf{S}_a$ can be formed by calculating the cumulants with different fixed measurement *m*. By selecting all possible values of *m* in (4.56) and (4.57), we have simply increased the computational stability, as all different cumulants were used to estimate the system responses.

4.2.3 Nonlinear optimisation of coarse estimates

The coarse approximations of the system-channel responses, given by (4.59), can further be improved by non-linear optimisation, as proposed in [61]. The core of the method is Newton-Gauss iteration which completes non-linear minimisation of cost function $V(\hat{\boldsymbol{\theta}})$:

$$V(\hat{\boldsymbol{\theta}})=\left[\mathbf{c}(\hat{\boldsymbol{\theta}})-\hat{\mathbf{c}}\right]^{T}\boldsymbol{\Sigma}^{-1}\left[\mathbf{c}(\hat{\boldsymbol{\theta}})-\hat{\mathbf{c}}\right] \tag{4.60}$$

where the $\mathbf{c}$ and $\hat{\mathbf{c}}$ vectors are obtained by placing the transposed rows of the matrix

$$\mathbf{C}=\begin{bmatrix} \mathbf{C}_{111},\mathbf{C}_{211},\ldots,\mathbf{C}_{M11} \\ \mathbf{C}_{122},\mathbf{C}_{222},\ldots,\mathbf{C}_{M22} \\ \vdots \\ \mathbf{C}_{1MM},\mathbf{C}_{2MM},\ldots,\mathbf{C}_{MMM} \end{bmatrix} \tag{4.61}$$

in column one bellow the other. $\mathbf{C}_{ijk}=\begin{bmatrix} \mathbf{C}_{ijk}^{a} \\ \mathbf{C}_{ijk}^{c} \end{bmatrix}$ in (4.61) denotes the entire cumulant matrix of measurements *i*, *j* and *k*. The vector $\hat{\mathbf{c}}$ is constructed with the cumulants $\mathbf{C}_{ijk}$ estimated according to (4.38) from the system outputs (these are considered the reference

values to be fitted in the optimisation procedure), while $\mathbf{c}(\hat{\boldsymbol{\theta}})$ corresponds to the cumulants computed from (4.39) at each step of iteration (their values are being optimized). The vector of parameters

$$\hat{\boldsymbol{\theta}} = \left[h_{11}(0),\ldots,h_{11}(L-1),h_{12}(0),\ldots,h_{12}(L-1),...,h_{MN}(0),\ldots,h_{MN}(L-1),\gamma_{3,s_1},...,\gamma_{3,s_N}\right] \quad (4.62)$$

combines the values of the system responses and input skewnesses estimated at a certain step of iteration, while $\boldsymbol{\Sigma}$ stands for the asymptotically normalized covariance matrix of errors [61, 68]:

$$\boldsymbol{\Sigma} = \lim_{T\to\infty} T \cdot \left[E\{\hat{\mathbf{c}}^T\hat{\mathbf{c}}\} - \mathbf{c}(\hat{\boldsymbol{\theta}})^T\mathbf{c}(\hat{\boldsymbol{\theta}})\right] \quad (4.63)$$

where E denotes mathematical expectation, and T the length of MIMO outputs. In practice, the matrix $\boldsymbol{\Sigma}$ is replaced by its estimate

$$\hat{\boldsymbol{\Sigma}} = N \cdot \left[\tfrac{1}{P}\textstyle\sum_{p=1}^{P}\{\hat{\mathbf{c}}_{p,P}{}^T\hat{\mathbf{c}}_{p,P}\} - \mathbf{c}(\hat{\boldsymbol{\theta}})^T\mathbf{c}(\hat{\boldsymbol{\theta}})\right] \quad (4.64)$$

where *indices p, P* indicate that cumulants in $\hat{\mathbf{c}}_{p,P}$ are calculated according to (4.33) only over the samples from $n = (p-1)\frac{T}{P}+1$ to $n = p\frac{T}{P}$, and P is a positive, arbitrary large integer, but much smaller than T.

Minimization of the cost function (4.60) is carried out by iterating the Newton-Gauss update formula until the convergence is reached [61]:

$$\hat{\boldsymbol{\theta}}_{i+1} = \hat{\boldsymbol{\theta}}_i + \delta_i \cdot \left[\frac{\partial\mathbf{c}(\hat{\boldsymbol{\theta}}_i)^T}{\partial\hat{\boldsymbol{\theta}}_i}\hat{\boldsymbol{\Sigma}}^{-1}\frac{\partial\mathbf{c}(\hat{\boldsymbol{\theta}}_i)}{\partial\hat{\boldsymbol{\theta}}_i}\right]^{-1}\frac{\partial\mathbf{c}(\hat{\boldsymbol{\theta}}_i)^T}{\partial\hat{\boldsymbol{\theta}}_i}\hat{\boldsymbol{\Sigma}}^{-1}(\hat{\mathbf{c}} - \mathbf{c}(\hat{\boldsymbol{\theta}}_i)), \quad (4.65)$$

where $\frac{\partial\mathbf{c}(\hat{\boldsymbol{\theta}})}{\partial\hat{\boldsymbol{\theta}}}$ denotes the Jacobian matrix of $\mathbf{c}(\hat{\boldsymbol{\theta}})$ and $\hat{\boldsymbol{\Sigma}}^{-1}$ the matrix inverse of projection of $\boldsymbol{\Sigma}$ onto the space of positive semi-definite matrices. $\hat{\boldsymbol{\Sigma}}^{-1}$ is calculated by setting all the negative eigenvalues of matrix $\boldsymbol{\Sigma}$ to zero, while the positive eigenvalues λ_i are replaced by their reciprocal values $\frac{1}{\lambda_i}$ [68, 115]. The step size δ_i is determined by using a line search algorithm at each iteration (e.g. the method of bisection, cubic interpolation, Newton's method, etc.), while the derivatives, forming the Jacobian matrix $\frac{\partial\mathbf{c}(\hat{\boldsymbol{\theta}})}{\partial\hat{\boldsymbol{\theta}}}$ are defined in Appendix 4A.

Input skewnesses γ_i cannot be optimized by Newton-Gauss iteration. The reason is that the derivatives of cumulants with respect to the skewnesses can be expressed as a linear combination of derivatives with respect to system responses:

$$\frac{\partial \mathbf{c}(\hat{\boldsymbol{\theta}})}{\partial \gamma_j} = \frac{1}{3\gamma_j} \sum_{m=1}^{M} \sum_{n=1}^{q} h_{mj}(n) \frac{\partial \mathbf{c}(\hat{\boldsymbol{\theta}})}{\partial h_{mj}}. \tag{4.66}$$

Including the unknown skewnesses γ_i into the optimisation iteration causes the Jacobian matrix $\frac{\partial \mathbf{c}(\hat{\boldsymbol{\theta}})}{\partial \hat{\boldsymbol{\theta}}}$ to become rank-deficient and prevents the existence of the inverse matrix $\left[\frac{\partial \mathbf{c}(\hat{\boldsymbol{\theta}}_i)^T}{\partial \hat{\boldsymbol{\theta}}_i} \hat{\boldsymbol{\Sigma}}^{-1} \frac{\partial \mathbf{c}(\hat{\boldsymbol{\theta}}_i)}{\partial \hat{\boldsymbol{\theta}}_i} \right]^{-1}$ in (4.60). Instead, the unknown γ_i should be estimated from the cumulants in (4.38) and (4.39):

$$\boldsymbol{\gamma} = \mathbf{S}_{ijk}(\hat{\boldsymbol{\theta}}, \tau_2)^{-1} \hat{\mathbf{c}}_{ijk}(0, \tau_2) \tag{4.67}$$

where vector $\boldsymbol{\gamma} = [\gamma_1 \ \cdots \ \gamma_N]^T$ combines the skewnesses of all inputs. Matrix $\mathbf{S}_{ijk}(\hat{\boldsymbol{\theta}}, \tau_2)$ consists of the auto- and cross-cumulants of system responses:

$$\mathbf{S}_{ijk}(\hat{\boldsymbol{\theta}}, \tau_2) = \begin{bmatrix} \sum_{l=0}^{L-1} h_{i1}(l) h_{j1}(l+0) h_{k1}(l+\tau_2) & \cdots & \sum_{l=0}^{L-1} h_{iN}(l) h_{jN}(l+0) h_{kN}(l+\tau_2) \\ \vdots & \ddots & \vdots \\ \sum_{l=0}^{L-1} h_{i1}(l) h_{j1}(l+L-1) h_{k1}(l+\tau_2) & \cdots & \sum_{l=0}^{L-1} h_{iN}(l) h_{jN}(l+L-1) h_{kN}(l+\tau_2) \end{bmatrix}, \tag{4.68}$$

while the vector $\hat{\mathbf{c}}_{ijk}(0, \tau_2)$ combines L cumulants of measurements:

$$\hat{\mathbf{c}}_{ijk}(0, \tau_2) = [cum(x_i(n) x_j(n+0) x_k(n+\tau_2)), ..., cum(x_i(n) x_j(n+L-1) x_k(n+\tau_2))], \tag{4.69}$$

with arbitrary measurements i, j and k ($i, j, k \in [1, ..., M]$). Although the input skewnesses $\boldsymbol{\gamma} = [\gamma_1 \ \cdots \ \gamma_N]^T$ can be estimated at the optional values of indices i, j, k and lag τ_2, it is advisable to repeat the calculation for several different sets of values and to average the results afterwards.

4.2.4 Influence of non-Gaussian noise and nonorthogonal sources

The w-slices and Newton-Gauss optimization are based on the Barlett-Brillinger-Rosenblatt equation. The latter supposes all the cross-cumulants of different sources equal to zero, i.e. it assumes the sources are independent. When the source pulses overlap, the cross-cumulant can not be neglected any more and the Barlett-Brillinger-Rosenblatt equation from (4.34) yields:

$$\begin{aligned} &C_{x_{i_1},...,x_{i_r}}(\tau_1,...,\tau_{r-1}) = cum(x_{i_1}(n), x_{i_2}(n+\tau_1),..., x_{i_r}(n-\tau_{r-1})) = \\ &= \sum_{j_1=1}^{N} \cdots \sum_{j_r=1}^{N} \sum_{l_1=0}^{L-1} \cdots \sum_{l_r=0}^{L-1} C_{s_{j_1},...,s_{j_r}}(l_2,...,l_r) h_{i_1}(l_1) h_{i_2}(l_1 - l_2 + \tau_1) \cdots h_{i_r}(l_1 - l_r + \tau_{r-1}), \end{aligned} \tag{4.70}$$

where $C_{s_{j_1},...,s_{j_M}}(l_2,...,l_M)$ denotes (cross-) cumulant of sources $s_{j_1},...,s_{j_M}$. An increase of the number of overlapped pulses increases the values of cross-cumulants and, consequently, the differences between the cumulants in (4.38) and (4.39). As a result, the w-slices and Newton-Gauss will try to match Eqs. (4.34) and (4.70), which will lead to a wrong estimation of the system responses.

A similar effect can be observed in the presence of non-Gaussian noise $\boldsymbol{\omega}(n) = [\omega_1(n),....,\omega_M(n)]^T$. Supposing $\boldsymbol{\omega}(n)^T$ are independent of sources the Barlett-Brillinger-Rosenblatt equation yields:

$$\begin{aligned} &C_{x_{i_1}+\omega_{i_1},...,x_{i_r}+\omega_{i_r}}(\tau_1,...,\tau_{r-1}) = \\ &\quad = cum(x_{i_1}(n) + \omega_{i_1}(n),..., x_{i_r}(n-\tau_{r-1}) + \omega_{i_r}(n-\tau_{r-1})) = \\ &\quad = \sum_{j=1}^{N} \gamma_{r,s_j} \cdot \left(\sum_{l=0}^{L-1} h_{i_1 j}(l) h_{i_2 j}(l+\tau_1) \cdots h_{i_r j}(l+\tau_{r-1}) \right) + \\ &\quad + C_{\omega_{i_1},...,\omega_{i_r}}(\tau_1,...,\tau_{r-1}) \end{aligned} \tag{4.71}$$

where $C_{\omega_{i_1},...,\omega_{i_r}}(\tau_1,...,\tau_{r-1})$ denotes the cumulant of noise.

Appendix 4A: Derivatives of cumulants

Let

$$c_{ijk}(\tau_1,\tau_2)=\sum_{q=1}^{N}\gamma_q\sum_{l=0}^{L-1}h_{iq}(l)h_{jq}(l+\tau_1)h_{kq}(l+\tau_2) \tag{4A.1}$$

denote the general (cross-) cumulant value at the arbitrary lags τ_1 and τ_2 and observe all possible derivatives $\frac{\partial c_{ijk}(\tau_1,\tau_2)}{\partial h_{mq}(n)}$, $m=i, j, k$. We distinguish three different cases:

a) derivatives of the auto-cumulant (i.e. $i=j=k$):

$$\begin{aligned}\frac{c_{iii}(\tau_1,\tau_2)}{\partial h_{iq}(n)}=\gamma_q[&h_{iq}(n-\tau_2)h_{iq}(n-\tau_2+\tau_1)+h_{iq}(n-\tau_1)h_{iq}(n-\tau_1+\tau_2)+\\&+h_{iq}(n+\tau_1)h_{iq}(n+\tau_2)];\end{aligned} \tag{4A.2}$$

b) derivatives of cross-cumulants based on two different measurements (i.e $i=j\neq k$, $i\neq j=k$, or $i=j\neq k$):

($i=j\neq k$)

$$\frac{c_{iik}(\tau_1,\tau_2)}{\partial h_{iq}(n)}=\gamma_q\left[h_{iq}(n-\tau_1)h_{mq}(n-\tau_1+\tau_2)+h_{iq}(n+\tau_1)h_{iq}(n+\tau_2)\right], \tag{4A.3}$$

$$\frac{c_{iik}(\tau_1,\tau_2)}{\partial h_{kq}(n)}=\gamma_q\left[h_{iq}(n-\tau_2)h_{iq}(n-\tau_2+\tau_1)\right], \tag{4A.4}$$

($i=k\neq j$)

$$\frac{c_{iji}(\tau_1,\tau_2)}{\partial h_{iq}(n)}=\gamma_q\left[h_{iq}(n-\tau_2)h_{jq}(n-\tau_2+\tau_1)+h_{iq}(n+\tau_2)h_{jq}(n+\tau_1)\right], \tag{4A.5}$$

$$\frac{c_{iji}(\tau_1,\tau_2)}{\partial h_{jq}(n)}=\gamma_q\left[h_{iq}(n-\tau_1)h_{iq}(n-\tau_1+\tau_2)\right], \tag{4A.6}$$

($j=k\neq i$)

$$\frac{c_{ijj}(\tau_1,\tau_2)}{\partial h_{iq}(n)} = \gamma_q\left[h_{jq}(n+\tau_1)h_{jq}(n+\tau_2)\right], \tag{4A.7}$$

$$\frac{c_{ijj}(\tau_1,\tau_2)}{\partial h_{jq}(n)} = \gamma_q\left[h_{iq}(n-\tau_1)h_{jq}(n-\tau_1+\tau_2)+h_{iq}(n-\tau_1)h_{jq}(n-\tau_1+\tau_2)\right]; \tag{4A.8}$$

c) derivatives of cross-cumulants based on three different measurements (i.e. $i\neq j\neq k$):

$$\frac{c_{ijk}(\tau_1,\tau_2)}{\partial h_{iq}(n)} = \gamma_q\left[h_{jq}(n+\tau_1)h_{kq}(n+\tau_2)\right], \tag{4A.9}$$

$$\frac{c_{ijk}(\tau_1,\tau_2)}{\partial h_{jq}(n)} = \gamma_q\left[h_{iq}(n-\tau_1)h_{kq}(n-\tau_1+\tau_2)\right], \tag{4A.10}$$

$$\frac{c_{ijk}(\tau_1,\tau_2)}{\partial h_{kq}(n)} = \gamma_q\left[h_{iq}(n-\tau_2)h_{jq}(n-\tau_2+\tau_1)\right]. \tag{4A.11}$$

5.

Inverse correlation based method

In previous chapters, two novel approaches to blind source separation of convolutive mixtures were introduced, the method of time-frequency distributions and w-slices. First we showed how joint diagonalization of STFD matrices can be utilized to estimate the mixing matrix and pulse sources up to a scalar factor. We identified the devastating influence of crossterms and explained how they can be avoided. The two main drawbacks of the presented method turned up to be its inability to confront the heavy noise, and the need to find at least one non-overlapped pulse of each source. While the first problem can partially be solved by averaging the STFD matrices, the second problem is much more serious. Keeping the number of sources small and sources orthogonal, the task of selecting the not-overlapped pulses is trivial. In the case of close-to-orthogonal sources and high signal-to-noise ratio the candidate pulses can be selected using the criterion in (4.21). But, by increasing the number of sources, the probability of finding the non-overlapped pulses drops drastically, and the efficiency of the presented approach decreases. Moreover, being sensitive to noise, the criterion (4.21) also causes a drop of performance in noisy environments.

On the other hand, being based on HOS, w-slices and Newton-Gauss optimisation are inherently resistant to the Gaussian noise. Unfortunately, they can only be used to reconstruct the system responses, while the information about the sources is lost in the decomposition process. The obvious drawback of all HOS-based decomposition techniques is also their high computation complexity. The theoretically expected behaviour of the cumulants calculated out of the measurements can only be observed if the processing signals are at least several 10000 samples long. As a result, HOS-based methods are both time and memory demanding.

The methods introduced in the previous chapter share another common drawback. They are based on the assumption of more measurements than sources. Although very common in BSS, this assumption cannot always be established in practice. Usually, we even don’t know the exact number of sources, and even if we did, it is sometimes hard to ensure enough measurements. In the case of surface EMG signals, the maximum size of the array of electrodes is limited by the size of the investigated muscle. Reducing the inter-electrode distance the differences among the detected MUAPs diminish and the mixing matrix becomes rank-deficient. On the other hand, there can easily be several tens of MUs active in the detection volume.

In this chapter, we discuss how the decomposition can be extended to the under-determined case, i.e. when the number of sources exceeds the number of measurements.

Realising it is much easier to estimate the binary sources than the mixing matrix in general, we first introduce a completely new decomposition concept. Throughout its derivation we ignore all the information contained in the mixing matrix and focus strictly on the properties of pulse sources. In other words, we try to cancel the interfering effect of the mixing matrix. As a result, we only manage to reconstruct the estimates of sources, while the mixing matrix is completely lost. We further show this is not a limitation at all, because the system responses can always be estimated by averaging the measurements in the vicinity of the reconstructed pulses.

The chapter is divided in four sections. In Section 5.1 we first consider the sources strictly orthogonal. We also assume more measurements than sources and derive a novel decomposition approach, the so called »inverse correlation (IC) based method«. The decomposition of non-orthogonal sources is introduced in Section 5.2, while Section 5.3 evaluates the influence of noise. The last section discusses the impact of the number of measurements and extends the novel approach to the case of under-determined MIMO systems, i.e. the systems with more sources than measurements.

5.1 Decomposition of orthogonal sources

Suppose the number of sources N smaller than the number of measurements M and extend the vector of noisy measurements by their delayed repetitions (Section 3.4.1). Further suppose the extended sources orthogonal and denote by $\mathbf{R}_{\bar{\mathbf{s}}}$ their correlation matrix:

$$\mathbf{R}_{\bar{\mathbf{s}}} = \lim_{T\to\infty} \frac{1}{T} \sum_{n=1}^{T} \bar{\mathbf{s}}(n)\bar{\mathbf{s}}^{H}(n) = diag[r_{1,1},...,r_{N(K+L-1),N(K+L-1)}] \tag{5.1}$$

where *diag[]* stands for the diagonal matrix with its parameters on the diagonal, $r_{j,j} = \lim_{T\to\infty} \frac{1}{T} \sum_{n=1}^{T} s_j(n)s_j^{*}(n)$ is the autocorrelation function of source $s_j(n)$, and $s_j^{*}(n)$ denotes the complex conjugate value of $s_j(n)$.

Let $\bar{\mathbf{y}}(n) = \bar{\mathbf{x}}(n) + \bar{\boldsymbol{\omega}}(n)$ denote the vector of noisy measurements and recall the mixing matrix $\mathbf{H}$ is a $KM \times N(L+K-1)$ rectangular matrix which corresponds to the over-determined system of equations.

According to (3.12) there are at least $KM - N(L+K-1)$ eigenvalues of $\mathbf{R}_{\bar{\mathbf{x}}}$ equal to zero. Consequently, the noise variance $\hat{\sigma}^2$ can be estimated by averaging the $KM - N(L+K-1)$ smallest eigenvalues of $\mathbf{R}_{\bar{\mathbf{y}}}$. Subtracting $\hat{\sigma}^2$ from $\mathbf{R}_{\bar{\mathbf{y}}}$ we obtain the estimation of correlation matrix of measurements:

$$\mathbf{R}_{\bar{\mathbf{x}}} = \mathbf{R}_{\bar{\mathbf{y}}} - \hat{\sigma}^2\mathbf{I} \approx \mathbf{H}\mathbf{R}_{\bar{\mathbf{s}}}\mathbf{H}^{T}. \tag{5.2}$$

Supposing the mixing matrix **H** of full column rank:

$$rang(\mathbf{H}) = N(L+K-1), \tag{5.3}$$

and multiplying the vector of the extended measurements with Moore-Penrose pseudo-inverse of $\mathbf{R}_{\bar{\mathbf{x}}}$, we introduce the so called activity index,

$$\begin{aligned} Ind(n) &= \bar{\mathbf{x}}(n)^T \mathbf{R}_{\bar{\mathbf{x}}}^{\#} \bar{\mathbf{x}}(n) + o_{\bar{\boldsymbol{\omega}}}(n,n) = \bar{\mathbf{s}}(n)^T \mathbf{H}^T (\mathbf{H}\mathbf{R}_{\bar{\mathbf{s}}}\mathbf{H}^T)^{\#} \mathbf{H}\bar{\mathbf{s}}(n) + o_{\bar{\boldsymbol{\omega}}}(n,n) = \\ &= \bar{\mathbf{s}}(n)^T \mathbf{H}^T (\mathbf{H}^T)^{\#} \mathbf{R}_{\bar{\mathbf{s}}}^{-1} \mathbf{H}^{\#} \mathbf{H}\bar{\mathbf{s}}(n) + o_{\bar{\boldsymbol{\omega}}}(n,n) = \bar{\mathbf{s}}(n)^T \mathbf{R}_{\bar{\mathbf{s}}}^{-1} \bar{\mathbf{s}}(n) + o_{\bar{\boldsymbol{\omega}}}(n,n) = \\ &= \sum_{j=1}^{N(L+K-1)} \bar{s}_j(n) r_{j,j}^{-1} \bar{s}_j(n) + o_{\bar{\boldsymbol{\omega}}}(n,n) \end{aligned} \tag{5.4}$$

where $^{\#}$ denotes the Moore-Penrose matrix pseudoinverse, and $o_{\bar{\boldsymbol{\omega}}}(n,n)$ replaces the impact of noise. In (5.4) we also used the equality $\mathbf{R}_{\bar{\mathbf{s}}}^{\#} = \mathbf{R}_{\bar{\mathbf{s}}}^{-1}$, which is generally valid only for non-singular matrices.

Neglecting the influence of noise, the activity index $Ind(n)$ could be thought of as an indicator of a global source activity. It differs from zero only at the time instants n where at least one source is active. Since the sources are orthogonal, there can be at most one active source in every time moment.

When a sample index n_0 is found in which the j-th source is active, the entire pulse train of the j-th source can be reconstructed as:

$$\begin{aligned} v_{n_0,j}(n) &= \bar{\mathbf{x}}(n_0)^T \mathbf{R}_{\bar{\mathbf{x}}}^{\#} \bar{\mathbf{x}}(n) + o_{\bar{\boldsymbol{\omega}}}(n_0,n) = \bar{\mathbf{s}}(n_0)^T \mathbf{R}_{\bar{\mathbf{s}}}^{-1} \bar{\mathbf{s}}(n) + o_{\bar{\boldsymbol{\omega}}}(n) = \\ &\quad r_{j,j}^{-1} \bar{s}_j(n) + o_{\bar{\boldsymbol{\omega}}}(n_0,n) \end{aligned} \tag{5.5}$$

assuming that $\bar{s}_j(n_0) = 1$. According to (5.5), the pulse trains of different sources can be reconstructed by simply choosing different time moments n_0.

Knowing the pulse sources, their system responses can be reconstructed by averaging corresponding measurements in the vicinity of their pulses [46]:

$$\hat{h}_{ij}(l) = \frac{1}{F_i} \sum_{k=1}^{F_i} \Psi(l) \cdot x_j(n_{i,k} + l), \tag{5.6}$$

where $n_{i,k}$ stands for the time moment corresponding to the k-th reconstructed pulse of the i-th source, F_i denotes the number of reconstructed pulses of the i-th source, while $\Psi(l)$ is an arbitrary window function. This procedure is commonly referred to as a spike triggered sliding window averaging technique [46].

5.2 Decomposition of non-orthogonal sources

Suppose now close-to-orthogonal sources, i.e. sources with small but significant number of overlapped pulses. This implies the correlation matrix $\mathbf{R}_{\bar{\mathbf{s}}}$ has dominant diagonal but is not strictly diagonal. The fact that there can be several sources active in the same time moment has a crucial impact on the activity index.

Suppose there are Q_{n_0} sources active in the same time moment n_0 and denote them by a set of indices $G_{n_0} = \{j_{n_0,1},..., j_{n_0,Q_{n_0}}\}$, where $G_{n_0} \subseteq \{1,2...,K(L+N)\}$. Neglecting the influence of noise, the value of the activity index in time moment n_0 yields:

$$\begin{aligned} Ind(n_0) = \bar{\mathbf{s}}(n_0)^T \mathbf{R}_{\bar{\mathbf{s}}}^{-1} \bar{\mathbf{s}}(n_0) = r^{inv}_{j_{n_0,1},j_{n_0,1}} \bar{s}^2_{j_{n_0,1}}(n_0) + r^{inv}_{j_{n_0,1},j_{n_0 2}} \bar{s}_{j_{n_0,1}}(n_0)\bar{s}_{j_{n_0,2}}(n_0) + \cdots \\ \cdots + r^{inv}_{j_{n_0,1},j_{n_0,Q_{n_0}}} \bar{s}_{j_{n_0,1}}(n_0)\bar{s}_{j_{n_0,Q_{n_0}}}(n_0) + r^{inv}_{j_{n_0,2},j_{n_0,1}} \bar{s}_{j_{n_0,2}}(n_0)\bar{s}_{j_{n_0,1}}(n_0) + r^{inv}_{j_{n_0,2},j_{n_0,2}} \bar{s}^2_{j_{n_0,2}}(n_0) + \\ + r^{inv}_{j_{n_0,2},j_{n_0,3}} \bar{s}_{j_{n_0,2}}(n_0)\bar{s}_{j_{n_0,3}}(n_0) + \cdots + r^{inv}_{j_{n_0,Q_{n_0}},j_{n_0,Q_{n_0}}} \bar{s}^2_{j_{n_0,Q_{n_0}}}(n_0) \end{aligned} \quad (5.7)$$

where $r^{inv}_{i,j}$ denotes the *(i,j)*-th element of matrix $\mathbf{R}_{\bar{\mathbf{s}}}^{-1}$. According to (5.7), the activity index at a given time moment n_0 comprises the contributions of all the sources that are active in that very moment. The non-diagonal matrix $\mathbf{R}_{\bar{\mathbf{s}}}^{-1}$ causes the activity index also contains all possible combinations $r^{inv}_{i,j}\bar{s}_i(n_0)\bar{s}_j(n_0)$ of active sources.

The same is true for the reconstructed pulse trains of sources. Using (5.5) in the case of the overlapping sources, we can generally reconstruct only the superimposition of all the sources:

$$v_{n_0,G_{n_0}}(n) = \bar{\mathbf{s}}(n_0)^T \mathbf{R}_{\bar{\mathbf{s}}}^{-1} \bar{\mathbf{s}}(n) = \sum_{q=1}^{Q_{n_0}} \bar{s}_{j_{n_0,q}}(n_0) \mathbf{r}^{inv}_{j_{n_0,q}} \bar{\mathbf{s}}(n) = \sum_{i=1}^{N(L+K)} \sum_{q=1}^{Q_{n_0}} r^{inv}_{j_{n_0,q},i} \bar{s}_i(n) \quad (5.8)$$

where $\mathbf{r}^{inv}_i$ denotes the *i*-th row of matrix $\mathbf{R}_{\bar{\mathbf{s}}}^{-1}$, while $\bar{s}_{j_{n_0,q}}(n_0)$ is supposed equal to 1.

Suppose now the matrix $\mathbf{R}_{\bar{\mathbf{s}}}$ diagonal dominant. Then, as proven in Appendix 5A, its inverse has a superior diagonal, while all the off-diagonal elements are much smaller than the diagonal ones. This implies the superimposition $v_{n_0,G_{n_0}}(n)$ has strong contributions only from the sources that are contained in set G_{n_0}. The contributions form all other sources are much smaller in amplitude and can be removed by a simple threshold operation. Eq. (5.8) now derives into:

$$v_{n_0,G_{n_0}}(n) \approx \sum_{j \in G_{n_0}} r^{inv}_{j,j} \bar{s}_j(n_0)\bar{s}_j(n) = \sum_{j \in G_{n_0}} r^{inv}_{j,j} \bar{s}_j(n). \quad (5.9)$$

According to (5.9), the entire pulse train of the *j*-th source can still be reconstructed, providing we have found any time moment n_0 with a contribution (pulse) from only the *j*-th source. But such moments are very hard to find. We can partially rely on the activity index as its value increases with the number of simultaneously active sources. However, the reverse is not generally true. A high value of the activity index can either be the consequence of a bigger number of simultaneously active sources or can indicate the activity of only one source with a low firing rate. Namely, according to (5.7) the contributions from different sources are multiplied by factors $r_{i,j}^{inv}$, which are, as shown in Appendix 5A, proportional to the inverses of the corresponding firing rates. Consequently, the activity index must be used with precaution and more formal procedure for separation of the source superimpositions is needed.

5.2.1 Separation of the superimposed sources

As explained in Section 3.1, the firing of a pulse source can be modelled by a random process. Each sample in the pulse train is considered a random variable with two possible outcomes (pulse, no pulse). Its probability density function depends on the outcome of the several previous samples and is, hence, very difficult to define. In the sequel we will rather operate with conditional probabilities. We will assume there is at least one source active in the observed time moment, and define the probability that two or more sources are active in the same moment. In other words, we will try to answer the question of how probable it is that the *k*-th pulse of *j*-th source overlaps with any other pulse from any other source.

Suppose now the overlap of two pulses from different sources a random event. Its probability is proportional to its relative frequency of occurrence and is, at least in the case of close-to-orthogonal pulse sources, relatively low (usually well below 0.1). This may be seen more readily by realizing that the relative frequency of overlaps between the *i*-th and *j*-th source is proportional to the (*i*,*j*)-th element of the matrix $\mathbf{R}_{\tilde{\mathbf{s}}}$.

The maximum number of simultaneously triggered pulses is usually much smaller from the number of all extended sources. In the case of biomedical signals the refractory period prevents the delayed repetitions of the same source to overlap. Consequently, the number of simultaneously triggered pulses cannot exceed the number of original sources *N* (having *N+1* overlapped pulses would necessarily imply that at least two repetitions of the same source overlap).

Consider a more general case and assume the sources can be arranged into *R* subsets, each comprising *Q* sources, whereas the sources from the same subset cannot overlap with each other. Define the overlap of two subsets as the overlap of any two sources from different subsets and assume all sources within the same subset are equally probable to overlap with sources from the other subsets. For the sake of simplicity, also suppose all possible overlaps of all possible subsets are independent, equally probable random events, and denote their probability by *p* (Appendix 5B).

We are now in position to discuss the separation of superimposed sources. Using (5.8), reconstruct the superimposition of the pulse trains of sources active in a given time moment n_0:

$$v_{n_0,G_{n_0}}(n) \approx \sum_{j \in G_{n_0}} r_{j,j}^{inv} \bar{s}_j(n) \tag{5.10}$$

where $G_{n_0} = \{j_{n_0,1},..., j_{n_0,Q_{n_0}}\}$ denotes the set of all sources that fired in n_0. Without loss of generality we can assume the set G_{n_0} is not empty. Randomly select the pulse in $v_{n_0,G_{n_0}}(n)$ and denote the moment of its occurrence with n_1. Further, denote by $G_{n_1} = \{j_{n_1,1},..., j_{n_1,Q_{n_1}}\}$ the set of sources that were active in moment n_1. Because the moment n_1 was chosen from the superimposition $v_{n_0,G_{n_0}}(n)$, there is at least one source that is active in both moments n_0 and n_1. This implies the intersection $G_{n_0} \cap G_{n_1}$ contains at least one source:

$$G_{n_0} \cap G_{n_1} \neq \{\} \tag{5.11}$$

where {} denotes the empty set. The conditional probability that $G_{n_0} \cap G_{n_1}$ comprises g_{01} sources, if there is at least one source in each set G_{n_0} and G_{n_1}, can be estimated as (Appendix 5B):

$$P\left(card(G_{n_0} \cap G_{n_1}) = g_{01} \,\middle|\, card(G_{n_0}) > 0, card(G_{n_1}) > 0\right) \leq \binom{R}{g_{01}} \frac{p^{2(g_0-1)}}{Q^{g_{01}}} \tag{5.12}$$

where $card(G_{n_0} \cap G_{n_1}) = g_{01}$ denotes the cardinal number of intersection $G_{n_0} \cap G_{n_1}$.

According to (5.12), the probability $P\left(card(G_{n_0} \cap G_{n_1}) = g_{01} \,\middle|\, card(G_{n_0}) > 0, card(G_{n_1}) > 0\right)$ decreases rapidly with the number of expected common sources g_{01} and also with the number of sources in each subset Q. Consequently, the sets G_{n_0} and G_{n_1} will most probably differ from each other (Appendix 5B). The same conclusion may apply for superimpositions

$$v_{n_0,G_{n_0}}(n) \approx \sum_{j \in G_{n_0}} r_{j,j}^{inv} \bar{s}_j(n) \tag{5.13}$$

and

$$v_{n_1,G_{n_1}}(n) \approx \sum_{i \in G_{n_1}} r_{i,i}^{inv} \bar{s}_i(n). \tag{5.14}$$

The superimpositions $v_{n_0,G_{n_0}}(n)$ and $v_{n_1,G_{n_1}}(n)$ overlap in all the pulses that belong to the sources contained in $G_{n_0} \cap G_{n_1}$. They will also share those pulses which are overlapped for the sources from $G_{n_0} - G_{n_1}$ by the pulses of sources from $G_{n_1} - G_{n_0}$, where $M_1 - M_2$ stands for the set difference (M_1 without M_2). Suppose the superimpositions $v_{n_0,G_{n_0}}(n)$ and $v_{n_1,G_{n_1}}(n)$ share a pulse at time moment n_2. There are two possible complementary explanations:

Assumption 5.1: At least one source from $G_{n_0} \cap G_{n_1}$ fired at n_2:

$$\exists i;\ i \in G_{n_0} \cap G_{n_1} :\quad \bar{s}_i(n_2) = 1. \tag{5.15}$$

Assumption 5.2: There is no source in $G_{n_0} \cap G_{n_1}$ with a pulse at n_2. However, at least two different sources fired at n_2; the first is contained in $G_{n_0} - G_{n_1}$, the second in $G_{n_1} - G_{n_0}$:

$$\begin{aligned} &\forall i; i \in G_{n_0} \cap G_{n_1} : \bar{s}_i(n_2) = 0 \quad \wedge \quad \exists j; j \in G_{n_0} - G_{n_1} : \bar{s}_j(n_2) = 1 \quad \wedge \\ &\exists k; k \in G_{n_1} - G_{n_0} : \bar{s}_k(n_2) = 1. \end{aligned} \tag{5.16}$$

Let the superimpositions $v_{n_0,G_{n_0}}(n)$ and $v_{n_1,G_{n_1}}(n)$ share another pulse at time moment n_3 and assume that the moment n_2 fulfils the Assumption 5.1, while at n_3 Assumption 5.2 is fulfilled. Then, according to Assumption 5.1, the intersection $G_{n_0} \cap G_{n_1} \cap G_{n_2}$ contains at least one source, such that:

$$\exists i; i \in G_{n_0} \cap G_{n_1} : \bar{s}_i(n_2) = 1 \ \Rightarrow \quad G_{n_0} \cap G_{n_1} \cap G_{n_2} \neq \{\}, \tag{5.17}$$

while Assumption 5.2 guarantees

$$\begin{aligned} &\forall i; i \in G_{n_0} \cap G_{n_1} : \bar{s}_i(n_3) = 0 \ \Rightarrow \quad G_{n_0} \cap G_{n_1} \cap G_{n_3} = \{\}, \\ &\exists j; j \in G_{n_0} - G_{n_1} : \bar{s}_j(n_3) = 1 \ \Rightarrow \quad G_{n_0} \cap G_{n_3} \neq \{\}, \\ &\exists k; k \in G_{n_1} - G_{n_2} : \bar{s}_k(n_3) = 1 \ \Rightarrow \quad G_{n_1} \cap G_{n_3} \neq \{\}. \end{aligned} \tag{5.18}$$

Using the conclusions from Appendix 5B we quickly realize the probability of having more than one source in any intersection $G_{n_i} \cap G_{n_j}; i \neq j$, becomes negligible as soon as Q (the number of sources in each subset) exceeds the number of subsets R. In similar

fashion, the intersections $G_{n_0} \cap G_{n_2} \cap G_{n_3}$ and $G_{n_1} \cap G_{n_2} \cap G_{n_3}$ are (with high portability) empty sets and the number of pulses in the product $v_{n_0,G_{n_0}}(n) \cdot v_{n_1,G_{n_1}}(n) \cdot v_{n_2,G_{n_2}}(n) \cdot v_{n_3,G_{n_3}}(n)$ can be estimated as (Appendix 5C, Case 1):

$$I(p_Q) \approx Fp^2 \frac{R}{Q} \tag{5.19}$$

where, for the sake of simplicity, we assumed the total number of pulses for all the sources equals *F* (Appendix 5C).

Assume now both time moments n_2 and n_3 fulfil Assumption 5.1 and let there be at least one source active in all four time moments: n_0, n_1, n_2 and n_3. Then the intersection $G_{n_0} \cap G_{n_1} \cap G_{n_2} \cap G_{n_3}$ comprises at least one source, while conclusions in Appendix 5B guarantee the probability of having more than one source in $G_{n_0} \cap G_{n_1} \cap G_{n_2} \cap G_{n_3}$ is negligible. Analogously to the case above, we can estimate the number of pulses in the product $v_{n_0,G_{n_0}}(n) \cdot v_{n_1,G_{n_1}}(n) \cdot v_{n_2,G_{n_2}}(n) \cdot v_{n_3,G_{n_3}}(n)$ as (Appendix 5C, Case 2):

$$I(p_Q) \approx F(1+3\frac{p}{Q}). \tag{5.20}$$

According to (5.19) and (5.20), the superimposed sources can be separated by simply comparing the pulses in $v_{n_0,G_{n_0}}(n) \cdot v_{n_1,G_{n_1}}(n) \cdot v_{n_2,G_{n_2}}(n)$ and $v_{n_0,G_{n_0}}(n) \cdot v_{n_1,G_{n_1}}(n) \cdot v_{n_3,G_{n_3}}(n)$. Using (5.8), we first reconstruct the superimpositions $v_{n_i,G_{n_i}}(n)$ in all time moments n_i which correspond to the pulses in the product $v_{n_0,G_{n_0}}(n) \cdot v_{n_1,G_{n_1}}(n)$. For all possible pairs (i, j), where $i \neq j$, we check the number of pulses in the product

$$v_{n_0,G_{n_0}}(n) \cdot v_{n_1,G_{n_1}}(n) \cdot v_{n_i,G_{n_i}}(n) \cdot v_{n_j,G_{n_j}}(n) \tag{5.21}$$

and combine all the moments n_i and n_j for which the number of pulses in the product (5.21) exceeds a predefined threshold *F/2* in a common set A_l. In a rather unlikely case when the time moments n_i and n_j share more than one common source, we get several sets A_l. Recall the probability of having more than one source active in several different time moments converges to zero with the number of different time moments (Appendix 5B). Hence, the set A_l will contain the time moments corresponding to the pulses of one single, say the *i*-th source. By averaging the vector of extended measurements over all time moments comprising the set A_l we get:

$$\bar{\mathbf{x}}(A_i) = \frac{1}{card(A_i)} \sum_{n_i \in A_i} \bar{\mathbf{x}}(n_i) \tag{5.22}$$

where $card(A_i)$ stands for the cardinal number of set A_i. Vector $\overline{\mathbf{x}}(A_i)$ will have a strong contribution from the *i*-th source, while the contributions of all other sources will tend to zero, as the number of elements in A_i will rise over all limits:

$$\lim_{card(A_i)\to\infty} \overline{\mathbf{x}}(A_i) = \mathbf{H}\mathbf{e}_i \tag{5.23}$$

where $\mathbf{e}_i$ stands for the *i*-th unit norm vector with a single nonzero element at the *i*-th position. Using the vector $\overline{\mathbf{x}}(A_i)$ in (5.8) we reconstruct the *i*-th source

$$\begin{aligned} v_{A_i,i}(n) &= \overline{\mathbf{x}}(A_i)^T \mathbf{R}_{\overline{\mathbf{x}}}^{\#} \overline{\mathbf{x}}(n) + o_{\overline{\boldsymbol{\omega}}}(A_i, n) = \mathbf{e}_i^T \mathbf{R}_{\overline{\mathbf{s}}}^{-1} \overline{\mathbf{s}}(n) + o_{\overline{\boldsymbol{\omega}}}(A_i, n) \approx \\ &\approx r_{i,i}^{inv} \overline{s}_i(n) + o_{\overline{\boldsymbol{\omega}}}(A_i, n). \end{aligned} \tag{5.24}$$

5.3 Noise influence

Using the matrix calculus, the noise influence $o_{\overline{\boldsymbol{\omega}}}(n_0, n)$ in (5.5) can be expressed as:

$$o_{\overline{\boldsymbol{\omega}}}(n_0, n) = \overline{\boldsymbol{\omega}}(n_0)^T (\mathbf{H}^T)^{\#} \mathbf{R}_{\overline{\mathbf{s}}}^{-1} \overline{\mathbf{s}}(n) + \overline{\boldsymbol{\omega}}(n_0)^T \mathbf{R}_{\overline{\mathbf{x}}}^{\#} \overline{\boldsymbol{\omega}}(n) + \overline{\mathbf{s}}(n_0)^T \mathbf{R}_{\overline{\mathbf{s}}}^{-1} \mathbf{H}^{\#} \overline{\boldsymbol{\omega}}(n). \tag{5.25}$$

Assume now the procedure for separation of superimposed sources, which was derived in the previous section, succeeded to find a set of time moments $A_i = \{n_1, n_2, ...\}$ corresponding to the pulses of the *i*-th source. Let's average the extended vector of noisy measurements over all time moments in A_i:

$$\overline{\mathbf{y}}(A_i) = \frac{1}{card(A_i)} \sum_{n_i \in A_i} \overline{\mathbf{y}}(n_i) \tag{5.26}$$

and recall the noise is modelled as a white, zero-mean Gaussian random process. Using the Cramer-Rao lower limit [92], it is easy to show the variance of the average in the presence of the white Gaussian noise is limited by

$$\sigma^2_{\overline{\mathbf{y}}(A_i)} = \frac{\sigma^2}{card(A_i)} \tag{5.27}$$

where σ^2 stands for the noise variance. It is trivial to see

$$\lim_{card(A_i)\to\infty} \sigma^2_{\overline{\mathbf{y}}(A_i)} = 0. \tag{5.28}$$

According to (5.28), the influence of the noise vector $\overline{\boldsymbol{\omega}}(n_0)^T$ in (5.25) will tend to zero when the vector $\overline{\mathbf{y}}(A_i)$ will be calculated over a large enough set of time moments A_i. Consequently, (5.5) suggests

$$v_{A_i,i}(n) = \overline{\mathbf{y}}(A_i)^T \mathbf{R}_{\overline{\mathbf{x}}}^{\#} \overline{\mathbf{y}}(n) \approx r_{i,i}^{inv} \overline{s}_i(n) + o_{\overline{\boldsymbol{\omega}}}(A_i, n) \tag{5.29}$$

and the noise influence $o_{\overline{\boldsymbol{\omega}}}(A_i, n)$ reads:

$$o_{\overline{\boldsymbol{\omega}}}(A_i, n) = \frac{\overline{\boldsymbol{\omega}}^T(A_i)}{card(A_i)} (\mathbf{H}^T)^{\#} \mathbf{R}_{\overline{\mathbf{s}}}^{-1} \overline{\mathbf{s}}(n) + \frac{\overline{\boldsymbol{\omega}}^T(A_i)}{card(A_i)} \mathbf{R}_{\overline{\mathbf{x}}}^{\#} \overline{\boldsymbol{\omega}}(n) + \overline{\mathbf{s}}(A_i)^T \mathbf{R}_{\overline{\mathbf{s}}}^{-1} \mathbf{H}^{\#} \overline{\boldsymbol{\omega}}(n). \tag{5.30}$$

While the first two factors in (5.30) converge to zero, the averaging in vector $\overline{\mathbf{y}}(A_i)$ hardly changes the last term $\overline{\mathbf{s}}(A_i)^T \mathbf{R}_{\overline{\mathbf{s}}}^{-1} \mathbf{H}^{\#} \overline{\boldsymbol{\omega}}(n)$. Its impact can be reduced by averaging the reconstructions of different repetitions of the same source. There is yet another, very efficient way towards a better signal-to-noise quality of the results. We present it in the next subsection.

5.3.1 Noise in the space of sources

The noise influence can also be observed in the space of sources. For the sake of simplicity, suppose the square mixing matrix **H** of full rank and define

$$\begin{aligned} \overline{\mathbf{s}}(n) &= \mathbf{H}^{-1} \overline{\mathbf{x}}(n) \\ \overline{\mathbf{s}}_{\overline{\boldsymbol{\omega}}}(n) &= \mathbf{H}^{-1} (\overline{\mathbf{x}}(n) + \overline{\boldsymbol{\omega}}(n)). \end{aligned} \tag{5.31}$$

Taking into account the standard norm inequality $\left\|\overline{\mathbf{x}}(n)\right\| \le \left\|\mathbf{H}\right\| \left\|\overline{\mathbf{s}}(n)\right\|$, the influence of noise in the space of sources can be estimated as [68, 115]:

$$\frac{\left\|\overline{\mathbf{s}}(n) - \overline{\mathbf{s}}_{\overline{\boldsymbol{\omega}}}(n)\right\|}{\left\|\overline{\mathbf{s}}(n)\right\|} = \frac{\left\|\mathbf{H}^{-1}\overline{\boldsymbol{\omega}}(n)\right\|}{\left\|\overline{\mathbf{s}}(n)\right\|} \le \frac{\left\|\mathbf{H}^{-1}\overline{\boldsymbol{\omega}}(n)\right\| \left\|\mathbf{H}\right\|}{\left\|\overline{\mathbf{x}}(t)\right\|} \le \frac{\left\|\mathbf{H}\right\| \left\|\mathbf{H}^{-1}\right\| \left\|\overline{\boldsymbol{\omega}}(n)\right\|}{\left\|\overline{\mathbf{x}}(n)\right\|} = \kappa \frac{\left\|\overline{\boldsymbol{\omega}}(n)\right\|}{\left\|\overline{\mathbf{x}}(n)\right\|} \tag{5.32}$$

where $\kappa = \left\|\mathbf{H}\right\| \left\|\mathbf{H}^{-1}\right\|$ stands for the condition number of matrix **H**. On the other hand, the inequality $\left\|\overline{\mathbf{s}}(n)\right\| \le \left\|\mathbf{H}^{-1}\right\| \left\|\overline{\mathbf{x}}(n)\right\|$ implies

$$\frac{\left\|\overline{\mathbf{s}}(n) - \overline{\mathbf{s}}_{\overline{\boldsymbol{\omega}}}(n)\right\|}{\left\|\overline{\mathbf{s}}(n)\right\|} \ge \frac{\left\|\overline{\boldsymbol{\omega}}(n)\right\|}{\left\|\mathbf{H}\right\| \left\|\overline{\mathbf{s}}(n)\right\|} \ge \frac{\left\|\overline{\boldsymbol{\omega}}(n)\right\|}{\left\|\mathbf{H}\right\| \left\|\mathbf{H}^{-1}\right\| \left\|\overline{\mathbf{x}}(n)\right\|} = \frac{1}{\kappa} \frac{\left\|\overline{\boldsymbol{\omega}}(n)\right\|}{\left\|\overline{\mathbf{x}}(n)\right\|}, \tag{5.33}$$

hence

$$\frac{1}{\kappa}\frac{\left\|\overline{\boldsymbol{\omega}}(n)\right\|}{\left\|\bar{\mathbf{x}}(n)\right\|} \le \frac{\left\|\bar{\mathbf{s}}(n)-\bar{\mathbf{s}}_{\overline{\boldsymbol{\omega}}}(n)\right\|}{\left\|\bar{\mathbf{s}}(n)\right\|} \le \kappa\frac{\left\|\overline{\boldsymbol{\omega}}(n)\right\|}{\left\|\bar{\mathbf{x}}(n)\right\|}. \tag{5.34}$$

According to (5.34), the influence of noise depends strongly on the condition number $\kappa = \left\|\mathbf{H}\right\|\left\|\mathbf{H}^{-1}\right\|$. If **H** is well-conditioned a small noise in the space of measurements can not produce large errors in the source space, but if **H** is ill-conditioned even a small noise can result in large corruption of the sources' space.

As there are many different norms there are also many different matrix condition numbers. We will focus only on the condition number derived from the second norm. It is defined as the ratio between the largest and the smallest singular value of matrix **H**. The singular values of matrix **H** indicate how much distortion can occur under the transformation by **H**. The largest singular value of matrix **H** reveals the upper limit to which the unit vector can be amplified by matrix **H**, while the smallest singular value defines its lower limit. The actual degree of distortion depends also on the orientation of the vector entering the transformation with respect to the right singular vectors of matrix **H**. Suppose the noise vector $\overline{\boldsymbol{\omega}}(n)$ aligned with the right singular vector which corresponds to the largest singular value, and the source vector $\bar{\mathbf{s}}(n)$ aligned with the right singular vector which corresponds to the smallest singular value. Then the relative impact of noise will be amplified by factor κ.

Denote by **D** the diagonal matrix of singular values of matrix **H**, and by **U** and **V** the matrices comprising the left and right singular vectors of matrix **H**, respectively

$$\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^T. \tag{5.35}$$

U and **V** are by definition unitary:

$$\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}, \tag{5.36}$$

$$\mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}. \tag{5.37}$$

For the sake of simplicity, suppose the correlation matrix of source $\mathbf{R}_{\bar{\mathbf{s}}}$ equal to the identity matrix and express the correlation matrix of measurements $\mathbf{R}_{\bar{\mathbf{x}}}$ as:

$$\mathbf{R}_{\bar{\mathbf{x}}} = \mathbf{H}\mathbf{R}_{\bar{\mathbf{s}}}\mathbf{H}^T = \mathbf{H}\mathbf{H}^T = \mathbf{U}\mathbf{D}\mathbf{V}^T\mathbf{V}\mathbf{D}\mathbf{U}^T = \mathbf{U}\mathbf{D}^2\mathbf{U}^T. \tag{5.38}$$

Because the matrix $\mathbf{R}_{\bar{\mathbf{x}}}$ is symmetric and positive semidefinite we can always uniquely factorize it to

$$\mathbf{R}_{\bar{\mathbf{x}}} = \mathbf{U}_L \mathbf{\Lambda} \mathbf{U}_L{}^T \tag{5.39}$$

where $\mathbf{U}_L$ denotes the unitary matrix of eigenvectors, and $\mathbf{\Lambda}$ the diagonal matrix with the eigenvalues of matrix $\mathbf{R}_{\bar{\mathbf{x}}}$. The uniqueness of the eigendecomposition [115] guarantees the following relations:

$$\mathbf{U}_L = \mathbf{U} \tag{5.40}$$

and

$$\mathbf{\Lambda} = \mathbf{D}^2 . \tag{5.41}$$

According to (5.41), altering the eigenvalues of the matrix $\mathbf{R}_{\bar{\mathbf{x}}}$ changes the singular values of the mixing matrix **H**. Suppose now the eigenvalues in $\mathbf{\Lambda}$ sorted in a descending order (from the largest to the smallest eigenvalue) and set the *k* smallest eigenvalues to zero:

$$\hat{\mathbf{\Lambda}} = \mathbf{\Lambda} \cdot \begin{bmatrix} \mathbf{I}_{KM-k,KM-k} & \mathbf{0}_{KM-k,k} \\ \mathbf{0}_{k,KM-k} & \mathbf{0}_{k,k} \end{bmatrix} = \mathbf{D}\mathbf{J}_k\mathbf{D} \tag{5.42}$$

where $\mathbf{J}_k = \begin{bmatrix} \mathbf{I}_{KM-k,KM-k} & \mathbf{0}_{KM-k,k} \\ \mathbf{0}_{k,KM-k} & \mathbf{0}_{k,k} \end{bmatrix}$ and $\mathbf{I}_{KM-k,KM-k}$ stands for the identity matrix of size $(KM-k)\times(KM-k)$, while $\mathbf{0}_{KM-k,k}$ is a matrix with all elements equal to zero. Construct the new correlation matrix $\hat{\mathbf{R}}_{\bar{\mathbf{x}}}$:

$$\hat{\mathbf{R}}_{\bar{\mathbf{x}}} = \mathbf{U}_L \hat{\mathbf{\Lambda}} \mathbf{U}_L{}^T \tag{5.43}$$

and rewrite (5.4) to

$$\begin{aligned} \bar{\mathbf{x}}(n)^T \hat{\mathbf{R}}_{\bar{\mathbf{x}}}{}^{\#} \bar{\mathbf{x}}(n) &= \bar{\mathbf{s}}(n)^T \mathbf{H}^T (\mathbf{U}\mathbf{D}\mathbf{J}_k\mathbf{D}\mathbf{U}^T)^{-1}\mathbf{H}\bar{\mathbf{s}}(n) = \\ &= \bar{\mathbf{s}}(n)^T \mathbf{V}\mathbf{D}\mathbf{U}^T\mathbf{U}\mathbf{D}^{-1}\mathbf{J}_k\mathbf{D}^{-1}\mathbf{U}^{-T}\mathbf{U}\mathbf{D}\mathbf{V}^T\bar{\mathbf{s}}(n) = \bar{\mathbf{s}}(n)^T \mathbf{T}\bar{\mathbf{s}}(n). \end{aligned} \tag{5.44}$$

$\mathbf{T} = \mathbf{V}\mathbf{J}_k\mathbf{V}^T$ is orthogonal projection matrix, which converges to identity matrix as *k* converges to zero [68, 115]:

$$\lim_{k\to 0} \mathbf{V}\mathbf{J}_k\mathbf{V}^T = \mathbf{I}. \tag{5.45}$$

Its influence will be assessed in the next section where the decomposition of more measurements than sources will be studied. For now it suffice to say that cutting the

eigenvalues of the matrix $\mathbf{R}_{\bar{\mathbf{x}}}$ introduces the effect similar to nonorthogonal sources, but significantly improves the condition number of the mixing matrix **H**. The optimal degree of cutting (the value of parameter *k*) depends on the original condition number of the mixing matrix **H** and on the signal-to-noise ratio: the stronger the noise, the higher the optimal degree of cutting.

5.4 Decomposition of under-determined systems

In section 5.1, where the IC method was derived, we assumed the number of sources exceeds the number of measurements. But the number of sources is usually hard to estimate. Supposing the number of sources unknown we cannot correctly estimate the extension factor *K* and we risk the underdetermined system. On the other hand, the number of sources can always exceed the number of measurements.

Suppose the number of sources *N* greater than the number of measurements *M* (*M*<*N*), or equally, the extension factor *K* too small to satisfy the relation $KM > N(L+K-1)$. The $KM \times N(L+K-1)$ mixing matrix **H** is now rectangular and has more columns than rows. Without loss of generality, we can suppose the mixing matrix **H** of full row rank:

$$rang(\mathbf{H}) = KM \,. \tag{5.46}$$

Its decomposition to singular values yields:

$$\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^T \tag{5.47}$$

where **U** and **V** denote unitary matrices of size $KM \times KM$ and $N(L+K-1) \times N(L+K-1)$, respectively. **D** is diagonal matrix with the singular values of matrix **H** on its diagonal.

Supposing the sources are orthogonal and matrix $\mathbf{R}_{\bar{\mathbf{s}}}$ equals the identity matrix, (5.4) and (5.5) transform into

$$\begin{aligned} Ind(n) &= \bar{\mathbf{y}}(n)^T \mathbf{R}_{\bar{\mathbf{x}}}^{\#} \bar{\mathbf{y}}(n) = \\ &= \bar{\mathbf{s}}(n)^T \mathbf{H}^T (\mathbf{H}\mathbf{H}^T)^{\#} \mathbf{H}\bar{\mathbf{s}}(n) + o_{\bar{\boldsymbol{\omega}}}(n,n) = \\ &= \bar{\mathbf{s}}(n)^T \mathbf{V}\mathbf{D}^T (\mathbf{D}\mathbf{D}^T)^{\#} \mathbf{D}\mathbf{V}^T \bar{\mathbf{s}}(n) + o_{\bar{\boldsymbol{\omega}}}(n,n) = \bar{\mathbf{s}}(n)^T \mathbf{T}\bar{\mathbf{s}}(n) + o_{\bar{\boldsymbol{\omega}}}(n,n) \end{aligned} \tag{5.48}$$

and

$$v_{n_0,j}(n) = \bar{\mathbf{y}}(n_0)^T \mathbf{R}_{\bar{\mathbf{x}}}^{\#} \bar{\mathbf{y}}(n_0) = \bar{\mathbf{s}}(n)^T \mathbf{T}\bar{\mathbf{s}}(n) + o_{\bar{\boldsymbol{\omega}}}(n_0,n)\,, \tag{5.49}$$

where $\mathbf{T} = \mathbf{V}\mathbf{D}^T(\mathbf{D}\mathbf{D}^T)^{\#}\mathbf{D}\mathbf{V}^T = \mathbf{V}^T\begin{bmatrix} \mathbf{I}_{KM\times KM} & \mathbf{O}_{KM\times N(L+K-1)-KM} \\ \mathbf{O}_{N(L+K-1)-KM\times KM} & \mathbf{O}_{N(L+K-1)-KM\times N(L+K-1)-KM} \end{bmatrix}\mathbf{V}^T$ is orthogonal projection matrix which projects the extended source vectors into the space of measurements. Its properties are described in Appendix 5D.

What can now be said about the activity index? Because the matrix **T** is positive semidefinite (Appendix 5D), the activity index is still non-negative in every possible time moment. But the interpretation of its value turns out to be tricky. In previous sections, we saw that having more measurements than sources the value of the activity index increases with the number of simultaneously active sources. This may not necessarily be true any more in the case of under-determined MIMO system. Let there be two sources s_i and s_j active in the time moment n_0. The value of the activity index in the time moment n_0 yields:

$$Ind(n_0) = \bar{s}_i(n_0)\bar{s}_i(n_0)t_{ii} + \bar{s}_j(n_0)\bar{s}_j(n_0)t_{jj} + 2\bar{s}_i(n_0)\bar{s}_j(n_0)t_{ij} \tag{5.50}$$

where t_{ij} denotes the (i,j)-th element of the matrix **T**. Having three sources – s_i, s_j and s_k – active in time moment n_1, the value of the activity index yields:

$$\begin{aligned} Ind(n_1) = {} & \bar{s}_i(n_1)\bar{s}_i(n_1)t_{ii} + \bar{s}_j(n_1)\bar{s}_j(n_1)t_{jj} + \bar{s}_k(n_1)\bar{s}_k(n_1)t_{kk} + \\ & + 2\bar{s}_i(n_1)\bar{s}_j(n_1)t_{ij} + 2\bar{s}_i(n_1)\bar{s}_k(n_1)t_{ik} + 2\bar{s}_j(n_1)\bar{s}_k(n_1)t_{jk}. \end{aligned} \tag{5.51}$$

Supposing now the pulses of all the sources equal to 1 and $t_{ii} > t_{jj} > t_{kk}$, we can always assume $t_{ik}/2 < -t_{kk}$ and $t_{jk} < 0$. Comparing (5.50) and (5.51),

$$t_{ii} + t_{jj} + t_{kk} + 2t_{ij} + 2t_{ik} + 2t_{jk} < t_{ii} + t_{jj} + 2t_{ij} - 2\left|t_{jk}\right| < t_{ii} + t_{jj} + 2t_{ij}\,, \tag{5.52}$$

we see how the value of the activity index can decrease, although the number of active sources increases. Although perfectly possible, this is not a very likely case, because many different conditions have to be met. The index activity can, therefore, still be used as an indicator of global activity of sources.

The experimental results and the theory in Appendix 5D show the matrix **T** will have at least a part of its diagonal elements superior to all off-diagonal elements, as long as $KM<N(L+K-1)<2KM$. The sources corresponding to the dominant diagonal elements of the matrix **T** still obey the theory from the over-determined case:

$$v_{G_{n_0},i}(n) = \bar{\mathbf{s}}(n_0)^T\mathbf{T}\bar{\mathbf{s}}(n) = \sum_{i\in G_{n_0}}\sum_{j=1}^{N(L+K)} t_{ij}\bar{s}_i(n_0)\bar{s}_j(n) \approx \sum_{i\in G_{n_0}} t_{ii}\bar{s}_i(n)\,. \tag{5.53}$$

The procedures derived in Subsections 5.2 and 5.3 can, hence, readily be applied also in the case of under-determined MIMO systems. The only difference between the over-

determined and the under-determined case turns up to be the degree to which the mixing matrix **H** is cancelled. While **H** is completely lost in the over-determined case, it is still partially present (in the form of matrix **T**) in the under-determined case. When **H** is close to the square matrix, **T** is close to the identity matrix. Rising the ratio between the number of sources and the number of measurements, **T** is gradually loosing its diagonal form and the impact of the mixing matrix increases (Appendix 5D). Nevertheless, keeping the ratio $KM : N(L+K-1)$ below 2, at least a part of all sources will be completely reconstructed. This conclusion will be further confirmed by comprehensive set of experiments which are described in the next chapter.

Appendix 5A: Inverse of diagonally dominant matrix

Suppose the source correlation matrix $\mathbf{R}_{\bar{\mathbf{s}}}$ of size $N \times N$ diagonally dominant, i.e. the matrix in which the magnitude of each diagonal entry exceeds the sum of the magnitudes of the off-diagonal entries in the corresponding row:

$$r_{i,i} > \sum_{j,\, j \neq i} \left| r_{i,j} \right| \tag{5A.1}$$

where $r_{i,j}$ denotes the *(i,j)*-th element of matrix $\mathbf{R}_{\bar{\mathbf{s}}}$. Supposing the sources close-to-orthogonal we also have

$$\forall i \forall j,\, i \neq j : \frac{\left| r_{i,j} \right|}{r_{i,i}} < \varepsilon; \;\; \varepsilon << 1 \tag{5A.2}$$

where ε is usually much smaller than 0.1.

The matrix $\mathbf{R}_{\bar{\mathbf{s}}}$ can always be written as a sum of the diagonal and non-diagonal matrices:

$$\mathbf{R}_{\bar{\mathbf{s}}} = \mathbf{D}_{\bar{\mathbf{s}}} + \mathbf{N}_{\bar{\mathbf{s}}}, \tag{5A.3}$$

where

$$\mathbf{D}_{\bar{\mathbf{s}}} = diag(\mathbf{R}_{\bar{\mathbf{s}}}) = \begin{bmatrix} r_{1,1} & 0 & \cdots & 0 \\ 0 & r_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & r_{N,N} \end{bmatrix} \tag{5A.4}$$

denotes the diagonal, and

$$\mathbf{N}_{\bar{\mathbf{s}}} = \mathbf{R}_{\bar{\mathbf{s}}} - diag(\mathbf{R}_{\bar{\mathbf{s}}}) = \begin{bmatrix} 0 & r_{1,2} & \cdots & r_{1,N} \\ r_{2,1} & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_{N-1,N} \\ r_{N,1} & \cdots & r_{N,N-1} & 0 \end{bmatrix} \tag{5A.5}$$

non-diagonal part of the matrix $\mathbf{R}_{\bar{\mathbf{s}}}$. Using the Neumann series, the inverse of $\mathbf{R}_{\bar{\mathbf{s}}}$ can be expressed as [115]:

$$\mathbf{R}_{\bar{\mathbf{s}}}^{-1} = (\mathbf{D}_{\bar{\mathbf{s}}} + \mathbf{N}_{\bar{\mathbf{s}}})^{-1} = (\mathbf{D}_{\bar{\mathbf{s}}}(\mathbf{I} + \mathbf{D}_{\bar{\mathbf{s}}}^{-1}\mathbf{N}_{\bar{\mathbf{s}}}))^{-1} = (\mathbf{I} + \mathbf{D}_{\bar{\mathbf{s}}}^{-1}\mathbf{N}_{\bar{\mathbf{s}}})^{-1}\mathbf{D}_{\bar{\mathbf{s}}}^{-1} = \mathbf{D}_{\bar{\mathbf{s}}}^{-1} + (\sum_{k=1}^{\infty}[-\mathbf{D}_{\bar{\mathbf{s}}}^{-1}\mathbf{N}_{\bar{\mathbf{s}}}]^k)\mathbf{D}_{\bar{\mathbf{s}}}^{-1}, \tag{5A.6}$$

providing the absolute value of each eigenvalue of the matrix

$$\mathbf{D}_{\bar{\mathbf{s}}}^{-1}\mathbf{N}_{\bar{\mathbf{s}}} = \begin{bmatrix} 0 & \frac{r_{1,2}}{r_{1,1}} & \cdots & \frac{r_{1,N}}{r_{1,1}} \\ \frac{r_{2,1}}{r_{2,2}} & 0 & \cdots & \frac{r_{2,N}}{r_{2,2}} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{r_{N,1}}{r_{N,N}} & \cdots & \frac{r_{N,N-1}}{r_{N,N}} & 0 \end{bmatrix} \tag{5A.7}$$

is smaller than 1. According to the Gerschgorin's theorem, the eigenvalues of $\mathbf{D}_{\bar{\mathbf{s}}}^{-1}\mathbf{N}_{\bar{\mathbf{s}}}$ are contained in the union of the Gerschgorin circles defined by [115]:

$$\forall k \;\; \exists i\colon \left|\lambda_k - 0\right| \le \sum_{\substack{j=1 \\ j\neq i}}^{N}\left|\frac{r_{i,j}}{r_{i,i}}\right| = \frac{1}{\left|r_{i,i}\right|}\sum_{\substack{j=1 \\ j\neq i}}^{N}\left|r_{i,j}\right| < 1, \tag{5A.8}$$

where λ_k stands for k-th eigenvalue of the matrix $\mathbf{D}_{\bar{\mathbf{s}}}^{-1}\mathbf{N}_{\bar{\mathbf{s}}}$ and the relationship from (5A.1) was used. Using (5A.8), it is easy to show

$$\lim_{k\to\infty}(\mathbf{D}_{\bar{\mathbf{s}}}^{-1}\mathbf{N}_{\bar{\mathbf{s}}})^k = \mathbf{0}. \tag{5A.9}$$

Thus, the first-order approximation of $\mathbf{R}_{\bar{\mathbf{s}}}^{-1}$ yields:

$$\mathbf{R}_{\bar{\mathbf{s}}}^{-1} \approx \mathbf{D}_{\bar{\mathbf{s}}}^{-1} - \mathbf{D}_{\bar{\mathbf{s}}}^{-1}\mathbf{N}_{\bar{\mathbf{s}}}\mathbf{D}_{\bar{\mathbf{s}}}^{-1} = \begin{bmatrix} \frac{1}{r_{1,1}} & -\frac{r_{1,2}}{r_{1,1}r_{2,2}} & \cdots & -\frac{r_{1,N}}{r_{1,1}r_{N,N}} \\ -\frac{r_{2,1}}{r_{2,2}r_{1,1}} & \frac{1}{r_{2,2}} & \cdots & -\frac{r_{2,N}}{r_{2,2}r_{N,N}} \\ \vdots & \ddots & \ddots & \vdots \\ -\frac{r_{N,1}}{r_{N,N}r_{1,1}} & \cdots & -\frac{r_{N,N-1}}{r_{N,N}r_{N-1,N-1}} & \frac{1}{r_{N,N}} \end{bmatrix}. \tag{5A.10}$$

Using inequality (5A.2), we obtain

$$\mathbf{R}_{\bar{\mathbf{s}}}^{-1} \approx \mathbf{D}_{\bar{\mathbf{s}}}^{-1} - \mathbf{D}_{\bar{\mathbf{s}}}^{-1}\mathbf{N}_{\bar{\mathbf{s}}}\mathbf{D}_{\bar{\mathbf{s}}}^{-1} \leq \begin{bmatrix} \frac{1}{r_{1,1}} & -\frac{\varepsilon}{r_{1,1}} & \cdots & -\frac{\varepsilon}{r_{1,1}} \\ -\frac{\varepsilon}{r_{2,2}} & \frac{1}{r_{2,2}} & \cdots & -\frac{\varepsilon}{r_{2,2}} \\ \vdots & \ddots & \ddots & \vdots \\ -\frac{\varepsilon}{r_{N,N}} & \cdots & -\frac{\varepsilon}{r_{N,N}} & \frac{1}{r_{N,N}} \end{bmatrix} \tag{5A.11}$$

and

$$\mathbf{R}_{\bar{\mathbf{s}}}^{-1} \approx \mathbf{D}_{\bar{\mathbf{s}}}^{-1} - \mathbf{D}_{\bar{\mathbf{s}}}^{-1}\mathbf{N}_{\bar{\mathbf{s}}}\mathbf{D}_{\bar{\mathbf{s}}}^{-1} \leq \begin{bmatrix} \frac{1}{r_{1,1}} & -\frac{\varepsilon}{r_{2,2}} & \cdots & -\frac{\varepsilon}{r_{N,N}} \\ -\frac{\varepsilon}{r_{1,1}} & \frac{1}{r_{2,2}} & \cdots & -\frac{\varepsilon}{r_{N,N}} \\ \vdots & \ddots & \ddots & \vdots \\ -\frac{\varepsilon}{r_{1,1}} & \cdots & -\frac{\varepsilon}{r_{N-1,N-1}} & \frac{1}{r_{N,N}} \end{bmatrix}, \tag{5A.12}$$

which, when the inequality $\varepsilon << 1$ is taken into account, proves the inverse of diagonally dominant matrix $\mathbf{R}_{\bar{\mathbf{s}}}$ has a superior diagonal.

According to comprehensive numerical simulations, similar conclusions apply also to the matrices which are not strictly diagonally dominant (in the sense of (5A.1)) but still fulfil the condition in (5A.2). Moreover, in each simulated matrix the ratio between the diagonal and off-diagonal elements was yet improved by the matrix inverse, pushing the superiority of the diagonal elements even higher. The exact proof of the observed phenomenon reaches beyond the scope of this dissertation.

Appendix 5B: Probability of overlapped pulses

Let the pulse sources be organized into R subsets, each comprising Q sources, and assume the sources from the same subset cannot overlap with each other. Define the overlap of the i-th source with the m-th subset of sources as a set of all possible overlaps of the i-th source with any source from the m-th subset. Further assume that overlaps of all possible sources with all possible subsets are independent, equally probable random events and denote their probability by p.

For the sake of simplicity, also assume that any source from a given subset is equally probable to overlap with any source from any other subset:

$$\forall j \forall k; j \neq k : p_{j,k} = \frac{1}{Q} p \tag{5B.1}$$

where $p_{j,k}$ denotes the probability that a single pulse of the j-th source overlapped with any pulse of the k-th source, whereas we assume the j-th and k-th source do not belong to the same subset.

Let n_0 be the time moment in which at least one source was active and denote the active source by index j. The probability of having g_0 ($g_0 \leq R$) sources active in given time moment n_0 equals the probability that the j-th source overlap with g_0-1 out of R-1 possible subsets of sources:

$$P(card(G_{n_0}) = g_0 \mid card(G_{n_0}) > 0) = \binom{R-1}{g_0-1} p^{g_0-1} (1-p)^{R-g_0} \tag{5B.2}$$

where the set G_{n_0} comprises the sources that were active in time moment n_0, $card(G_{n_0})$ is its cardinal number, while $\binom{m}{r}$ denotes the number of combinations of m elements taken r elements at a time.

The number of all possible combinations of g_0 (out of RQ) simultaneously active sources is

$$C_Q^R(g_0) = \binom{R}{g_0} Q^{g_0}, \tag{5B.3}$$

and the probability of having g_0 selected sources active in given time moment n_0 yields:

$$\frac{1}{C_Q^R(g_0)} P(card(G_{n_0}) = g_0 \mid card(G_{n_0}) > 0). \tag{5B.4}$$

Finally, excluding g_1 selected sources from the set of all possible sources, the number of all possible combinations of g_0 overlapped sources reads:

$$C_Q^R(g_0,g_1)=\sum_{i=0}^{\min(g_0,g_1)}\binom{R-g_1}{g_0-i}\binom{g_1}{i}Q^{g_0-i}(Q-1)^i. \tag{5B.5}$$

Observe now two time moments, n_0 and n_1, with the pulses from at least one (not necessarily the same) source. Denote by g_0 the number of sources that were active in moment n_0, and by g_1 the number of sources that were active in moment n_1. Using the denotations

$$\begin{aligned} p_{G_{n_0}}(g_0)&=P\left(card(G_{n_0})=g_0 \,\middle|\, card(G_{n_0})>0\right) \text{ and}\\ p_{G_{n_1}}(g_1)&=P\left(card(G_{n_1})=g_1 \,\middle|\, card(G_{n_1})>0\right)\end{aligned} \tag{5B.6}$$

the probability of having g_{01} common sources active in both time moments can be obtained as:

$$\begin{aligned} p_{G_{n_0}\cap G_{n_1}}(g_{01})&=P\left(card(G_{n_0}\cap G_{n_1})=g_{01} \,\middle|\, card(G_{n_0})>0, card(G_{n_1})>0\right)=\\ &=\sum_{g_0=g_{01}}^{R}\sum_{g_1=g_{01}}^{R}\binom{g_0}{g_{01}}\frac{C_Q^{R-g_{01}}(g_1-g_{01},g_0-g_{01})}{C_Q^R(g_1)}p_{G_{n_0}}(g_0)p_{G_{n_1}}(g_1)\end{aligned} \tag{5B.7}$$

where the fact that at least g_{01} sources had to fire in both time moments n_0 and n_1: $g_0\ge g_{01}$ and $g_1\ge g_{01}$ is taken into consideration. Using (5B.5), we obtain:

$$\begin{aligned} p_{G_{n_0}\cap G_{n_1}}(g_{01})&=\sum_{g_0=g_{01}}^{R}\sum_{g_1=g_{01}}^{R}p_{G_{n_0}}(g_0)p_{G_{n_1}}(g_1)\binom{g_0}{g_{01}}\frac{C_Q^{R-g_{01}}(g_1,g_{01})}{C_Q^R(g_1)}=\\ &\sum_{g_0=g_{01}}^{R}\sum_{g_1=g_{01}}^{R}p_{G_{n_0}}(g_0)p_{G_{n_1}}(g_1)\frac{g_0!(R-g_0)!(R-g_1)!g_1!}{g_{01}!R!}\cdot\\ &\left[\sum_{i=0}^{\min(g_0-g_{01},\,g_1-g_{01})}\frac{1}{(R-g_0-g_1+g_{01}+i)!(g_0-g_{01}-i)!(g_1-g_{01}-i)!i!}\cdot\frac{Q^{g_1-g_{01}-i}(Q-1)^i}{Q^{g_1}}\right].\end{aligned} \tag{5B.8}$$

Owing to

$$\begin{aligned}&\sum_{i=0}^{\min(g_0-g_{01},\,g_1-g_{01})}\frac{1}{(R-g_0-g_1+g_{01}+i)!(g_0-g_{01}-i)!(g_1-g_{01}-i)!i!}=\\ &\qquad\qquad\frac{(R-g_{01})!}{(R-g_0)!(R-g_1)!(g_0-g_{01})!(g_1-g_{01})!},\end{aligned} \tag{5B.9}$$

(5B.8) yileds:

$$p_{G_{n_0}\cap G_{n_1}}(g_{01}) \le \sum_{g_0=g_{01}}^{R}\sum_{g_1=g_{01}}^{R} p_{G_{n_0}}(g_0)p_{G_{n_1}}(g_1)\frac{g_0!(R-g_0)!(R-g_1)!g_1!}{g_{01}!R!}\cdot \left[\frac{(R-g_{01})!}{(R-g_0)!(R-g_1)!(g_0-g_{01})!(g_1-g_{01})!}\cdot\frac{1}{Q^{g_{01}}}\right]= \tag{5B.10}$$

$$\frac{(R-1)!}{g_{01}!R(R-g_{01})!Q^{g_{01}}}\sum_{g_0=g_{01}}^{R}\sum_{g_1=g_{01}}^{R} g_0g_1\frac{(R-g_{01})!}{(R-g_0)!(g_0-g_{01})!}\cdot\frac{(R-g_{01})!}{(R-g_1)!(g_1-g_{01})!}\cdot$$

$$p^{g_0-1}(1-p)^{R-g_0}\cdot p^{g_1-1}(1-p)^{R-g_1}.$$

Because $g_0g_1 \le R^2$, the probability $p_{G_{n_0}\cap G_{n_1}}(g_{01})$ can be estimated as follows:

$$p_{G_{n_0}\cap G_{n_1}}(g_{01}) \le \frac{(R-1)!R^2}{g_{01}!R(R-g_{01})!Q^{g_{01}}}\sum_{g_0=g_{01}}^{R}\binom{R-g_{01}}{g_0-g_{01}}p^{g_0-1}(1-p)^{R-g_0}\cdot \tag{5B.11}$$

$$\sum_{g_1=g_{01}}^{R}\binom{R-g_{01}}{g_1-g_{01}}p^{g_1-1}(1-p)^{R-g_1}.$$

Finally, using

$$\sum_{g_i=g_{01}}^{R}\binom{R-g_{01}}{g_i-g_{01}}p^{g_i-1}(1-p)^{R-g_i} = p^{g_{01}-1}\cdot\sum_{m=0}^{R-g_{01}}\binom{R-g_{01}}{m}p^m(1-p)^{R-g_{01}-m} = p^{g_{01}-1}, \tag{5B.12}$$

we get

$$p_{G_{n_0}\cap G_{n_1}}(g_{01}) \le \frac{R!p^{2(g_0-1)}}{g_{01}!(R-g_{01})!Q^{g_{01}}} = \binom{R}{g_{01}}\frac{p^{2(g_0-1)}}{Q^{g_{01}}}. \tag{5B.13}$$

The estimation (5B.13) represents the upper bound which is never attained and is accurate only for large *Q*. Nevertheless we can see that the probability of having g_{01} common sources in two time moments drops rapidly with the number of common sources, g_{01}, and also with the cardinal number of subsets, *Q* (Figs. 5B.1, 5B.2 and 5B.3).

In the case of EMG signals, the number of subsets corresponds to the number of active MUs, while *Q* represents the number of the delayed repetitions of each innervation pulse train. Taking into account that *Q* is usually greater than 20, while the probability *p* hardly exceeds 0.5, we can conclude the probability of having two common sources active in two different time moments is negligibly small. This fact can readily be seen also from Fig. 5B.2.

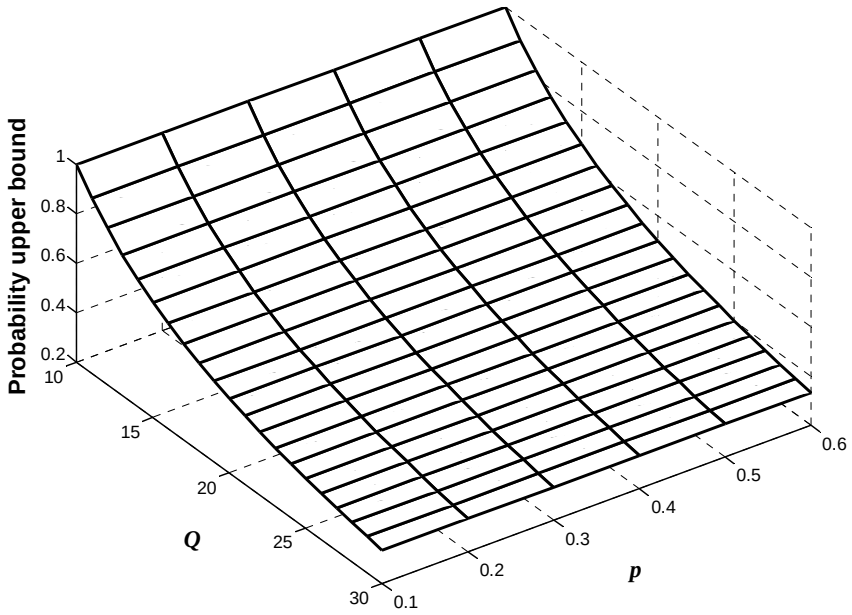


Figure 5B.1: The upper bound (5B.13) of probability $p_{G_{n_0} \cap G_{n_1}}(g_{01})$ at $g_{01}=1$ and $R=10$ in dependence of the number of sources in each subset, Q, and the probability of overlapped subsets, p.

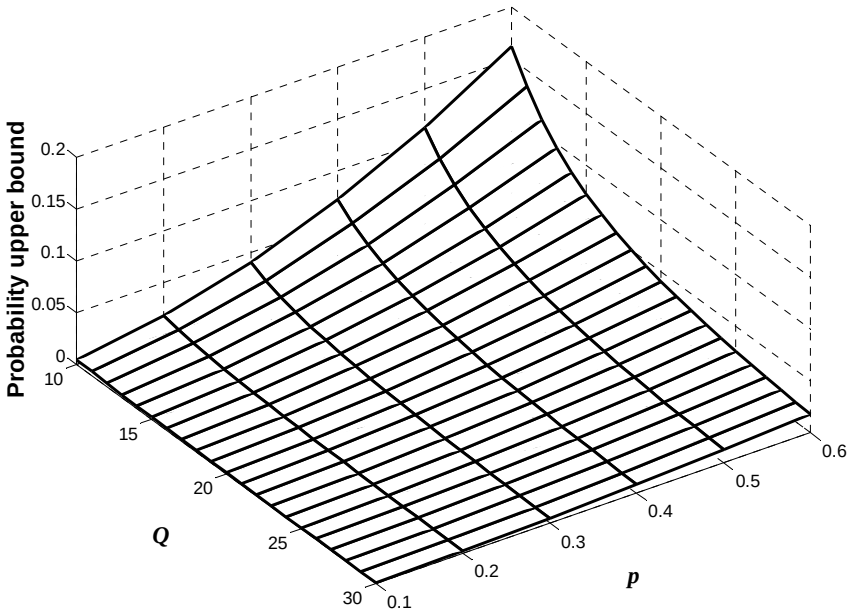


Figure 5B.2: The upper bound (5B.13) of probability $p_{G_{n_0} \cap G_{n_1}}(g_{01})$ at $g_{01}=2$ and $R=10$ in dependence of the number of sources in each subset, Q, and the probability of overlapped subsets, p.

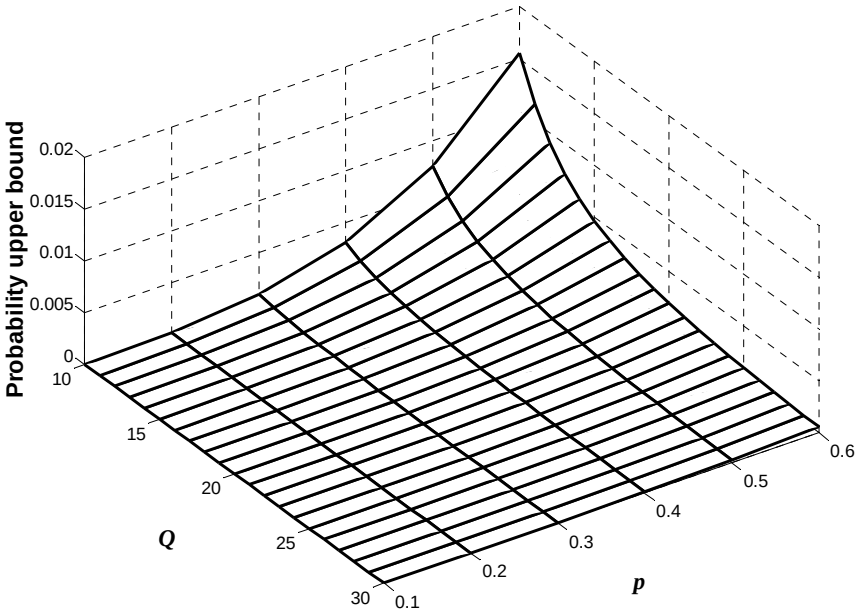


Figure 5B.3: The upper bound (5B.13) of probability $p_{G_{n_0} \cap G_{n_1}}(g_{01})$ at $g_{01}=3$ and $R=10$ in dependence of the number of sources in each subset, Q, and the probability of overlapped subsets, p.

The probability of having g_{01} common sources active in three or more time moments are much smaller than $p_{G_{n_0} \cap G_{n_1}}(g_{01})$ and can thus be neglected – at least when Q (p) is big (small) enough.

Appendix 5C: Average number of pulses in the product of the superimposed pulse trains of sources

In this appendix we use the assumptions and conclusions from Appendix 5B to estimate the average number of pulses in a product of the superimposed pulse trains of close-to-orthogonal sources.

Suppose four time moments, n_0, n_1, n_2 and n_3, with at least one (not necessarily the same) active source and use (5.5) to reconstruct the corresponding superimpositions of sources' pulse trains:

$$v_{n_i,G_{n_i}}(n) \approx \sum_{j \in G_{n_i}} r_{j,j}^{inv} \bar{s}_j(n),\ i=0,1,2,3. \tag{5C.1}$$

Denote by $G_{n_i} = \{j_{n_i,1},..., j_{n_i,g_i}\}$ the set of sources that were active in time moment n_i, $i=0,1,2,3$. According to Appendix 5B, the probability that the set G_{n_i} differ from other sets is high.

For the sake of simplicity, assume the number of pulses in each source equals F and observe the average number of pulses in the element-wise product $v_{n_i,G_{n_i}}(n) \cdot v_{n_j,G_{n_j}}(n)$, $i \neq j$. The product $v_{n_i,G_{n_i}}(n) \cdot v_{n_j,G_{n_j}}(n)$ will certainly contain the pulses of all the sources from the intersection $G_{n_i} \cap G_{n_j}$. Considering their mutual overlapping, the average number of pulses that the sources from $G_{n_i} \cap G_{n_j}$ contribute to the product $v_{n_i,G_{n_i}}(n) \cdot v_{n_j,G_{n_j}}(n)$ can be estimated as:

$$F \cdot f(g_{ij},0) \tag{5C.2}$$

where, following the convention from Appendix 5B, g_{ij} is the cardinal number of $G_{n_i} \cap G_{n_j}$,

$$f(n_1,0) = \left(n_1 - \sum_{i=2}^{n_1} (-1)^i \binom{n_1}{i} p_Q^{i-1} \right) \tag{5C.3}$$

denotes the correction factor introduced by the fact that the pulses of the sources in $G_{n_i} \cap G_{n_j}$ overlap, and $p_Q = \frac{p}{Q}$ stands for the probability that the pulses of two arbitrary sources overlap (Appendix 5B).

The product $v_{n_i,G_{n_i}}(n)\cdot v_{n_j,G_{n_j}}(n)$ will also contain all those pulses of any source from $G_{n_i}-G_{n_j}$ which randomly overlap with the pulses of any source from $G_{n_j}-G_{n_i}$, where M_1-M_2 denotes the set difference (M_1 without M_2). Their number can be estimated as:

$$F\cdot p_Q\cdot f(g_i-g_{ij},g_{ij})\cdot f(g_j-g_{ij},g_{ij}) \tag{5C.4}$$

where g_i denotes the cardinal number of G_{n_i}, and

$$f(n_1,n_2)=\left(n_1-\sum_{i=2}^{n_1+n_2}(-1)^i\binom{n_1+n_2}{i}p_Q^{i-1}\right) \tag{5C.5}$$

is a generalised correction factor (5C.3) which accumulates the influence of all possible overlaps among the sources from subsets $G_{n_i}-G_{n_j}$ and $G_{n_j}-G_{n_i}$, respectively, and excludes the influences of sources form $G_{n_i}\cap G_{n_j}$. Taking into account the fact that not all sources can overlap (Appendix 5B), the total average number of pulses in the product $v_{n_i,G_{n_i}}(n)\cdot v_{n_j,G_{n_j}}(n)$ yields:

$$I(p_Q,g_i,g_j,g_{ij})\le F\cdot[g_{ij}\cdot f(g_{ij},0)+p_Q\cdot f(g_i-g_{ij},g_{ij})\cdot f(g_j-g_{ij},g_{ij})\,]. \tag{5C.6}$$

We can follow the same route to estimate the average number of pulses in the product $v_{n_0,G_{n_0}}(n)\cdot v_{n_1,G_{n_1}}(n)\cdot v_{n_2,G_{n_2}}(n)\cdot v_{n_3,G_{n_3}}(n)$. For the clarity reasons, define first the following mutually disjunctive sets:

$$\begin{aligned}
M_{0,1,2,3}&=G_{n_0}\cap G_{n_1}\cap G_{n_2}\cap G_{n_3}\\
M_{i,j,k}&=G_{n_i}\cap G_{n_j}\cap G_{n_k}-M_{0,1,2,3}\\
M_{i,j}&=G_{n_i}\cap G_{n_j}-(\bigcup_k M_{i,j,k}\cup M_{0,1,2,3})\\
M_i&=G_{n_i}-(\bigcup_{j,k}M_{j,k}\cup\bigcup_k M_{i,j,k}\cup M_{0,1,2,3})
\end{aligned} \tag{5C.7}$$

where $i,j,k\in\{0,1,2,3\}$ and $i\ne j$, $i\ne k$, $j\ne k$. By definition, the unions of all the *M* sets equal the union of all the *G* sets:

$$\bigcup_i G_{n_i}=\bigcup_i M_i\cup\bigcup_{i,j}M_{i,j}\cup\bigcup_{i,j,k}M_{i,j,k}\cup M_{0,1,2,3}, \tag{5C.8}$$

and all possible intersections of the M sets are empty. Following the convention from Appendix 5B, denote by $m=card(M)$ the cardinal number of the set M and estimate the average number of pulses in the product $v_{n_0,G_{n_0}}(n)\cdot v_{n_1,G_{n_1}}(n)\cdot v_{n_2,G_{n_2}}(n)\cdot v_{n_3,G_{n_3}}(n)$ as:

$$
\begin{aligned}
&I(p_Q,g_0,g_1,g_2,g_3,m_{01},m_{02},m_{03},m_{12},m_{13},m_{23},m_{012},m_{013},m_{023},m_{123},m_{0123})\le \\
&\quad F\cdot[f(m_{0123},0)+p_Q\cdot f(m_{012},m_{0123})\cdot f(m_3,g_3-m_3)+ \\
&\quad p_Q\cdot f(m_{013},m_{0123})\cdot f(m_2,g_2-m_2)+p_Q\cdot f(m_{023},m_{0123})\cdot f(m_1,g_1-m_1)+ \\
&\quad p_Q\cdot f(m_{123},m_{0123})\cdot f(m_0,g_0-m_0)+ \\
&\quad p_Q\cdot f(m_{01},m_{012}+m_{013}+m_{0123})\cdot f(m_{23},m_{023}+m_{123}+m_{0123})+ \\
&\quad p_Q\cdot f(m_{02},m_{012}+m_{023}+m_{0123})\cdot f(m_{13},m_{013}+m_{123}+m_{0123})+ \\
&\quad p_Q\cdot f(m_{03},m_{013}+m_{023}+m_{0123})\cdot f(m_{12},m_{012}+m_{123}+m_{0123})+ \\
&\quad p_Q^2\cdot f(m_{01},m_{012}+m_{013}+m_{0123})\cdot f(m_2,g_2-m_2)\cdot f(m_3,g_3-m_3)+ \\
&\quad p_Q^2\cdot f(m_{02},m_{012}+m_{023}+m_{0123})\cdot f(m_1,g_1-m_1)\cdot f(m_3,g_3-m_3)+ \\
&\quad p_Q^2\cdot f(m_{03},m_{013}+m_{023}+m_{0123})\cdot f(m_1,g_1-m_1)\cdot f(m_2,g_2-m_2)+ \\
&\quad p_Q^2\cdot f(m_{12},m_{012}+m_{123}+m_{0123})\cdot f(m_0,g_0-m_0)\cdot f(m_3,g_3-m_3)+ \\
&\quad p_Q^2\cdot f(m_{13},m_{013}+m_{123}+m_{0123})\cdot f(m_0,g_0-m_0)\cdot f(m_2,g_2-m_2)+ \\
&\quad p_Q^2\cdot f(m_{23},m_{023}+m_{123}+m_{0123})\cdot f(m_0,g_0-m_0)\cdot f(m_1,g_1-m_1)+ \\
&\quad p_Q^3\cdot f(m_0,g_0-m_0)\cdot f(m_1,g_1-m_1)\cdot f(m_2,g_2-m_2)\cdot f(m_3,g_3-m_3)].
\end{aligned}
\tag{5C.9}
$$

Supposing the probability $p_Q=\frac{p}{Q}$ small (Appendix 5B) and neglecting all the factors which are multiplied by the second or the higher power of p_Q, (5C.9) yields:

$$
\begin{aligned}
&I(p_Q,g_0,g_1,g_2,g_3,m_{01},m_{02},m_{03},m_{12},m_{13},m_{23},m_{012},m_{013},m_{023},m_{123},m_{0123})\approx \\
&\quad F\cdot[f(m_{0123},0)+p_Q\cdot f(m_{012},m_{0123})\cdot f(m_3,g_3-m_3)+ \\
&\quad p_Q\cdot f(m_{013},m_{0123})\cdot f(m_2,g_2-m_2)+p_Q\cdot f(m_{023},m_{0123})\cdot f(m_1,g_1-m_1)+ \\
&\quad p_Q\cdot f(m_{123},m_{0123})\cdot f(m_0,g_0-m_0)+ \\
&\quad p_Q\cdot f(m_{01},m_{012}+m_{013}+m_{0123})\cdot f(m_{23},m_{023}+m_{123}+m_{0123})+ \\
&\quad p_Q\cdot f(m_{02},m_{012}+m_{023}+m_{0123})\cdot f(m_{13},m_{013}+m_{123}+m_{0123})+ \\
&\quad p_Q\cdot f(m_{03},m_{013}+m_{023}+m_{0123})\cdot f(m_{12},m_{012}+m_{123}+m_{0123})].
\end{aligned}
\tag{5C.10}
$$

According to (5C.10), the average number of pulses in the product $v_{n_0,G_{n_0}}(n)\cdot v_{n_1,G_{n_1}}(n)\cdot v_{n_2,G_{n_2}}(n)\cdot v_{n_3,G_{n_3}}(n)$ depends mainly on the number of sources in the

sets G_{n_0}, G_{n_1}, G_{n_2} and G_{n_3} and their division in the corresponding M sets. In the sequel we are going to study just two of many possible cases.

Case 1: Suppose the set $M_{0,1,2}$ contains a single source, while sets $M_{0,1,3}$, $M_{0,2,3}$, $M_{1,2,3}$ and $M_{0,1,2,3}$ are empty: $m_{012}=1$ and $m_{013}=m_{023}=m_{123}=m_{0123}=0$. Further assume m_{01}, m_{02}, m_{03}, m_{12}, m_{13} and m_{23} are all equal to or less than 1. Then (5C.10) yields:

$$I(p_Q) \approx F\, p_Q[\ f(1,0) f(m_3, g_3 - m_3) + 3 f(1,1) f(1,0)]. \tag{5C.11}$$

Recall that $m_3 = g_3 - m_{03} - m_{13} - m_{23} - m_{013} - m_{023} - m_{123} - m_{0123}$ and that on average the number of sources in the set G_{n_3} can be estimated as $g_3 = Rp$. Calculate the values of all functions $f(\cdot,\cdot)$ and neglect the factors which are multiplied by the second or higher power of p_Q to get the final estimate

$$I(p,R,Q,F) \approx Fp^2 \frac{R}{Q}. \tag{5C.12}$$

Case 2: Suppose the set $M_{0,1,2,3}$ contains a single source, while sets $M_{0,1,2}$, $M_{0,1,3}$, $M_{0,2,3}$ and $M_{1,2,3}$ are empty. Further assume (in like fashion to Case 1) that m_{01}, m_{02}, m_{03}, m_{12}, m_{13} and m_{23} are equal to or less than 1. Then (5C.10) simplifies to

$$I(p_Q) \approx F[\,1 \cdot f(1,0) + 3 p_Q f(1,1) f(1,0)]. \tag{5C.13}$$

Neglecting all the factors which are multiplied by the second or higher power of p_Q we get the final estimate

$$I(p,R,Q,F) \approx F(1 + 3\frac{p}{Q}). \tag{5C.14}$$

Appendix 5D: Properties of orthogonal projection matrix

Define the matrix **T** as

$$\mathbf{T} = \mathbf{H}^T (\mathbf{H}\mathbf{H}^T)^{\#}\mathbf{H} \tag{5D.1}$$

where $^{\#}$ denotes the Moore-Penrose matrix pseudoinverse [68, 115], and **H** is arbitrary mixing matrix of size $KM \times N(L+K-1)$, with *KM<N(L+K-1).* Without loss of generality, we can suppose the mixing matrix **H** of full row rank. Using the singular value decomposition:

$$\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^T \tag{5D.2}$$

the matrix **T** can be expressed as:

$$\mathbf{T} = \mathbf{V}\mathbf{D}^T (\mathbf{D}\mathbf{D}^T)^{\#}\mathbf{D}\mathbf{V}^T = \mathbf{V}\begin{bmatrix} \mathbf{I}_{KM\times KM} & \mathbf{O}_{KM\times N(L+K-1)-KM} \\ \mathbf{O}_{N(L+K-1)-KM\times KM} & \mathbf{O}_{N(L+K-1)-KM\times N(L+K-1)-KM} \end{bmatrix}\mathbf{V}^T, \tag{5D.3}$$

where **V** is $N(L+K-1)\times N(L+K-1)$ unitary matrix, and **D** is $KM \times N(L+K-1)$ diagonal matrix with the singular values of matrix **H** on its diagonal.

It is easy to show that:

$$\mathbf{T} = \mathbf{T}^T = \mathbf{T}^2, \tag{5D.4}$$

which proves Property 1.

Property 1: The matrix **T** is orthogonal projector mapping the $N(L+K-1)$-dimensional space of sources onto *KM*-dimensional space of measurements.

Taking into consideration the common properties of the orthogonal projection matrices we have [68, 115]:

$$trace(\mathbf{T}) = \sum_{i=1}^{N(L+K-1)} t_{ii} = KM \tag{5D.5}$$

and

$$0 \le t_{ii} = \sum_{j=1}^{KM} v_{ij}^2 = 1 - \sum_{j=KM+1}^{N(L+K-1)} v_{ij}^2 \le 1, \tag{5D.6}$$

where t_{ii} stands for the *i*-th diagonal element of matrix **T**, v_{ij} denotes the (*i*,*j*)-th element of the unitary matrix **V**, while *trace*(.) is the matrix trace (the sum of all diagonal elements).

As an orthogonal projector the matrix **T** has many useful properties. In the sequel only those which are crucial for the IC method will be outlined.

Property 2: T is by definition positive semidefinite [115]:

$$\forall \bar{\mathbf{s}};\ \ \bar{\mathbf{s}} \in \Re^{N(L+K)} :\ \ \bar{\mathbf{s}}^T \mathbf{T} \bar{\mathbf{s}} \geq 0 \tag{5D.7}$$

and

$$\det(\mathbf{T}_s) \geq 0 \tag{5D.8}$$

where $\mathbf{T}_s$ denotes any principle submatrix of matrix **T**, i.e. the matrix obtained from **T** by deleting the same set of columns and rows. Applying (5D.8) to the arbitrary principal submatrix $\mathbf{T}_s$ of size 2×2 we obtain:

$$\det\left(\begin{bmatrix} t_{ii} & t_{ij} \\ t_{ij} & t_{jj} \end{bmatrix}\right) = t_{ii} t_{jj} - {t_{ij}}^2 \geq 0 . \tag{5D.9}$$

Moreover, using (5D.7) and setting $\bar{\mathbf{s}} = [0,...,0,1,0,...,0,-1,0,...,0]^T$ (the vector whose *i*-th element equals *1*, *j*-th element equals *-1*, while all the other elements equal zero) we get

$$\bar{\mathbf{s}}^T \mathbf{T} \bar{\mathbf{s}} = t_{ii} + t_{jj} - 2t_{ij} \geq 0 . \tag{5D.10}$$

Changing the sign of the *j*-th element in vector $\bar{\mathbf{s}}$ i.e. $\bar{\mathbf{s}} = [0,...,0,1,0,...,0,1,0,...,0]^T$ (5D.7) yields:

$$\bar{\mathbf{s}}^T \mathbf{T} \bar{\mathbf{s}} = t_{ii} + t_{jj} + 2t_{ij} \geq 0 . \tag{5D.11}$$

By merging the inequalities (5D.6), (5D.9), (5D.10) and (5D.11) we get

$$\left| t_{ij} \right| \leq \frac{t_{ii} + t_{jj}}{2} \leq \max(t_{ii}, t_{jj}) \tag{5D.12}$$

and

$$\begin{aligned} {t_{ij}}^2 &\leq t_{ii} t_{jj} \leq t_{ii} , \\ {t_{ij}}^2 &\leq t_{ii} t_{jj} \leq t_{jj} . \end{aligned} \tag{5D.13}$$

Finally, from (5D.12) and (5D.13) it follows:

Property 3: The absolute value of the off-diagonal element t_{ij} is limited to the interval

$$\left|t_{ij}\right| \in \left[0, \min\left(\sqrt{t_{ii}}, \sqrt{t_{jj}}, \frac{t_{ii}+t_{jj}}{2}\right)\right]. \tag{5D.14}$$

Property 4: According to (5D.5) and (5D.6) there are at least *KM* diagonal elements t_{ii} different from zero.

Assume now the matrix **T** has *j*-th diagonal element equal to *0*. According to (5D.4) we have

$$\mathbf{T} = \mathbf{T}^2 \rightarrow t_{jj} = \sum_k t_{jk} t_{kj} \tag{5D.15}$$

and

$$\mathbf{T} = \mathbf{T}^T \rightarrow t_{jk} = t_{kj}\,. \tag{5D.16}$$

Finally, we have

$$0 = t_{jj} = \sum_k t_{jk} t_{kj} = \sum_k t_{jk}^2 \geq 0 \rightarrow \forall k : t_{kj} = 0\,. \tag{5D.17}$$

which proves Property 5.

Property 5: The orthogonal projection matrix with the *j*-th diagonal element equal to *0* must have all the elements of the *j*-th row and *j*-th column all *0*.

The last property we are going to outline is based on a less formal proof. From (5D.1) it follows:

$$\lim_{KM \rightarrow N(L+K-1)} \mathbf{T} = \mathbf{I}_{N(L+K-1)\times N(L+K-1)} \tag{5D.18}$$

Observe now the *i*-th column of the matrix **V**. Denote it by $\mathbf{v}_i$ and consider three different cases:

1. The majority of the energy of $\mathbf{v}_i$ is concentrated in the first *KM* elements. This is the ideal case as orthogonality of the matrix **V** guarantees the diagonal element t_{ii} will be close to 1, while all off-diagonal elements t_{ij} and t_{ji} will be close to *0*.

2. The energy of $\mathbf{v}_i$ is uniformly distributed among all elements. This is most probable case and implies the value of the diagonal element t_{ii} will slowly decrease when the ratio $\frac{N(L+K-1)}{KM}$ increases, while the off-diagonal elements t_{ij} and t_{ji} will slowly increase.

3. The majority of the energy of $\mathbf{v}_i$ is concentrated in the last $KM - N(L+K-1)$ elements. This is the worst case as (5D.6) guarantees the diagonal element t_{ii} will be close to zero, whereas according to Properties 3 and 4 the off-diagonal elements t_{ij} and t_{ji} will be much smaller than at least several diagonal elements.

The last conclusions can be further justified by noticing that the orthogonality of the matrix **V** guarantees the number of the columns $\mathbf{v}_i$ concomitant with the case 3 is limited. Namely, being orthogonal, the matrix **V** preserves the norms and the angles between arbitrary vectors. Supposing that all its columns correspond to case 3, it would necessarily imply the matrix **V** is similar to the projection matrix which maps $N(L+K-1)$-dimensional space onto $N(L+K-1)-KM$-dimensional subspace. This directly contradicts the preservation of angles between the arbitrary vectors [68, 115].

Property 6: Supposing the ratio $\frac{N(L+K-1)}{KM}$ small enough, the matrix **T** will have at least a part of its diagonal elements superior to all off-diagonal elements. The numerical simulations reveal the dominant diagonal elements are at least several times higher than the corresponding off-diagonal elements, as long as the number of sources, $N(L+K-1)$, does not exceed the number of measurements, KM, by factor 2.

6.

Surface EMG decomposition results

In previous chapters we derived three novel decomposition approaches. We explained the theory behind the influences of noise and of non-orthogonal sources, and in the last chapter, we evaluated the impact of the ratio between the number of sources and number of measurements. We further estimated the probability of overlapped sources and stated very general conclusions regarding the efficiency of the introduced methods.

Supposing general pulse sources, the derived methods depend neither on the type of the pulse sources nor on the nature of the mixing process. They can, hence, be used for separation of both the convolutive and multiplicative mixtures. The only assumption in which they follow the example of biomedical signals concerns the minimal length of the inter-pulse interval which must be longer than the corresponding system responses. This assumption was further moderated in the case of the IC method where we supposed the sources are divided into R subsets, each comprising Q sources. In this way the decomposition can easily be extended to include more general cases. In the case of biomedical signals we take $R=N$ and $Q = L + K - 1$ [77], while, on the other hand, when allowing the repetitions of the same source to overlap, we take $R = N(L + K - 1)$ and $Q=1$. Following the common convention from the theory of BSS the mixing matrix was further supposed constant in time.

In this chapter, the performance of the developed decomposition approaches will be tested on the synthetic and real SEMG signals. As already explained in the Chapter 3, the MU can be treated as a pulse sources. The refractory period and maximal possible firing frequency guarantee the delayed repetitions of the same innervation pulse train cannot overlap, while the innervation trains of different MUs do not overlap significantly (at least at low muscle contraction levels). On the other hand, the assumption of the stationary mixing matrix imposes some crucial restrictions on the surface EMG measurements. The latter must be taken in steady, controlled conditions during isometric muscle contractions and without any observable effect of the fatigue.

While the decomposition of the EMG signals to the MU firing patterns has proved to be a very important clinical issue, there is yet another, more technical reason for testing the developed decomposition approaches on the surface EMG. Namely, the physiological properties of MUs put some severe limitations on both the innervation pulse trains and on the shape of MUAPs. Evaluating the results on the real signals these limitations can be readily used in the verification and validation procedure. This issue is crucial, as the original innervation trains are usually unknown.

The chapter is divided into two large sections. Section 6.1 evaluates the decomposition on the synthetic surface EMG signals, while Section 6.2 studies the decomposition on the real surface EMG signals. Both sections are further divided into three subsections, each providing the results of different decomposition approach. The results will be discussed and mutually compared in the next chapter.

6.1 Results on synthetic surface EMG signals

Decomposing the real surface EMG signals the original innervation pulse trains and even the number of active MUs are unknown. We can partially rely on the indirect measures which will be described more in detail in the next section. Nevertheless, the different decomposition procedures should first be evaluated on the synthetic signals. This way the impact of the number of active MUs, their firing frequency, depth in the muscle tissue, etc., as well as the influence of noise can be readily evaluated. Moreover, the experience gained on the synthetic signals can prove crucial when evaluating the performance on the real signals. However, to be able to reason from the thier results, the synthetic signals must resemble the real ones as much as possible.

To ensure their representativeness, all the synthetic EMG signals were generated using the advanced surface EMG simulator built by the world recognised experts from the bioengineering laboratory LISiN (Laboratorio di Ingegneria del Sistema Neuromuscolare) at Politecnico di Torino, Italy. The simulator allows simulation of multi-channel spatially filtered surface EMG signals generated during voluntary contractions by the activity of a large number of motor units [56]. All the main features of the surface EMG signal are modelled, including the fatigue, the generation and extinction phenomena of the action potentials at the end-plate and tendon regions, the size and shape of the recording electrodes, and the inclination of the fibres with respect to the detection system. The volume conductor is described as an anisotropic layered medium with muscle, fat and skin layers. Motor units are placed randomly in the detection volume and are active at selectable firing rates. The detection systems are placed either along with, or transversal to, the fibre direction.

The influence of two factors was evaluated during our simulation: a) the number of active MUs, and b) the signal-to-noise ratio (SNR). The number of active MUs was set to 5, 10 and 20, respectively, while SNR ranged from 0 dB to 20 dB, in steps of 5 dB. The other important SEMG parameters were set as follows:

- Skin was simulated as 1 mm thick isotropic layer and subcutaneous fatty tissue as 3 mm thick isotropic layer.
- Active MUs consisted of a random number of fibres (uniformly distributed between 50 and 300) with the circular MU territories of 20 fibres/mm^2. The depth of active MUs in the anisotropic muscle layer varied uniformly from 3 mm to 10 mm, while additional random shift (uniformly distributed between -10 and 10 mm) from the centre of the electrode array in the direction transversal to the muscle fibres was applied to each MU.

- MU firing rate was normally distributed around the mean of 15 Hz with standard deviation of 4 Hz. The IPI variability was modelled zero-mean Gaussian variable with the variance equal to 20 % of the IPI mean.
- Conduction velocity was assumed normally distributed with the mean of 4 m/s and standard deviation of 1 m/s.
- The innervation zones were assumed placed in the middle between the tendons, and the semi-fibre length was set to 70 mm. The spread of innervation zones was limited to 5 mm.
- The average length of simulated MUAPs was estimated to 25 ms.
- 11×5 array detection system with 11 rows and 5 electrodes per row was centred over the distal half of muscle fibres, columns aligned with the direction of fibres.
- Rectangular 1 by 1 mm electrodes with the inter-electrode distance of 5 mm were simulated.
- Measurements were supposed longitudinal single-differential.
- Synthetic SEMG signals in duration of 30 s were sampled at 1024 Hz.

10 simulations were performed for each number of active MUs (5, 10 and 20). In each of the simulation runs the depth of the active MUs, their firing rate, number of fibres, shift in the direction transversal to the muscle fibres, and conduction velocity were randomly selected. In addition, signals from each simulation run were corrupted by additive noise (5 realisations of noise for each SNR) resulting in 250 test signals for each number of active MUs.

6.1.1 Method based on time-frequency distributions

We first tested the method based on time-frequency distributions. Employing all 50 single-differential measurements of synthetic surface EMG, and having the average length of MUAPs estimated to 25 samples, the vector of measurements was extended by factor K=3 in the case of 5 active MUs, and by factor K=7 in the case of 10 active MUs. According to (3.12), the number of extended sources was estimated to 135 and 310, respectively. In order to ensure the number of measurements exceeds the number of sources, the signals with 20 active MUs would have to be extended by the factor K=17. As a result, the set of STFD $\mathbf{D}_{\bar{\mathbf{z}}\bar{\mathbf{z}}}(n, f)$ matrices entering the joint-diagonalization would comprise at least 820 matrices of size 820×820. Having its computational complexity proportional to the cube of the matrix size [74], the joint-diagonalization would require enormous amount of processing time to complete. Hence, the method based on time-frequency distributions was tested only on the signals with 5 and 10 active MUs. This decision is further justified by the results, which, as we will see, prove the efficiency of the method decreases with the number of active MUs.

The method based on time-frequency distributions assumes the number of sources (active MUs) is known. When this is not the case, the number of sources can be estimated from the rank of correlation matrix of measurements $\mathbf{R}_{\bar{\mathbf{x}}}$. Note also, the number of sources can be slightly overestimated. As a result, the method reconstructs some extra sources, but

they all have amplitude close to zero. Nevertheless, testing the method on the synthetic signals we supposed the number of sources was known.

The exact decomposition procedure used can be summarized in the following six steps:

1. Using the eigen-decomposition of the correlation matrix of extended measurements, reconstruct the whitening matrix **B** and estimate the noise variance σ^2. The exact procedure is described in [13, 17].

2. Whiten the extended vector of noisy measurements and, according to (4.15). calculate the matrices of pseudo Wigner-Ville distribution $\mathbf{D}_{\overline{\mathbf{zz}}}(n, f)$.

3. Apply the criterion (4.21) to all $\mathbf{D}_{\overline{\mathbf{zz}}}(n, f)$ matrices and join the matrices which correspond to the single-autoterms of the same, say the *i*-th source to the set A_i. If the number of sets A_i reaches the estimated number of sources, go to step 4. Otherwise use (4.21) and (4.24) to find the matrices which correspond to the missing sources and contain as little contributions from other sources as possible.

4. Average the matrices in each set A_i and use (4.28) to remove the influence of noise. Denote the resulting matrices by M_i.

5. Joint-diagonalize all matrices M_i to estimate the missing unitary matrix **U** and use (4.16) to reconstruct the innervation pulse trains.

6. Mutually compare the obtained innervation trains and classify all the repetitions of the same innervation train into the same group. Align in time all the delayed repetitions of the same innervation train and calculate their average.

Evaluating the results, only the reconstructed innervation trains containing at most 30 % misplaced pulses were taken into consideration. All other reconstructions were discarded. The results are summarized in Table 6.1 and exemplified in Figs. 6.1 and 6.2.

The results in Table 6.1 show the decomposition efficiency decreases with the SNR and the number of active MUs. At SNR of 20 dB, the average number of the reconstructed MUs yields 3.8 and 7.3 in the case of 5 and 10 MUs, respectively. Decreasing SNR to 0 dB the average number of MUs drops to 1.6 and 0.7, respectively. The drop of performance in the case of 10 active MUs can partially be explained by realizing the method based on TF distributions strongly depends on the number of non-overlapped pulses. The latter decreases rapidly with the number of active MUs. Using the results of Appendix 5B and setting *R=5* (5 active MUs), $Q = 27$ and $p = 0.25$, the probability of non-overlapped pulse yields 0.24. On the other hand, setting *R=10* (10 active MUs) and $Q = 31$, the probability of non-overlapped pulse drops to 0.06. Increasing the number of active MUs by factor 2, the number of matrices corresponding to the single-autoterms decreases by factor 4. The number of non-ideal STDF matrices entering the joint diagonalization increases, and the performance drops (Subsection 4.1.4).

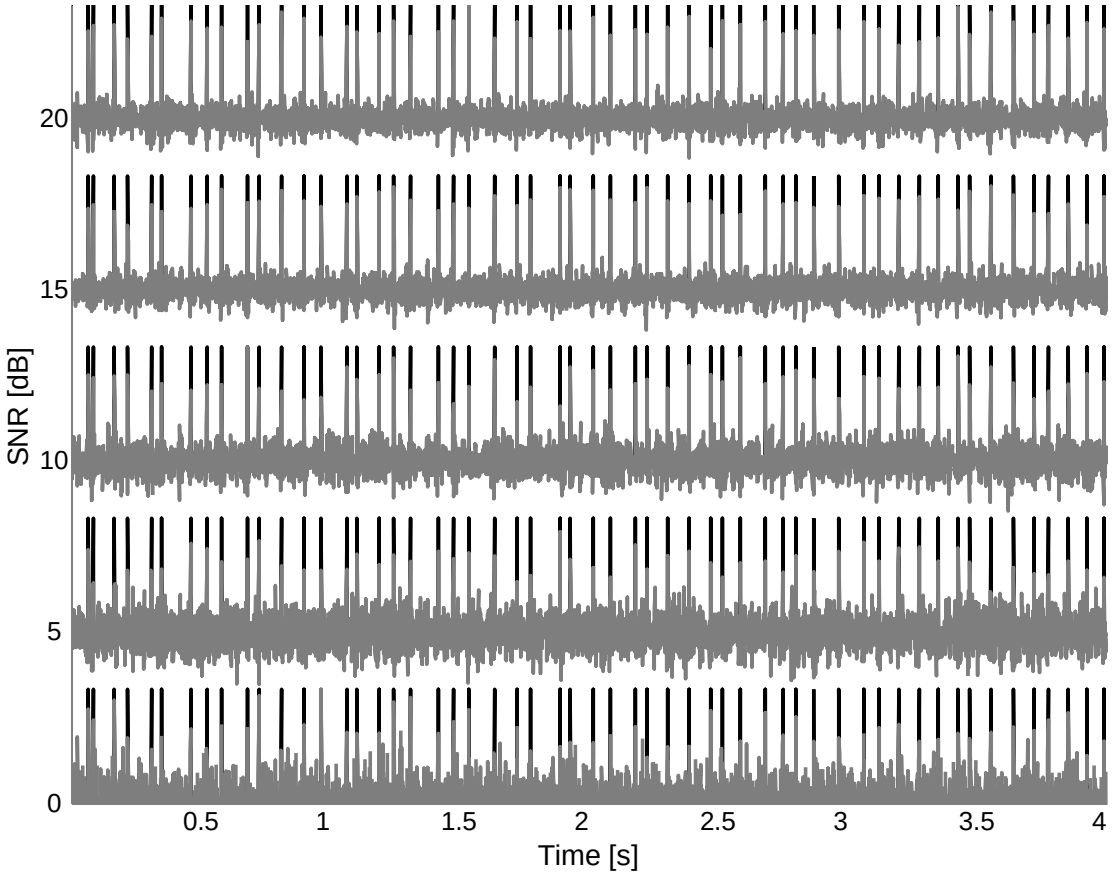


Figure 6.1: Original synthetic innervation pulse trains (black) and pulse trains reconstructed by the method based on TF distributions (grey) in the case of 5 active MUs (simulation 2, MU 5) and at SNR = 20, 15, 10, 5 and 0 dB.

The performance also drops with noise. We can partially cancel its influence by averaging the STFD matrices and applying (4.28). However, the noise also influences the selection of STFD matrices, which, as already explained, is crucial for the estimation of the mixing matrix. Small probability of overlapped pulse in the case of 5 active MUs compensates the drop in performance of criterion (4.21). In the case of 10 active MUs the number of overlapped pulses increases and the effect of noise becomes evident.

The TF-based method uses the reconstructed mixing matrix to estimate the pulse sources. According to (3.16), the mixing matrix comprises K delayed repetitions of each system response. As a result, the MUAPs can be reconstructed by averaging the corresponding rows of the estimated mixing matrix. However, according to BSS indeterminacy 3.1, the reconstructed sources and corresponding columns of the reconstructed mixing matrix can be arbitrarily permuted (with respect to the original sources). Hence, before the averaging, the columns of the mixing matrix must be rearranged to match the original order. Note the original order of sources (and matrix columns) is already determined in step 6 of the decomposition procedure described above. While the reconstructed sources can be averaged over $K+L-1$ delayed estimations of each source, there are only K delayed reconstructions of each MUAP. In our test, we used K=3 in the case of 5 active

MUs and K=7 in the case of 10 active MUs. Averaging over only 7 different values is not very efficient and we can readily expect the reconstructed MUAPs will significantly differ form the original ones. The results are described in Table 6.2, while the norm of the MUAP difference, which is used as a measure of fit, will be defined in the next subsection.

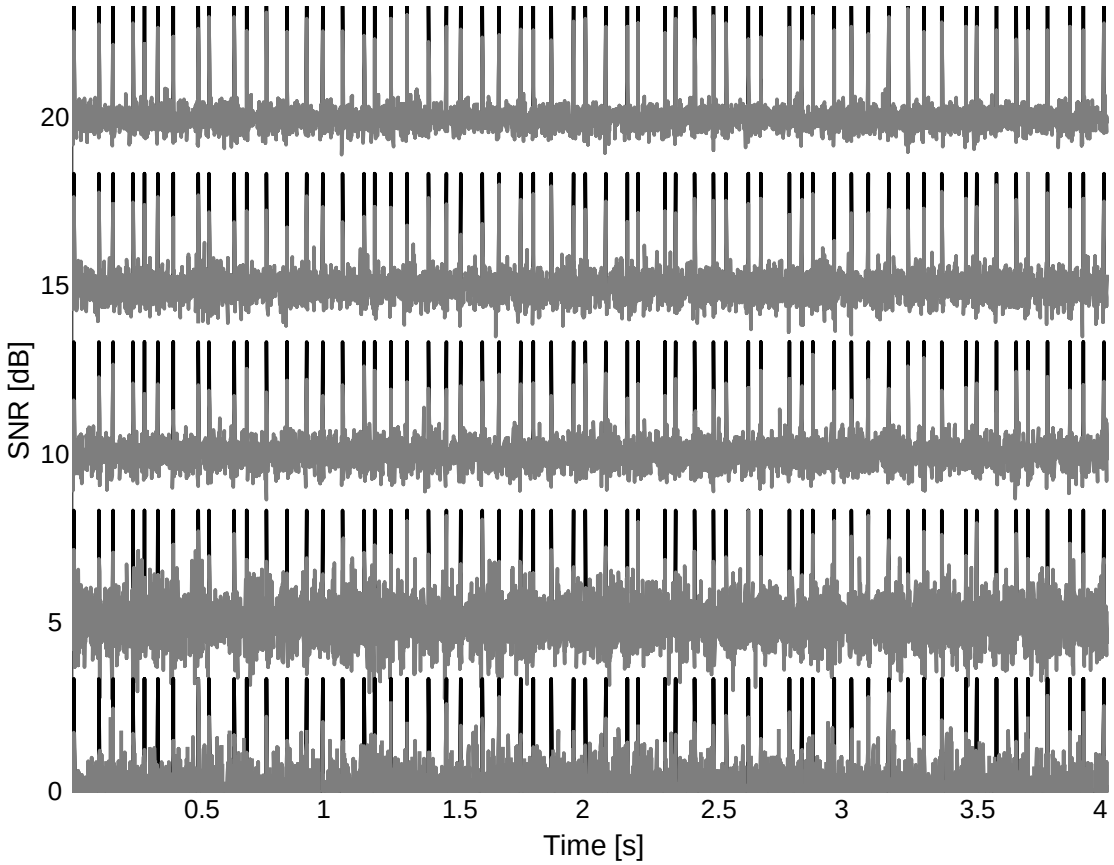


Figure 6.2: Original synthetic innervation pulse trains (black) and the pulse trains reconstructed by the method based on TF distributions (grey) in the case of 10 active MUs (simulation 3, MU 8) and at SNR = 20, 15, 10, 5 and 0 dB.

The results in Table 6.2 will be further discussed in the next chapter, where they will be compared to the results of other two methods. For the time being, it suffices to realize the innervation pulse trains were reconstructed much more accurately than the corresponding MUAPs.

Table 6.1. The number of reconstructed innervation pulse trains (mean ± standard deviation), the percentage of reconstructed pulses with respect to the number of pulses in the original synthetic pulse train (true positive statistics) and the percentage of misplaced pulses with respect to the number of reconstructed pulses (false positive statistics). The results are classified according to the number of active MUs and the simulated SNR.

Number of active MUs	**SNR [*dB*]**	**Average number of reconstructed innervation pulse trains**	**Portion of reconstructed pulses [%]**	**Portion of misplaced pulses [%]**
5	20	3.8 ± 0.9	90.0 ± 12.0	5.9 ± 6.0
	15	3.3 ± 1.3	81.6 ± 14.6	7.5 ± 7.0
	10	2.8 ± 0.9	77.3 ± 16.0	12.7 ± 8.7
	5	2.3 ± 0.5	66.4 ± 11.4	14.3 ± 8.0
	0	1.6 ± 0.5	60.8 ± 5.9	18.7 ± 9.1
10	20	7.3 ± 1.1	89.0 ± 14.7	5.0 ± 8.7
	15	6.1 ± 0.8	83.9 ± 14.2	5.9 ± 8.9
	10	4.7 ± 1.0	76.7 ± 14.5	9.4 ± 11.0
	5	2.2 ± 0.8	68.6 ± 10.4	13.6 ± 8.4
	0	0.7 ± 0.8	57.6 ± 4.1	21.5 ± 9.6

Table 6.2. The number of MUs (mean ± standard deviation) reconstructed by the method based on TF distributions, and the average norm of the difference (6.1) between the MUAPs reconstructed from the mixing matrix and their reference values. The results were obtained on the synthetic signals in the case of 5 and 10 active MUs, respectively.

Number of active MUs	**SNR [*dB*]**	**Average number of reconstructed MUs**	**Average norm d_{ij} [%]**
5	20	3.8 ± 0.9	52.6 ± 5.6
	15	3.3 ± 1.3	56.9 ± 9.2
	10	2.8 ± 0.9	58.3 ± 9.7
	5	2.3 ± 0.5	60.4 ± 6.4
	0	1.6 ± 0.5	67.5 ± 7.5
10	20	7.3 ± 1.1	68.0 ± 12.9
	15	6.1 ± 0.8	70.0 ± 12.8
	10	4.7 ± 1.0	71.4 ± 13.2
	5	2.2 ± 0.8	72.3 ± 11.8
	0	0.7 ± 0.8	78.2 ± 13.1

6.1.2 Method based on higher-order statistics

The efficiency of all possible decomposition methods which are based on higher-order statistics decreases whenever:

- the noise is not Gaussian,
- the sources are not independent and identically distributed (i.i.d.),
- the processed signals are too short to provide the accurate estimation of cumulants (4.34).

In the case of synthetic EMG signals the noise was simulated white and Gaussian and we can readily expect the w-slices will prove their high resistance to it. Having the signals limited to length of 30720 samples we can also assume (4.34) will relatively accurately estimate the values of cumulants. The main source of errors will, therefore, be the non-orthogonal sources. The latter are in the case of synthetic surface EMG signals still independent, but exhibit the overlapping pulses which contribute the non-zero values to the cross-cumulants of sources. As a results, the Barlett-Brillinger-Rosenblatt equation fails to model the system accurately and the Newton-Gauss optimisation converges to the wrong solution. However, w-slices do not suppose the sources extended by their delayed repetitions, which guarantees the number of overlapped pulses is much smaller than in the case of other two methods (the IC method and TF-based method). Therefore, we can assume the decomposition method based on the HOS will still manage to reconstruct good estimates of the true MUAPs.

A nice property of HOS–based decomposition methods is their ability to reconstruct the system responses also when the number of sources by far exceeds the number of measurements. Given the number of sources and the length of the corresponding system responses, the required number of measurements can be calculated by (4.42). Setting the number of sources to 20 and the length of system responses to L=25 we see that all the system responses can be reconstructed from a single measurement. The derived estimation is highly theoretical. In practice, various calculation errors hinder the decomposition and the number of sources reconstructed from a single measurement drops drastically. HOS-based decomposition technique, presented in this dissertation, comprises two steps. While the Newton-Gauss optimisation allows more sources than measurements, the w-slices assume the number of sources does not exceed the number of measurements.

All the properties of HOS are not favourable. One of the most obvious drawbacks is their computational complexity. In the case of w-slices, the computational complexity increases with the square of the length of system responses and also with the square of the estimated number of sources. Consequently, the decomposition method was tested only on the signals with 5 active MUs. The computational complexity of Newton-Gauss optimisation is also responsible for our decision to test the method only on 5 out of 50 available surface EMG measurements. In all tests on the synthetic signals, only the measurements corresponding to the relative simulated uptake positions (3,3), (3,4), (3,5),

(3,6) and (3,7) were decomposed, where pair (*i*, *j*) denotes the *i*-th row and *j*-th column. All the other measurements were simply discarded.

When w-slices are used to coarsely estimate the MUAPs, we first need to know their length. Setting the sampling frequency to 1 kHz, the MUAP length may be expected at around 25 samples. There is yet the condition from (4.59) to be met:

$$h_{ij}(0) = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$$

In the case of surface EMG signals this condition is relatively difficult to attain. The closest possible approximation is achieved when the surface EMG measurements are shifted in time. Taking the travelling of AP into consideration this implies a longitudinal placement of electrodes and also justifies our decision to take 5 measurements form the central column of electrodes. On the other hand, one must be aware that each decomposed MUAP is going to be normalized by its starting sample set to 1. Hence, the amplitudes of MUAPs will be lost in the decomposition process. The fact that we cannot entirely met the condition from (4.59) does not hinder the decomposition. As already explained, the w-slices are used to only coarsely estimate the shapes of MUAPs, while their final form is gained through the Newton-Gauss optimisation.

Applying our HOS-based method to the synthetic EMG signals with 5 active MUs, the decomposition process can be summarized in the following 4 steps:

1. Take measurements of surface EMG which correspond to the longitudinal positions of electrodes. In our case, the measurements from the central columns of electrodes (the measurements corresponding to uptake positions (3,3), (3,4), (3,5), (3,6) and (3,7) were processed.

2. Calculate the third-order cumulants of these measurements only up to the lags τ_1 and τ_2 which correspond to the estimated MUAP length *L=25*.

3. Use (4.57) and (4.58) to estimate the casual and anti-casual cumulant matrices $\mathbf{S_a}$ and $\mathbf{S_c}$, and calculate the coarse estimates of MUAPs according to (4.59) and (4.60).

4. Optimise the coarse estimation by the Newton-Gauss optimisation. In order to reduce the computational complexity as much as possible, mainly because of huge dimensions of matrix $\boldsymbol{\Sigma}$, it is wise to select an appropriate subset of the cumulant values in matrix $\mathbf{C}$ (4.62), such that the conditions given by (4.42) and (4.43) are just fulfilled. In our experiments, we considered only the first quadrants of cumulants $\mathbf{C}_{klm}$, i.e. $0 \le \tau_1 \le L-1 = 24$, $0 \le \tau_2 \le L-1 = 24$. The size of step δ_i in (4.66) was set by the method of bisection, whereas the vector of source cumulants $\boldsymbol{\gamma} = [\gamma_1 \ \cdots \ \gamma_K]^T$ was, according to (4.68), averaged over all possible combinations of indices *k*, *l* and *m*.

In all our experiments, the optimisation was terminated whenever all the possible 2-norms of pair-wise differences among the MUAPs reconstructed in three successive optimisation steps calculated below 5 % (with respect to the 2-norm of the reconstructed MUAP). On average, the optimisation converged in 30 optimisation steps.

The reconstructed MUAPs are partially depicted in Figs. 6.3 and 6.4, where only two of many reconstructed MUAPs are illustrated in dependence of SNR. Quick visual inspection of the results confirms the expected resistance to the Gaussian noise. Although the shape of MUAPs changes with SNR, the noise influence is relatively small.

In order to quantify the decomposition efficiency, the reconstructed MUAPs were also numerically compared against their reference shapes. In all our experiments, the 2-norm of difference between the reconstructed and the reference MUAP defined in (6.1) was used as a measure. Since the value of the MUAP difference depends also of the norm on the MUAPs, the compared MUAPs (6.1) were first normalized. Expressing the difference referring to the 2-norm of the reference MUAPs and taking into consideration that the norm of a normalized vector equals to 1, the introduced comparison criterion takes its final form:

$$d_{ij} = \frac{\left\| \frac{\mathbf{h}_{ij}}{\|\mathbf{h}_{ij}\|} - \frac{\hat{\mathbf{h}}_{ij}}{\|\hat{\mathbf{h}}_{ij}\|} \right\|}{\left\| \frac{\mathbf{h}_{ij}}{\|\mathbf{h}_{ij}\|} \right\|} = \left\| \frac{\mathbf{h}_{ij}}{\|\mathbf{h}_{ij}\|} - \frac{\hat{\mathbf{h}}_{ij}}{\|\hat{\mathbf{h}}_{ij}\|} \right\| \qquad (6.1)$$

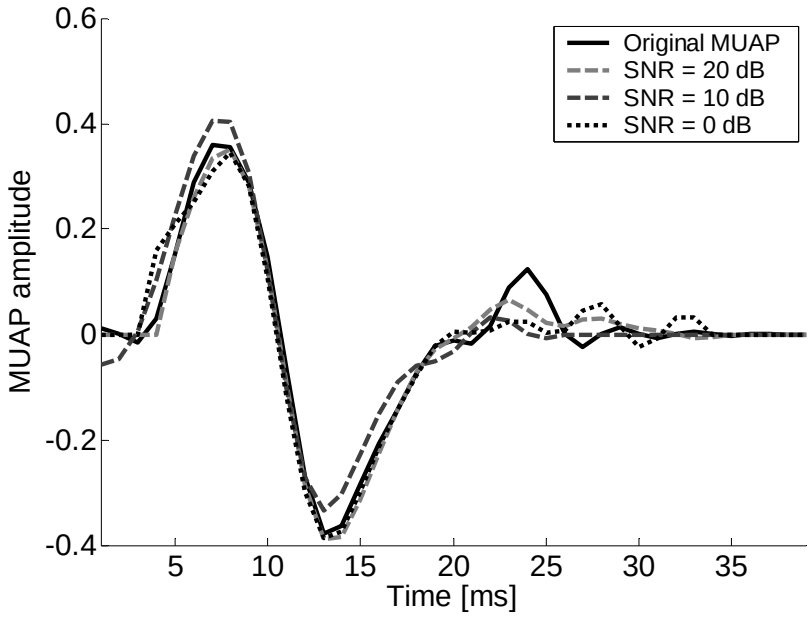
,

where we write $\mathbf{h}_{ij} = [h_{ij}(0),...,h_{ij}(L-1)]$ and $\hat{\mathbf{h}}_{ij} = [h_{ij}(0),...,h_{ij}(L-1)]$ to denote the reference and the reconstructed MUAP $h_{ij}(n)$.

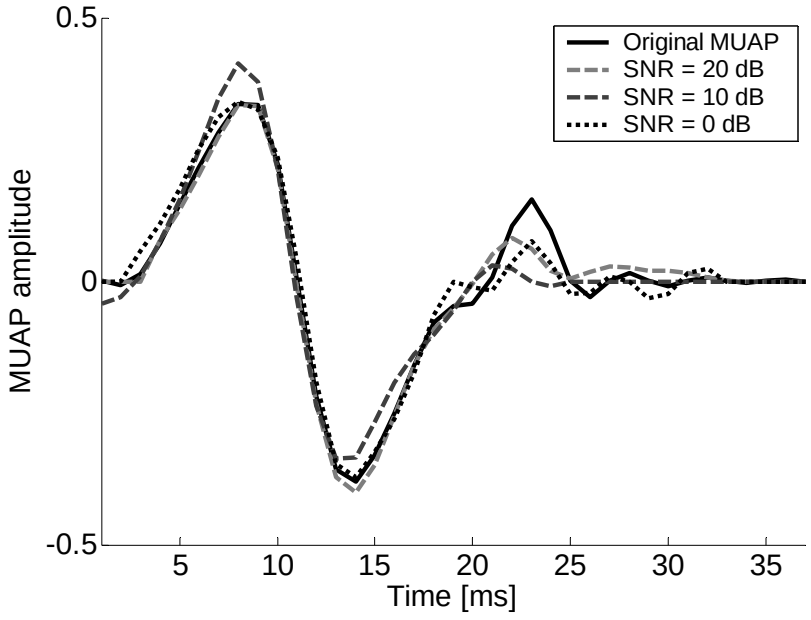
The results were averaged over all the simulation runs with the same SNR and are summarized in Table 6.3. Only the MUAPs whose norm differences (6.1) were below the threshold of 30 % were classified as identified. The robustness to the Gaussian noise was also confirmed by the numeric results. The mean 2-norm difference (6.1) yields 12.0 % at SNR = 20 dB and increases only for 3.7 % at SNR = 0 dB.

Table 6.3. The number of MUAPs (mean ± standard deviation) reconstructed by w-slices in the case of 5 active MUs, and their comparison to the reference shape in the dependence on the SNR. The applied norm of difference d_{ij} is defined by (6.1).

The number of active MUs	**SNR [*dB*]**	**Average number of reconstructed MUs**	**Average norm difference d_{ij} [%]**
5	20	3.6 ± 0.8	12.0 ± 10.0
	15	3.4 ± 0.6	12.0 ± 11.0
	10	3.2 ± 0.5	11.3 ± 10.0
	5	2.7 ± 0.8	13.1 ± 12.5
	0	2.0 ± 0.9	15.7 ± 12.9

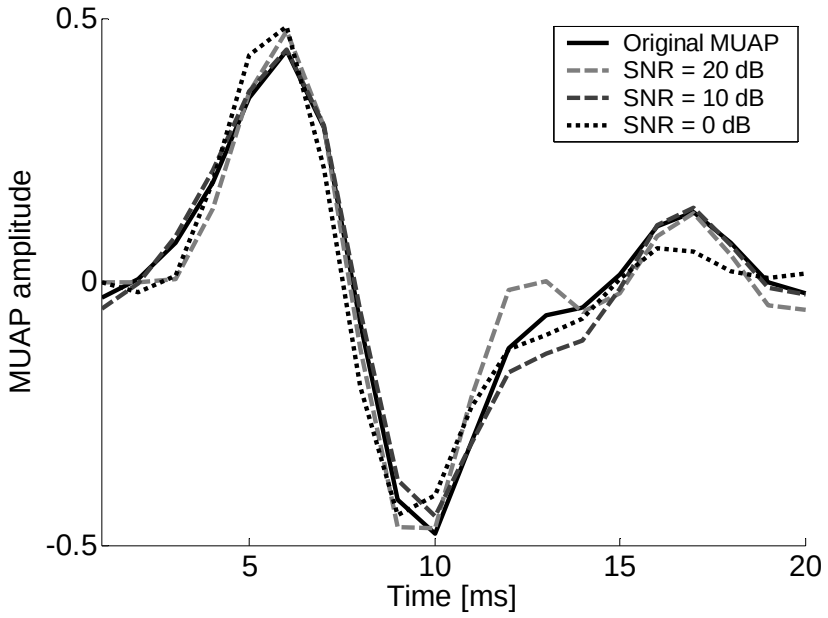


a)

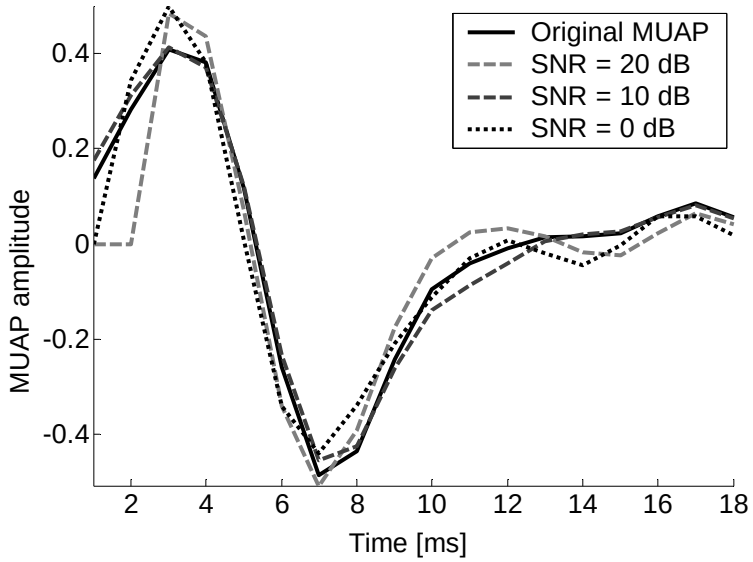


b)

Figure 6.3: The MUAPs, reconstructed by w-slices from the synthetic surface EMG signals of length 30720 samples in the case of 5 active MUs and at SNR=20 dB (light grey dashed line), SNR=10 dB (dark grey dashed line), SNR=0 dB (black dotted line): a) simulation 1, contribution of the second MU to the measurement (3,4), b) simulation 1, contribution of the second MU to the measurement (3,5). Reference values are depicted with black solid lines.



a)



b)

Figure 6.4: The MUAPs, reconstructed by w-slices from the synthetic surface EMG signals of length 30720 samples in the case of 5 active MUs and at SNR=20 dB (light grey dashed line), SNR=10 dB (dark grey dashed line), SNR=0 dB (black dotted line): a) simulation 2, contribution of the first MU to the measurement (3,5), b) simulation 2, contribution of the fourth MU to the measurement (3,7). Reference values are depicted with black solid lines.

6.1.3 Inverse correlation based method

In Chapter 5, we derived a completely novel decomposition approach, the so called inverse correlation (IC) based method. We demonstrated its efficiency in the case of orthogonal sources, and assessed the influences of noise and non-orthogonal sources. We also argued the IC method can be used for the decomposition of under-determined MIMO systems. Finally, we explained how the performance drops with the ratio between the number of measurements and the number of sources. We concluded that, even when the number of sources exceeds the number of measurements by factor 2, the method will still be able to completely reconstruct at least a part of the sources.

In this section we will test its performance on the synthetic surface EMG signals. We will assess the impact of noise and check the hypothesis that the MUs with low firing rates are more likely to be decomposed. Finally, we will evaluate the reconstruction of MUAPs by the spike triggered sliding window averaging technique.

To verify how the method's efficiency depends on the ratio between the number of sources and the number of measurements we fixed the number of measurements. The latter were, regardless to the number of active MUs, always extended by the factor K=10. In other words, we added 9 delayed repetitions to each measurement, which increased the number of measurements to 500. The number of the extended sources depends on the number of active MUs and reaches 170, 340 and 680 for 5, 10 and 20 active MUs, respectively. In the last case, the number of sources exceeds the number of measurements by 1.36.

The applied decomposition procedure can be summarized in the following 8 steps:

1. Calculate the correlation matrix of extended measurements, estimate the noise variance, $\hat{\sigma}^2$, and use (5.4) to calculate the activity index. Using the estimated noise variance determine the noise threshold:

$$t_{\hat{\sigma}^2} = \hat{\sigma}^2 \left\| \mathbf{R}_{\bar{\mathbf{x}}}^{-1} \right\|_1 \tag{6.2}$$

   where $\left\| \mathbf{R}_{\bar{\mathbf{x}}}^{-1} \right\|_1$ denotes 1-norm of matrix $\mathbf{R}_{\bar{\mathbf{x}}}^{-1}$. Set the values in which the activity index does not exceed the noise threshold to zero. When $\mathbf{R}_{\bar{\mathbf{x}}}^{-1}$ is ill-conditioned use the procedure described in 5.3 to cut the eigenvalues of $\mathbf{R}_{\bar{\mathbf{x}}}$ and to improve its sensitivity to noise. The optimal degree of cutting can be determined from the eigenvalues of matrix $\mathbf{R}_{\bar{\mathbf{x}}}$. The higher the degree of cutting, the lower the sensitivity to noise. Empirically, at least 80 % of the energy of original matrix $\mathbf{R}_{\bar{\mathbf{x}}}$ must be preserved after cutting in order to prevent the loss of information on the mixing matrix.

2. Calculate the median value of the activity index and denote by n_0 an arbitrary time moment (the sample of the activity index) in which the median value is reached.

3. Use (5.8) to reconstruct a superimposition of innervation pulse trains $v_{n_0,G_{n_0}}(n)$ and randomly choose one of its pulses exceeding the noise threshold $t_{\hat{\sigma}^2}$. Denote by n_1 the time moment of the selected pulse and reconstruct the superimposition $v_{n_1,G_{n_1}}(n)$.

4. In the product $v_{n_0,G_{n_0}}(n)\cdot v_{n_1,G_{n_1}}(n)$, select all the pulses exceeding the noise threshold $t^2_{\hat{\sigma}^2}$ and denote their number by *Z*. Denote by n_i, *i=2,3,..Z+1*, the time moments in which the *i*-th selected pulse appears and reconstruct the superimpositions $v_{n_i,G_{n_i}}(n)$. Multiply (element-wise) each superimposition $v_{n_i,G_{n_i}}(n)$ by $v_{n_0,G_{n_0}}(n)\cdot v_{n_1,G_{n_1}}(n)$.

5. Mutually compare all the products $v_{n_0,G_{n_0}}(n)\cdot v_{n_1,G_{n_1}}(n)\cdot v_{n_i,G_{n_i}}(n)$ and classify them according to the number of overlapped pulses. Put all the time moments n_i and n_j, for which the products $v_{n_0,G_{n_0}}(n)\cdot v_{n_1,G_{n_1}}(n)\cdot v_{n_i,G_{n_i}}(n)$ and $v_{n_0,G_{n_0}}(n)\cdot v_{n_1,G_{n_1}}(n)\cdot v_{n_j,G_{n_j}}(n)$ overlap at least in *J* pulses, to the same set A_i. The threshold *J* can be computed as a product of the signal length and the lowest possible firing frequency (approximately 6 Hz in the case of EMG signals).

6. For all the sets A_i whose cardinal number exceeds the threshold 2*J*/3, calculate the average vector of extended measurements:

$$\bar{\mathbf{x}}_{A_i} = \frac{1}{card(A_i)} \sum_{n_i \in A_i} \bar{\mathbf{x}}(n_i)$$

and use (5.29) to reconstruct a single innervation pulse train. Set to zero the activity index values in all the time moments which correspond to the pulses in the reconstructed innervation pulse train.

7. Compare the obtained innervation train to all pervious reconstructed pulse sequences and classify it either as a new pulse sequence or already detected pulse sequence.

8. Repeat steps 2 to 7 until all the values of the activity index are set to zero.

In our experiments the cutting of the eigenvalues in matrix $\mathbf{R}_{\bar{\mathbf{x}}}$ (Section 4.3) was limited to 50 % (50 % of eigenvalues were set to zero). Although the selected degree of cutting may seem high, the eliminated eigenvalues represented only 10 % of the original matrix

energy, on average. Following the advice in step 5, the threshold J was estimated to 6·30=180, where 30 stands for the length of the synthetic signals (in seconds) and the 6 is the lowest expected MU firing rate (in Hz). The decomposition results are summarized in Table 6.4. Some of the reconstructed innervation pulse trains are further illustrated in Figs. 6.5, 6.6 and 6.7.

Table 6.4: The number of reconstructed MUs' innervation pulse trains (mean ± standard deviation), the percentage of reconstructed pulses with respect to the number of pulses in the original synthetic pulse trains (true positive statistics), and the percentage of misplaced pulses with respect to the number of reconstructed pulses (false positive statistics), in dependence of the number of active MUs and SNR.

Number of active MUs	**SNR [*dB*]**	**Average number of reconstructed innervation pulse trains**	**Percentage of accurately reconstructed pulses**	**Percentage of misplaced pulses**
5	20	4.7 ± 0.6	99 ± 0.6	0.0 ± 0.0
	15	4.2 ± 0.6	99 ± 0.9	0.0 ± 0.0
	10	3.9 ± 0.7	97 ± 2.3	0.0 ± 0.0
	5	3.2 ± 0.6	95 ± 3.8	1.0 ± 2.0
	0	2.0 ± 0.9	93 ± 4.1	2.0 ± 3.0
10	20	8.3 ± 1.3	99 ± 0.8	0.0 ± 0.0
	15	7.0 ± 1.6	99 ± 1.1	0.0 ± 0.0
	10	6.4 ± 1.9	97 ± 2.6	0.0 ± 0.0
	5	4.6 ± 1.7	94 ± 4.7	1.0 ± 2.0
	0	2.6 ± 1.4	91 ± 5.1	3.0 ± 4.0
20	20	10.4 ± 1.4	99 ± 0.7	0.0 ± 0.0
	15	8.6 ± 1.7	98 ± 1.3	1.0 ± 3.0
	10	5.7 ± 1.3	96 ± 3.5	1.0 ± 3.0
	5	3.5 ± 1.3	93 ± 5.1	3.0 ± 5.0
	0	2.3 ± 1.9	91 ± 6.3	4.0 ± 5.0

Evaluating the reconstructed innervation pulse trains, only the trains containing at most 30 % of misplaced pulses were considered identified. All other reconstructions were simply discarded. In the case of 5 and 10 active MUs (the over-determined case) and SNR of 20 dB, almost all the simulated MUs were completely reconstructed. Simulating 20 active MUs (the under-determined case), the portion of reconstructed MUs dropped to 50 %. Increasing the noise influence, both over-determined cases (5 and 10 active MUs) suffered approximately the same drop in performance. At SNR = 10 dB, the method reconstructed approximately 80 % of the MUs which were identified at SNR of 20 dB, while at SNR of 0 dB the portion of the recognised MUs dropped to 35 %. The drop in performance was more obvious in the case of 20 active MUs (the under-determined case). At SNR of 10 and 0 dB, only 55 % and 25 % of the MUs identified at SNR = 20 dB were

recognised, respectively. Regardless the number of active MUs and the SNR, the reconstructed innervation pulse trains exhibited almost perfect match with their reference values, while the number of misplaced pulses was negligible.

The theory predicts the IC method should reconstruct the MUs with lower firing frequencies more successfully. In the decomposition process, the MUs with high firing frequencies are multiplied by the smaller diagonal elements of the inverse correlation matrix of sources (Appendix 5A) and are less likely to exceed the noise threshold. Following the same conclusions, the MUs with low firing frequencies have higher contributions in the reconstructed pulse trains and are more likely to surpass the noise level. To test this hypothesis we observed the firing rates of the reconstructed MUs. The firing rates of the simulated MUs were normally distributed around the mean of 15 Hz, while the standard deviation was set to 4 Hz. The lowest simulated firing rates were around 10 Hz, whereas the highest ones outreached 19 Hz.

The decomposition efficiency also depends on other parameters of MUs, mainly on the depth of MU in the muscle tissue and on the number of muscle fibres comprising each MU. Both parameters are strongly correlated because they both affect the amplitude of a MUAP as detected on the skin surface. Superficial MUs with a high number of muscle fibres produce stronger MUAPs and are expected to resist the influence of noise with more success. Therefore, evaluating the reconstructed MUs, the influence of all three parameters (the firing rate, depth in the muscle tissue, and the number of fibres) must be taken into consideration. Due to the clarity reasons, we combined them in the following morphological index:

$$I_{f,A} = \frac{\frac{1}{M}\sum_{m=1}^{M} Amp_m}{f_{fir}} \tag{6.3}$$

where Amp_m stands for the average peak-to-peak amplitude of a MUAP as detected by the m-th electrode, and f_{fir} denotes the average MU firing rate. The higher the value of the morphological index, the stronger the contribution of the MU to the SEMG recordings. The decomposition results are presented in Table 6.5.

The results in Table 6.5 disagree with the theory. No significant impact of the firing rate was noticed. The reconstructed MUs clearly show their correlations with the depth of MUs in the muscle tissue and with the number of muscle fibres. As expected, the method most successfully recognises the superficial MUs with a high number of muscle fibres, while all the others are treated as a background noise. Increasing the noise, the average depth of the reconstructed MUs decreases, while the number of their fibres increases. Recall all the parameters were chosen randomly in each simulation run, independently from each other. Therefore, it is highly unlikely for the MUs with low firing rates only to be deeper in the muscle tissue or to be composed out of smaller number of fibres. This fact is also confirmed by the average firing rates of the identified and missed MUs (Table 6.5, columns 7 and 8). The average firing rate in both groups yields 15.5 Hz on average,

whereas the standard deviation is always around 3.5 Hz. To conclude, the results clearly demonstrate the influence of the MU firing rate is negligible.

Table 6.5: Values of parameters (mean ± standard deviation) of the identified (+) and missed (–) MUs: the MU depth in the muscle tissue, the number of fibres, and the average firing rate, versus the number of active MUs and SNR. The normalized values of the morphological index are depicted in the right-most column. Results were obtained on the synthetic SEMG signals.

Number of active MUs	**SNR [*dB*]**	**Average depth in the muscle tissue [*mm*]**		**Average number of fibres**		**Average firing frequency [*Hz*]**		**Average value of index $I_{f,A}$**	
		+	–	+	–	+	–	+	–
5	20	6.1 ± 1.0	8.3 ± 0.6	156 ±53	129 ± 65	15.5 ± 3.5	15.3 ± 3.9	0.55 ± 0.24	0.23 ± 0.14
	15	6.1 ± 1.0	7.9 ± 1.5	168 ± 45	98 ± 41	15.3 ± 3.0	15.9 ± 3.1	0.61 ± 0.21	0.25 ± 0.18
	10	6.1 ± 1.0	7.1 ± 1.4	175 ± 40	96 ± 34	15.6 ± 3.2	15.4 ± 3.4	0.63 ± 0.20	0.28 ± 0.19
	5	5.9 ± 0.8	6.8 ± 1.4	173 ± 41	116 ± 53	15.5 ± 3.2	15.7 ± 3.4	0.65 ± 0.21	0.32 ± 0.12
	0	5.8 ± 0.4	6.8 ± 1.4	188 ± 40	124 ± 43	15.3 ± 3.3	15.5 ± 3.6	0.70 ± 0.22	0.39 ± 0.15
10	20	6.6 ± 1.0	7.8 ± 1.0	161 ± 47	93 ± 50	16.0 ± 2.8	14.9 ± 3.8	0.57 ± 0.22	0.29 ± 0.18
	15	6.6 ± 1.0	6.8 ± 0.9	165 ± 45	100 ± 51	15.9 ± 2.9	14.9 ± 2.9	0.59 ± 0.22	0.30 ± 0.19
	10	6.4 ± 1.0	6.5 ± 0.8	170 ± 46	109 ± 46	16.0 ± 2.7	15.2 ± 2.8	0.63 ± 0.19	0.31 ± 0.17
	5	6.0 ± 0.9	6.5 ± 0.9	178 ± 48	124 ± 47	15.9 ± 2.8	15.6 ± 2.8	0.67 ± 0.16	0.39 ± 0.22
	0	6.1 ± 0.8	6.4 ± 1.0	200 ± 36	132 ± 49	15.9 ± 2.8	15.8 ± 2.9	0.68 ± 0.17	0.46 ± 0.24
20	20	6.0 ± 0.6	7.2 ± 0.9	171 ± 40	125 ± 52	15.1 ± 3.2	15.7 ± 2.5	0.49 ± 0.18	0.26 ± 0.14
	15	6.0 ± 0.7	7.0 ± 1.0	176 ± 40	128 ± 50	15.8 ± 2.7	15.0 ± 3.6	0.53 ± 0.17	0.27 ± 0.13
	10	5.9 ± 0.5	6.9 ± 1.0	176 ± 48	138 ± 49	15.0 ± 3.2	15.9 ± 2.9	0.55 ± 0.18	0.31 ± 0.16
	5	5.7 ± 0.5	6.8 ± 1.0	203 ± 25	139 ± 49	15.9 ± 3.0	16.0 ± 2.7	0.69 ± 0.13	0.33 ± 0.15
	0	5.7 ± 0.3	6.5 ± 1.0	206 ± 24	142 ± 50	16.0 ± 2.8	15.9 ± 2.7	0.70 ± 0.14	0.34 ± 0.17

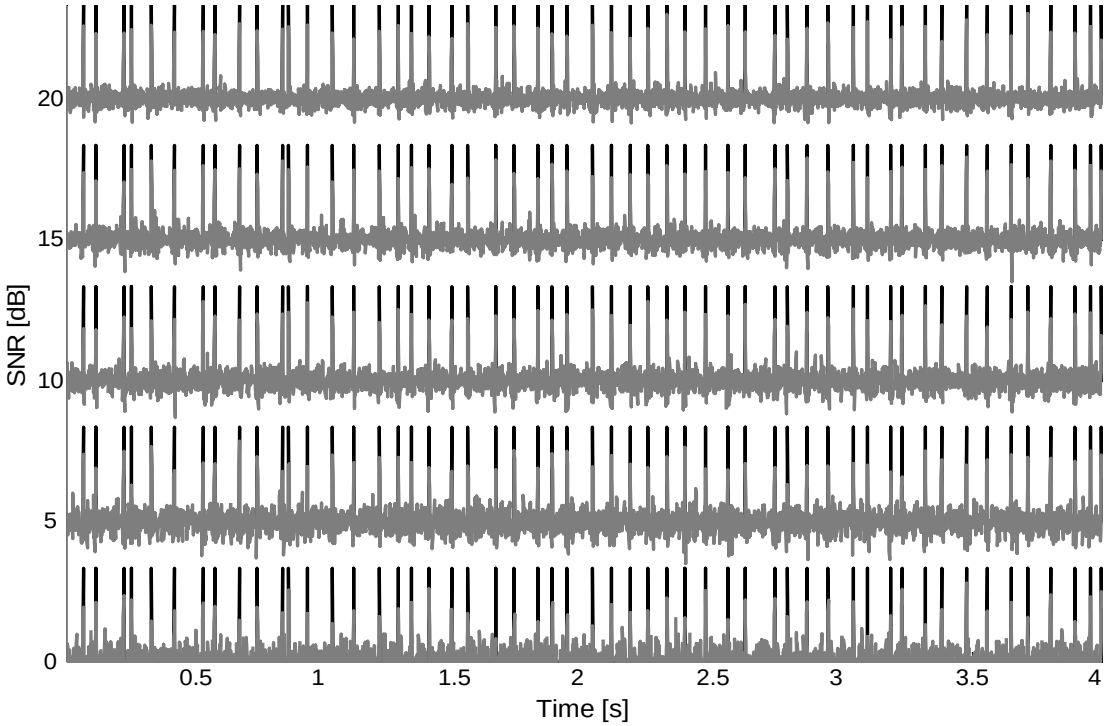


Figure 6.5: Original synthetic innervation pulse trains (black) and the pulse trains reconstructed by the IC method (grey) in the case of 5 active MUs (simulation 1, MU 3) and at SNR = 20, 15, 10, 5 and 0 dB.

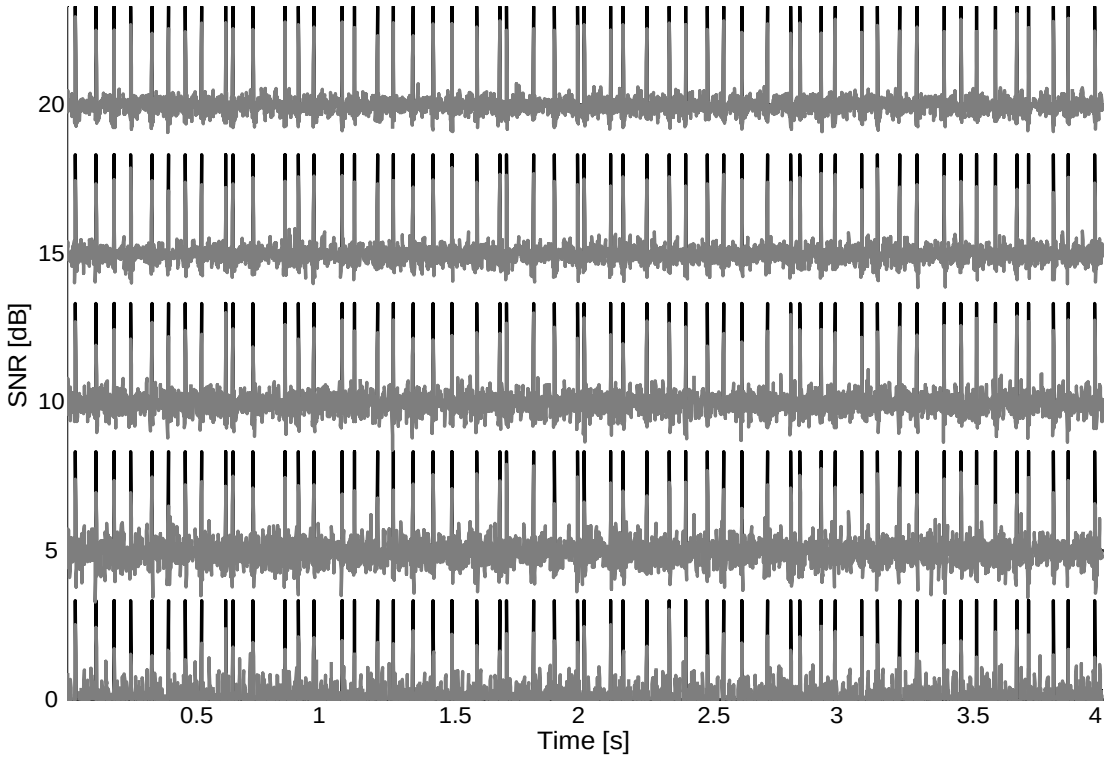


Figure 6.6: Original synthetic innervation pulse trains (black) and the pulse trains reconstructed by the IC method (grey) in the case of 10 active MUs (simulation 5, MU 6) and at SNR = 20, 15, 10, 5 and 0 dB.

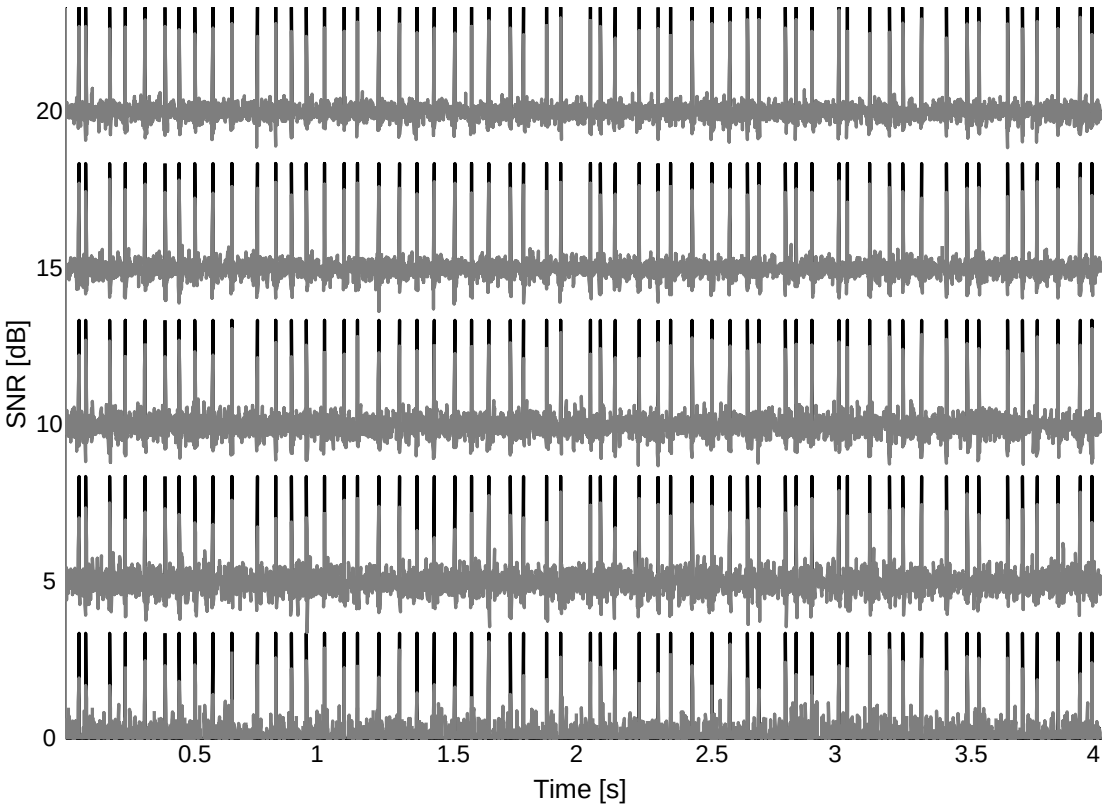


Figure 6.7: Original synthetic innervation pulse trains (black) and the pulse trains reconstructed by the IC method (grey) in the case of 20 active MUs (simulation 2, MU 16) and at SNR = 20, 15, 10, 5 and 0 dB.

Finally, let us evaluate the results of the spike triggered sliding window averaging technique, described in Section 5.1. Reconstructing the MUAPs form the synthetic signals, the rectangular window $\Psi(l)$ of length 50 samples was centred to the position of every identified pulse of the corresponding MU. Averaging 50 measurements, 50 different MUAPs of each MU were reconstructed. Using the norm of differences (6.1), the reconstructed MUAPs were numerically compared to their reference values. The results are presented in Table 6.6, while a few examples of the reconstructed MUAPs are depicted in Figs. 6.8, 6.9 and 6.10.

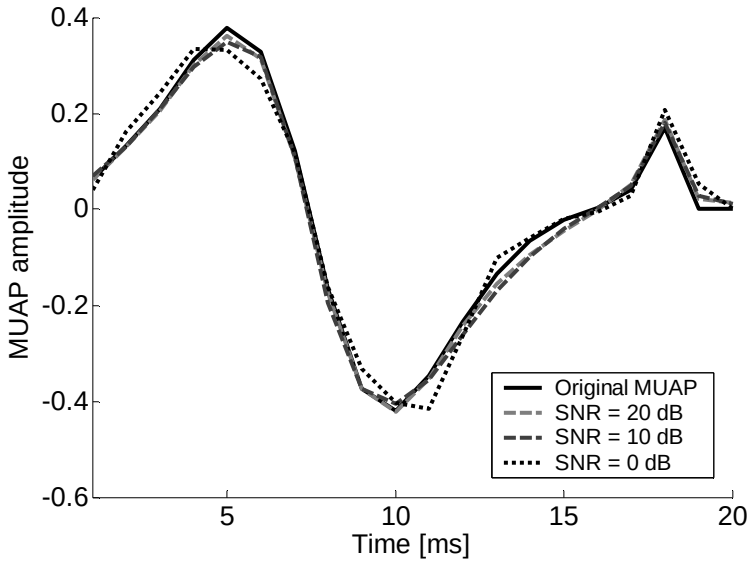


Figure 6.8: The MUAPs (contribution of MU 3 to the measurement (3,5)) reconstructed by the spike triggered sliding window averaging technique (390 averages) from the synthetic surface EMG signals of length 30720 samples in the case of 5 active MUs (simulation 1) and at SNR=20 dB (light grey dashed line), SNR=10 dB (dark grey dashed line), SNR=0 dB (black dotted line). Reference values are depicted black solid.

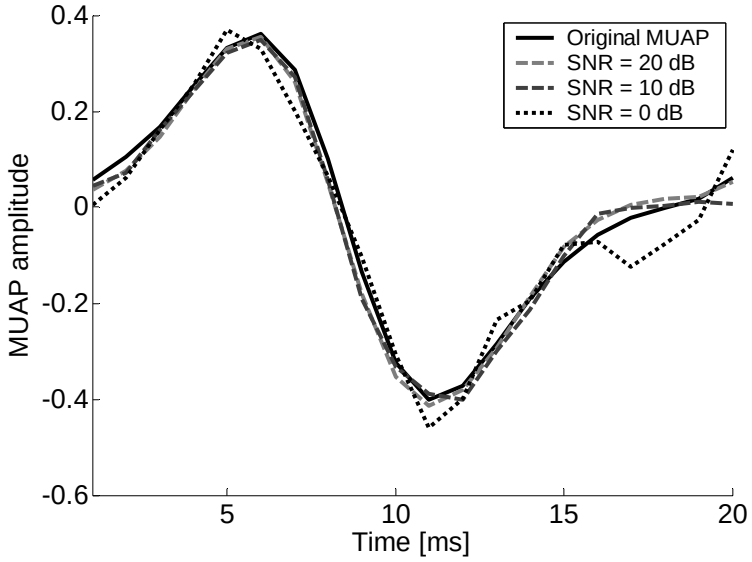


Figure 6.9: The MUAPs (contribution of MU 6 to the measurement (3,5)) reconstructed by the spike triggered sliding window averaging technique (386 averages) from the synthetic surface EMG signals of length 30720 samples in the case of 10 active MUs (simulation 5) and at SNR=20 dB (light grey dashed line), SNR=10 dB (dark grey dashed line), SNR=0 dB (black dotted line). Reference values are depicted black solid.

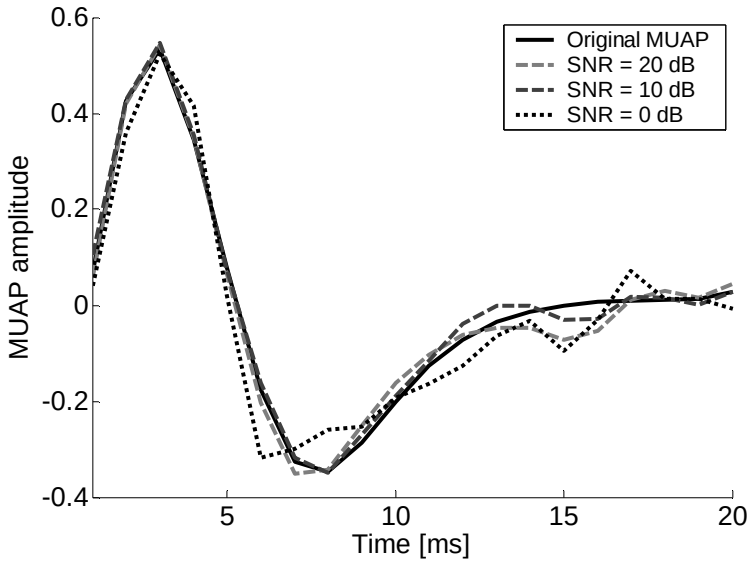


Figure 6.10: The MUAPs (contribution of MU 16 to the measurement (3,5)) reconstructed by the spike triggered sliding window averaging technique (418 averages) from the synthetic surface EMG signals of length 30720 samples in the case of 20 active MUs (simulation 2) and at SNR=20 dB (light grey dashed line), SNR=10 dB (dark grey dashed line), SNR=0 dB (black dotted line). Reference values are depicted black solid.

Table 6.6: The number of reconstructed MUs (mean ± standard deviation) and a comparison of the reconstructed MUAPs to their reference values (the 2-norm of the MUAP differences d_{ij} defined in (6.1)) versus the number of active MUs and SNR.

Number of active MUs	**SNR [*dB*]**	**Average number of reconstructed MUs**	**Average norm d_{ij} [%]**
5	20	4.7 ± 0.6	8.5 ± 5.3
	15	4.2 ± 0.6	7.5 ± 4.6
	10	3.9 ± 0.7	7.6 ± 4.7
	5	3.2 ± 0.6	7.2 ± 4.8
	0	2.0 ± 0.9	6.7 ± 5.1
10	20	8.3 ± 1.3	9.3 ± 5.4
	15	7.0 ± 1.6	8.8 ± 5.5
	10	6.4 ± 1.9	8.3 ± 5.4
	5	4.6 ± 1.7	9.3 ± 6.5
	0	2.6 ± 1.4	6.8 ± 4.9
20	20	10.4 ±1.4	13.9 ± 8.6
	15	8.6 ± 1.7	14.2 ± 9.3
	10	5.7 ± 1.3	12.7 ± 8.5
	5	3.5 ± 1.3	11.7 ± 9.7
	0	2.3 ± 1.9	11.3 ± 8.1

6.2 Results on real surface EMG signals

The experiments were conducted with signals from the dominant biceps brachii, recorded in Laboratorio di Ingegneria del Sistema Neuromuscolare (LISiN), Politecnico di Torino, Italy. Nine healthy male subjects of age 26.8 ± 2.2 years, height 179 ± 7 cm and weight of 72.1 ± 8.3 kg participated in our study. All subjects gave their informed consent. Surface EMG signals were detected by an array of 13×5 electrodes (without the four corner electrodes) of size 1×1 mm and of inter-electrode distance of 5 mm (Fig. 6.11).

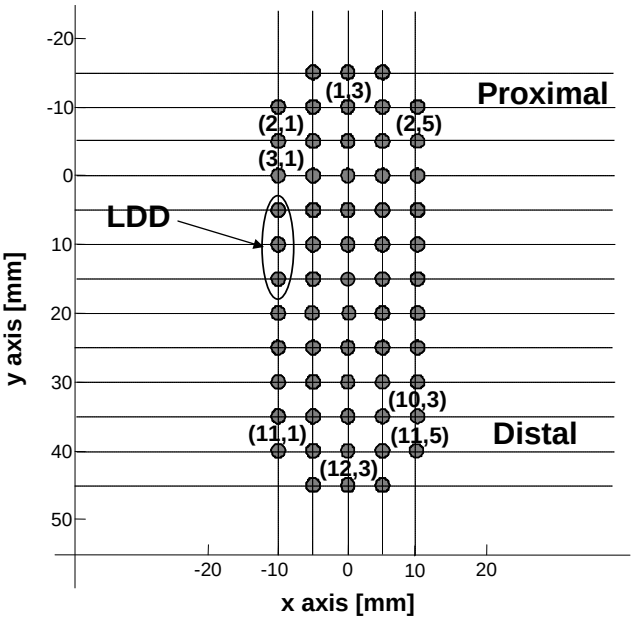


Figure 6.11: Array of electrodes (grey circles) used for detection of real surface EMG signals and its relative position with respect to the investigated muscle. The y axis is aligned with the muscle fibres, while its origin is lined up with the estimated location of the dominant innervation zone. Proximal and distal half of the muscle fibres are also indicated. The coordinates **(x,y)** below the corresponding electrodes denote their relative position in the array, while **LDD** exemplifies how three different electrodes comprise the longitudinal double differential spatial filter (source: LISiN, Politecnico di Torino, Italy).

The experimental protocol was designed by world-recognized bioengineering experts Prof. Roberto Merletti, Dr. Dario Farina and Marco Gazzoni and consisted of:

1. The dominant arm of the subject was placed into the isometric brace at 120° (Fig. 6.12).

2. Three five-second contractions at maximum voluntary contraction (MVC) force were performed separated by 2 minutes. Using the torque sensors (Fig. 6.12), the maximum contraction force was measured and averaged over all three measurements. Afterwards, a 5-minute rest was given to the subject.

3. The skin was slightly abraded with abrasive paste and moistened to improve the electrode-skin contact.

4. The location of the innervation zone in the dominant biceps brachii of the subject was determined from the travelling AP detected during voluntary low force muscle contraction by the linear array of 16 electrodes of size 10×1 *mm* and the inter-electrode distance of 10 *mm*. Afterwards, the linear array of electrodes was removed and the skin remoistened.

5. The array of 61 electrodes (Fig. 6.11) was placed over the distal half of dominant biceps brachii with its third electrode row centred over the estimated innervation zone and columns aligned with the muscle fibres.

6. Surface EMGs were recorded during the 30 *s* long isometric voluntary muscle contraction sustained at 5 % and 10 % of MVC. After each contraction the subject relaxed for 5 minutes. The contraction force was measured by the torque sensor and displayed on the oscilloscope to provide the visual feedback to the subjects.

The detected signals were amplified (gain set to 10000) by a 64-channel EMG amplifier (LISiN; Prima Biomedical & Sport, Treviso, Italy), band-pass filtered (-3 dB bandwidth, 10 Hz – 500 Hz), and sampled at 2500 *Hz* by 12-bit A/D converter. Longitudinal double-differential recording technique was used with the adjacent electrode pairs along the columns in the electrode array (Fig. 6.11), resulting in 51 SEMG recordings. The noise and movement artefacts were visually controlled and reduced by applying water to the skin surface. Before any further processing, all the measurements were digitally filtered to suppress the power-line interference.

In the case of real SEMG signals, of course, there are no reference innervation trains available. Hence, the performance must be evaluated by other, indirect measures. The ones used in this study were:

- the regularity of the reconstructed firing patterns,
- the nature of IPI variability, and
- the shape of reconstructed MUAPs.

Supposing short-term low level contractions, the MU innervation trains are well known to follow regular, relatively slowly changing patterns. IPI irregularities are often modelled as realizations of a Gaussian random variable whose maximum value in normal conditions should not exceed the 20 % limit of the mean IPI [56, 60, 90]. Any larger irregularity (in the sense of discussion above), therefore, must be taken as an early warning of possibly wrong decomposition. Similarly, the reconstructed innervation pulses can be evaluated with respect to the IPI distribution. Several statistical tests (Jarque-Bera, Kolmogorov-Smirnov, Lilliefors) can be applied to test the probability that the obtained IPI realisation follow Gaussian distribution (so called *p*-value in the hypothesis test terminology). Possible non-stationarities of the MU firing rate can be compensated by modelling the mean IPI as a function of time.

Figure 6.12: The isometric brace used to keep the elbow angle constant at 120 degrees during the acquisition of surface EMG signals. Torque sensors are located at the joint of the rigidly connected parts of isometric brace (source: LISiN, Politecnico di Torino, Italy).

The last criterion is based on the MUAP shape and is the most intuitive one. Double-differential recording technique implies relatively firm limitations to the shapes of MUAPs. Variability is only expected in the peak-to-peak value, the length of MUAPs and minor shape details (mainly in the MUAP's tail regions). Furthermore, spatially adjoining electrodes are expected to detect very similar MUAP shapes. Finally, by examining the MUAP's shape in different channels, the relative positions of the detected MU with respect to the array of electrodes, and also the travelling of APs should be traceable.

Decomposing the real signals by three different decomposition techniques the results can also be mutually compared. For the clarity reasons, it is then appropriate to first provide the results of the most efficient decomposition method. Afterwards, the remaining decomposition techniques can be evaluated with respect to the superior method. The results on the synthetic signals proved the IC method by far outperforms the method based on TF distributions as well as the HOS-based method. The similar results can be expected also on the real signals. Therefore, we will reverse the order of subsections (with respect to the subsection in Section 6.1). We will first evaluate the results of the IC method. The method based on TF distributions will be evaluated in Subsection 6.2.2, while the results of the HOS-based decomposition method will be described in Subsection 6.2.3. Evaluating the results of the last two methods, we will rely upon the results of the IC method. Taking into account all the described indirect measures, the number of successfully reconstructed MU trains will be considered the most significant performance criterion.

6.2.1 Inverse correlation based method

Let us first evaluate the results of the IC method. Decomposing the real signals we followed the procedure described in Subsection 6.1.3. All 51 measurements of surface EMG were used in our experiments. The measurements with 5 % muscle contraction were extended by the factor K=10 (9 delayed repetitions of each measurement were added) resulting in 510 extended measurements. In the case of 10 % muscle contraction, the measurements were extended by the factor K=15, yielding 765 extended measurements. The average length of MUAPs was estimated to 20 ms (50 samples). Supposing the over-determined MIMO system, the assumed number of sources yields 8 in the case of 5 % contraction, and 11 in the case of 10 % muscle contraction. Although the number of active MUs can be expected much higher, there are at least three crucial reasons for limiting the number of extended measurements. The first and the most obvious is the computational complexity. The IC method is based on a relatively high number of matrix multiplications and their computational complexity increases with the square of the matrix size. The second reason originates in the used double-differential recording technique. The latter will filter out the superficial MUs. Due to their mutual resemblance the contributions of deeper MUs to different electrodes constituting the same spatial filter will be cancelled. The third and the final reason is obvious form the results on the synthetic signals. They proved the IC method is quite capable of decomposing under-determined MIMO systems. Hence, the assumptions of 9 and 12 active MUs, respectively, should be used only as an orientation and are not restrictive.

Evaluate now the reconstructed innervation pulse trains. It is well known that the IPI variability is normally limited to 10-20 % of the mean IPI. Furthermore, supposing low muscle contraction levels only slow changes of MUs' firing rates in time are possible. The reconstructed innervation pulse trains are, hence, expected to follow relatively strict firing patterns. Calculating the IPI variability of the reconstructed pulse trains over a longer time interval, the slow changes of firing rate must be taken into consideration. In our experiments, the reconstructed innervation pulses were first divided into 6 epochs of equal length. In each epoch the mean firing rate and the IPI variability were calculated. The results were averaged over all 6 epochs and are depicted in Tables 6.7 and 6.8.

The results in Tables 6.7 and 6.8 prove all the reconstructed innervation pulse trains agree with theoretical expectations. The firing rates of the reconstructed MUs are estimated between 8 and 16 Hz. The IPI variability of the reconstructed pulse trains is concentrated around the mean of 10 % and rarely reaches its upper limit of 20 %. Statistical moments enable only limited conclusions, but s detailed analysis quickly reveals the regularity of the reconstructed innervation pulse trains. Examples of the reconstructed innervation pulse trains are illustrated in Figs. 6.13 to 6.16. The IPI variability is depicted in Figs. 6.17, 6.18 and 6.19.

Table 6.7: The number of detected innervation trains, their average firing rate (mean ± standard deviation), and variability of inter-pulse interval (with respect to the mean firing rate of the corresponding MU). The innervation pulse trains were reconstructed from 30 s long SEMG signals recorded during the isometric 5 % MVC measurements of the dominant biceps brachii of nine healthy male subjects (age 26.8 ± 2.2 years, height 179 ± 7 cm, weight 72.1 ± 8.3 kg).

Subject	Number of reconstructed MUs	Average firing rate [Hz]	Relative IPI variability [%]
1	4	12.1 ± 1.7	10.7 ± 2.8
2	4	15.1 ± 1.5	11.3 ± 1.2
3	2	11.4 ± 1.7	15.0 ± 2.5
4	2	11.4 ± 0.1	10.3 ± 2.8
5	3	14.0 ± 0.3	18.1 ± 4.7
6	5	12.8 ± 1.8	10.6 ± 2.0
7	3	13.0 ± 0.3	14.5 ± 0.7
8	5	9.7 ± 0.9	8.9 ± 2.6
9	2	12.9 ± 0.7	12.0 ± 5.5

Table 6.8: The number of detected innervation trains, their average firing rate (mean ± standard deviation), and variability of inter-pulse interval (with respect to the mean firing rate of the corresponding MU). The innervation pulse trains were reconstructed from 30 s long SEMG signals recorded during the isometric 10 % MVC measurements of the dominant biceps brachii of nine healthy male subjects (age 26.8 ± 2.2 years, height 179 ± 7 cm, weight 72.1 ± 8.3 kg).

Subject	Number of reconstructed MUs	Average firing rate [Hz]	Relative IPI variability [%]
1	5	10.8 ± 1.4	15.3 ± 4.3
2	10	13.9 ± 2.4	12.9 ± 3.6
3	5	13.2 ± 1.6	18.8 ± 2.5
4	3	9.5 ± 3.8	14.5 ± 5.6
5	4	11.19 ± 4.9	16.8 ± 2.2
6	8	14.13 ± 1.5	18.7 ± 4.0
7	7	14.0 ± 1.3	19.0 ± 2.9
8	8	10.6 ± 1.3	10.7 ± 3.7
9	6	13.5 ± 1.9	15.7 ± 1.9

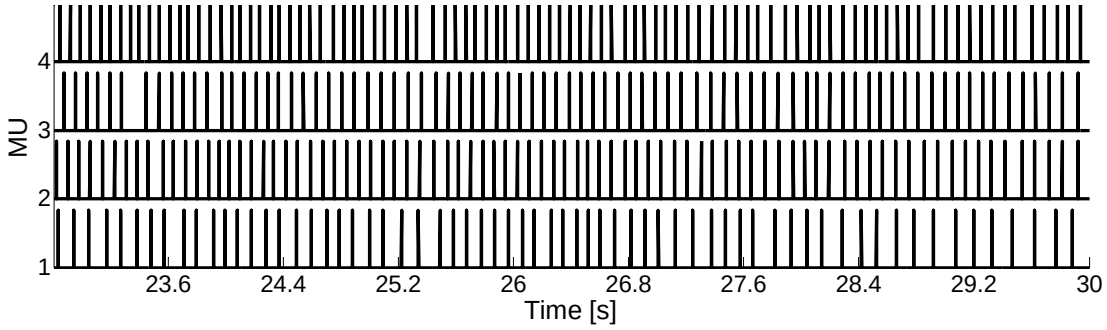


Figure 6.13: The MU innervation trains reconstructed by the IC method from 30 s long real SEMG signal. SEMG was recorded during an isometric 5 % MVC measurement of the dominant biceps brachii of subject 1 (age 26 years, height 176 cm, weight 68 kg). Only a part of the reconstructed innervation trains is depicted.

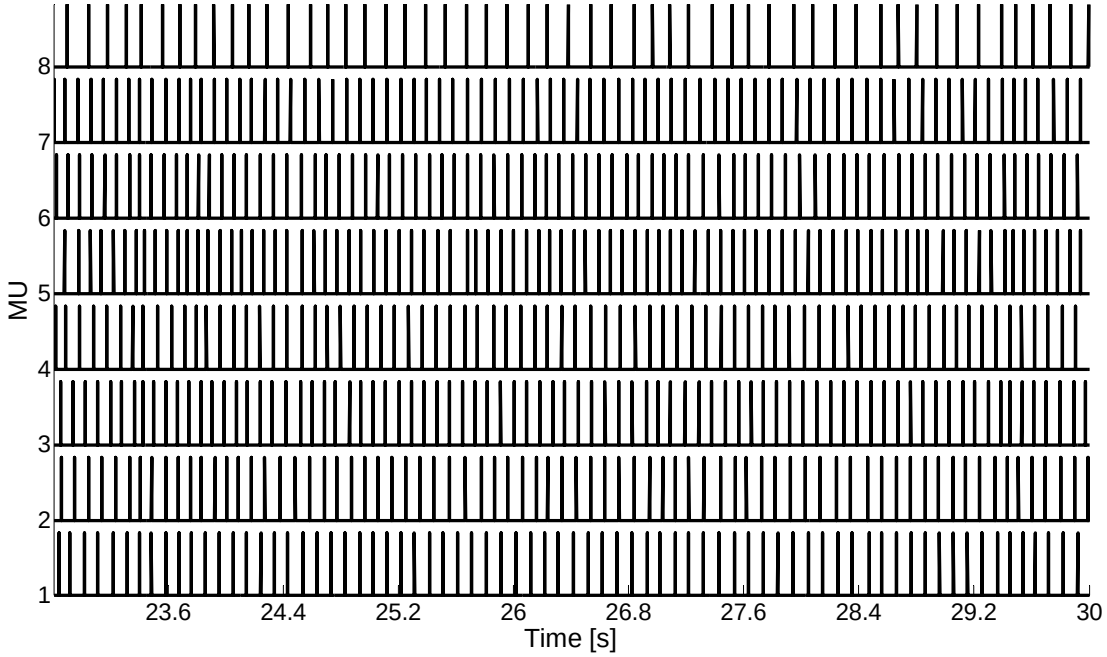


Figure 6.14: The MU innervation trains reconstructed by the IC method from 30 s long real SEMG signal. SEMG was recorded during an isometric 10 % MVC measurement of the dominant biceps brachii of subject 8 (age 31 years, height 180 cm, weight 65 kg). Only a part of the reconstructed innervation trains is depicted.

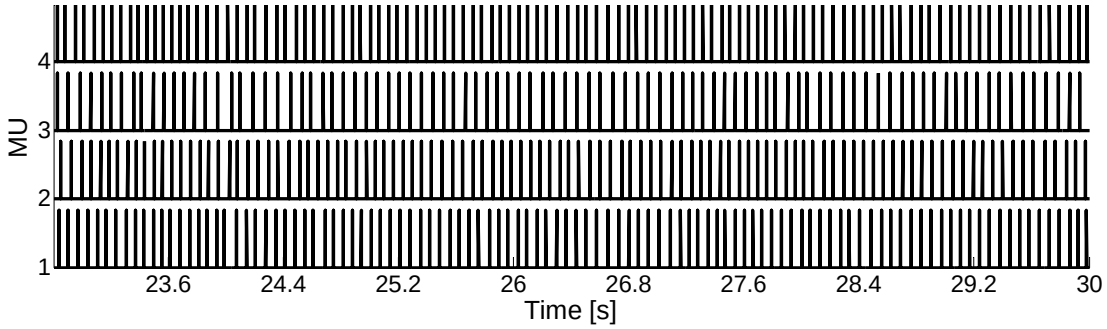


Figure 6.15: The MU innervation trains reconstructed by the IC method from 30 s long real SEMG signal. SEMG was recorded during an isometric 5 % MVC measurement of the dominant biceps brachii of subject 2 (age 25 years, height 170 cm, weight 63 kg). Only a part of the reconstructed innervation trains is depicted.

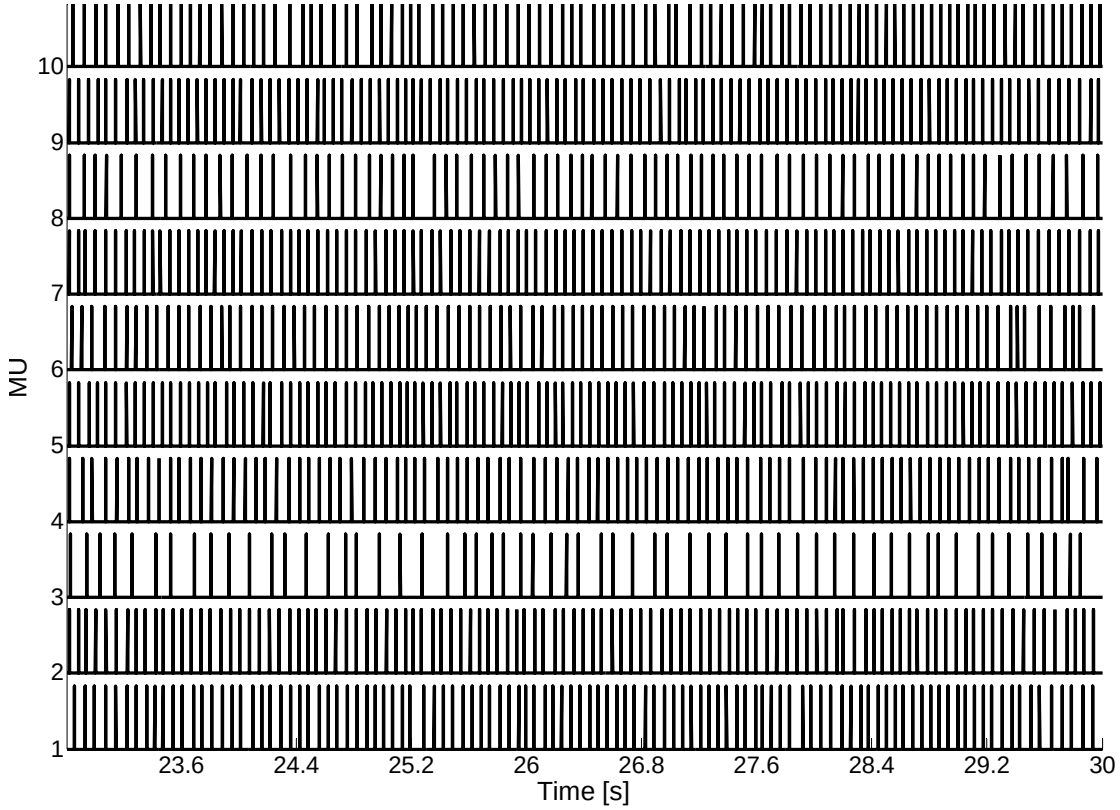


Figure 6.16: The MU innervation trains reconstructed by the IC method from 30 s long real SEMG signal. SEMG was recorded during an isometric 10 % MVC measurement of the dominant biceps brachii of subject 2 (age 25 years, height 170 cm, weight 63 kg). Only a part of the reconstructed innervation trains is depicted.

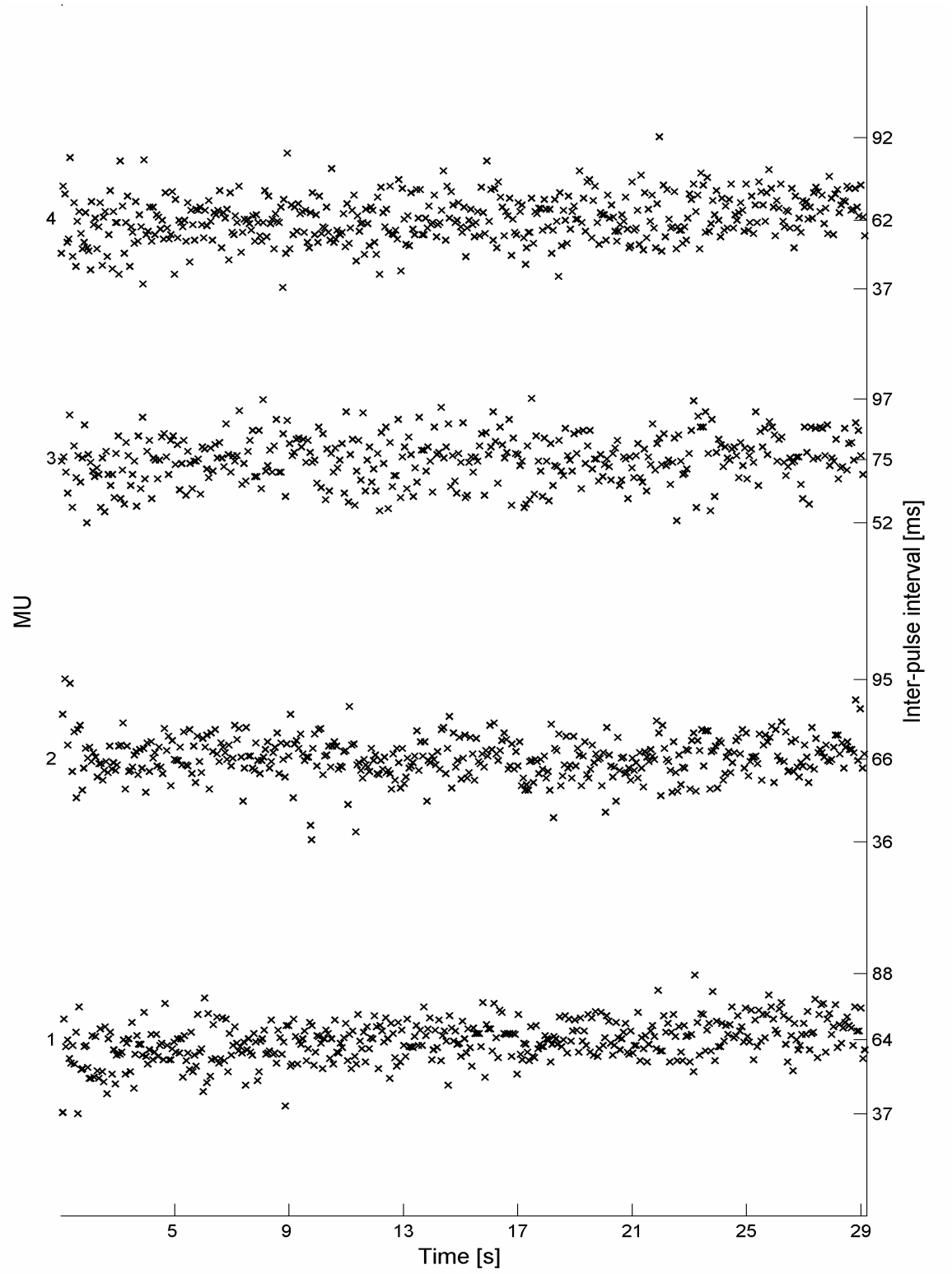


Figure 6.17: Inter-pulse interval in innervation pulse trains reconstructed by the IC method from 30 s long real SEMG signal. SEMG was recorded during an isometric 5 % MVC measurement of the dominant biceps brachii of subject 2 (age 25 years, height 170 cm, weight 63 kg).

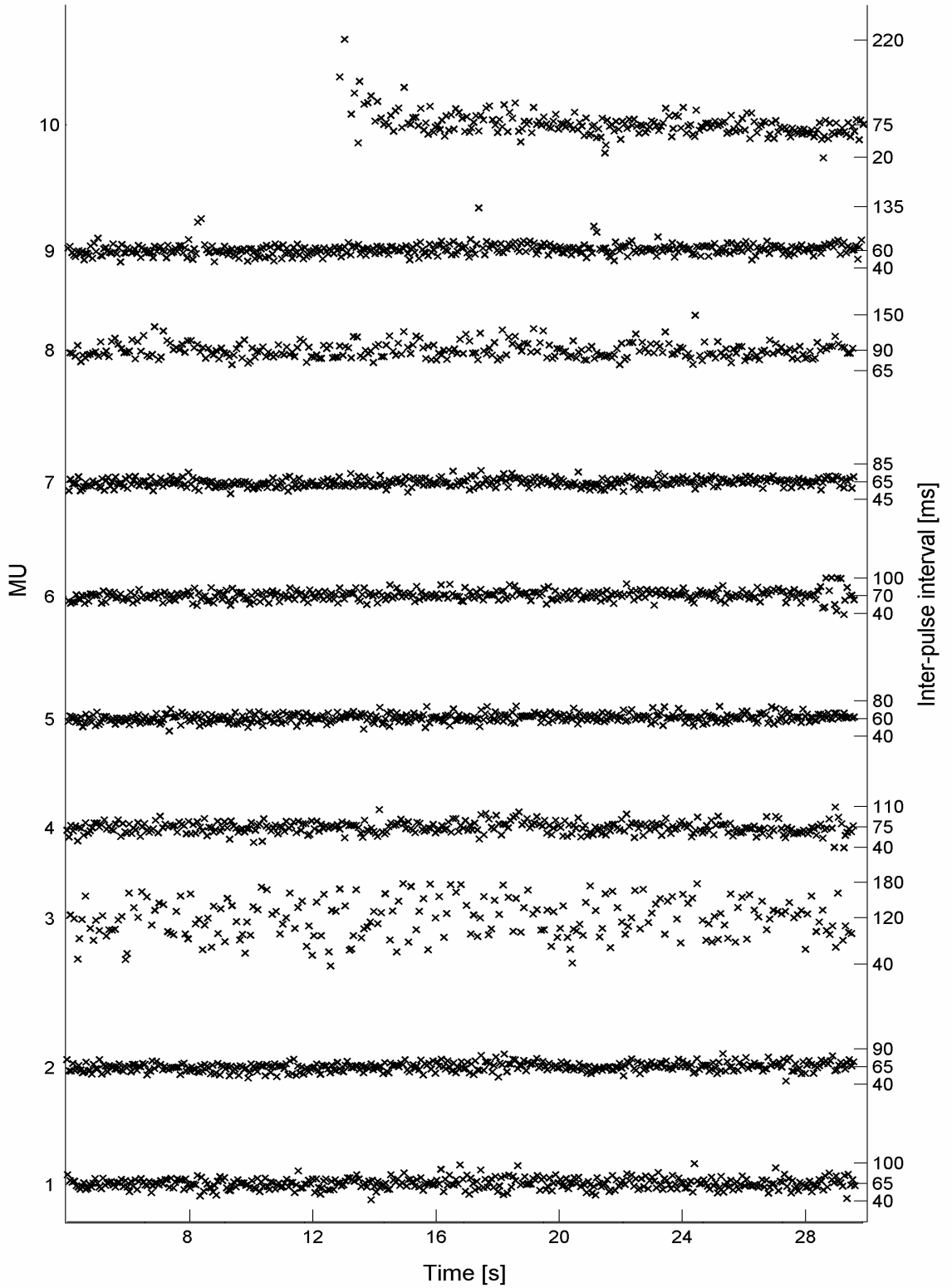


Figure 6.18: Inter-pulse interval in innervation pulse trains reconstructed by the IC method from 30 s long real SEMG signal. SEMG was recorded during an isometric 10 % MVC measurement of the dominant biceps brachii of subject 2 (age 25 years, height 170 cm, weight 63 kg).

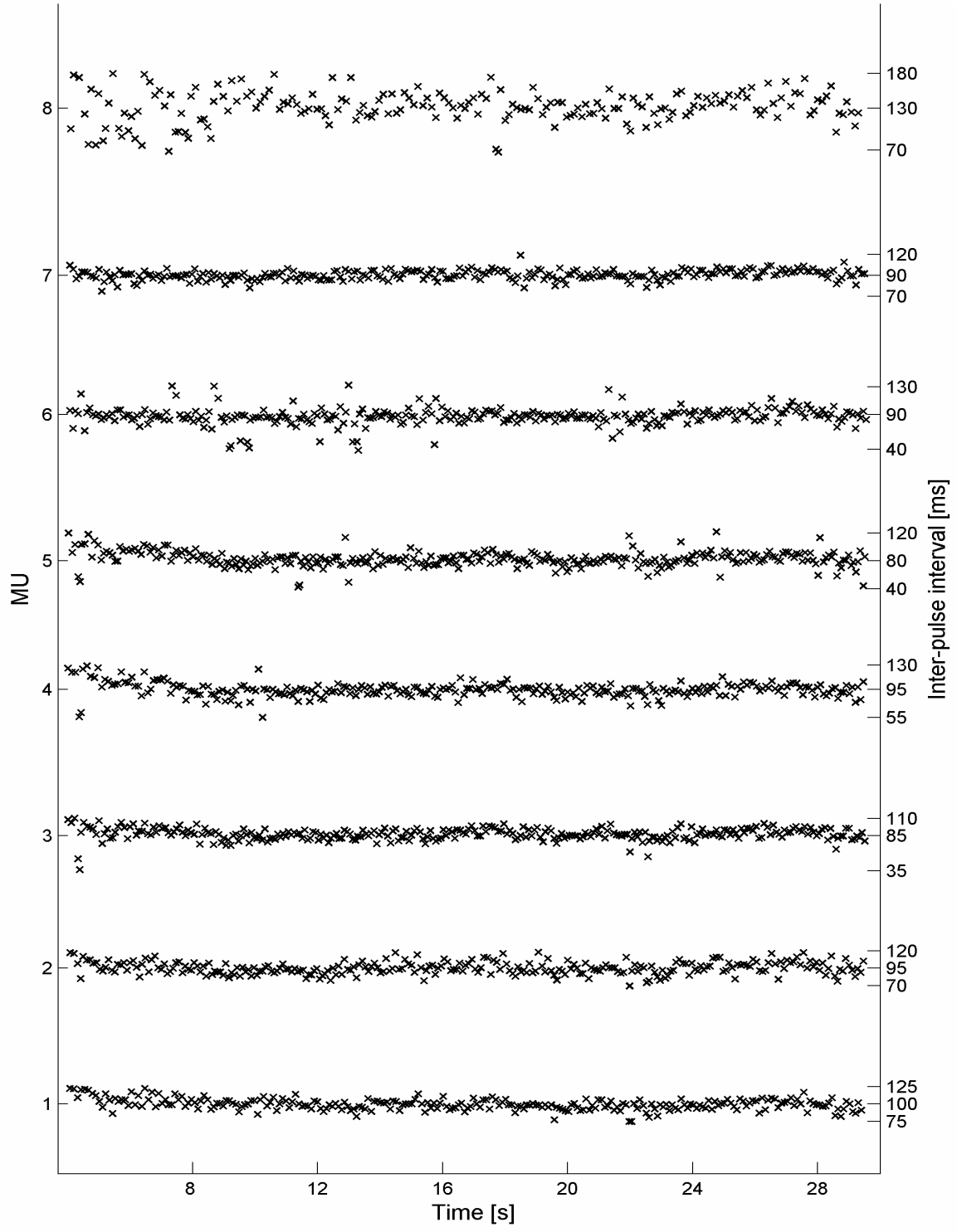


Figure 6.19: Inter-pulse interval in innervation pulse trains reconstructed by the IC method from 30 s long real SEMG signal. SEMG was recorded during an isometric 10 % MVC measurement of the dominant biceps brachii of subject 8 (age 31 years, height 180 cm, weight 65 kg).

Figs. 6.13 to 6.19 demonstrate how IPI in the reconstructed pulse trains obey the physiologically induced restrictions. They also prove almost all pulses were identified. Missed pulses in Figs. 6.17 to 6.19 correspond to unusually high IPIs and we can quickly conclude there number is negligible. Depending on the MU firing rate, the number of the identified pulses varies between 300 and 500, while the estimated number of missed pulses varied between 5 and 15.

Finally, the reconstructed innervation pulse trains were tested against the hypothesis that the IPIs are normally distributed. Several statistical tests (Jarque-Bera, Kolmogorov-Smirnov, Lilliefors) were applied. In order to increase the size of the tested sample the time changes in MU firing rates were neglected. As a result, the tested sample included from 300 to 500 elements. None of the statistical tests applied could reject the hypothesis of Gaussian IPI variability. Although this fact can not be taken as a proof that the IPI truly follows the Gaussian distribution, it can be understood as a strong suggestion that the reconstructed results do not disagree with expectations. The quantile-quantile plots of the sample quantiles of reconstructed IPIs versus theoretical quantiles from a normal distribution are depicted in Figs. 6.20 and 6.21. Note that when the distribution of IPI is normal, the plot should be linear.

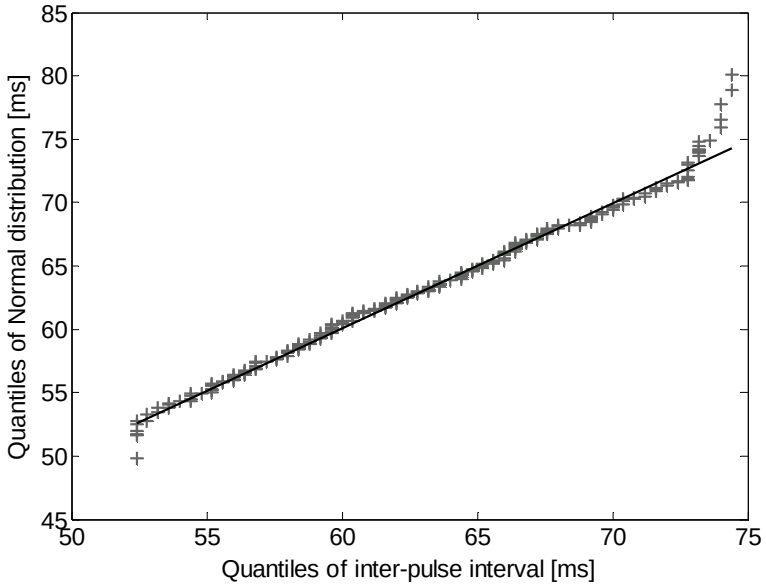


Figure 6.20: The quantile-quantile plots of the sample quantiles of reconstructed IPIs (grey +) versus theoretical quantiles from a normal distribution with the same mean and standard deviation (black solid). The IPIs were calculated from the second reconstructed MU of subject 2 (age 25 years, height 170 cm, weight 63 kg). Surface EMG was recorded during an isometric 5 % MVC measurement of the dominant biceps brachii.

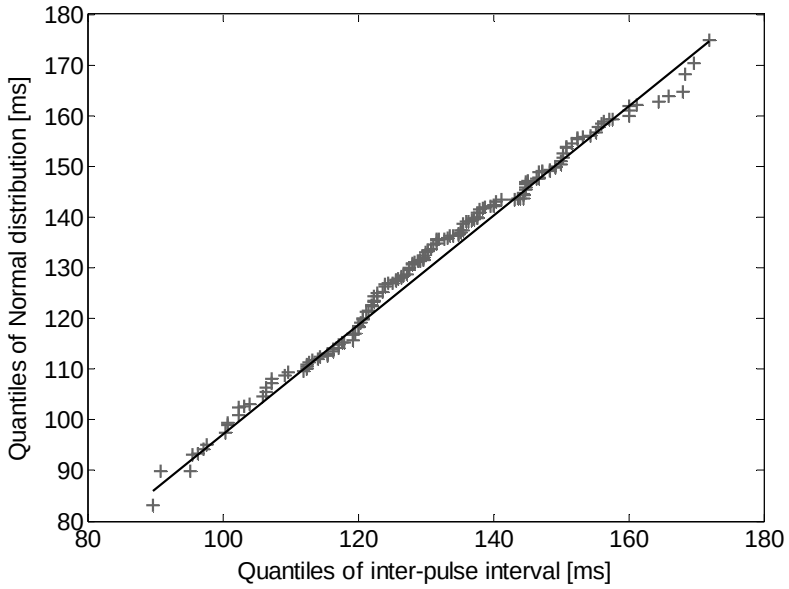


Figure 6.21: The quantile-quantile plots of the sample quantiles of reconstructed IPIs (grey +) versus theoretical quantiles from a normal distribution with the same mean and standard deviation (black solid). The IPIs were calculated from the second reconstructed MU of subject 8 (age 31 years, height 180 cm, weight 65 kg). Surface EMG was recorded during an isometric 10 % MVC measurement of the dominant biceps brachii.

The last criterion applied was the shape of the MUAPs reconstructed by the spike triggered sliding window averaging technique. The length of rectangular window was set to 100 samples, while the pulses reconstructed by the IC method were used as triggering spikes. The results are depicted in Figs. 6.22 to 6.25. The propagation of MUAPs along the columns of the electrode array and, hence, the location of the MU innervation zone can be recognized clearly. Finally, spatially adjoining electrodes detected very similar MUAP shapes, which is yet another indirect proof of accurately reconstructed MU innervation pulse trains.

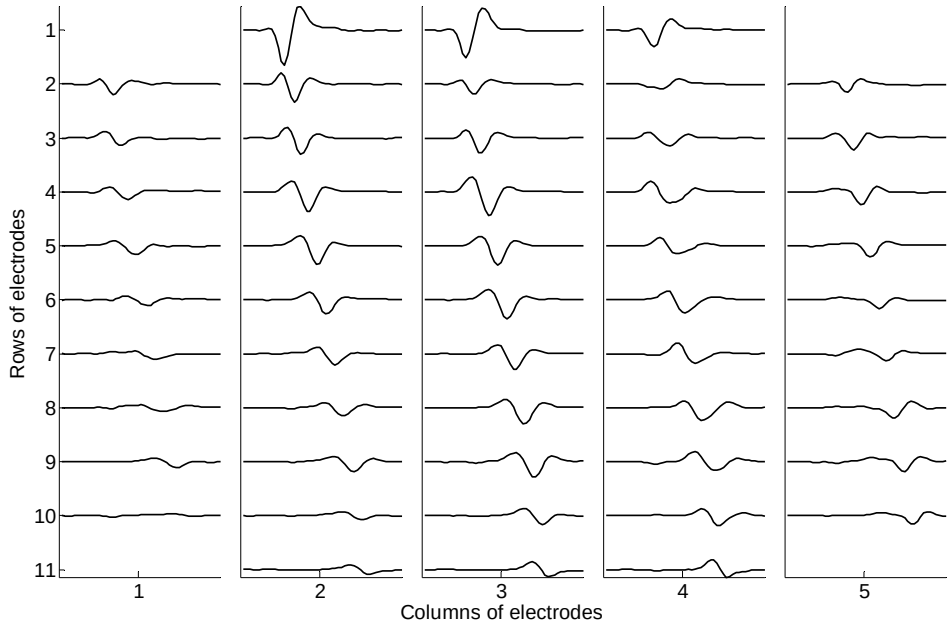


Figure 6.22: The MUAPs corresponding to MU 4 reconstructed by the spike triggered sliding window averaging technique (473 averages) from 30 s long SEMG signals. The MUAPs are depicted with respect to the relative position of the corresponding spatial filter. Surface EMG signals were recorded during an isometric 5 % MVC measurement of the dominant biceps brachii of subject 2 (age 25 years, height 170 cm, weight 63 kg).

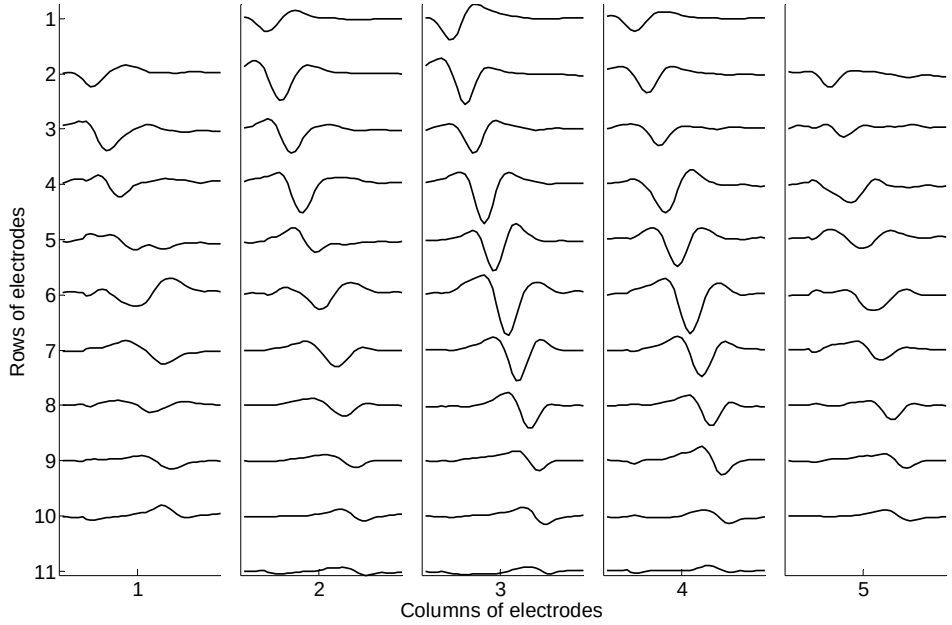


Figure 6.23: The MUAPs corresponding to MU 1 reconstructed by the spike triggered sliding window averaging technique (293 averages) from 30 s long SEMG signals. The MUAPs are depicted with respect to the relative position of the corresponding spatial filter. Surface EMG signals were recorded during an isometric 5 % MVC measurement of the dominant biceps brachii of subject 8 (age 31 years, height 180 cm, weight 65 kg).

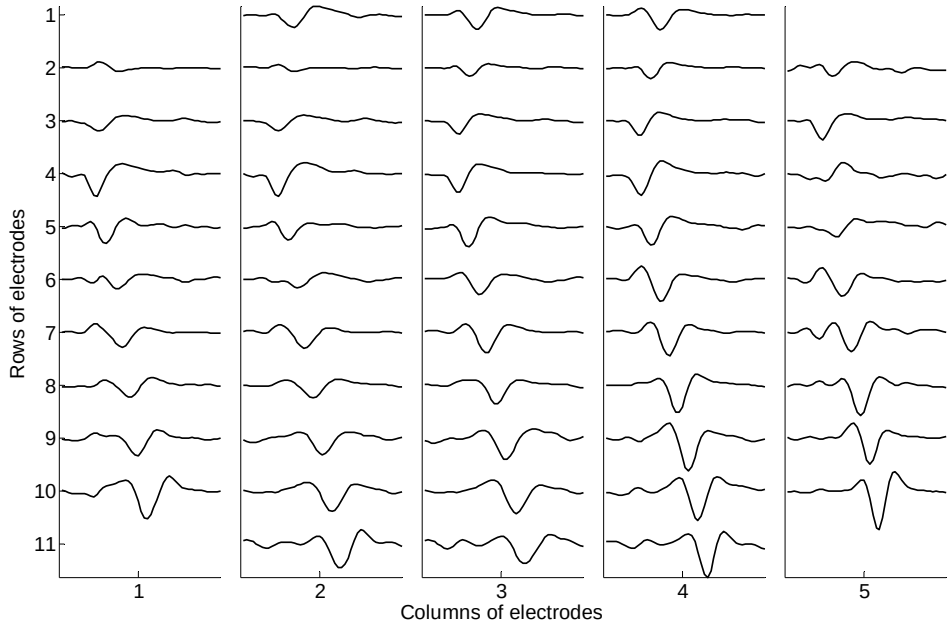


Figure 6.24: The MUAPs corresponding to MU 2 reconstructed by the spike triggered sliding window averaging technique (210 averages) from 30 s long SEMG signals. The MUAPs are depicted with respect to the relative position of the corresponding spatial filter. Surface EMG signals were recorded during an isometric 10 % MVC measurement of the dominant biceps brachii of subject 1 (age 26 years, height 176 cm, weight 68 kg).

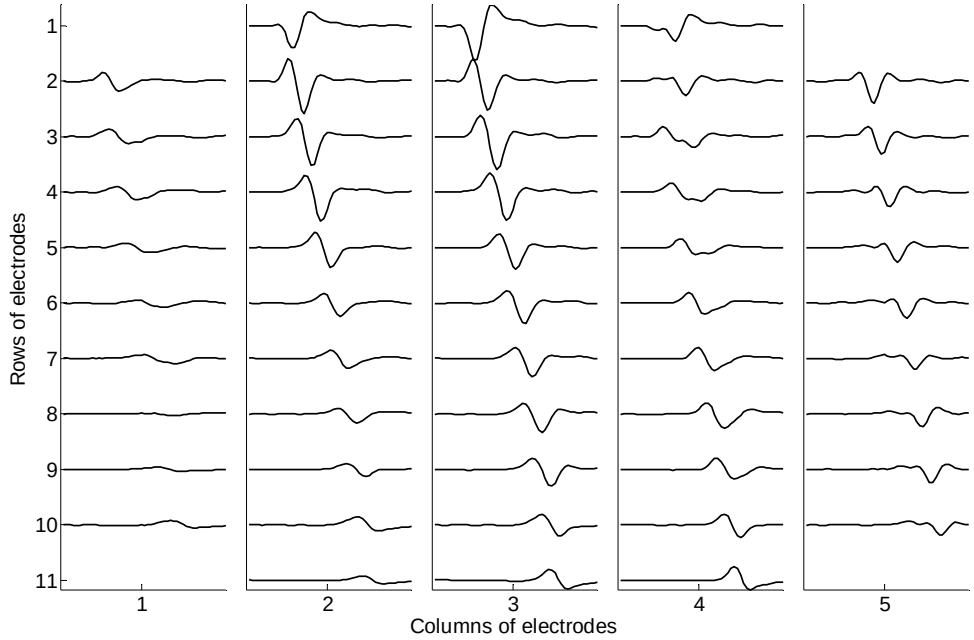


Figure 6.25: The MUAPs corresponding to the MU 9 reconstructed by the spike triggered sliding window averaging technique from 30 s long SEMG signals. The MUAPs are depicted with respect to the relative position of the corresponding spatial filter. Surface EMG signals were recorded during an isometric 10 % MVC measurement of the dominant biceps brachii of subject 2 (age 25 years, height 170 cm, weight 63 kg).

The results in Tables 6.7 and 6.8 show the average number of reconstructed MUs reached 3.4 ± 1.3 (mean ± standard deviation) at 5 % MVC and 6.2 ± 2.2 at 10 % MVC. The exact number of active MUs is, of course, unknown. Nevertheless, we can approximately estimate the percentage of the information that was extracted from the surface EMG signals. The coarse estimation is given by the ratio between the energy of the reconstructed MUAP sequences and the energy of measurements. In our experiments, the MUAP sequences were constructed by convolving the reconstructed innervation pulse trains by the estimated MUAPs. All the MUAP sequences reconstructed from the same measurement were first summed together. The energy of the summation was compared to the energy of the corresponding measurement. Finally, the energy ratios were averaged over all measurements of the same muscle contraction. The results are depicted in Table 6.9.

The energy ratios show surprising results. According to Table 6.9, the IC method was significantly more efficient in the case of 10 % muscle contraction. At 10 % MVC, the IC method identified the origin of 70 % of energy, while at 5 % MVC the portion of the identified energy dropped to 45 %. This outcome disagrees with the theory and also with the results on the synthetic signals. They all predict the IC method is more efficient in the case of fewer sources. The observed phenomenon can be partially explained with the influence of noise. Namely, the measurements that were recorded at 5 % MVC have significantly lower energy than the measurements recorded at 10 % MVC (the average energy ratio was 65 ± 10 %). It is therefore possible that approximately the same noise contributed much bigger relative portion to the measurements at 5 % MVC than to the measurements at 10 % MVC. This hypothesis remains open for future investigations.

Table 6.9: The ratio between the energy of the reconstructed MUAP sequences and the energy of the surface EMG measurements (mean ± standard deviation). Surface EMG signals were recorded during an isometric 5 % and 10 % MVC measurement of the dominant biceps brachii of nine healthy male subjects (age 26.8 ± 2.2 years, height 179 ± 7 cm, weight 72.1 ± 8.3 kg).

Subject	**The energy ratio between the reconstructed MUAP sequences and 5 % MVC measurements [%]**	**The energy ratio between the reconstructed MUAP sequences and 10 % MVC measurements [%]**
1	36.0 ± 12.0	53.0 ± 8.0
2	77.7 ± 13.4	93.8 ± 14.7
3	56.7 ± 18.5	67.2 ± 10.2
4	43.7 ± 17.3	55.0 ± 30.0
5	35.7 ± 16.3	67.4 ± 17.9
6	46.4 ± 14.3	78.6 ± 11.7
7	54.0 ± 21.4	69.8 ± 7.0
8	35.0 ± 13.0	74.9 ± 16.8
9	43.7 ± 8.7	82.6 ± 6.9

6.2.2 Method based on time-frequency distributions

In the previous subsection we evaluated the results obtained by the IC method. Applying the knowledge about the properties of MUs, we provided various indirect evidences for the accuracy of the reconstructed innervation pulse trains. The efficiency of the IC method was also demonstrated on the synthetic signals. After such extensive verification we can readily trust its results. On the other hand, the method based on TF distributions proved less reliable on the synthetic signals. Its sensitivity to the noise and to non-orthogonal sources was also predicted in theory. Proving its accuracy in the same way as for the IC method would be very time-consuming and meaningless. We will rather rely on the results of the IC method and compare the performances of both methods. For the sake of clarity, we will only focus on the main decomposition results. All possible differences between the both methods will be discussed in detail in the next chapter.

Following the example of the IC method, we first supposed the number of active MUs limited to 8 in the case of 5 % MVC measurements, and to 11 in the case of 10 % MVC measurements. Estimating the average MUAP length to 50 samples, the measurements were extended by the factor *K=10* at 5 % MVC, and *K=15* at 10 % MVC. As result the number of extended sources increased to 472 and 765 for 5 % and 10 % MVC measurements, respectively.

The extended measurements entered the same decomposition procedure as in the case of synthetic signals. First, the correlation matrix of extended measurements was used to estimate the whitening matrix **B** and the noise variance σ^2. The extended measurements were whitened and pseudo Wigner-Ville STFD matrices $\mathbf{D}_{\overline{\mathbf{zz}}}(n, f)$ calculated. Using the criterion (4.21), all $\mathbf{D}_{\overline{\mathbf{zz}}}(n, f)$ matrices corresponding to the single-autoterms of the same source were first grouped into different sets, averaged and, afterwards, joint-diagonalized to estimate the unitary mixing matrix **U**. The sources were reconstructed by applying the Moore-Penrose pseudoinverse of the estimated mixing matrix to the extended measurements. Finally, the reconstructed repetitions of the same source were aligned in time and averaged to produce the final estimates. The exact decomposition procedure is described in Subsection 6.1.1.

Comparing the pulse trains only those reconstructed trains were taken into consideration, which contain less then 30 % of misplaced pulses referring to the results obtained by the IC method. All other reconstructions were tested against the regularity of IPI in order to exclude the possibility of discarding a novel MU (the MU which was not discovered by the IC method). The results are summarized in Tables 6.10 and 6.11. As expected from the results on synthetic signals no new MUs were identified. Some of the reconstructed innervation pulse trains are compared with the results of the IC method and depicted in Figs. 6.26 and 6.27.

Table 6.10: The number of innervation pulse trains reconstructed by the method based on TF distributions and their comparison to the results of the IC method: the percentage of the pulses, identified by both methods with respect to the number of pulses identified by the IC method (true positive statistics), and the percentage of the pulses that were identified only by the method based on TF distributions with respect to the number of all identified pulses (false positive statistics). The innervation pulse trains were reconstructed from 30 s long SEMG signals recorded during the isometric 5 % MVC measurements of the dominant biceps brachii of nine healthy male subjects (age 26.8 ± 2.2 years, height 179 ± 7 cm, weight 72.1 ± 8.3 kg).

Subject	Number of reconstructed MUs	Percentage of the pulses identified by both methods [%]	Percentage of pulses identified only by the method based on TF distributions [%]
1	4	90.9 ± 6.3	7.8 ± 6.8
2	3	98.3 ± 1.4	1.6 ± 1.5
3	1	95.7 ± 0.0	9.4 ± 0.0
4	2	85.0 ± 1.7	24.8 ± 2.1
5	2	71.3 ± 0.7	8.9 ± 2.0
6	3	88.4 ± 6.6	11.4 ± 12.6
7	3	96.5 ± 2.1	6.6 ± 5.5
8	2	77.0 ± 3.0	3.6 ± 4.0
9	2	90.1 ± 8.0	16.8 ± 12.5

Table 6.11: The number of innervation pulse trains reconstructed by the method based on TF distributions and their comparison to the results of the IC method: the percentage of the pulses, identified by both methods with respect to the number of pulses identified by the IC method (true positive statistics), and the percentage of the pulses that were identified only by the method based on TF distributions with respect to the number of all identified pulses (false positive statistics). The innervation pulse trains were reconstructed from 30 s long SEMG signals recorded during the isometric 10 % MVC measurements of the dominant biceps brachii of nine healthy male subjects (age 26.8 ± 2.2 years, height 179 ± 7 cm, weight 72.1 ± 8.3 kg).

Subject	Number of reconstructed MUs	Percentage of the pulses identified by both methods [%]	Percentage of pulses identified only by the method based on TF distributions [%]
1	4	82.5 ± 4.9	14.6 ± 7.2
2	3	80.0 ± 7.8	14.1 ± 8.3
3	2	68.0 ± 8.4	19.6 ± 7.9
4	1	62.0 ± 0.0	14.3 ± 0.0
5	2	71.7 ± 7.3	16.8 ± 15.6
6	4	96.1 ± 5.9	6.9 ± 3.9
7	4	72.3 ± 6.4	20.4 ± 12.7
8	4	97.2 ± 0.6	5.2 ± 3.6
9	2	93.3 ± 4.3	4.4 ± 4.0

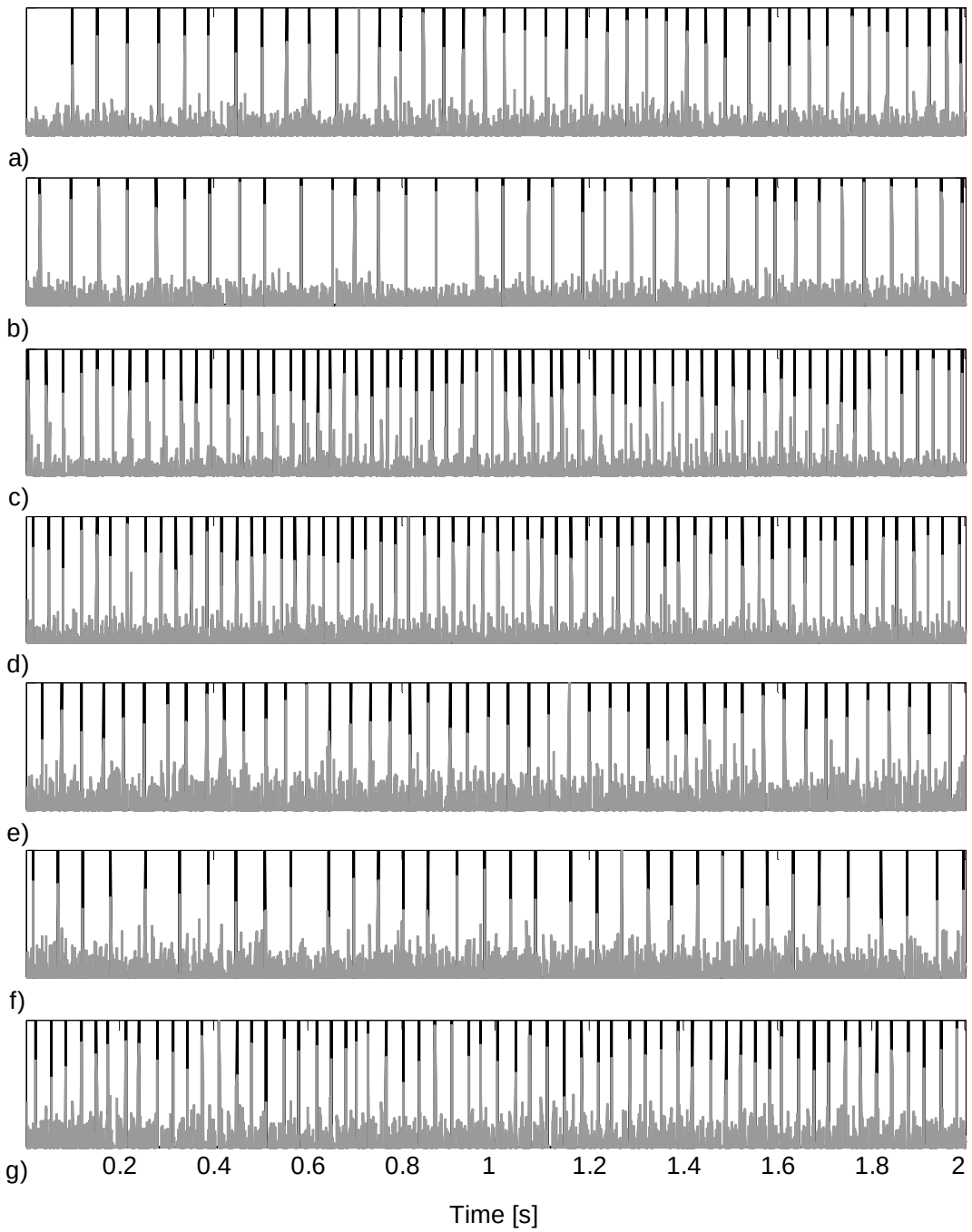


Figure 6.26: Innervation pulse trains reconstructed by the method based on TF distributions (grey) and their comparison to the results of the IC method (black): a) subject 1, MU 1 b) subject 1, MU 2, c) subject 2, MU 1, d) subject 2, MU 2, e) subject 6, MU 1, f) subject 6, MU 2 in g) subject 9, MU 1. The SEMG signals were recorded during the isometric 5 % MVC measurements of the dominant biceps brachii of nine healthy male subjects (age 26.8 ± 2.2 years, height 179 ± 7 cm, weight 72.1 ± 8.3 kg). The pulses reconstructed by the IC method are normalised to 1.

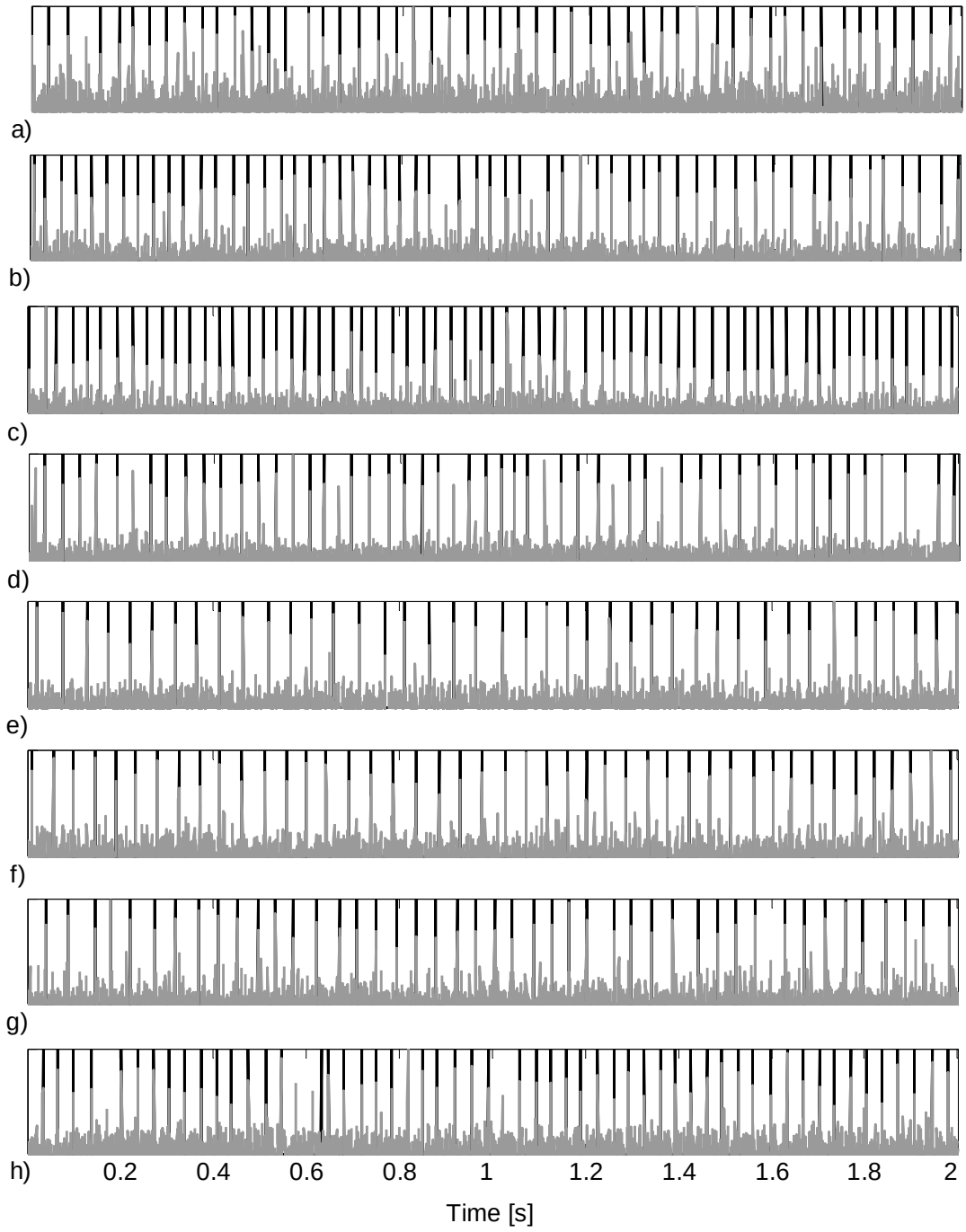


Figure 6.27: Innervation pulse trains reconstructed by the method based on TF distributions (grey) and their comparison to the results of the IC method (black): a) subject 1, ME 3 b) subject 6, MU 1, c) subject 6, MU 3, d) subject 6, MU 5, e) subject 8, MU 2, f) subject 8, MU 3, g) subject 9, MU 2 in h) subject 9, MU 4. The SEMG signals were recorded during the isometric 10 % MVC measurements of the dominant biceps brachii of nine healthy male subjects (age 26.8 ± 2.2 years, height 179 ± 7 cm, weight 72.1 ± 8.3 kg). The pulses reconstructed by the IC method are normalised to 1.

6.2.3 Method based on higher-order statistics

At the end, let us evaluate the results obtained by w-slices and the Newton-Gauss optimisation. Following the example of the previous subsection, the decomposition results will be compared to those, obtained by the IC method. To be more exact, the MUAPs reconstructed by the HOS-based method will be compared with the MUAPs reconstructed by the spike triggered sliding window averaging technique. As in the case of synthetic signals, the results will be expressed in the form of the norm of the MUAP differences defined by Eq. (6.1).

We already explained the calculation of cumulates is computationally complex and time-demanding operation. In practice, the cumulants are usually expressed by the corresponding higher-order moments. The accuracy of such calculation depends mainly on the length of the processed signals. When w-slices are used to coarsely estimate the MUAPs, the cumulant values must be calculated up to the lags τ_1 and τ_2 which correspond to the estimated MUAP length. As a result, the computational complexity increases with both the square of the MUAP length and the length of the signals.

In the case of real surface EMG signals the length of MUAPs was estimated to 50 samples, on average, while the length of signals yielded 75000 samples. To speed up the calculation, all the measurements were undersampled by factor 2, reducing the sampling frequency form 2500 Hz to 1250 Hz. The undersampled measurements still fulfil the Nyquist theorem (the frequency content of the surface EMG signals is usually limited to the interval from 0 to 500 Hz). Regarding the results of the IC method, the number of active MUs was estimated to 5 and 10 at 5 % and 10 % MVC measurements, respectively. Due to enormous computational complexity of Newton-Gauss optimisation method (Subsection 6.1.2), the HOS-based method was tested only with 5 % MVC measurements. Similar to the case of synthetic signals, only 5 measurements from the central column of electrodes (the measurements recorded by the spatial filters (4,3), (5,3), (6,3), (7,3) and (8,3)) were used in our experiments.

The decomposition procedure used follows the one with synthetic signals. Coarse estimations obtained by w-slices were optimized by the Newton-Gauss optimisation. Due to its computational complexity, only the first quadrants of the cumulant matrix $\mathbf{C}_{klm}$ were considered. The size of the step δ_i in (4.66) was calculated by bisection, while the vector of skewnesses $\boldsymbol{\gamma} = [\gamma_1 \quad \cdots \quad \gamma_K]^T$ was calculated over all possible combinations of indices *k, l* and *m*. The decomposition procedure is described more in detail in Subsection 6.1.2.

The reconstructed MUAPs were compared to the MUAPs reconstructed by the IC method (Subsection 6.2.1) according to Eq. (6.1). Only the MUAPs which differed less than 30 % from their reference values (the MUAPs detected by the IC method) were considered identified. The MUs were considered detected if at least 4 out of 5 MUAPs were identified. The results are reported in Table 6.12. The examples of perfectly reconstructed MUAPs are depicted in Figs. 6.28 and 6.29, while Figs. 6.30 and 6.31 illustrate the

examples of less accurate decomposition. The latter differ significantly form their reference values in the tail regions of MUAPs, while the central parts of the MUAPs are relatively accurately reconstructed. Approximately 40 % of the reconstructed MUs suffer fluctuations in the MUAPs' tail regions. This phenomenon will be further discussed in the next chapter.

The results of HOS-based decomposition were also visually inspected in order to exclude the possibility of additionally reconstructed MU (the MU which was not reconstructed by the IC method). Many decomposition results indicated existence of travailing components (components delayed in time in different EMG channels), nevertheless, no result was classified as a novel MU. Although there is no direct evidence, we can not exclude the possibility that some of the discarded results represented novel MUs.

Table 6.12: The number of the MUs reconstructed by the HOS-based method and their comparison to the results of the IC method using the norm of the MUAP difference d_{ij} defined by (6.1) (mean ± standard deviation). The MUAPs were reconstructed from 30 s long SEMG signals recorded during the isometric 5 % MVC measurements of the dominant biceps brachii of nine healthy male subjects (age 26.8 ± 2.2 years, height 179 ± 7 cm, weight 72.1 ± 8.3 kg).

Subject	**The number of MUs, reconstructed by both methods**	**The number of MUs, reconstructed only by HOS-based method**	**Average norm of differences d_{ij} between the MUAPs, reconstructed by the HOS-based method, and the MUAPs, reconstructed by the IC method [%]**
1	3	0	27.1 ± 9.7
2	4	0	16.9 ± 7.5
3	1	0	28.6 ± 0.0
4	2	0	33.4 ± 4.7
5	2	0	24.2 ± 15.8
6	3	0	18.9 ± 13.4
7	2	0	12.8 ± 10.6
8	3	0	21.5 ± 6.4
9	2	0	25.3 ± 11.7

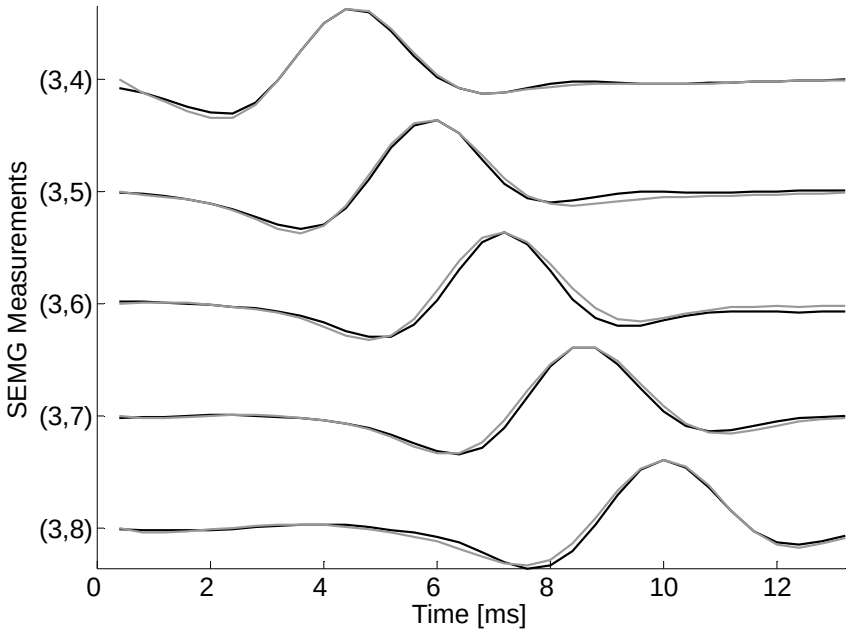


Figure 6.28: MUAPs of the third MU reconstructed by HOS-based method (grey) and by the IC method (black). Surface EMG signals were recorded during the isometric 5 % MVC measurements of the dominant biceps brachii muscle of subject 6 (age 28 years, height 182 cm, weight 70 kg).

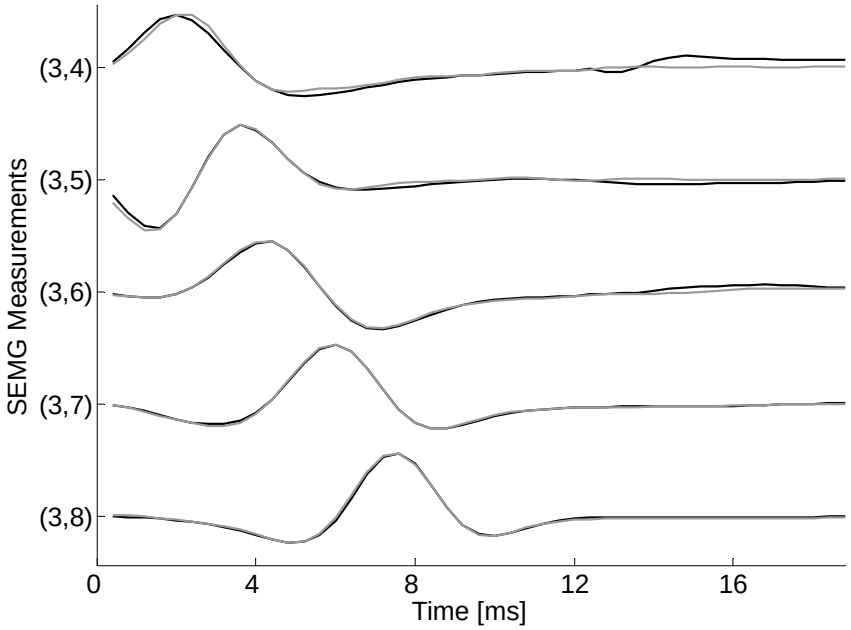


Figure 6.29: MUAPs of the second MU reconstructed by HOS-based method (grey) and by the IC method (black). Surface EMG signals were recorded during the isometric 5 % MVC measurements of the dominant biceps brachii muscle of subject 7 (age 26 years, height 191 cm, weight 90 kg).

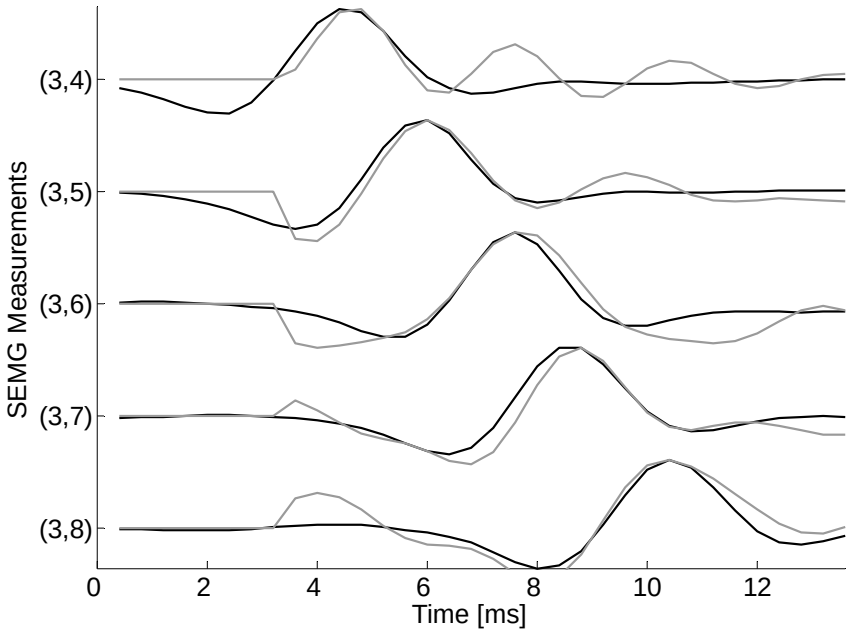


Figure 6.30: MUAPs of the third MU reconstructed by HOS-based method (grey) and by the IC method (black). Surface EMG signals were recorded during the isometric 5 % MVC measurements of the dominant biceps brachii muscle of subject 6 (age 28 years, height 182 cm, weight 70 kg).

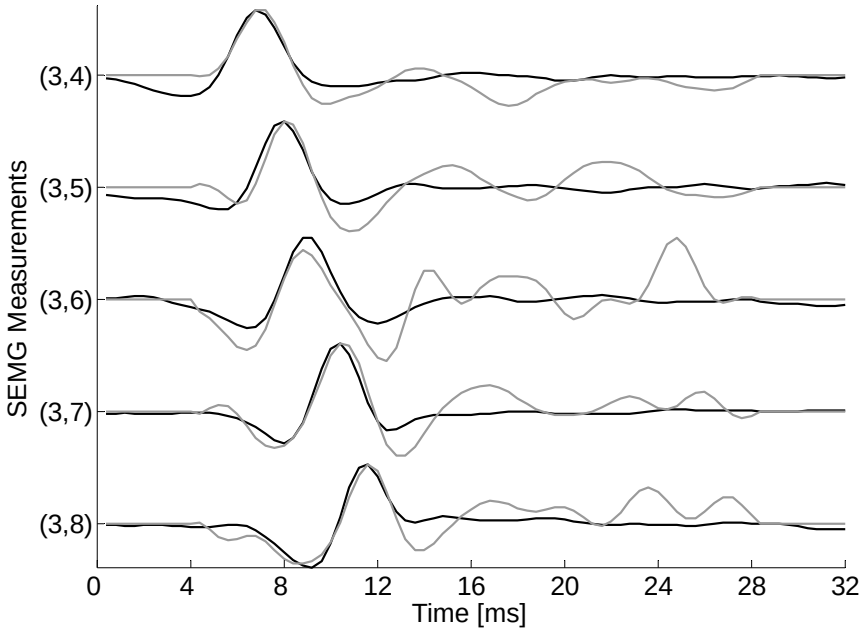


Figure 6.31: MUAPs of the second MU reconstructed by HOS-based method (grey) and by the IC method (black). Surface EMG signals were recorded during the isometric 5 % MVC measurements of the dominant biceps brachii muscle of subject 5 (age 25 years, height 170 cm, weight 70 kg).

7.

Discussion

In Chapters 4 and 5 three different approaches to decomposition of convolutive mixtures of pulse sources were introduced. All three approaches suppose close-to-orthogonal pulse sources and can be used for MIMO system identification whenever the minimum inter-pulse interval exceeds the length of system responses. In other words, they can be considered a general decomposition tool and applied to wide range of both convolutive and multiplicative pulse source mixtures.

For the sake of generality, only basic ideas behind the decomposition algorithms were explained in Chapters 4 and 5. The exact decomposition procedures were introduced in Chapter 6, where we applied the approaches to the synthetic and real surface EMG signals and justified the choice of different parameters. In the case of synthetic signals, the decomposition results were compared against their reference values. The main attention was paid to the influence of noise and number of sources, whereas the IC method was also tested against the MU depth in the muscle tissue, the number of fibres in each MU, and their firing frequencies. Decomposing the real surface EMG signals, the reference values were unknown and we were forced to use different indirect quality measures.

In this chapter, the results of all three decomposition approaches will be mutually compared. The pros and cons of every method will be identified and the deviations from the theoretical predictions explained. The most interesting and surprising results will also be clarified. Finally, the computational complexities of approaches will be estimated.

7.1 Comparison of the results on synthetic surface EMG signals

Although being based on the same assumptions, the approaches introduced in Chapters 4 and 5 differ significantly form each other. The first difference can be observed by comparing their goals. The method based on time-frequency distributions enables the reconstruction of both MUAPs and innervation pulse trains. W-slices can only be used to reconstruct the MUAPs, while the information about the innervation trains is lost in the very first step where the higher-order cumulants are calculated. On the other hand, the IC method eliminates the impact of MUAPs and identifies only the innervation pulse trains.

The introduced methods also differ in their theoretical basis. The method based on the TF distributions is derived form the time-frequency analysis, which is primarily used for investigation and characterisation of nonstationary signals. W-slices are based on higher-order statistics which are manly focused on identification of non-Gaussian random processes, whereas the IC method can be classified to a large class of linear algebra-based decomposition techniques.

Comparing both methods, based on second-order statistics, i.e. the method based on TF distributions and the IC method, one may claim they both follow the same theoretical postulates. Firstly, they are both based on the extension of the vector of measurements. Secondly, generally speaking, they both comprise the whitening of the measurements. Nevertheless, they were derived completely independently from each other. Moreover, the IC method does not rely on non-overlapped pulses and is much more noise-resistant. Finally, the method based on TF distributions simultaneously reconstructs all the pulse trains, while using the IC method the pulse trains are reconstructed independently, one after the other.

Theoretical expectations are also confirmed by the results in Section 6.1. Keeping the SNR high, the numbers of MUs reconstructed by both methods show no significant differences. In the case of 5 and 10 active MUs and at SNR of 20 dB, 3.8 and 7.3 MUs were identified by the method based on TF distributions, respectively (Table 6.1). Processing the same signals, the IC method reconstructed 4.7 and 8.3 MUs (Table 6.4). Superiority of the IC method becomes obvious at low SNR. At SNR = 0 dB, the average number of MUs reconstructed by the method based on TF distributions yields 1.6 (5 active MUs) and 0.7 (10 active MUs), whereas the same results for the IC method yield 2.0 and 2.6 MUs. Still higher differences can be expected in the case of 20 active MUs. Although the number of sources highly exceeded the number of measurements, the IC method managed to identify 2.3 MUs. Because of high computation complexity the TF based method was only tested on signals with 5 and 10 active MUs. However, quick extrapolation of its results demonstrates that decomposing the signals with 20 active MUs at SNR = 0 dB, the TF based method would most probably fail to reconstruct any MU.

The fact that the performance of TF based method drops with the number of active MUs should not come as a surprise. This was already foreseen in Subsection 4.1.5, where we explained how the probability of overlapped pulses increases with the number of active sources. At high SNR, the criterion (4.21) introduced in Subsection 4.1.4 ensures that only the STFD matrices comprising the contribution from a single source enter the joint-diagonalization. At low SNR and large number of active MUs, the selection of STFD matrices fails, the impact of crossterms in STFD matrices increases and the joint-diagonalization produces incorrect result.

In other words, the noise and non-orthogonality of sources prevent the TF-based method to separate the sources completely. Even if we suppose the measurements completely white it is perfectly possible (but not strictly necessary) that the joint-diagonalization will converge to the wrong unitary matrix $\mathbf{U}_{JD}$. Accordingly, the reconstructed sources will still be linear combinations of all the original sources and we will still have to identify the

unknown unitary mixing matrix $\mathbf{U}_{JD}^{T}\mathbf{U}$ which rotates the base of $N(L+K-1)$-dimensional space of extended sources for the unknown angles around the unknown axes to yield the space of reconstructed sources.

Following the example from Subsection 5.2.1, one could argue the reconstructed sources can further be separated by applying the procedure for separation of the superimposed sources. This is generally true, but at least two major problems appear. Firstly, the averaging of the delayed repetitions of sources reconstructed by the TF based method has to be omitted and the resistance to noise drops considerably. Namely, by adding together the delayed repetitions of the same source, the number of overlapped pulses increases enormously and the separation described in Subsection 5.2.1 fails. Secondly, such a solution would considerably increase the computation complexity of the TF based method. Both the joint-diagonalization in the case of the TF based method and the separation of superimposed sources in the case of the IC method are computationally demanding steps. Joining them together would necessarily result in a time-demanding, but not necessarily effective decomposition.

Another proof of superiority of the IC method can be observed by comparing the percentage of accurately identified pulses in Tables 6.1 and 6.4. At high SNR, almost all original pulses were reconstructed by the IC method, while the number of misplaced pulses was negligible. On the other hand, averaging the results at SNR of 20 and 15 dB, the method based on TF distribution reconstructed only 85 % of original pulses, whereas more than 5 % of identified pulses were misplaced. The differences are even more obvious at low SNR. At SNR of 0 dB, the TF based method reconstructed only 60 % of original pulses, while the percentage of misplaced pulses increased to 20 %, on average. Decomposing the same signals, the IC method identified more than 90 % of original pulses and misplaced less than 4 % of pulses.

Finally, using the measure (6.1) defined in Subsection 6.1.2 the MUAPs reconstructed by both methods can be mutually compared. As already explained, using the TF based method the MUAPs can be estimated directly from the mixing matrix. However, the results in Table 6.2 demonstrate this is not the optimal route because the reconstructed MUAPs differ significantly from their reference values. At SNR of 20 dB, the average 2-norm of differences between the reconstructed and the original MUAPs yields 60 %, and is further increased to 70 % at SNR of 0 dB. This is relatively interesting result, implying that (at least at high SNR) the relatively inaccurate estimations of the mixing matrix can produce relatively accurate estimations of pulse sources (using the TF based method the sources are estimated by multiplying the measurements by Moore-Penrose pseudoinverse of the estimated mixing matrix). The observed phenomenon can be partially explained by noticing that not all the estimates of the delayed source repetitions exhibit the same accuracy. In particular, the source repetitions corresponding to the elements with small absolute value in the mixing matrix are very sensitive to the noise and, hence, poorly reconstructed. Averaging the repetitions of the same source, all inaccurate estimations are left out (they are even not classified as repetitions of any particular source) and do not influence the final source estimation. On the other hand, the estimation of MUAPs

depends on all reconstructed repetitions of the original sources, including the inaccurate ones, and is hence much more sensitive to the noise.

Using the IC method, the MUAPs have to be reconstructed by spike triggered sliding window averaging technique described in Section 5.1. The latter is (according to the estimation theory) an unbiased estimator. Applying the Cramer-Rao lower bound [92], it can easily be shown that its variance decreases proportionally with the number of detected MU firings [77]. In our experiments on the synthetic signals, the IC method reconstructed between 300 and 600 pulses of each MU, on average. Therefore, the theory predicts the noise will have negligible impact on the MUAP estimation. However, the noise is not the only factor which may hinder the reconstruction of MUAPs. Decomposing the signals with a large number of active MUs also the influence of other MUs has to be considered. However, supposing the zero-mean measurements and the number of identified pulses large enough we can readily expect the contributions of other active MUs will average to zero.

The results on synthetic signals coincide with theoretical expectations (Table 6.6). In the case of 5 active MUs and at SNR = 20 dB, the average 2-norm of differences between the reconstructed MUAPs and their reference values yields only 8.5 % and further decreases to 6.7 % at SNR = 0 dB. The fact that the errors decrease with the noise is not surprising. At low SNR, only the strongest MUs (MUs contributing the strongest MUAP conrtibutions to the measurements) are reconstructed by the IC method (Table 6.5). While the noise and other active MUs contribute the same absolute errors to all reconstructed MUAPs, their relative contributions depend on the power of the reconstructed MUAPs. The strongest MUAPs exhibit the smallest relative errors and vice versa. As expected, the errors in reconstructed MUAPs slightly increase with the number of active MUs. At SNR of 0 dB the 2-norm difference yields 6.8 % in the case of 10 active MUs, and 11.3 % in the case of 20 active MUs.

The spike triggered sliding window averaging technique can also be combined with the results of the method based on TF distributions. However, in the case of the TF based method, the number of accurately identified pulses significantly decreases with the SNR, while the influence of misplaced pulses cannot be neglected any more. Consequently, the sliding window averaging technique will provide less accurate MUAP estimations when combined with the method based on TF distributions.

Let us now evaluate the results of the HOS-based method (Table 6.3). Due to its computational complexity the method was only tested on the signals with 5 active MUs. Nevertheless, it proved the expected resistance to the Gaussian noise. On average, 3.6, 3.4 and 2.0 MUs were reconstructed at SNR of 20, 10 and 0 dB, respectively. Although inferior at high SNR, the HOS-based method becomes superior to the TF-based method and comparable to the IC method at low SNR. Comparing the reconstructed MUAPs the HOS-based method ranks second. It gives much more accurate estimations than the TF-based method, but falls behind the IC method. In similar fashion to the IC method, only the strongest MUs are reconstructed at low SNR. Unlike the errors of the IC method, the errors of the HOS-based method slightly increase with the noise. The average 2-norm of

the MUAP errors yields 12 % at SNR = 20 dB and increases to 16 % at SNR = 0 dB. This proves that the increase in third-order cumulants of the noise outranks the decrease of the ratio between the 2-norm of the noise and the 2-norm of the reconstructed MUAP.

Finally, recall the HOS-based method used only 5 out of 50 synthetic surface EMG signals, whereas the other two methods employed all possible measurements. In spite of that, the HOS-based method proved to be on a par with the IC method (at least at low SNR).

7.2 Comparison of the results on the real surface EMG signals

Evaluating the results obtained on the real surface EMG signals the original MU firing patterns and the reference MUAP shapes are unknown. However, encouraged by the results on the synthetic signals we can rely on less strict measures. In our experiments, the results of the IC method were first tested against the well-known properties of MUs (Chapter 3). We proved the reconstructed innervation pulse trains agree with all physiologically induced measures: all of the identified MUs exhibit the regular firing patterns, the IPI variability is almost in all cases limited to 10 % of the IPI mean, while none of several statistical tests was able to reject the hypothesis of normally distributed IPIs. The most intuitive evidence was the shape of the MUAPs reconstructed by spike triggered sliding window averaging technique. Not only the similar shapes detected by the adjacent electrodes, but also the clear MUAP propagation patterns (Figs. 6.22, 6.23, 6.24 and 6.25) prove the accuracy of the detected innervation pulse trains.

Based on numerous direct and indirect evidences, we supposed the IC method can be trusted and assessed its efficiency first. As in the case of synthetic signals, the main performance criteria were the number of reconstructed MUs, and the regularity of their firing patterns (the number of accurately reconstructed pulses). Strict assessment of the reconstructed innervation trains is, of course, impossible. Due to the original innervation pulse train being unknown, we can only coarsely estimate the number of missed and misplaced pulses. In our experiments, we assumed all the MUs strictly follow their regular firing patterns, hence, every IPI which exceeded the IPI mean by factor 1.5 was classified as a missed pulse, while every IPI that was shorter than the half of the mean IPI was classified as a misplaced pulse. In reality, this may not necessarily be the case as both the double firings as well as the missing pulses are perfectly possible. Nevertheless, this was the only way to coarsely assess the efficiency of the IC method.

Decomposing 30 s long real surface EMG measurements the IC method reconstructed approximately 400 pulses, on average. Both the average number of missed pulses and the average number of misplaced pulses were estimated to 10 (Figs. 6.17, 6.18 and 6.19). Again, the estimation of misplaced pulses is critical and should only be used as a coarse indication. Nevertheless, the regularity of the reconstructed firing patterns proves with high probability that the number of misplaced pulses can be neglected.

Evaluating the number of reconstructed MUs, the IC method was only partially successful. A total of 30 MUs (3.4 ± 1.3 MUs per subject) were reconstructed in the case

of 5 % MVC measurements. A coarse comparison of the reconstructed MUAP sequences to the original measurements revealed the identified MUs contribute only 45 % of total energy. The efficiency was significantly improved in the case of 10 % MVC measurements, where a total of 56 MUs (6,2 ± 2,2 MUs per subject) were identified, while the energy ratio between the reconstructed MUAP trains and the original measurements increased to approximately 70 %. The partial success of the IC method can be explained by considering the impact of the noise. Nevertheless, this phenomenon should be assessed more critically in the future.

In order to assess their accuracy, the other two methods were compared to the IC method. According to the results on the synthetic signals the TF-based method proved to be less reliable. In the case of 5 % MVC real measurements a total of 22 MUs (2.4 ± 0.9 MUs per subject) was identified, while 87 ± 9 % of pulses reconstructed by the IC method were also identified by the TF-based method (Table 6.10). On the other hand, 10 ± 9 % of the pulses reconstructed by the TF-based method was not identified by the IC method. Decomposing the 10 % MVC measurements, the total number of identified MUs reached 26 (2.7 ± 1.2 MUs per subject). On average, 80 ± 13 % of pulses identified by the IC method was also reconstructed by the TF-based method, while the percentage of the pulses reconstructed only by the TF-based method increased to 13 ± 6 % (Table 6.11). Not all of the pulses that were detected only by the TF-based method seem to be misplaced (see, for example, Fig. 6.27 h). But such cases were extremely rare and a detailed analysis of the reconstructed IPIs revealed the vast majority of pulses identified only by the TF-based method were misplaced.

There is yet another parallel which can be drawn when comparing the results on synthetic signals to the results on real signals. In both cases, the efficiency of the TF-based method significantly decreases with the number of active MUs. Comparing the decompositions of the real 5 % and 10 % MVC measurements, the total number of identified MUs differs only by 4 (in comparison with the IC method where the number of identified MUs almost doubled). On the other hand, the percentage of accurately identified pulses decreases by almost 10 %. These results coincide perfectly with the results on synthetic signals, and also with the theoretical expectations (Subsection 4.1.4).

We are now in a position to assess the efficiency of the HOS-based method. This is not an easy task as the HOS-based method reconstructs only the MUAPs. The only possible route is to mutually compare the results of HOS-based method to the MUAPs reconstructed by the spike triggered sliding window averaging technique, which already proved to be very efficient. However, interpreting the results in Table 6.12, we must always bear in mind that we only compare two different estimates of the same MUAP.

Due to its high computational complexity, the HOS-based method was only applied to the 5 % MVC measurements. The total number of reconstructed MUs reached 22 (2.4 ± 0.9 MUs per subject) and is perfectly comparable to the results of other two methods. The central regions of the reconstructed MUAPs exhibit an almost perfect match with the estimates of the IC method (Figs. 6.28 and 6.29). The significant differences were detected mainly in both marginal regions (Figs. 6.30 and 6.31). The average 2-norm of

the MUAP differences (6.1) yields 23 ± 6 % (Table 6.12) and should be carefully interpreted. The results on synthetic signals reveal the average 2-norm error of the reconstructed MUAPs is 7 % in the case of IC method, and 14 % in the case of HOS-based method. Adding the errors together, the 2-norm of differences between both estimates yields 21 %, which agrees perfectly with the results on the real surface EMG signals.

How can we explain the differences among the marginal regions of reconstructed MUAPs (Figs. 6.30 and 6.31)? One of many possible reasons lies in the non-orthogonality of MU innervation trains. The latter contribute nonzero values to the cross-cumulants of sources and induce the differences between the cumulant values calculated from measurements and the cumulant values estimated from the Barlett-Brillinger-Rosenblatt equation. Consequently, the Newton-Gauss optimisation converges towards the wrong solutions (Subsection 4.2.4). Another possible explanation is that the marginal fluctuations are mainly cased by non-Gaussian noise.

7.3 Computational complexity

In previous two sections, the accuracy and efficiency of three novel decomposition approaches were assessed. The results proved the IC method is superior on both synthetic and real surface EMG signals. In this section, the experimentally measured computational complexities of all three approaches are mutually compared. For the clearly reasons, only the results on synthetic signals are reported, while the computation complexities are always expressed in the amount of time required to completely decompose one second of all simultaneously processed measurements of synthetic surface EMG (50 measurements whenever the TF-based or IC method is applied, and 5 measurements whenever the HOS-based method is used). Actually, we talk about the duration of computation.

All three decomposition methods were implemented in the Matlab programming tool and tested on personal computer with Pentium IV 2 GHz processor and 1 GB of memory. Decomposing 5 synthetic measurements of 5 active MUs, the HOS-based method required approximately 2 minutes per every second of measurements. In the case of 10 signals with 10 active MUs the duration of computation increased to 32 minutes. On the other hand, the TF-based method spent approximately 45 and 90 seconds to process every second of the 50 surface EMG signals with 5 and 10 active MUs, respectively. The detailed analysis revealed that the time-critical sections lie in the Newton-Gauss optimisation of the HOS-based method and in the joint-diagonalization of the TF-based method.

Using the parallel version of the joint-diagonalization [74], the time for decomposition improves enormously. Supposing the homogenous cluster of processing units, a coarse speedup ratio can be estimated by multiplying the number of processing units by factor 0.9 [74]. For example, decomposing 50 synthetic measurements of 10 active MUs on 10 homogenous personal computers with 2 GHz Pentium IV processors, the average

duration of computation decreases to 10 seconds per every second of 50 EMG signals [74].

Exhibiting superior accuracy, the IC method also proved to be most time-efficient. Processing 50 measurements of 30 s long synthetic surface EMG with 5 active MUs, the duration of computation per every second of measurements yielded approximately 10 s, and was increased to 13 and 15 s in the case of 10 and 20 active MUs, respectively. The vast majority of time was spent searching for possibly unidentified MUs, while all the strongest MUs were completely reconstructed within 1 to 4 seconds (depending on the SNR).

The IC method exhibits another nice property. Its computation complexity is almost independent of the number of active MUs. The only step which is affected by the number of active MUs is the selection of the time moments in which the activity index exceeds the noise threshold. The larger the number of active MUs the smaller the relative influence of noise and the larger the number of time moments in which the IC method searches for the candidate pulses. The computation complexity can further be reduced by replacing the brute force searching of the candidate MU firings by heuristic search methods. Their detailed explanation and evaluation reaches beyond the scope of this discussion. The main goal in this thesis was to assess the efficiency and accuracy of the developed approaches, while the time complexity was mainly ignored.

8.

Conclusions

In this dissertation, three novel approaches to blind source separation of convolutive mixtures of close-to-orthogonal pulse sources are introduced. The first approach is based on joint-diagonalization of spatial time-frequency distribution matrices and enables the reconstruction of both the mixing matrix and pulse sources. Our tests on synthetic surface EMG signals revealed its sensitivity to the noise and overlapped pulses. Although special measures were derived to avoid their devastating influence, the noise and non-orthogonal sources remain the biggest problem of the TF-based approach. The tests also pointed out the reconstruction of impulse responses based on the estimated mixing matrix is very prone to errors, even when the reconstructed sources show their perfect match with reference values. Knowing the sources, the system responses can also be estimated by employing the spike triggered sliding window averaging technique (Section 5.1). It is very robust to noise, but its performance is degraded when the number of misplaced pulses becomes significantly large. The TF-based method is, hence, suitable only for the MIMO systems identification in low-noise environments. Moreover, being sensitive to overlapped pulses, the TF-based method suffers a sever drop in performance when the number of sources increases. Supposing high SNR and orthogonal sources, the presented approach proves very successful – it outperforms the HOS-based method and even matches the performance of the IC method. Thus, in some specific environments the TF-based method remains an important tool for blind deconvolution of close-to-orthogonal pulse sources.

The second approach introduced in this dissertation is based on higher-order statistics and is a generalization of the so call w-slices method. Neglecting the information about the source pulse trains, it can only be applied to estimate the system responses. The tests on synthetic surface EMG signals demonstrated its high resistance to the Gaussian noise, while the main drawback can be found in its high computational complexity, which often hinders the decomposition.

High computational complexity was also the reason why the HOS-based method was only tested on the signals with the smallest possible number of sources. Nevertheless, the presented approach proved very robust and can be expected to work well also on MIMO systems with more sources. In the case of synthetic surface EMG signals, the reconstructed MUAPs showed a perfect match with their reference values, especially at low SNRs. Applying the HOS-based method to the real surface EMG signals, slight deviations in the marginal regions of reconstructed MUAPs appeared. One could argue the marginal regions of the reconstructed MUAPs can simply be neglected. Since the corresponding reference values are small, no information would be lost, while the decomposition results would significantly improve. But the reference shape of system

responses is unknown and (at least in general) the marginal regions can not simply be neglected. Yet, there are at least two more reasons for keeping the marginal regions of reconstructed MUAPs intact. The first one is the indicated difference between the synthetic and real signals. Although closely inspected, no oscillations were observed on synthetic MUAPs. The second one is the unbiased and critical comparison of HOS-based method to the other two decomposition approaches. Cutting the errors would certainly blur the true picture.

Deriving the third decomposition approach (the so called inverse correlation based method), the decomposition of slightly under-determined MIMO systems was investigated. Our tests on synthetic and real surface EMG signals proved the IC method is superior in both the number and the accuracy of reconstructed innervation pulse trains. In the synthetic case, the reconstructed MUs' firing patterns perfectly matched their reference values, while the number of misplaced pulses was negligible, also at low SNR. In the case of real signals, no reference values can be used, but using the statistics and estimation theory numerous indirect measures showing the agreement of the reconstructed innervation trains with physiologically induced limitations were derived.

There are several assumptions and limitations which have to be considered when applying the decomposition approaches to the surface EMG signals. The first assumption is inherently induced by the theory of blind source separation. When processing the convolutive mixtures it is common to suppose sources mutually independent [4, 21, 22, 23, 78, 79, 84, 97, 106]. The approaches introduced in this dissertation follow the same assumption. However, in order to emphasize their practical value the small number of overlapped pulses is also allowed. Both the TF-based and IC method comprise special algorithms which suppress the devastating influences of noise and non-orthogonal sources. The IC method even enables the decomposition of highly superimposed sources (when two arbitrary sources overlap in 10 % of their pulses). This modification proves crucial in the case of SEMG signals. Although the MUs should fire asynchronously at low contraction forces [56, 114, 139, 158], the short-term MU synchronization has been frequently observed [93, 139, 158]. While limited in duration (usually up to a few firing pulses), this phenomenon clearly violates the assumption of mutually independent sources and could considerably affect the separation process. Supposing close-to-orthogonal sources, any temporal pulse overlapping, such as in the case of short-term MU synchronization, are allowed and, moreover, implicitly included into the data model. Moreover, as depicted in Fig. 6.18, the IC method is perfectly capable of detecting both recruitment of a new MUs and drop outs of the already detected ones. From the medical point of view this proves to be a very important result.

Also important are the assumptions concerning the mixing matrix. It is, even in the case of more sources than measurements, assumed of a full row rank [22, 23, 68]. We further supposed the system impulse responses finite, casual and time invariant [56, 114, 166, 100, 33]. This implies certain practical limitations. The phenomenon of muscle contraction, for example, can be modelled as time invariant only if the EMG recording has been taken in a stable and controllable measurement session, i.e. during an isometric muscle contraction, without the appearance of fatigue-induced changes [56, 114, 166,

167]. The latter can not always be established, especially when long fatigable contractions are observed. In such cases, the measurements should be divided into consecutive partially overlapped epochs which have to be processed separately. In this fashion, also the long-term changes in the shape of MUAPs can be temporarily tracked. The rapid changes of MUAPs, such as the one induced by dynamic contractions [116], are very hard to cope with and are not recognised by the approaches introduced in this thesis.

The last major limitation is initiated by the ratio between the number of measurements and the number of sources. While the TF-based and HOS-based methods assume more measurements than sources, the IC method is expected to work successfully only when the number of extended sources does not exceed twice the number of extended measurements. Applying the TF-based and IC methods, the number of extended sources is proportional to the length of the system responses. Therefore, it is wise to limit the sampling frequency to its smallest possible value still fulfilling the Nyquist condition. Due to the filtering effects of subcutaneous tissues the highest frequency components constituting the surface EMG signals hardly exceed 500 Hz. Hence, it usually suffices to set the sampling frequency to 1024 Hz. Then the length of MUAPs can be estimated to approx. 25 samples which implies the number of active MUs is increased by factor 25.

Surface electromyography (SEMG) has recently become a rather developed and matured measuring technique. High-density multi-electrode arrays enabling acquisition of more than 100 surface EMG channels and providing all the required reliability, robustness, and repeatability of surface EMG measurements are being developed [136]. Suppose now 100 measurements of surface EMG and estimate the number of MUs that can theoretically be reconstructed by the IC method. This method works also with slightly underdetermined systems, but as a rule, the number of reconstructed sources hardly exceeds the number of measurements. Keeping the number of measurements limited to 100 (by not additionally extending the measurements) and setting the sampling frequency to 1000 Hz, this implies that only 4 MUs can be completely reconstructed. Extending the measurements by factor K=10, the number of reconstructed MUs increases to approx. 25, while setting the extension factor to K=20 approx. 35 active MUs can be identified. Of course, this is just a coarse extrapolation of the results which are presented in this thesis. Deriving more accurate estimation at least two additional factors have to be considered. Firstly, the size of the correlation matrix of measurements and the computational complexity of the IC method are proportional to the square of the extension factor K. Therefore, multiplying the K by 2 the computational complexity increases by 4. Secondly, when the extension factor K is increased the condition number of the mixing matrix and its robustness to noise usually decrease. Therefore, the optimal extension factor depends on the signal's quality and usually varies between 5 and 15. When the number of active MUs exceeds the number of surface EMG channels ($N>M$), the optimal ratio between the number of extended sources and the number of extended measurements is achieved by setting extension factor to 1.

In the case of surface EMG signals the number of active MUs is usually high [19, 56, 113, 114, 146]. Even at low contraction forces there are several tens of MUs active in the

detection volume. Increasing the level of muscle contraction, the number of active MUs rises and it is quite common for SEMG to comprise MUAPs of more than a hundred MUs. Improving the spatial selectivity of the detection system, i.e. reducing the detection volume, can efficiently limit the number of active sources. Nevertheless, the optimal decomposition of SEMG signals is expected at the contractions not exceeding 30 % MVC. Preliminary tests prove the IC method can be also used at much higher contractions to reconstruct at least a few MUs. However, in this thesis only slightly underdetermined MIMO systems with close-to-orthogonal sources are studied, while the assessment of the IC method at high muscle contractions is left for future investigations.

We conclude the thesis by checking the hypotheses introduced in Chapter 1. According to the results we can readily focus on the IC method. In the first hypothesis, we supposed that the complete decomposition of convolutive mixtures is possible if and only if the number of measurements exceeds the number of sources. When this is not the case, only a limited number of sources can be completely reconstructed. Decomposing the synthetic surface EMG signals at SNR = 20 dB, the IC method completely reconstructed almost all MUs (Table 6.4). In the underdetermined case, the number of extended sources exceeded the number of measurements by factor 1.4. Consequently, only a half of the simulated MUs where completely reconstructed. Hence, the first hypothesis can be considered confirmed.

The second hypothesis predicts the influence of non-orthogonal sources. It assumes that non-orthogonal sources cause the impulse responses in statistical moments and higher-order cumulants to interfere, what introduces additional errors to the decomposition process. Non-orthogonality of sources was extensively studied throughout this dissertation. We explained that overlapped pulses hinder the TF-based decomposition (Subsection 4.1.4), induce the non-zero cross cumulants of sources (Subsection 4.2.4) and cause the superimpositions of sources reconstructed by the IC method (Section 5.2). In the case of the TF-based and IC method, special countermeasures suppressing the influence of non-orthogonal sources were also derived. Therefore, also the second hypothesis can be considered confirmed.

The third hypothesis comprises two different assumptions. We first suppose that the exact number of sources which can be completely reconstructed from noisy measurements depends on their firing rate. The results in Table 6.5 clearly reject this assumption. The decomposition results exhibit high positive correlation with the number of MU fibres and high negative correlation with the MU depth in the muscle tissue. Surprisingly, no significant influence of the MUs' firing rate was detected. The second assumption predicted the superiority of the HOS-based approaches in the noisy environments. Although the HOS-based method proved to be highly resistant to noise it was not superior. Significantly better results were produced by the IC method (Chapters 6 and 7). Hence, also the second assumption of the third hypothesis must be rejected.

The principal thesis of our research, as stated in Chapter 1 summarises the assumptions and limitations from the first and second abovementioned hypotheses and can, thus, also be considered fully proved.

References

[1] E. W. Abel, P. C. Zacharia, A. Forster, T. L. Farrow: Neural network analysis of the EMG interference pattern, Med. Eng. & Physics, Vol. 18, No. 1, 1996, pp. 12-17.

[2] S. I. Amari, A. Cichoki, H. H. Yang: A new learning algorithm for blind signal separation, Neural Information Processing System 8, D. S. Toureyzky et al. (Ed.), 1995, pp. 757-763.

[3] S. I. Amari, A. Cichocki, H. H. Yang: Blind signal separation and extraction: neural and information theoretic approaches, in S. Haykin (Ed.): Unsupervised adaptive filtering, Vol. 1, Wiley, 2000, pp. 63-138.

[4] S. I. Amari, A. Hyvarinen, T. W. Lee, V. D. A. Sánchez: Blind signal Separation and Independent Component Analysis, Neurocomputing, Vol. 49, 2002, pp. 1-5.

[5] P. O. Amblard, J. M. Brossier, N. Charkani: New adaptive estimation of the fourth-order cumulant: application to transient detection, blind deconvolution and timing recovery in communication, in Signal Processing VII, Theories and applications, M. J. J. Holt, C. F. N. Cowan, P. M. Grant, W. A. Sandham (Eds.), Elsevier, Edinburgh, Scotland, 1994, pp. 466-469.

[6] F. Auger, P. Flandrin, P. Goncalves, O. Lemoine: Time-Frequency toolbox for use with MATLAB, CNRS, France, 1996.

[7] G. Balestra, S. Frassinelli, M. Knaflitz, F. Molinari: Time-frequency analysis of surface myoelectric signals during athletic movement, IEEE Eng. in Med. and Biol., Vol. 20, No. 6, 2001, pp. 106-115.

[8] A. K. Barros, A. Mansour, and N. Ohnishi: Removing artifacts from ECG signals using independent components analysis, Neurocomputing, Vol. 22, 1999, pp. 173-186.

[9] E. A. Bartnik, K. J. Blinowska: Wavelets – new method of evoked potential analysis, Med. and Biolog. Eng. and computing, Vol. 30, 1992, pp. 125-126.

[10] A. J. Bell, T. J. Sejnowski: An information-maximization approach to blind separation and blind deconvolution, Neural computation, Vol. 7, No. 6, 1995, pp. 1129-1159.

[11] A. Belouchrani, K. Abed-Meraim, Y. Hua: Jacobi-like algorithms for joint block diagonalization: application to source localization, ISPACS, Melbourne, Australia, 1998, pp. 645-649.

[12] A. Belouchrani, K. Abed-Meraim, M. G. Amin, A. Zoubir: Joint-antidiagonalization for blind source separation, in Proc. ICASSP, Utah, USA, 2001, pp. 456-459.

[13] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, E. Moulines: A blind source separation technique using second-order statistics, IEEE Trans. Sig. Proc., Vol. 45, No. 2, 1997, pp. 434-444.

[14] A. Belouchrani, M.G. Amin: Blind source separation based on time-frequency signal representations, IEEE Trans. Sig. Proc., Vol. 46, No. 11, 1998, pp. 2888-2897.

[15] A. Belouchrani, J. F. Cardoso: Maximum likelihood source separation for discrete sources, Signal Processing VII, Theories and applications, M. J. J. Holt, C. F. N. Cowan, P. M. Grant, and W. A. Sandham (Eds.), Elsevier, Edinburgh, Scotland, 1994, pp. 768-771.

[16] A. Belouchrani, A. Cichocki: A robust whitening procedure in blind source separation context, Electronics Letters, Vol. 36, 2000, pp. 2050-2051.

[17] B. Boashash: Time-frequency signal analysis and processing, Prentice Hall PTR, Englewood Cliffs, New Jersey, 2001.

[18] H. Broman: Knowledge-based signal processing in the decomposition of myoelectric signals, IEEE Trans. on Eng. Med. Biol., Vol. 7, No. 2, 1988, pp. 24-28.

[19] H. S. M. Brown, R. B. Stein: The relation between the surface electromyogram and muscular force, J. of Physiol., No. 246, 1975, pp. 549-569.

[20] A. Burden, R. Bartlett: Normalisation of EMG amplitude: an evaluation and comparison of old and new methods, Med. Eng. & Physics, Vol. 21, No. 4, 1999, pp. 247-257.

[21] J. F. Cardoso: Separation of non stationary sources: achievable performance, in Proc. SSAP 2000, 2000, pp. 359-363.

[22] J. F. Cardoso: On the performance of orthogonal source separation algorithms, in Proc. EUSIPCO 1995, Edinburgh, Scotland, 1995, pp. 365-370.

[23] J. F. Cardoso: On the stability of source separation algorithms, Journal of VLSI Sig. Proc. Systems, Vol. 26, No. 1/2, 2000, pp. 7-14.

[24] J. F. Cardoso: Blind signal separation: statistical principles, in Proc. of the IEEE, Special issue on blind identification and estimation, R. W. Liu in L. Tong (Eds.), Vol. 90, No. 8, 1998, pp. 2009-2026.

[25] J. F. Cardoso: A least-squares approach to joint diagonalization, IEEE Sig. Proc. Lett., Vol. 4, 1997, pp. 52-53.

[26] J. F. Cardoso: Perturbation of joint diagonalizers. Technical report ref.# 94d027, Télécom Paris, 1994.

[27] J. F. Cardoso: Infomax and maximum likelihood for source separation, IEEE Letters on Signal Processing, Vol. 4, 1997, pp. 112-114.

[28] J. F. Cardoso: the three easy routes to independent component analysis; contrasts and geometry, in Proc. ICA 2001, San Diego, California, USA, 2001, on CD.

[29] J. F. Cardoso, P. Comon: Independent component analysis, a survey of some algebraic methods, in Proc. ISCAS'96, Vol. 2, 1996, pp. 93-96.

[30] J. F. Cardoso, B. Laheld: Equivariant adaptive source separation, IEEE Trans. on Sig. Proc., Vol. 44, No. 12, 1996, pp. 3017-3030.

[31] J. F. Cardoso, A. Souloumiac: Jacobi angles for simultaneous diagonalization, SIAM Journal on Matrix Analysis and Applications, Vol. 17, No. 1, 1996, pp. 161-164,

[32] E. Chaumette, P. Common, D. Muller: Application of ICA to airport surveillance, in Proc. HOS 93, South Lake Tahoe, California, 1993, pp. 210-214.

[33] B. Chen, P. Petropulu: Frequency domain blind MIMO system identification on second- and higher-order statistics, IEEE Trans. on Signal Processing, Vol. 49, No. 8, 2001, pp. 1677-1688.

[34] J. J. Chen, T. Sun, T. Lin, T. Lin: Spatio-temporal representation of multichannel EMG firing patterns and its clinical applications, Med. Eng. & Physics, Vol. 19, No. 5, 1997, pp. 420-430.

[35] A. Cichoki, W. Kasprzak, S. I. Amari: Multi-layer neural networks with a local adaptive learning rule for blind separation of source signals, in Proc. NOLTA '95, Tokyo, Japan, 1995, pp. 61-65.

[36] E. A. Clancy, N. Hogan: Probability Density of the surface electromyogram and its relation to amplitude detectors, IEEE Trans. Biomed. Eng., Vol. 46, No. 6, 1999, pp. 730-739.

[37] L. Cohen: Time-frequency analysis, Prentice Hall PTR, Englewood Cliffs, New Jersey, 1995.

[38] P. Common: Independent component analysis, a new concept?, Sig. Proc., Vol. 36, 1994, pp. 287-314.

[39] C. J. De Luca, A. Adam: Decomposition and analysis of intramuscular electromyographic signals, in Modern Techniques in Neuroscience Research, U. Windhorst , H. J. Heidelberg (Eds.), Springer, 1999, pp. 757-776.

[40] C. J. De Luca, P. J. Foley, Z. Erim: Motor unit control properties in constant-force isometric contractions, J. of Neurophysiol., Vol. 76, No. 3, 1996, pp. 1503-1516.

[41] D. Dembele, G. Favier: Recursive estimation of fourth-order cumulants with application to identification, Signal Processing, Vol. 68, 1998, pp. 127-139.

[42] G. Desodt, D. Muller: Complex independent components analysis applied to the separation of radar signals, in Signal Processing V, Theories and applications, L. Torres, E. Masgrau, M. A. Lagunas (Eds.), Elsevier, Barcelona, Španija, 1994, pp. 665-668.

[43] P. Devijver, J. Kittler: Pattern recognition: a statistical approach, Prentice Hall International, Inc., Englewood Cliffs, New Jersey, 1982.

[44] Y. Deville, N. Charkani: Analysis of the stability of time-domain source separation algorithms for convolutively mixed signals, in Proc. ICASSP, München, Germany, 1997, pp. 1835-1838.

[45] Y. Deville N. Damour, J. Charkani: Improved multi-tag radio-frequency identification systems based on new source separation neural networks, in Proc. International Workshop on Independent Component Analysis and Blind Separation of Signals, Aussois, France, 1999, pp. 449-454.

[46] C. Disselhorst-Klug, G. Rau, A. Schmeer, J. Silny: Non-invasive detection of the single motor unit action potential by averaging the spatial potential distribution triggered on a spatially filtered motor unit action potential, J. of Electromyogr. and Kinesiol., Vol. 9, 1999, pp. 67-72.

[47] K. Drašlar, M. Kočevar: Skripta za občo fiziologijo, predavanja iz obče fiziologije, Filozofska fakulteta, Oddelek za psihologijo, 2004.

[48] S. C. Douglas and S. I. Amari: Natural-gradient adaptation, in S. Haykin (Ed.): Unsupervised Adaptive Filtering, Vol. 1, pp. 13-61, Wiley & sons, Inc., New York, 2000.

[49] K. Englehart, B. Hudgins, P. A. Parker, M. Stevenson: Classification of the myoelectric signal using time-frequency based representations, Medical Engineering & Physics, Vol. 21, No. 6-7, 1999, pp. 431-438.

[50] N. Fallentin, K. Jorgensen, E. B. Simonsen: Motor unit recruitment during prolonged isometric contractions, Eur. J. of Appl. Physiol. and Occupat. Physiol., Vol. 67, 1993, pp. 335-341.

[51] D. Farina, R. Colombo, R. Merletti, H. B. Olsen: Evaluation of intra-muscular EMG signal decomposition algorithms, J. of Electromyogr. and Kinesiol., No. 11, 2001, pp. 175-187.

[52] D. Farina, C. Fevotte, C. Doncarli, R. Merletti: Blind separation of linear instantaneous mixtures of nonstationary surface myoelectric signals, IEEE Trans. on Biomedical Eng., Vol. 51, No. 9, 2004, pp. 1555-1567.

[53] D. Farina, E. Fortunato, R. Merletti: Noninvasive estimation of motor unit conduction velocity distribution using linear electrode arrays, IEEE Trans. on Biomed. Eng., Vol. 47, No. 3, 2000, pp. 380-388.

[54] D. Farina, M. Fosci, R. Merletti: Motor unit recruitment strategies investigated by surface EMG variables, Journ. Appl. Physiol., Vol. 92, 2002, pp. 235-247.

[55] D. Farina, R. Merletti: Effect of electrode shape on spectral features of surface detected motor unit action potentials, Acta Physiol. Pharmacol. Bulg., Vol. 26, 2001. pp. 63-66.

[56] D. Farina, R. Merletti: A novel approach for precise simulation of the EMG signals detected by surface electrodes, IEEE trans. Biomed. Eng., Vol. 48, 2001, pp. 637-646.

[57] D. Farina, R. Merletti, B. Indino, M. Nazzaro, A. Bottin, M. Pozzo, I. Caruso: Surface EMG cross-talk evaluated from experimental recordings and simulated signals, reflections on cross-talk interpretation, quantification and reduction, in Proc. 4th International Workshop on Biosignal Interpretation BSI'02, 2002, Como, Italy, pp. 163-166.

[58] D. Farina, A. Rainoldi: Compensation of the effect of sub-cutaneous tissue layers on surface EMG: a simulation study, Med. Eng. & Physics, No. 21, 1999, pp. 487-496.

[59] C. Fevotte, S. J. Godsill, P. J. Wolfe: Bayesian approach for blind separation of underdetermined mixtures of sparse sources, in Proc. ICA 2004, Granada, Spain, 2004, pp. 398-405.

[60] J. Finsterer: EMG-interference pattern analysis, J. of Electromyogr. and Kinesiol., Vol. 11, 2001, pp. 231–246.

[61] B. Friedlander, B. Porat: Asymptotically optimal estimation of MA and ARMA parameters of non-Gaussian processes from high-order moments, IEEE Trans. on Automatic Control, Vol. 35, No. 1, 1990, pp. 27-35.

[62] G. A. Garcia, R. Okuno, and K. Akazawa: Decomposition algorithm for surface electrode-array electromyogram in voluntary isometric contraction, IEEE BME Magazine, Vol. 23, No. 5, 2004.

[63] M. Gazzoni, D. Farina, R. Merletti: Motor unit recruitment during constant low force and long duration muscle contractions investigated with surface EMG, in Proc. IX International Symposium on Motor Control, Varna, Bulgaria, 2000, pp. 67-71.

[64] A. Gerber, R. M. Studer, R. J. P. de Figueiredo, G. S. Moschytz: A new framework and computer program for quantitative EMG signal analysis, IEEE Trans. Biomed. Eng., Vol. 31, No. 12, 1984, pp. 857-863.

[65] S. Van Gerven, D. Van Compernolle: Signal separation by symmetric adaptive decorrelation: stability, convergence, and uniqueness, IEEE Trans. on Sig. Proc., Vol. 43, No. 7, 1995, pp. 1602-1612.

[66] S. Van Gerven, D. Van Compernolle, L. Nguyen Thi, C. Jutten: Blind separation of sources: a comparative study of a 2nd and a 4th order solution, in Signal Processing VII, Theories and applications, M. J. J. Holt, C. F. N. Cowan, P. M. Grant, W. A. Sandham (Eds.), Elsevier, Edinburgh, Scotland, 1994, pp. 1153-1156.

[67] G. B. Giannakis, Y. Inouye, J. M. Mendel: Cumulant based identification of multichannel moving-average models, IEEE Trans. on Automatic Control, Vol. 34, No. 7, 1989, pp. 783-787.

[68] G. H. Golub, and C. F. V. Loan: Matrix computations, John Hopkins Univ. Press 1989.

[69] I. F. Gorodnitsky, B. D. Rao: Sparse signal reconstruction from limited data using FOCUSS: a reweighted minimum norm algorithm. IEEE Trans. on Sig. Proc., Vol. 45, No. 3, 1997, pp. 600-616.

[70] A. Gorokhov, P. Loubaton: Blind identification of MIMO-FIR systems: A generalized linear prediction approach, Signal Processing, Vol. 73, 1999, pp. 105-124.

[71] S. Grigis, A. Holobar, D. Zazula: Improved recognition performance for orthogonal sources, in Proc. The IEEE Region 8 EUROCON 2003: computer as a tool, Ljubljana, Slovenia, 2003, Vol. 2, pp. 153-157.

[72] Z. He, L. Yang, J.Liu, Z.Lu, C. He, Y. Shi: Blind source separation using clustering-based multivariate density estimation algorithm, IEEE Trans. on Signal Processing, Vol. 48, No. 2, 2000, pp. 575-579.

[73] A. Holobar, C. Fevotte, C. Doncarli in D. Zazula: Single autoterm selection for blind source separation in time-frequency plane, in Proc. EUSIPCO'02, Toulouse, France, 2002, on CD.

[74] A. Holobar, M. Ojsteršek, D. Zazula: A new approach to parallel joint diagonalization of symmetric matrices, in Proc. The IEEE Region 8 EUROCON 2003: computer as a tool, Ljubljana, Slovenia, 2003, Vol. 2, pp. 16-20.

[75] A. Holobar, D. Zazula: A new approach for blind source separation of convolutive mixtures of pulse trains, 4th International Workshop on Biosignal Interpretation BSI ‘02, Como, Italy, 2002, pp. 163-166.

[76] A. Holobar, D. Zazula: Surface EMG Decomposition using a novel approach for blind source separation, Informatica Medica Slovenica, Vol. 8, 2003, pp. 2-14.

[77] A. Holobar, D. Zazula: Verjetnostni model površinskih elektromiografskih signalov, technical report AHIRD6, System Software Laboratory, Faculty of Electrical Engineering and Computer Science, Maribor, Slovenia, 2003.

[78] A. Hyvarinen: Survey on independent component analysis, Neural Computing Surveys, Vol. 2, 1999, pp. 94-128.

[79] A. Hyvarinen: Independent component analysis: algorithms and applications, Neural Networks, Vol. 13 No. 4-5, 2000, pp. 411-430.

[80] A. Hyvärinen: New Approximations of differential entropy for independent component analysis and projection pursuit, in Proc. Advances in Neural Information Processing Systems 10 (NIPS’97), MIT Press, 1998, pp. 273-279.

[81] A. Hyvärinen M. Inki: Estimating overcomplete independent component bases for image windows, Journal of Mathematical Imaging and Vision, Vol. 17, 2002, pp. 139-152.

[82] A. Hyvarinen, E. Oja: A fast fixed-point algorithm for independent component analysis, Neural Computation, Vol. 9, No. 7, 1997, pp. 1483-1492.

[83] A. Hyvarinen, E. Oja: Independent component analysis by general nonlinear Hebbian-like learning rules, Sig. Proc., Vol. 78, 1999, pp. 277-287.

[84] A. Hyvarinen, J. Karhunen, E. Oja: Independent Component Analysis, John Wiley & sons, Inc. New York, 2001.

[85] M. Z. Ikram, D. R. Morgan: A beamforming approach to permutation alignment for multichannel frequency-domain blind speech separation, Proc. ICASSP, 2002, pp. 881-884.

[86] S. C. Jun, B. A. Pearlmutter, G. Nolte: EMG source localization using an MLP with a distributed output representation. IEEE Trans. Biomed. Eng., Vol. 50, No. 6, 2003, pp. 786-789.

[87] T. P. Jung, C. Humphries, T. W. Lee, M. J. McKeown, V. Iragui, S. Makeig, T. J. Sejnowski: Removing electroencephalographic artifacts by blind source separation, Psychophysiology, Vol. 37, 2000, pp. 163-178.

[88] C. Jutten, J. Herault: Independent components analysis versus principal components analysis, in Signal Processing IV, Theories and Applications, J. L. Lacoume, A. Chehikian, N. Martin, J. Malbos (Eds.), Elsevier, Grenoble, France, 1988, pp. 643-646.

[89] K. J. Kang, C. K. Cheng, J. S. Lai, J. R. Shiu, T. S. Kuo: A comparative analysis of various EMG pattern recognition methods, Med. Eng. & Physics, Vol. 18, No. 5, 1996, pp. 390-395.

[90] K. Kanosue, M. Yoshida, K. Akazawa, K. Fujii: The number of active motor units and their firing rates in voluntary contraction of human brachialis muscle, Jp. Journal of Physiology, Vol. 29, 1979, pp. 427-443.

[91] M. F. Kelly, P. A. Parker, R. N. Scott: The application of neural networks to myoelectric signal analysis: a preliminary study, IEEE Trans. Biomed. Eng., Vol. 37, No. 3, 1990, pp. 221-230.

[92] S. M. Key: Fundamentals of statistical signal processing: estimation theory, Prentice Hall International, Inc., Englewood Cliffs, New Jersey, 1993.

[93] J. M. Kilner, S. N. Baker, R. N. Lemon: A novel algorithm to remove electrical cross-talk between surface EMG recordings and its application to the measurement of short-term synchronisation in humans, J. of Physiol., Vol. 8, 2002, pp. 919 - 930.

[94] M. Knaflitz, G. Balestra: Computer analyisis of the myoelectric signal, IEEE Micro, No. 10, 1991, pp. 12-58.

[95] V. M. Koch, H. A. Loeliger: A new EMG signal decomposition approach using factor graphs, in Proc. XVth ISEK Congres, Boston, USA, 2004, pp. 65.

[96] T. Kohonen: Self-organizing maps, Springer, 1995.

[97] J. L. Lacoume, F. Harroy: Performances in blind source separation, in Proc. HOS 95, Girona, Spain, 1995, pp. 25-29.

[98] R. H. Lambert, C. L. Nikias: Blind deconvolution of multipath mixtures, in S. Haykin (Ed.): Unsupervised adaptive filtering, Vol. 1, Wiley, 2000, pp. 377-436.

[99] F. Laterza, L. Lo Conte, M. Matacchione, R. Merletti: Information in the surface EMG spectral content, in Deliverable of the SENIAM project, The State of the art on modelling methods for surface electromyography, H. J. Hermens, D. Stegeman, J. Blok, B. Freriks (Eds.), Vol. 6, 1998, pp. 109-116.

[100] L. D. Lathauwer, P. Comon: Blind source separation and multichannel deconvolution, Sig. Proc., Vol. 73, No. 1-2, 1999, pp. 3-4.

[101] T. W. Lee, A. J. Bell, R. H. Lambert: Blind separation of delayed and convolved sources, Advances in Neural Information Processing Systems, Vol. 9, 1997, pp. 758-764.

[102] R. S. LeFever, C. J. De Luca: A procedure for decomposing the myoelectric signal into its constituent action potentials. Part I. Technique, theory and implementation, IEEE Trans. on Biomed. Eng., Vol. 29, 1982, pp. 149–157.

[103] R. S. LeFever, A. P. Xenakis, C. J. De Luca: A procedure for decomposing the myoelectric signal into its constituent action potentials. Part II. Execution and test for accuracy, IEEE Trans. Biomed. Eng., Vol. 29, 1982, pp. 158–164.

[104] T. W. Lee, S. Lewicki, M. Girolami, T. J. Sejnowski: Blind source separation of more sources than mixtures using overcomplete representations, IEEE Sig. Proc. letters, Vol. 6, No. 4, 1999, pp. 87-90.

[105] H. Liang: Adaptive independent component analysis of multichannel electrogastrograms, Med. Eng. & Physics, Vol. 23, No. 2, 2001, pp. 91-97.

[106] E. Mansour, A. K. Barros, N. Ohnishi: Blind separation of sources: methods, assumptions and applications, IEICE trans. Fundamentals, Vol. E83-A, No. 8, 2000, pp. 1498-1512.

[107] A. Mansour and C. Jutten: Fourth order criteria for blind separation of sources, IEEE Trans. Sig. Proc., Vol. 43, No. 8, 1995, pp. 2022-2025.

[108] A. Mansour and C. Jutten: A direct solution for blind separation of sources, IEEE Trans. Sig. Proc., Vol. 44, No. 3, 1996, pp. 746-748.

[109] A. Mansour, C. Jutten, P. Loubaton: An adaptive subspace algorithm for blind separation of independent sources in convolutive mixture, IEEE Trans. on Signal Processing, Vol. 48, No. 2, 2000, pp. 583-586.

[110] A. Mansour, N. Ohnishi: Multichannel blind separation of sources algorithm based on cross-cumulant and the levenberg-marquardt method., IEEE Trans. Sig. Proc., Vol. 47, No. 11, 1999, pp. 3172-3175.

[111] J. M. Mendel: Tutorial on higher-order statistics (spectra), signal processing and system theory: theoretical results and some applications, in Proc. the IEEE, Vol. 79, No. 3, 1991, pp. 278-305.

[112] K. A. Meraim, P. Loubaton, E. Moulines: A subspace algorithm for certain blind identification problems, IEEE Trans. on Information Theory, Vol. 43, No. 3, 1997, pp. 499-511.

[113] R. Merletti: Surface electromyography: possibilities and limitations, J. of Rehab. Sci., Vol. 7, No. 3, 1994, pp. 25-34.

[114] R. Merletti, L. Lo Conte, E. Avignone, P. Guglielminotti: Modelling of surface myoelectric signals; part I: model implementation, IEEE Trans. Biomed. Eng., Vol. 46, No. 7, 1999, pp. 810-820.

[115] C. D. Meyer: Matrix Analysis and applied Linear Algebra, SIAM, 2001.

[116] S. Micera, G. Vannozzi, A. M. Sabatini, P. Dario: Improving detection of muscle activation intervals, IEEE Eng. in Med. and Biol., Vol. 20, No. 6, 2001, pp. 38-47.

[117] E. Moreau and O. Macchi: New self-adaptive algorithms for source separation based on contrast functions, in Proc. IEEE Signal Processing Workshop on Higher-Order Statistics, South Lake Tahoe, Kalifornija, USA, 1993, pp. 215-219.

[118] J. C. Mosher, R. M. Leahy: Source localization using recursively applied projected (RAP) MUSIC, IEEE Trans. Sig. Proc., Vol. 47, No. 2, 1999, pp. 332-340.

[119] R. Mukai, H. Sawada, S. Araki, S. Makino: Frequency domain blind source separation for many speech signals, in Proc. ICA 2004, Granada, Spain, 2004, pp. 461-469.

[120] L. T. Nguyen, A. Belouchrani, K. A. Meraim, B. Boashash: Separating more sources than sensors using time-frequency distributions, in Proc. 6th ISSPA, No. 2 , Kuala-Lumpur, Malaysia, 2001, pp. 583-586.

[121] L. T. Nguyen, C. Jutten: Blind sources separation for convolutive mixtures, Sig. Proc., Vol. 45, No. 2, 1995, pp. 209-229.

[122] L. Nguyen Thi, C. Jutten, J. Caelen: Speech enhancement: Analysis and comparison of methods in various real situations, in Signal Processing VI, Theories and Applications, J. Vandewalle, R. Boite, M. Moonen, A. Oosterlinck (Eds.), Elsevier, Brussels, Belgium, 1992, pp. 303-306.

[123] C. L. Nikias, J. M. Mendel: Signal processing with higher-order spectra, IEEE Sig. Proc., Vol. 10, No. 3, 1993, pp. 10-37.

[124] C. L. Nikias, A. P. Petropulu: Higher-order spectra analysis: a nonlinear signal processing framework, Englewood Cliffs, Prentice Hall, signal processing series, New Jersey, 1993.

[125] B. A. Olshausen, K. J. Millman: Learning sparse codes with a mixture-of-gaussians prior, in Advances in neural information processing systems, Vol. 12, 2000, pp. 841-847.

[126] A. Papoulis: Probability, random variables, and stochastic processes, McGraw-Hill, New York, 1991.

[127] L. Parra, C. Spence: Convolutive blind source separation on nonstationary sources, IEEE Trans. on Speech and Audio Proces., 2000, pp. 320-327.

[128] L. C. Parra, C. V. Alvino: Geometric source separation: Merging convolutive source separation with geometric beamforming, IEEE Trans. Speech Audio Processing, Vol. 10, 2002, pp. 352-362.

[129] S. C. Pattichis, A. G. Elia: Autoregressive and cepstral analyses of motor unit action potentials, Med. Eng. & Physics, Vol. 21, No. 6-7, 1999, pp. 405-419.

[130] B. A. Pearlmutter, L. C. Parra: A context-sensitive generalization of ICA, in Proc. International Conference on Neural Information Processing, Hong Kong, 1996, pp. 151-157.

[131] D. T. Pham: Blind separation of instantaneous mixture sources via an independent component analysis, IEEE Trans. on Signal Processing, Vol. 44, No. 11, 1996, pp. 2768-2779.

[132] D. T. Pham, J. F. Cardoso: Blind separation of instantaneous mixtures of non stationary sources, IEEE Trans on Sig, Proc, Vol. 49, No. 9, 2001, pp. 1837-1848.

[133] D. T. Pham, P. Garat, C. Jutten: Separation of a mixture of independent sources through a maximum likelihood approach, in Signal Processing VI, Theories and Applications, J. Vandewalle, R. Boite, M. Moonen, A. Oosterlinck (Eds.), Brussels, Belgija, Elsevier, 1992, pp. 771-774.

[134] E. Plévin, D. Zazula: Decomposition of surface EMG signals using non-linear LMS optimisation of higher-order cumulants, in Proc. 15th IEEE symposium on computer-based medical systems CBMS '02, Maribor, Slovenia, 2002, pp. 149-154.

[135] A. Prieto, C. G. Puntonet, B. Prieto: A neural algorithm for blind separation of sources based on geometric properties, Signal Processing, Vol. 64, No. 3, 1998, pp. 315-331.

[136] D. Prutchi: A high-resolution large array (HRLA) surface EMG system, Med. Eng. & Physics, Vol. 17, No. 6, 1995, pp. 442-454.

[137] K. Roeleveld: Surface motor unit potentials, the building stones of surface electromyography, PhD thesis, 1997.

[138] K. Roeleveld, D. F. Stegeman, H. M. Vingerhoets, M. J. Zwarts: How inter-electrode distance and motor unit depth influence surface potentials, in Deliverable of the SENIAM project, The state of the art on sensors and sensor placement procedures for surface electromyography, H. J. Hermens, B. Freriks (Eds.), Vol. 5, 1997, pp. 55-59.

[139] T. A. Sears, D. Stagg: Short-term synchronization of intercostal motoneurone activity. J. of Physiol., No. 263, 1976, pp. 357-381.

[140] O. Shalvi, E. Weistein: New criteria for blind deconvolution of on-minimum-phase systems (channels), IEEE Trans. on Information Theory, Vol. 36, No. 2, 1990, pp. 312-321.

[141] P. Smaragdis: Blind separation of convolved mixtures in frequency domain, Neurocomputing, Vol. 22, 1998, pp. 21-34.

[142] E. Stålberg and B. Falck: The role of electromyography in neurology, Electroencep. and Clin. Neurophys., Vol. 103, 1997, pp. 579–598.

[143] E. Stålberg, J. Trontelj: Single Fiber Electromyography, Raven Press, New York, 1994.

[144] D. W. Stashuk: EMG signal decomposition: how can it be accomplished and used?, J. of Electromyogr. Kinesiol., Vol. 11, 2001, pp. 151-173.

[145] D. W. Stashuk: Decomposition and quantitative analysis of clinical electromyographic signals, Med. Eng. and Phy., Vol. 21, 1999, pp. 389–404.

[146] G. Staude, W. Wolf: Objective motor response onset detection in surface myoelectric signals, Med. Eng. & Physics, Vol. 21, No. 6-7, 1999, pp. 449-467.

[147] M. Suojanen, S. Andreassen, K. G. Olesen: The EMG diagnosis – an interpretation based on partial information, Med. Eng. & Physics, Vol. 21, No. 6-7, 1999, pp. 517-523.

[148] A. Šoštarič, D. Zazula, C. Doncarli: Time-scale decomposition of compound (EMG) signal, Elektrotehniški vestnik, Vol. 67, 2000, pp. 69-75.

[149] A. C. Tang, B. A. Pearlmutter, N. A. Malaszenko, D. B. Phung, B. C. Reeb: Independent components of magnetoencephalography: localization, Neural Computation, Vol. 14, No. 8, 2002, pp. 1827-1858.

[150] A. C. Tang, B. A. Pearlmutter, M. Zibulevsky, S. A. Carter: Blind separation of multichannel neuromagnetic responses, Neurocomputing, No. 32-33, 2000, pp. 1115-1120.

[151] N. Thirion, J. Mars, J. L. Boelle: Separation of seismic signals: a new concept based on a blind algorithm, in Signal Processing VIII, Theories and Applications, Trieste, Italy, 1996, pp. 85-88.

[152] L. Tong, R. Liu, Y. H. V. C. Soon: Indeterminacy and identifiability of blind identification, IEEE Trans. Circuits Syst., Vol. 38, 1991, pp. 499-509.

[153] L. Tong, R. Liu: Blind estimation of correlated source signals, in Proc. Asilomar Conf., 1990, pp. 161-164.

[154] J. K. Tugnait: Identification and deconvolution of multichannel nongaussian processes using higher-order statistics and inverse filter criteria, IEEE Trans. on Signal Processing, Vol. 45, 1997, pp. 658-672.

[155] J. Vidal, J. A. R. Fonollosa: Adaptive blind system identification using weighted cumulant slices, International Journal on Adaptive Control, Vol. 10, No. 2-3, 1996, pp. 213-237.

[156] R. Vigário, J. Särelä, V. Jousmäki, M. Hämäläinen, and E. Oja: Independent component approach to the analysis of EEG and MEG recordings, IEEE Trans. on Biomed. Eng., Vol. 47, No. 5, IEEE, New York, 2000, pp. 589-593.

[157] W. Wang, J. A. Chambers, S. Sanei: A novel hybrid approach to the permutation problem of frequency domain blind source separation, in Proc. ICA 2004, Granada, Spain, 2004, pp. 532-539.

[158] J. L. F Weytjens, D. van Steenberghe: The effects of motor unit synchronization on the power spectrum of the electromyogram, Biol. Cybern., No. 51, 1984, pp. 71-77.

[159] R. Wotiz, C. J. De Luca, S. H. Nawab: Second generation precision decomposition: solutions, in Proc. XVth ISEK Congres, Boston, USA, 2004, pp. 159.

[160] G. Wubbeler, A. Ziehe, B. M. Mackert, K. R. Muller, L. Trahms, G. Curio: Independent component analysis of non-invasively recorded cortical magnetic DC-fields in humans, IEEE Trans. on Biomed. Eng., Vol. 47, No. 5, 2000, pp. 594-599.

[161] Z. Xu, S. Xiao, Z. Chi: ART2 neural network for surface EMG decomposition, Neural Computing & Applications, Vol. 10, No. 1, 2001, pp. 29-38.

[162] H. Szu, T. Yamakawa, C. Hsu: Visual image communication using advanced neural networks, in Proc. ICA99, Aussois, France, 1999, pp. 121-126.

[163] D. Yellin, E. Weinstein: Criteria for multichannel signal separation, IEEE Trans. on Sig. Proc., Vol. 42, No. 8, 1994, pp. 2158-2167.

[164] D. Yellin, E. Weinstein: Multichannel signal separation: methods and analysis, IEEE Trans. on Sig. Proc., Vol. 44, No. 1, 1996, pp. 106-118.

[165] E. Zalewska, I. H. Petrusewicz: Quantitative evaluation of the motor unit potential complexity, Electromyog. and Clin. Neurophys., Vol. 31, 1991, pp. 373–376.

[166] D. Zazula: Experience with surface EMG decomposition using higher-order cumulants, Signal Processing 2001, in Proc. Poznañ, Poland, 2001, pp. 19-24.

[167] D. Zazula, A. Holobar: An approach to surface EMG decomposition based on higher-order cumulants, International Journal of Medical Informatics, in press.

[168] D. Zazula, D. Korže, A. Šoštarič, D. Korošec: Study of methods for decomposition of superimposed signals with application to electromyograms, in Neuroprosthetics, Pedotti et al. (Ed.), Berlin: Springer Verlag, 1996.

[169] D. Zazula, E. Plévin: An approach to decomposition of muscle and nerve signals, in Proc. Int. Conf. on Signal, Speech, and Image Processing ICOSSIP '02, Skiathos, Greece, 2002, on CD.

[170] D. Zazula, A. Šoštarič: Possible approaches to surface EMG decomposition, SENIAM, Vol. 7, 1999, pp. 169-176.

[171] D. Zennaro, P. Wellig, V. M. Koch, G. S. Moschytz, T. Läubli: A software package for the decomposition of long-term multichannel EMG signals using wavelet coefficients, IEEE Trans on Biomed. Eng., Vol. 50, No. 1, 2003, pp. 58-69.

[172] P. Zhou, M. Lowery, W. Z. Rymer: Independent component analysis of the surface electromyogram signals: simulation study, in Proc. 5. International conference on Neural Networks and Expert Systems in Medicine and Health Care, Sheffield, Great Britain, 2003, on CD.

[173] J. Zhu, X. R. Cao, Z. Ding: An algebraic principle for the blind separation of white non-Gaussian sources, Sig. Proc., Vol. 76, 1999, pp. 105-115.

[174] C. Ziegaus, E. W. Lang: Independent component extraction of natural images based on fourth-order cumulants, in Proc. ICA99, Aussois, France, 1999, pp. 115-120.

[175] A. Ziehe, K. R. Muller, G. Nolte, B. M. Mackert, G. Curio: Artefact reduction in magnetoneurography based on time-delayed second order correlations, IEEE Trans. on Biomed. Eng., Vol. 47, No. 1, 2000, pp. 75-87.

Biography

First and family name: Aleš Holobar.

Date of birth: 12. 8. 1976.

Birthplace: Celje, Slovenia

Father's name: Anton Holobar.

Mother's name: Milena Holobar.

Education:	1983 – 1991	Primary school Peter Šprajc – Jur, Žalec, Slovenia
	1991 – 1995	Grammar school I. Gimnazija Celje, Slovenia
	1995 – 2000	Faculty of Electrical Engineering and Computer Science, University of Maribor, Slovenia
	2000 – 2004	PhD study at Faculty of Electrical Engineering and Computer Science, University of Maribor, Slovenia