JYVÄSKYLÄ STUDIES IN COMPUTING
109

Andriy Ivannikov

Extraction of Event-Related Potentials from Electroencephalography Data

JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ STUDIES IN COMPUTING 109

Andriy Ivannikov

Extraction of Event-Related Potentials from Electroencephalography Data

Esitetään Jyväskylän yliopiston informaatioteknologian tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston Agora-rakennuksen Gamma-salissa
joulukuun 11. päivänä 2009 kello 14.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Information Technology of the University of Jyväskylä,
in the Building Agora, Gamma-hall, on December 11, 2009 at 2 p.m.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2009

Extraction of Event-Related Potentials from Electroencephalography Data

JYVÄSKYLÄ STUDIES IN COMPUTING 109

Andriy Ivannikov

Extraction of Event-Related Potentials from Electroencephalography Data

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2009

Editors
Timo Männikkö
Department of Mathematical Information Technology, University of Jyväskylä
Pekka Olsbo, Marja-Leena Tynkkynen
Publishing Unit, University Library of Jyväskylä

URN:ISBN:978-951-39-3771-3
ISBN 978-951-39-3771-3 (PDF)

ISBN 978-951-39-3742-3 (nid.)
ISSN 1456-5390

Jyväskylä University Printing House, Jyväskylä 2009

ABSTRACT

Ivannikov, Andriy
Extraction of Event-Related Potentials from Electroencephalography Data
Jyväskylä: University of Jyväskylä, 2009, 108 p.(+included articles)
(Jyväskylä Studies in Computing
ISSN 1456-5390; 109)
ISBN 978-951-39-3742-3
Finnish summary
Diss.

The research work reported in this thesis addresses the issues related to denoising of event-related potentials (ERP) in multichannel electroencephalography (EEG) data. The main idea behind the ERP denoising methods presented in this thesis lies in separating ERP and noise subspaces according to the linear instantaneous mixing model. When subspaces are extracted, the denoising of the channels of measurements is reached by the inverse transformation of the previously obtained ERP components ignoring components related to the noise subspace. The emphasis of the thesis is on finding appropriate problem-specific criteria, which allow ERP and noise components in multidimensional EEG data space to be reliably distinguished and, thus, for the separation of ERP and noise subspaces by finding the basis vectors that span them. The criteria that have been studied are based on exposing the data to some modification that influences signal and noise subspaces or, more precisely, signal and noise constituents of the data, differently. We explore those subspace-specific changes that are seen on the level of second-order statistical properties of the data. Namely, the two covariance matrices of data before and after the modification are compared. Moreover, we concentrate our attention on those modifications that exploit the data which have three dimensions of variability: channels, time samples, and trials. Yet the scope of this thesis goes beyond a sole proposition of the subspace separation criteria and touches also on such topics as (1) practical aspects of application of the denoising methods, (2) validation of the data and results, (3) developing a comparison framework, (4) analysis and interpretation of the results, (5) elaboration of suggestions and recommendations for improving the performance of the denoising methods.

Keywords: Component analysis, Electroencephalography, Event-related potential, Signal processing, Signal-to-noise ratio, Source separation, Spatial denoising

Author Andriy Ivannikov
Department of Mathematical Information Technology
University of Jyväskylä
Finland

E-mail: `andriy.v.ivannikov@jyu.fi`

Supervisors Professor Tommi Kärkkäinen
Department of Mathematical Information Technology
University of Jyväskylä
Finland

Professor Tapani Ristaniemi
Department of Mathematical Information Technology
University of Jyväskylä
Finland

Professor Heikki Lyytinen
Department of Psychology
University of Jyväskylä
Finland

Reviewers Professor Karen Eguiazarian
Institute of Signal Processing
Tampere University of Technology
Finland

Professor Tarmo Lipping
Department of Information Technology, Pori
Tampere University of Technology
Finland

Opponents Professor Dennis L. Molfese
Department of Molecular, Cellular & Craniofacial Biology
University of Louisville
USA

Professor Tarmo Lipping
Department of Information Technology, Pori
Tampere University of Technology
Finland

PREFACE

The work reported in this thesis has been carried out at the Department of Mathematical Information Technology of the Faculty of Information Technology of the University of Jyväskylä in close collaboration with the Department of Psychology of the Faculty of Social Sciences of the University of Jyväskylä during the years 2003-2009.

The work presented in this thesis is at the interface of two disciplines: signal processing technology and psychophysiology, thereby implementing the strategy of multidisciplinary research designed to produce innovative approaches by merging advanced knowledge from two different domains of research. This work promotes the problem-specific approach to the research, where the stage of investigating the relations between the concepts and the parameters of the problem acquires critical importance. This thesis comprises six articles which have been published in international conferences and journals presenting original contributions to the area of ERP denoising.

ACKNOWLEDGEMENTS

First of all I wish to express my sincere gratitude to Professor Tommi Kärkkäinen for his truly invaluable support and guidance provided in the course of my work. It is hard to overestimate his positive influence on my formation as a scientist. His constant belief in my abilities has encouraged me all this time.

It would be difficult to imagine this thesis materializing without the substantial contribution provided by Professor Heikki Lyytinen to whom I give my profound gratitude. He was the one who took me under his supervision at the very beginning of this scientific journey, when I was looking for a topic for my Master's Thesis. His bright ideas delivered in his unique manner in the warm friendly atmosphere of our discussions in the office, with the beautiful view of the Jyväsjärvi lake, strongly influenced my scientific ideology and always inspired me.

I am greatly indebted to Professor Tapani Ristaniemi for his immeasurable encouragement and help during the work on the dissertation. He also supervised my Master's Thesis. After a break he became my supervisor again at a critical moment in 2006, when I had reached an impasse in the research progress. His return marked a new milestone in the development of the thesis, when the research accelerated significantly and improved in both quality and amplitude.

My special gratitude is addressed to colleagues from the Department of Psychology and particularly to Professor Paavo H.T. Leppänen, Jarmo Hämäläinen, and Tiina Huttunen for providing me with data, expressing their interest in my research, and for valuable comments on various aspects of the data processing and interpretation.

I would also like to thank to all the staff members of the Department of Mathematical Information Technology and the Faculty of Information Technology for helping me with all the administrative issues arising during my work and being very friendly and interesting interlocutors to meet. The financial support from the Agora Center, Jyväskylä Graduate School in Computing and Mathematical Sciences (COMAS), and Department of Mathematical Information Technology is gratefully acknowledged.

I am grateful to the reviewers of the thesis, Professor Karen Eguiazarian and Professor Tarmo Lipping for their careful manuscript reading and for useful suggestions that helped to improve the quality of the thesis. I deeply appreciate their work that was done to a rather tight schedule.

My special thanks are due to Narciso González Vega for sharing experience with me during friendly informal communications and coffee breaks.

I am especially grateful to Professor Vagan Terziyan and Dr. Helen Kaykova who initiated the exchange program between the University of Jyväskylä and the Kharkov National University of Radioelectronics. Moreover, I appreciate very much their great support during the early period of my stay and studies in Jyväskylä and their continuing always interesting and informative discussions.

It gives me great pleasure to say thank you to my friends who were also my colleagues, Igor Kalyakin and Daniil Ivanov. Our expressive and emotionally rich debates during which a lot of bright ideas were born will forever remain in my memory. I greatly value our time-tested friendship.

Of no less importance for me are my friends who were not my colleagues and whom I feel privileged to have met. My hearty thanks to Sergey, Liuda, Misha, Zhenja, Natasha, Sasha, Anna, and many others for sharing time with me and making my life richer.

I am very thankful to my dear Lenochka for being in my life and being the endless source of inspiration for me. Finally, I wish to show everlasting gratitude to my family who love me so much for who I am. My warmest thanks go to my younger brother Aleksandr, who always encouraged and stood by me. My heartfelt 'spasibo' to my mama and papa who laid the foundations of my being and aspirations, taught me all the important things I need to know, and were always ready to help me.

ACRONYMS

AEP	Auditory Evoked Potential
BCI	Brain Computer Interface
BSS	Blind Source Separation
DBI	Direct Brain Interface
EEG	Electroencephalography
ECG, EKG	Electrocardiography
EMG	Electromyography
EOG	Electro-Oculography
ERP	Event-Related Potential
EP	Evoked Potential
fMRI	Functional Magnetic Resonance Imaging
GA	Grand Average
HOS	Higher-Order Statistics
ICA	Independent Component Analysis
ISA	Independent Subspace Analysis
ISI	Interstimulus Interval
LDN	Late Discriminative Negativity
MCA	Minor Component Analysis
MEG	Magnetoencephalography
MMN	Mismatch Negativity
MSE	Mean-Square Error
NSR	Noise-to-Signal Ratio
PCA	Principal Component Analysis
pdf	Probability Density Function
SEP	Somatosensory Evoked Potential
SNR	Signal-to-Noise Ratio
SOS	Second-Order Statistics
VEP	Visual Evoked Potential

LIST OF FIGURES

CONTENTS

LIST OF INCLUDED ARTICLES

PI Andriy Ivannikov, Tommi Kärkkäinen, Tapani Ristaniemi, and Heikki Lyytinen. Extraction of ERP from EEG data. *Proceedings of the 9th International Symposium on Signal Processing and its Applications*, 2007.

PII Andriy Ivannikov, Tommi Kärkkäinen, Tapani Ristaniemi, and Heikki Lyytinen. SNR Criterion Maximization for the Extraction of ERP from EEG Data. *Proceedings of the Finnish Signal Processing Symposium (FINSIG'07)*, 2007.

PIII Andriy Ivannikov, Tommi Kärkkäinen, Tapani Ristaniemi, and Heikki Lyytinen. Traditional Averaging, Weighted Averaging, and ERPSUB for ERP Denoising in EEG Data: A Comparison of the Convergence Properties. *Proceedings of the International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2008): International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS 2008)*, 1, 2008, 195–202.

PIV Andriy Ivannikov, Tommi Kärkkäinen, Tapani Ristaniemi, and Heikki Lyytinen. A Method for Extracting Subspace of Deterministic Sources from EEG Data. *Proceedings of the 3rd International Symposium on Communications, Control and Signal Processing (ISCCSP 2008)*, 2008, 1361–1365.

PV Andriy Ivannikov, Tommi Kärkkäinen, Tapani Ristaniemi, and Heikki Lyytinen. Separation of ERP and Noise Subspaces in EEG Data without Whitening. *Proceedings of the 21th IEEE International Symposium on Computer-Based Medical Systems (CBMS 2008)*, 2008, 65–69.

PVI Andriy Ivannikov, Igor Kalyakin, Jarmo Hämäläinen, Paavo H.T. Leppänen, Tapani Ristaniemi, Heikki Lyytinen, and Tommi Kärkkäinen. ERP Denoising in Multichannel EEG Data Using Contrasts between Signal and Noise Subspaces. *Journal of Neuroscience Methods*, 180(2), 2009, 340–351.

1 INTRODUCTION

The main issue addressed in this thesis is the denoising of Event-Related Potentials (ERP) in ElectroEncephaloGraphy (EEG) data. The EEG data are measurements of the electric fields produced by the brain. These fields are captured with the aid of electrodes positioned on the scalp. A channel of measurement is associated with each pair of active electrode and its reference electrode, by means of which a signal is measured. Therefore, put formally, channel is a dimension of data variability physically related to space. Analysis of EEG provides invaluable opportunities for both basic brain research and clinical applications, like diagnosis, monitoring of the patient's condition, and assessment of drug effects [dSvLR87, Lip01, LJ04, NdS05].

ERPs are electric potentials which originate in the brain and arise as a consequence of brain activity related to the processing of stimuli. Event-related potentials are an interesting object for investigation, because they reveal an important and efficient tool for studying cognitive processes [Nää92, Han04, Luc05, CTB07]. The understanding of mechanisms of cognition and the forms of their manifestation provides a basis for studies in applied fields related to, for example, reading, learning, and language skills development [Mol00, LAE$^+$04a, KMR06, HS07, MMB$^+$08]. ERPs are time- and heavily phase-locked to a stimulus. Besides ERPs, EEG measurements are contributed to also by spontaneous EEG activity and artifacts (potentials related to various biological and external phenomena). All of these artifactual sources plus spontaneous EEG activity mixed with time-phase-locked signals complicate the determination of the parameters of the ERP waveform and, therefore, are treated as noise in the context of ERP investigation. Typical magnitudes of ERP signals range from less than a microvolt to tens of microvolts. The level of noise may exceed that of ERP by hundreds of microvolts resulting in a poor signal-to-noise ratio (SNR) in a single measurement of brain response to a stimulating event.

The conventional method to improve the accuracy of the ERP estimate is the averaging of repeated measurements from the same electrode (see Section 2.5). However, the performance of traditional averaging is often not satisfactory, especially when the number of single measurements is decreased. Consequently, a number of more efficient methods built on top of the basic averaging were developed (see Section 3.1). Although these techniques are effective in denoising of ERP, the accuracy of ERP estimates produced by them may still be insufficient.

For the past few decades, since blind spatial decomposition methods gained significant attention and became widespread, EEG analysis research has developed rapidly (see Section 3.2). In particular, methods of this type can be applied to separating spatio-temporally distinct sources contributing EEG measurements and ERP denoising. Compared to the methods which consider channels only separately, the spatial component separation techniques benefit also from the information contained in the relationships between channels. These methods offer a good performance in source separation averaged over a wide-range of real-life problems. Nevertheless, there have been relatively few studies carried out with the focus on developing the specialized methodologies for spatial ERP extraction. Traditionally, source separation techniques used in the context of ERP denoising are applied to raw EEG segments in order to isolate artifacts prior to the averaging. Another target application of these methods appears in the context of ERP analysis and consists in decomposing ERP waveform into simpler and psychophysiologically interpretable subcomponents. This problem usually assumes that the estimates of the channel-wise ERP waveforms obtained after averaging have already sufficiently good accuracy. Our study focuses on extracting two subspaces spanned by ERP and noise sources.

Since ERP denoising is a many-sided and rapidly developing area, the following considerations are limited mainly, but not only, to denoising approaches based on spatial component extraction in the framework of the linear instantaneous mixing model. Moreover, the denoising on the level of a single channel without exploiting the spatial information is also discussed. Even more we further narrow the focus of our attention to seeking for consistent and reliable problem-specific criteria that are able to discriminate between signal and noise components or constituents in the framework of the assumed models. The effectiveness of the implementations of the denoising algorithms based on these criteria has been paid less attention. When it is possible, our realizations rely on explicit solution methods, like Eigenvalue decomposition. Partially they are done using basic gradient descent/ascent search procedure. The problem-specific criteria are found based on analysis of the problem area and underlying brain processing mechanisms. Furthermore, we generalize on the ideas contained in the attached articles in the form of a framework for separating the ERP and noise components in multidimensional EEG data using the linear transformation and the subsequent dimension reduction by neglecting the components related to the noise subspace during inverse transformation.

Often, when a new real problem arises, standard and well-known techniques designed for a class of similar problems are tried in order to solve it. For example, it is commonly accepted to apply Independent Component Analysis (ICA; see, e.g., [Hyv99c, HKO01]) techniques to EEG data for source separation [MBJS96, MJB+97, MWT+99, JMH+00, JMM+01, Vig97, MVMZ04], which are also applied for solving wide range of other real-life source separation problems including acoustics, speech processing, telecommunications etc. In general, ICA methods try to separate maximally independent sources. However in practice assumptions made about the ICA model and input data are usually not completely valid, and the criteria stressing independence do not necessarily lead to the most unmixed sources, i.e. when each component contains the largest part possible of one source and minor possible contributions from the other sources. Nevertheless, in almost every real-life problem some specific additional features and data properties exist. Such extra domain knowledge can be taken into account as part of the search criteria by the optimization procedure. This has an effect of regularization or specialization of the problem and, finally, may increase the accuracy of the solution.

Therefore in the given research we adhere to the following strategy of the research process. First, we consider how similar problems are solved in the current source separation research. Then the specific features of EEG/ERP data are plugged into the existing source separation methodology to result in new problem-specific algorithms.

The structure of the work is organized as follows. First, in Chapter 2 we explain the problem area and introduce the notations. This creates a basis in terms of which the following discussion is built. Then, in Chapter 3 an overview of the literature is made presenting the issues, state-of-the-art methodologies, and obtained achievements related to (1) denoising methods (Sections 3.1 and 3.2), (2) their application to EEG data (Section 3.3), and (3) the validation of the obtained results (Section 3.4). In Chapter 4 a short review of the results obtained in the research reported in this thesis is presented first (Section 4.1). It is then followed by a brief description of the attached articles (Section 4.2). Then, an extended overview of the contributions of the thesis is presented (Section 4.3). A discussion of issues that are encountered in practice that may prevent optimal subspace separation as well as some suggestions to overcome potential problems are given in Section 4.3.3. The conclusions and possible extensions of the work for further research are described in Chapter 5.

2 DESCRIPTION AND FORMALIZATION OF RESEARCH AREA

2.1 Brain functioning and EEG components

The brain is the anatomical organ of the nervous system responsible for control of the human organism and management of the unconscious and conscious operations comprising higher nervous activity [Nol01, Res01, BCP06]. Brain matter is composed of biological cells called neurons. On a higher level of the structural model, the brain can be divided into many areas specialized for performing their designated functions. The structural model of brain functioning is theoretically explained in terms of functional brain segregation, specialization and integration [Zek78, Bla90, You90]. According to this common neurophysiology assumption, certain cortical areas are specialized for processing of some specific aspects of perception or motion, and this specialization is segregated within the cortex. The functional integration mediates different specialized areas to form the cortical infrastructure responsible for a single function.

The EEG and ERP data are measurements of the electric potentials (usually measured in microvolts, μV) along the scalp arising as a consequence of brain activity. Thus, analysis of EEG data provides an understanding of the underlying brain processes if data properties are interpreted correctly. At the same time, certain numerical indicators computed from EEG can be used to quantify psychophysiological phenomena [FLA^{+}06, TT09]. For this purpose, pre-processing, to extract relevant information, is often required [LJYHH95]. The EEG data are collected with the aid of electrodes positioned on the scalp. The standardized International 10/20 Electrode Placement System developed by Jasper in [Jas58] is recommended by the International Federation of Electroencephalography and Clinical Neurophysiology and provides guidance for placement of 21 electrodes on the scalp(see Figure 1). As the need for higher spatial resolution has increased, other high-density electrode positioning configurations have been proposed to extend the original 10/20 system. Among these are the 10/10 system that defines locations for 81 electrodes [CLN85] and 10/5 configuration enabling the use of

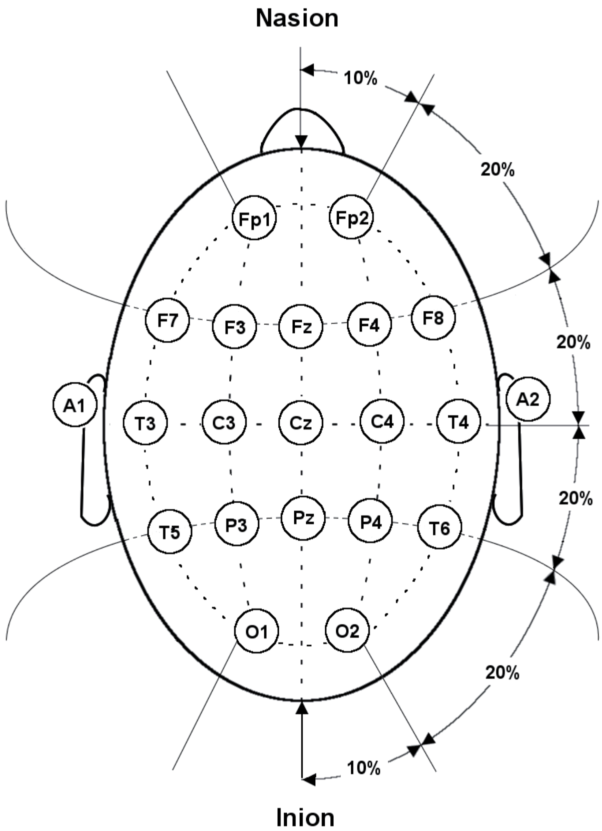

FIGURE 1 Schematic illustration of the International 10/20 System of Electrode Placement. Here the capital letters stand for: F - Frontal lobe, T - Temporal lobe, C - Central lobe, P - Parietal lobe, O - Occipital lobe. The small letter 'z' refers to an electrode placed on the mid-line.

more than 300 electrode locations [OP01]. We will also use the term *channel of measurement* or just *channel* to associate a measured signal with the pair of electrodes, with the aid of which the signal is acquired. As in this thesis only unipolar measurements are used, the channels are always labeled by the name of the corresponding non-reference electrode. For the same reason we may use the term *electrode location* instead of the *channel*.

As the *EEG signal*, we may assume one of the three following interpretations: (1) the real electric signal captured by the electrode, (2) its abstraction, when the physical meaning is ignored, or (3) its sampled, recorded, and visualized versions. However, a specific meaning should either be clear from the context or mentioned explicitly. Normally, if we talk about processing and visualization of a signal, we assume that it is digitized, i.e., discrete. Otherwise, when we discuss the nature of a signal as the derivative of a process, we assume that the signal has a continuous form, i.e., whatever fine resolution in time-frequency domain is available.

Besides the electrode cap, the measuring system includes a number of other components: amplifiers, hardware filters, recording devices, etc (see Figure 2). First, the differential amplifiers amplify the voltage between each active electrode and its reference. The amplified signal is then usually digitized via an analog-to-digital converter, after being passed through an analog anti-aliasing filter.

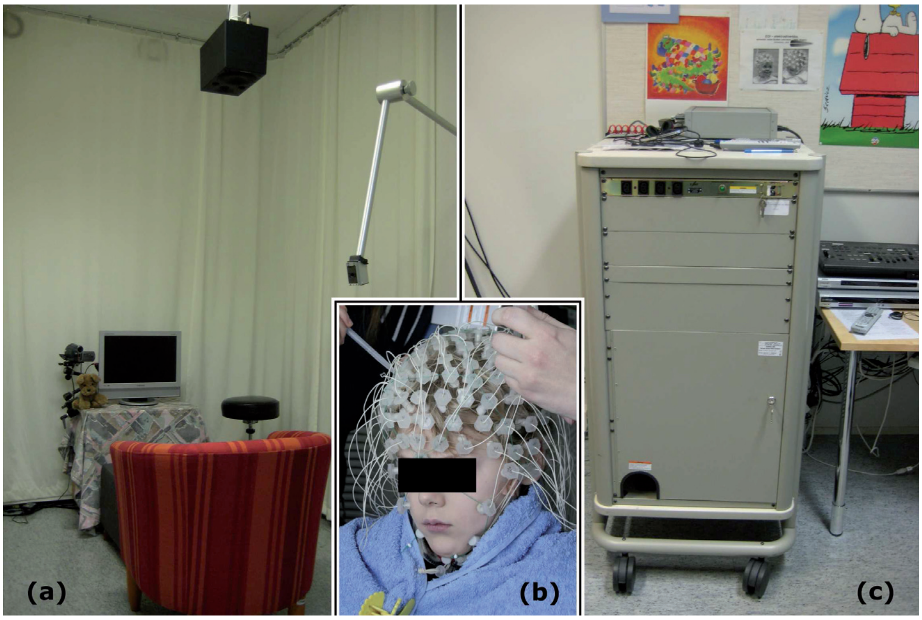

FIGURE 2 Elements of the measuring system: (a) participant's place, (b) electrode cap being fitted on a participant, (c) main unit of the data acquisition system including amplifiers, hardware filters, recording devices, etc.

The spontaneous EEG signals are composed of different brain activities, which can be categorized in particular by their frequency content [Dau02, SJ05, Buz06] [1]:

Delta (δ) waves occupy $0.5 - 4$ Hz frequency band. They are normally observed in adults during slow wave sleep and in infants. Pathologically, they are also observed from different lesions and metabolic encephalopathy hydrocephalus.

Theta (θ) waves have the frequency range $4 - 8$ Hz. In normal conditions, they arise in young children and during drowsiness or arousal in adults and older children. Abnormally theta oscillations may occur during certain lesions and disorders, metabolic encephalopathy and hydrocephalus.

[1] Some differences in the definition of these bands may appear in the literature.

Alpha (α) waves are found in $8-12$ Hz frequency range. They are normally seen during eyes closed and relaxation. Alpha waves can be abnormally observed during coma.

Beta (β) waves appear in $12-30$ Hz frequency range. Normally, beta waves manifest when the participant is in a state of active concentration and the process of excited, anxious, or busy thinking. Abnormally, beta waves accompany different pathologies and drug effects.

Gamma (γ) waves: reveal themselves in frequency band $26-100$ Hz. Gamma oscillations are supposed to reflect certain cognitive and motor functions. Gamma waves are also associated with stress.

Since the time courses of these brain activities reveal certain oscillations, they are called oscillatory activities, or brain oscillations, or brain rhythms. As follows from the above brief descriptions of the oscillatory activities, the actual frequency content of real EEG signals depends heavily on the level of attentiveness and concentration of the participant [2].

From the psychophysiological perspective, the brain activities can also be classified according to induced and ongoing (spontaneous) brain activities.

The ongoing brain activity is a sort of brain activity that is only weakly or not at all related to processing of external stimuli and is not a consequence of other specific events according to our consciousness. Although this type of activity is considered as noise in the stimulus-related brain research, it still possesses the information with regard to the current mental state of a patient (e.g., alertness, wakefulness). Some types of oscillatory activity, like alpha waves, are characterized as ongoing brain activity.

The induced brain activity is associated with the processing of a stimulus or the occurrence of some event coming from the external environment or body parts. The induced activity is thus always time-locked to a stimulus or event, but is not always phase-locked. For example, event-related potentials (ERPs) related to the induced brain activity are both time- and heavily phase-locked by definition [Nää92, Han04, Luc05, CTB07] (see Figure 3). In contrast, certain types of oscillatory activity, like gamma waves, are time-locked to the event but not phase-locked.

Event-related potentials are defined as electrophysiological brain responses associated with processing of an external event. The event can be, for example, a sensory stimulus or an absence of an expected sensory stimulus. The ERP components are divided into brainstem components (occurring approximately at the range $0-12$ msec after stimulus onset), middle-latency components (occurring approximately at the range $12-50$ msec after stimulus onset), and long-latency or cognitive ERP components (occur starting approximately at 50 msec and typically last to 500 msec after stimulus onset). Long-latency ERPs are associated with cognitive processes, like execution of memory, emotion, or attention tasks. The evoked potentials (EPs) reveal a class of ERPs that are not directly related

[2] In the literature, synonyms *patient*, *subject*, *probationer*, *individual*, and *person* may occur.

to cognition but to the execution of basic functions of perception, e.g., forming an internal representation of a stimulus and passing the retrieved information to cortex centers involved in the execution of cognitive functions. Therefore EPs include the ERP components that occur first with the latencies shorter than with those of ERPs associated with the execution of cognitive functions. The EPs are originated at the brainstem level. The EPs can not be elicited by the omission of an awaited stimulus, i.e. the physical presence of the stimulus is the necessary prerequisite for EPs to occur. The definition of EPs may vary, but in many cases it is appropriate to say that EPs encompass brainstem components and middle-latency components. By the stimulated sensory system the EPs are divided into the following types: auditory evoked potentials (AEP), visual evoked potentials (VEP), and somatosensory evoked potentials (SEP).

As our attention in this thesis is focussed on event-related potentials, let us discuss the nomenclature, nature, and properties of ERPs in more detail. Event-related potentials occur in a series of successively appearing deflections forming a so-called *ERP waveform*. Therefore the ERP waveform is comprised of several peaks or *ERP components* which have different functional explanations. In the following, we call the *ERP component* in abbreviated form simply as *ERP*. However, when a sentence is related to all ERP components, we tend to use the term *ERP waveform*. A schematic illustration of the ERP waveform typical for auditory modality is presented in Figure 3.

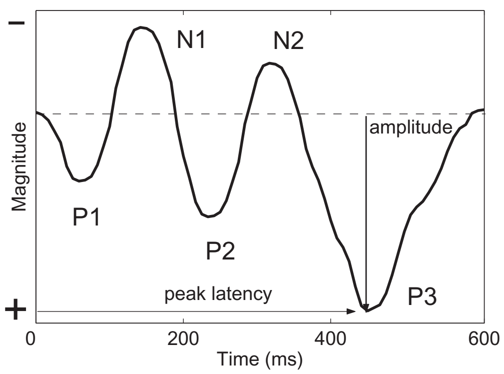

FIGURE 3 Schematic view of the averaged ERP waveform typical for auditory modality and indication of the related concepts

The three important characteristics of the ERP components are amplitude, latency, and polarity. The *amplitude* is the maximum voltage level of the ERP component with respect to a baseline. The *latency* is defined as the time elapsed from the stimulus onset to the place where the maximum voltage of the ERP component is reached. The *polarity* corresponds to a sign of the voltage of the ERP component. Moreover, besides the amplitude, latency, and polarity, ERPs are characterized by the spatial distribution [3] of the underlying activated areas over the cortex.

There is continuing debate regarding the nature of ERPs in the scientific literature. One group of researchers adhere to the opinion that ERPs are obtained, when many neurons are activated simultaneously. Another group hold the opinion that ERPs reveal themselves due to the phase resetting and synchronization of numerous ongoing oscillatory activities. Recent studies suggest that the nature of ERPs is more likely comprised of both models playing a significant role in generating ERPs [FDG$^+$04]. For more information on this issue see [RC96, Baş99a, Baş99b, SS08] and follow the included references. Note, however, that both physiological bases or physical explanations of ERP phenomena fit the mathematical model commonly assumed for ERPs (see Section 2.5) and thus the real nature of ERPs does not matter in the framework of our research goals.

To facilitate the exchange of knowledge, the nomenclature of ERP components is standardized and used in the scientific literature accordingly. The polarity of the potential is denoted by the leading letter 'N' for negative and 'P' for positive. The following number denotes either ordinal number of the component in the ERP waveform along the time or its typical observed latency. For example, N1 (or N100) is the negative potential that manifests at latency around 100 ms (70 to 200 ms), P3 (or P300) is a positive component that exhibits latencies in the time window from 300 to 500 ms. For some potentials, special names are attributed. Thus, mismatch negativity (MMN) is the negative component of ERP waveform that is found mostly at the latencies 150-250 ms. The EEG data used in numerical studies described in this thesis were collected in the scope of experiments designed to elicit MMN in particular. A comparable typology associated with positive/negative deflections at latencies from 50 to 500 msec is used in, for example, auditory and visual domains, although the processes they reflect may not be similar. Moreover, all data sets used in this thesis come from experiments related to auditory modality. Therefore the presentation has a tendency to discuss phenomena in the context of this modality. However the results should not be limited only to auditory ERPs.

Some of the ERP characteristics, such as amplitude and latency, are modulated by the properties of the stimuli and the internal state of the nervous system. The two groups of ERP subcomponents - exogenous and endogenous - are distinguished by the character of the conditions stipulating their properties. The exogenous components are modulated by physical properties of external stimuli, whereas the properties of endogenous components are affected by the internal state and capabilities of the subject (e.g., attention, arousal, memory performance etc). Since the ERP components can be contributed by several generators, they can be partially both exogenous and endogenous at the same time. The earlier ERP components (e.g., N1 and P1) are believed to be more exogenous in contrast to the components coming later in time (e.g., P3 and N4). Thus, properties of the ERPs can be important indicators of internal conditions of a participant including also those that are pathological in nature and therefore are potentially useful for

[3] In the literature, terms *topography* and *spatial map* may appear.

the purpose of medical diagnosis and monitoring.

The EEG data are collected during an experiment under some conditions called an experimental paradigm. Referring to ERP experiments the use of the term *paradigm* usually refers to stimulus presentation methods and parameters. The *oddball paradigm* is one of the most often used and characteristic to experiments in which mismatch negativity is recorded [Nää92]. This consists of the infrequent and unpredictable presentation of a target (deviant) stimulus within a relatively long series of non-target (standard) stimuli (see, for example, Figures 5 and 7 illustrating the schemes of some oddball paradigms). The target stimulus deviates from the standards in a property such as duration, magnitude, rise time, frequency, color, intensity related to some sensory modality, auditory, visual, or somatosensory, - the auditory being the most "native" modality of MMN theory and experiments. For more detailed description of ERP elicitation paradigms used in data collection experiments with outcomes used in this thesis and the following data processing see Section 2.3.

Below we explain in more detail some of the ERP components that are elucidated in measurements used in the research presented in this thesis:

N1 or N100: The N1 wave is a fronto-central negative deflection that is elicited by an abruptly commencing stimulus [NP87]. In particular, within the oddball paradigm, N1 appears as a response to any change of the stimulation independently of whether it occurs in the standard sequence or reveals the exposition of a deviant. It is not a unitary component but rather is composed of several functionally distinct subcomponents. In part, the properties of N1 are determined by specific physical or sensory properties of a stimulus (this is related to transient detection of an event) but also by a so-called modality non-specific subcomponent that is associated with the level of arousal and attention. This subcomponent participates in initiation of the involuntary orienting response (OR) to a novel stimulus.

P3 or P300: The P3 wave can be elicited during the selective attention experiment carried out within the framework of an oddball paradigm. During this discriminating task, the participant is instructed to detect a deviant stimulus within a sequence of standard stimuli and to react in some way to this event (pressing a keyboard or mouse button or simply mental counting and concentration on the event). From the psychophysiological perspective, the P3 is thought to represent activity related to evaluation of the target stimuli with the performance conditional on attentional capacity (see [PK95, HM03] for the theories on the rationale behind the P3). The latency of P3 reflects the speed of cognitive processing and the amplitude corresponds to the amount of allocated brain resources [Kok97]. The P3 itself is composed of the two subcomponents P3a and P3b, which represent different aspects of the information processing. The P3a is related to the attention mechanisms (cognitive equivalent of an orienting response), and P3b is the canonical P3 wave.

MMN: is an important tool for studying mechanisms of cognition process in its early stages. MMN is viewed as the manifestation of pre-attentive information processing related to the formation of memory traces representing incoming information and involuntary attention switching mechanisms [Nää92, WS95, Sch96, EAWN98, NW99, NJW05]. For more detailed discussion on interpretations and the rationale behind MMN one is referred to [WC98, Win07]. A review on the MMN and its role in basic research of central auditory processing can be found in [NPRA07]. The MMN elicitation experiments are built on top of the oddball paradigm. In principle, MMN can be elicited by stimulation of any type of the human sensory systems. However, most studies associated with MMN have focused on the auditory modality. The auditory MMN was first discovered by Näätänen et al. [NGM78]. Lately visual MMN was disclosed by Cammann [Cam90]. MMN arises in response to a presentation of infrequent random deviant stimulus within a sequence of standard stimuli. For example, in auditory paradigms the deviant tone can differ from the standards by frequency (pitch), duration, intensity, etc [PAO+00]. The deviance between the stimuli in the visual case can be expressed through such perceptual features as size, color, or duration of the presented images. It is commonly accepted that the increase of deviancy (from the standard) produces MMN with larger amplitude and smaller latency (see, e.g., Figure 5). However, the statement regarding the magnitude of deviance effect on the MMN amplitude was questioned by [HCJ+08], with authors suggested that this relationship does not hold. Referring to Figure 3 MMN should be thought to appear approximately at the same latency window as N1. In fact MMN and N1 are usually temporally overlapped. As MMN is a response to a deviant stimulus, and N1 occurs in response to both standard and deviant stimuli, MMN can be obtained by subtracting the ERP response to standard from the ERP response to deviant (see Section 2.2 for more details).

At this point, it is useful to discuss in more detail the clinical and psychophysiological relevance of ERP recordings. An illustrative example is related to studies in language skills development and speech processing. ERPs are shown to be characteristic indicators of auditory processes involved in speech perception. Therefore, they provide means to uncover neuronal prerequisites of language problems such as dyslexia, impairment of reading acquisition [LGH+05]. Dyslexia is entailed by abnormal processing of speech and visual language at their interface. It was shown that, even at a very early age, years before reading skills are developed, ERPs to speech sounds can be used to discriminate between children with and without risk for dyslexia and reliably determine predisposition to later language development and reading acquisition [LAE+04a, LAE+04b]. Since ERPs reveal the interface to understanding the underlying auditory processes, systematic differences between ERP characteristics among groups may reflect certain deficits or differences in associated processes. In the language-related psychophysiology studies, long-latency or cognitive ERP-responses to auditory and speech stimuli has received more attention. The auditory ERP components associated with language-related processes, which are most relevant to reveal differences between participants with and without language or reading impairment are N1, MMN, and P3. Among these MMN earned a priority attention in recent psychophysiological studies of language problems, in part due to the apparent potential to use it as an attention-independent measure of auditory cognition. As mentioned above, MMN is assumed to reflect the detection of the deviation from the memory representation of preceding repeated standard stimuli. Hence, if MMN does not appear, this can be interpreted as a consequence of problems in the auditory sensory memory, or in the auditory discrimination process. As seen from this example, MMN can be used to make early diagnosis of potential language problems to prescribe a preceding treatment.

Moreover, EEG measurements are contributed to by many other sources, including biological and non-biological ones, which are called artifacts (see Figure 4) [SJ05]. Below we present a list of artifacts with short descriptions of their nature and artifactual character.

Biological artifacts:

Cardiac: electric potentials that appear due to the heart beats and propagate to the electrodes mounted on the head. They are also analyzed in electrocardiography (EKG, ECG). These artifacts are typically represented by a series of repeated spikes synchronized to the contraction of the heart muscle and to the electrocardiogram. In fact, they are ECG measured by the electrodes placed on the scalp.

Ocular: electric potentials that are caused by the movements of eyes, saccades, and blinking. The electric activity due to the contraction of eye and eyelid muscles is measured also in electro-oculography (EOG). These potentials usually reveal very large peaks (up to hundreds of microvolts) and thus are one of the major sources of artifactual behavior in EEG measurements.

Myographic: electrical signals produced by muscles that contaminate the EEG data. These disturbances occur when body parts being moved result in high frequency fluctuations comparable in amplitude to the ongoing background EEG. These potentials are also studied in electromyography (EMG) recordings.

Respiratory: artifacts result from body movement due to breathing. Respiratory artifacts originate in phase with respiration, and thus reveal a cyclic-appearing low-frequency baseline sway.

External artifacts:

Equipment: artifacts are stipulated by properties of components of the measuring system. For example, if an EEG signal exceeds the dynamic range of the amplifier, then the resulting measurement appears to be cut at its lower or upper bound. Also, the internal electrical circuits of the measuring equipment may cause distortions to the signal while it is being passed through and processed.

Electrode: artifacts arising due to the poor attachment or sudden disconnection of an electrode that significantly increases the impedance. These artifacts are identified as rapid and large changes in the voltage of EEG measurements. Therefore, they look like high magnitude and often flattened signals. The flattening occurs as a result of the signal exceeding the dynamic range of the amplifier. Such artifacts have single-channel affiliation since only one channel is influenced by each of these corruptions.

Power supply noise: delivered by the alternating currents of the main power supply, e.g. 50 Hz noise.

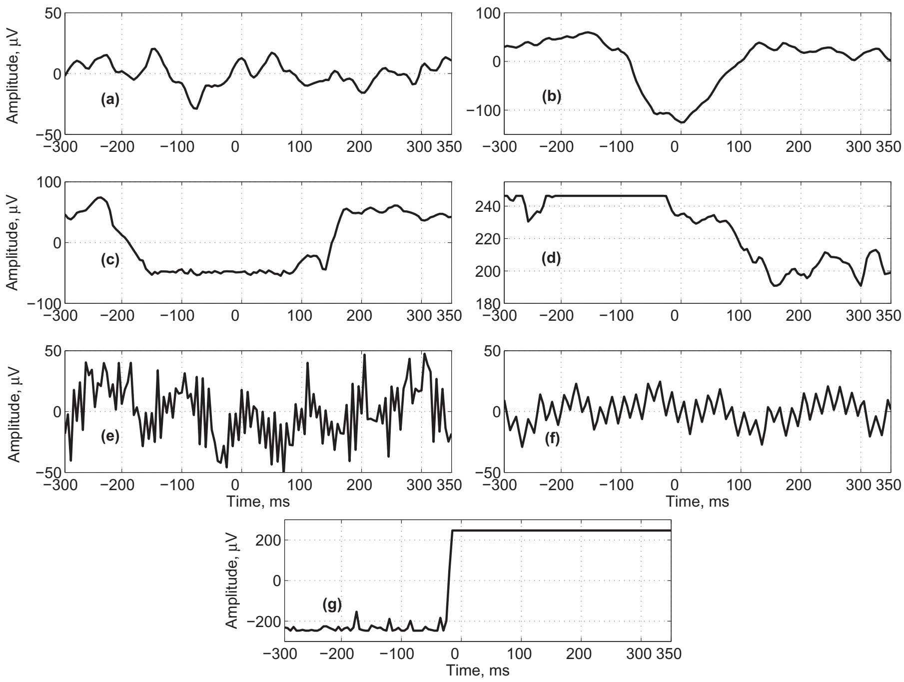

FIGURE 4 Examples of biological and other artifacts: (a) normal EEG signal; EEG signals corrupted with (b) eye blinking artifact, (c) eye movement artifact, (d) amplifier artifact, (e) motion artifact, (f) power line noise 50 Hz, and (g) electrode artifact. All signals are centered, except (d) and (g), which are shown with their original means.

Typically, ERPs have amplitudes much lower than those of the spontaneous EEG activity (from less than a microvolt to several microvolts against the tens of microvolts) [Luc05]. Moreover, the summed magnitude of ongoing EEG activity and artifacts may exceed that of the ERPs by hundreds of microvolts.

2.2 Preprocessing of ERP data

In the context of psychophysiology research, one is often interested in reliable estimation and the following analysis of ERP constituents of EEG data. Therefore, all other signals emitted by either brain, body, or external sources are treated as noise. In this thesis, we hold the same view point.

As has been mentioned above, the voltage levels of ERPs are significantly lower than those of spontaneous EEG and artifacts and thus the ERP characteristics are usually indeterminable reliably from a single measurement. The conventional procedure to increase the SNR of ERP estimate consists of collecting many single measurements, also called trials [4] obtained while repeating the stimuli and the following averaging of the EEG-segments stored during such trials for the study (see Section 2.5). Since all other activities except ERP associated with the stimulus eliciting the recorded changes of EEG should obey random temporal structure across repeated measurements, they are decreased by averaging. The reliability of the ERP estimate obtained after averaging depends heavily on the number of trials used for averaging.

In psychophysiology studies, the averaging is usually followed by a baseline correction, during which the mean of a pre-stimulus interval in the averaged response is subtracted from the estimate of the ERP waveform. Since in the present work we are not going to make any psychophysiological interpretations of the obtained results, the baseline correction is not used in the following discussion and computations. However, we always correct estimates of ERP waveform to zero mean, i.e., center them.

Moreover, some of the ERP components can only be seen when the averaged response to standard stimuli is subtracted from the averaged response to deviant stimuli. This procedure is called the difference wave approach and is used when the component of interest is temporally overlapped with some other components. For example, MMN deflection appears in the response to the deviant and is overlapped with the N1 component. Therefore, when MMN is studied, researchers typically proceed with the difference wave approach to cancel the N1 deflection after averaging (see, e.g., Figure 5(c,d)).

Some experimental settings may reduce the overlap between the MMN and N1 or reduce the N1 itself. For example, in contrast to N1, which is the response to the stimulus onset, MMN arises in response to detection of deviance. Therefore, MMN can be obtained beyond the latency window of the basic N1 peak by using duration deviants (by varying interstimulus interval (ISI)) [NPR89].

The problem of MMN and N1 overlapping can also be partially solved by using so-called uninterrupted sound paradigms, in which there are no silent periods between stimuli, with short (e.g., 100 ms) ISI. In this case, it is possible to significantly reduce the N1 amplitude that is useful for MMN analysis. One of the data sets used in this thesis and proposed by Pihko and colleagues in [PLL95] belongs to this type of paradigm (see Figure 5 and the description of the paradigm

[4] In the literature, synonyms *sweep*, *epoch*, *segment*, *trace*, and *episode* may occur.

in the next section).

2.3 ERP experiments and data

In our computations, we used two sets of EEG data that were collected in the scope of ERP-related psychophysiology studies. These data were introduced and studied in [PLL95] (see Figures 5 and 6) and [HLGL07, HLGL08] (see Figures 7 and 8).

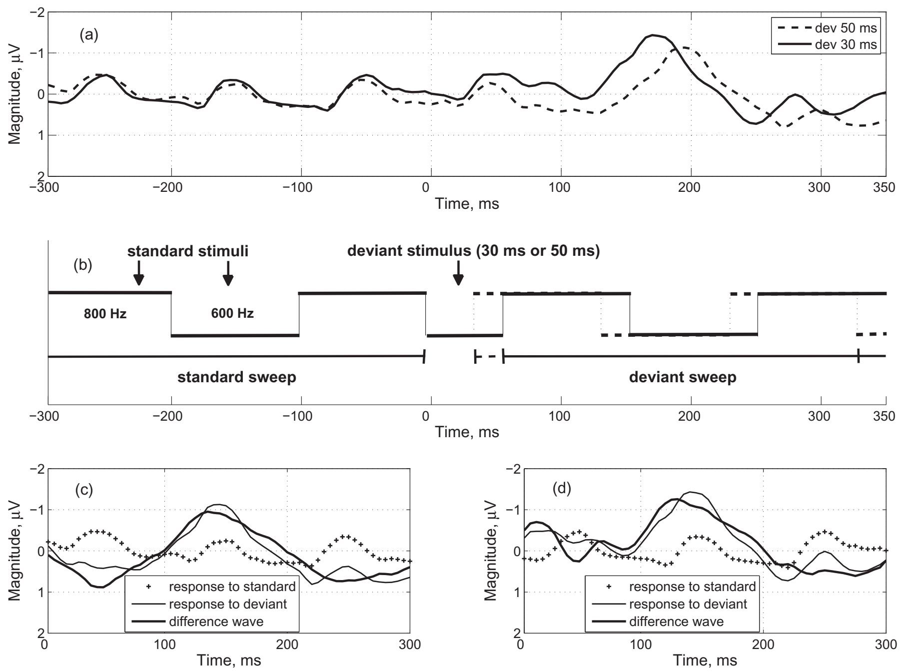

FIGURE 5 Uninterrupted sound paradigm for MMN elicitation proposed by Pihko and colleagues in [PLL95]: (a) grand average ERP waveforms in Cz channel for two deviants 30 ms and 50 ms, (b) schematic illustration of the experimental paradigm, (c) difference wave approach for 50 ms deviant, (d) difference wave approach for 30 ms deviant. In (c) and (d) time for the deviant responses starts at deviant tone offset.

The first data set was primarily collected for the purpose of analysis of mismatch negativity. The paradigm by Pihko and colleagues [PLL95] is based on a sequence of standard stimuli consisting of uninterrupted alternated sounds of 600 Hz and 800 Hz, each lasting 100 ms (see Figure 5(b)). Two types of deviant stimuli are randomly presented at a probability of 0.15 (each 0.075) in this sequence by shortening the 600 Hz tone to a duration of 30 ms or 50 ms. The measured trials contain 300 ms of recordings before the start of the deviant tone and 350 ms after the start of the deviant tone. Measurements were collected with the sampling rate 200 Hz, thus giving 130 time points for each trial. There were 102 participants involved in the data collection experiment. Overall the 700 trials data were collected for each of the 102 participants (350 trials for each of the two deviants). Measurements were recorded using a nine-electrode scheme, i.e., C3, C4, Cz, F3, F4, Fz, Pz, M1, M2 electrode locations were used. The nose electrode was used as reference.

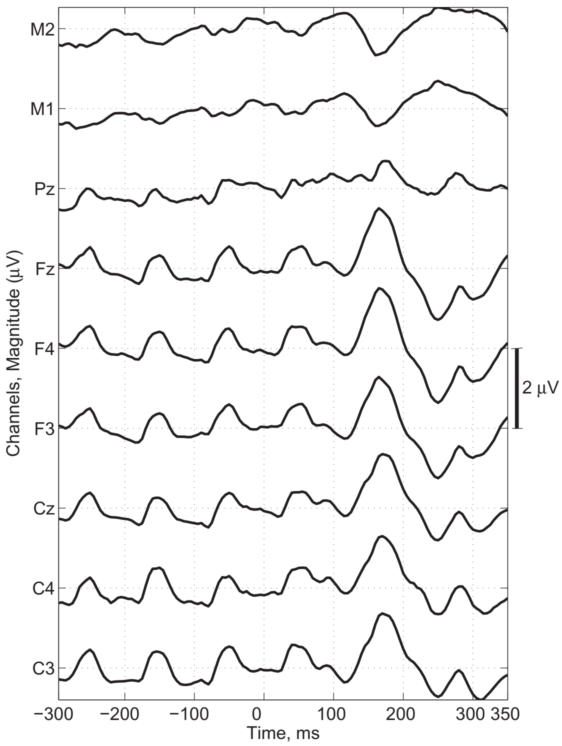

FIGURE 6 The grand averages of the ERP waveforms for nine EEG channels in the paradigm proposed by Pihko and colleagues in [PLL95]. Negativity is up.

In Figure 5(a) the grand averages (GA; the average across all participant averages) of 102 participants for 30 ms and 50 ms deviants in Cz channel are shown. GA reveals the average shape of the time-locked brain responses in a channel for all participants involved in the experiment. In contrast to the participant average, GA is almost noise-free if the number of participants is sufficiently large. However, the participant-specific features such as amplitudes and latencies of the components of ERP waveform are lost in GA, which is only an average approximation of the ERP waveform across all participants. The large peak around 175-200 ms is the sum of MMN deflection and the response to standard stimulus [KGJ$^+$07].

The series of repeated responses to standard stimuli is a so-called steady-state response (SSR) [Nää92]. In Figure 6 the typical spatio-temporal distribution of ERP activity for Pihko's paradigm is shown by using the GA waveforms for nine EEG channels. One can see that the polarities of MMN and SSR are positive in M1 and M2 electrode locations that is opposite to the negative polarities observed in the other channels. Moreover, even the grand average shows poor quality of the ERP waveform estimate in the Pz channel that has probably the lowest initial single trial SNR among channels.

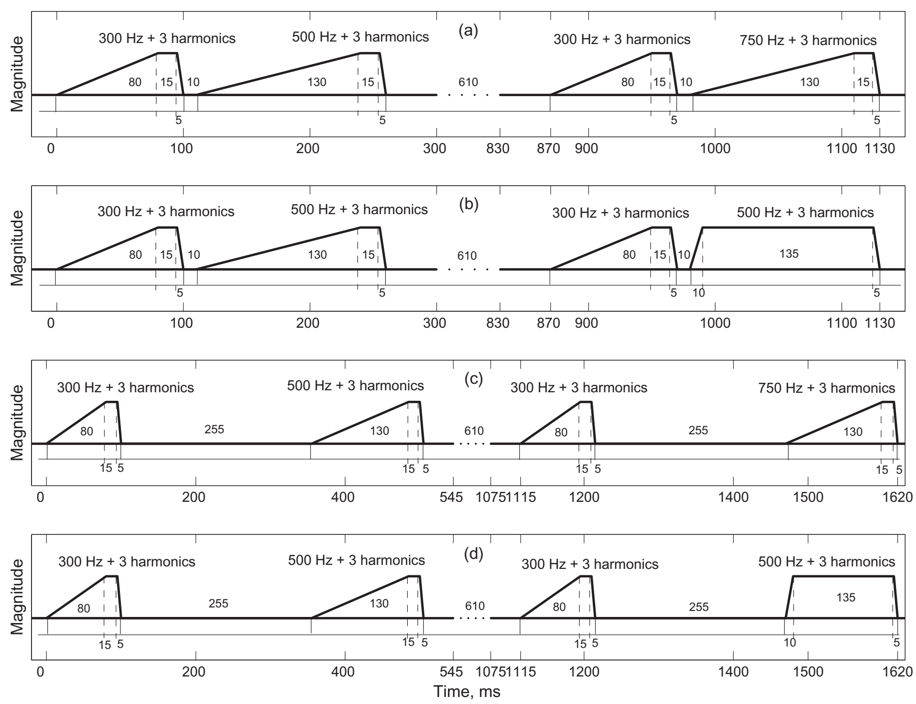

FIGURE 7 Oddball paradigm for ERP elicitation from [HLGL07, HLGL08]: (a) paradigm with short 10 ms ISI and pitch deviant, (b) paradigm with short 10 ms ISI and rise time deviant, (c) paradigm with long 255 ms ISI and pitch deviant, (d) paradigm with long 255 ms ISI and rise time deviant. In all subfigures, the left pair of tones represents the standard stimulus and the right pair of tones is the deviant stimulus.

The second data collection experiment was intended to elicit the following ERP components: MMN, P3a, and late discriminative negativity (LDN) obtained from difference waves, and P1 and N250 that can be seen in the standard waveform [HLGL07, HLGL08]. The two oddball paradigms were used in the framework of this study, both remarkably similar and differing only by the ISI between the pairs of tones (see Figure 7). In both paradigms, the standard and deviant stimuli consist of the pairs of tones. In the first paradigm, the interval between

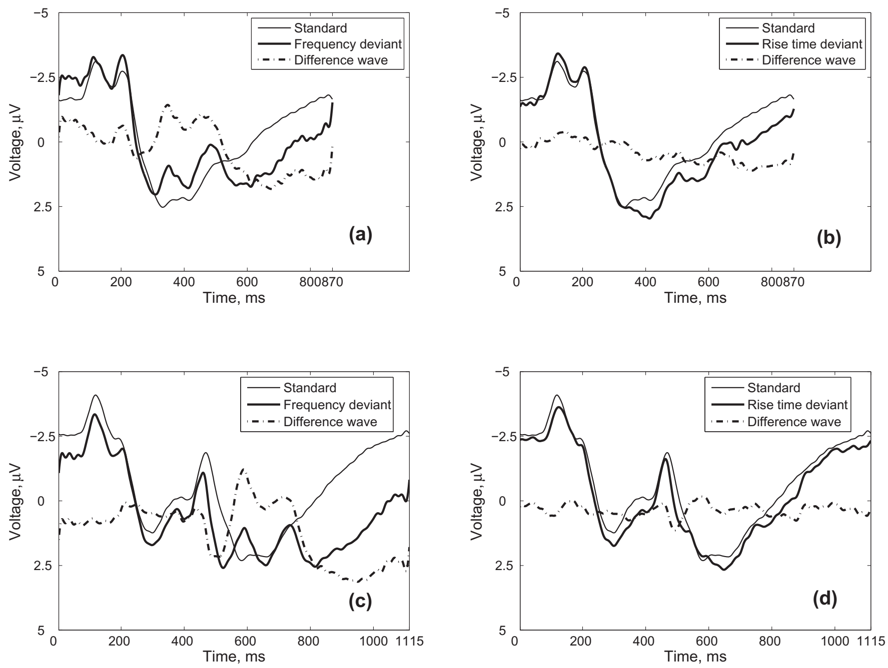

FIGURE 8 Grand averages of ERP waveforms and difference waves in Fz channel (average referencing) for the paradigm from [HLGL07, HLGL08]: (a) short ISI, frequency deviant, (b) short ISI, rise time deviant, (c) long ISI frequency deviant, (d) long ISI rise time deviant.

tones in every pair was set to 10 ms, and in the second paradigm, a 255 ms interval separates the two tones. The pairs of tones are separated by an interval of 610 ms. The first tone in all pairs was identical: with 300 Hz fundamental frequency and composed of four sinusoids 300, 600, 900, 1200 Hz, and was 100 ms in duration (linear 80 ms rise time, 5 ms fall time). The second tone in the standard pair was 150 ms long (linear 130 ms rise time and 5 ms fall time) with 500 Hz fundamental frequency and composed of four sinusoids 500, 1000, 1500, and 2000 Hz. The number of standard sweeps collected for each participant was 1010. In each of the two paradigms, two types of deviant stimuli were used: rise time change (second tone 150 ms long with linear 10 ms rise time and 5 ms fall time) and pitch change (the second tone has a fundamental frequency of 750 Hz with three harmonics). For every deviant stimuli, 125 trials were collected. The data were sampled with frequency 500 Hz. The length of the sweep was 870 ms or 435 samples for the paradigm with a short 10 ms ISI, and 1115 ms or 557 samples for the long 255 ms ISI paradigm. The recordings were collected using 128-channel sensor net. The locations of electrodes were set approximately according to 10/5 electrode placement system. The vertex (Cz) was used as the reference electrode during recording. In Figure 8 the ERP waveforms and the corresponding difference waves are demonstrated for each of the four combinations of standard and deviant stimuli.

2.4 Problem formulation from psychophysiology viewpoint and the road map

The ultimate goal of our research is to shorten the time for data collection experiments that are carried out within the electroencephalography studies intended to elicit and analyze ERP appropriately for observing brain and mind activity. Typical experiments are usually relatively long because SNR is very low in a single measurement of a brain activity following the stimulus event and therefore, many trials synchronized to the same stimulus need to be collected in order to increase SNR by the following trial averaging procedure done channel-wise (see Sect. 2.5). The shortening of the treatment time is important in order to keep the participant within an approximately stationary internal physiological state, - an additional aspect that influences the results of the experiment. If the experiment is not too long, the stationarity assumption of the background physiological conditions is closer to truth and they can be considered as not overly dependent on the time variable. Furthermore, some groups of participants, such as infants and clinical patients, may not be able to tolerate long-lasting experiments.

Another goal of the research is to increase the reliability of the obtained estimates of ERP parameters for a fixed number of trials. In fact, increasing SNR or the reliability of the ERP estimate is, actually, equivalent to shortening the time of the experiment. This is because, when SNR is increased, the number of trials necessary for reliable ERP identification can be reduced. One way to increase the reliability of the estimates of ERP characteristics is to apply denoising/filtering methods, which are able to extract the ERP from EEG data more efficiently than does the conventional averaging technique. This will allow for extraction of the desired accuracy of the estimates of ERP features using fewer trials and hence, shorten the time of the experiment.

Another set of recommendations for improving the determinability of ERP relates to the equipment and experimental design as pre-processing mechanisms. In addition to the qualitative equipment, which can predetermine a higher level of SNR in the raw EEG data, the carefully-designed experimental paradigms may provide data with desired and known patterns and properties which can be utilized as certain indicators for the methods to distinguish the sources of interest from the others. For example, the repeatable responses to standard stimuli, which the ERP waveform possesses in a paradigm from [PLL95], as depicted in Figure 5(a), reveal a good ERP-affiliated criterion in the spatio-temporal domain.

In the present work, we focus our attention on the advanced denoising methods that are able to reliably separate ERP from noise. In particular, such methods should effectively operate across different data representation domains, where signal and noise are separable (e.g., temporal, frequency, and spatial). In addition, these methods preferably should use as many dimensions of data variability (channels, time samples, trials, deviants) and inter-dimensional relations as possible in order to capture the larger amount of information on separability of ERP and noise and thus, improve the results of averaging. According to the exploited data dimensions, the separation methods can be classified as one-channel or multi-channel. The examples of one-channel techniques are traditional averaging and weighted averaging (see, e.g., [**PIII**]), i.e., those that do not use spatial information. The denoising algorithms developed in the scope of our research (e.g., [**PI**,**PIV**,**PV**,**PVI**]) mainly belong to the class of multi-channel techniques, i.e., those that utilize the inter-channel discriminating information.

2.5 Single channel data model

In this subsection, (1) the basic data notations are introduced, (2) the one-channel model of EEG recording containing the ERP signal is considered, (3) the traditional averaging approach to increase the SNR of an ERP estimate is described, and (4) the underlying assumptions and constraints imposed on data are discussed.

The overall data obtained during the EEG/ERP experiment is usually five-dimensional and, thus, can be indexed as $x_{ik}^{pq}(t)$, where $i = 1, \ldots, N$, $t = 1, \ldots, T$, $k = 1, \ldots, K$, $p = 1, \ldots, P$, and $q = 1, \ldots, Q$. Here N denotes the number of measured trials, T is the number of time points per trial, K stands for the number of measured channels, P denotes the number of the participants, and, finally, Q is the number of different types of stimuli including standard and deviant. Basically, for different stimuli q different numbers of trials N can be collected during the experiment. However, we do not indicate this relation explicitly in order to facilitate the notations.

Since the discussed methods will operate subject-wise and stimulus-wise and, thus, do not use the subject and the stimulus dimensions, in the following discussion we skip the p and q indices. Moreover, we omit the channel index k, when between-channel relationships are not concerned, and only within-channel data are considered. Likewise, the use of the trial index i is avoided, if we are interested in phenomena related to spatial dimension and not in relations between trials. A specific case should either be clear from the context or, otherwise, mentioned explicitly. If, however, both trial and channel dimensions are important we prefer to use a vector notation as $\mathbf{x}_i(t) = [x_{i1}(t)\ x_{i2}(t)\ \ldots\ x_{iK}(t)]^{\mathrm{T}}$. In general, we use the vectorial form to hide the channel index whenever it is appropriate.

One commonly accepted model, which is also assumed here, assumes that each recorded trial $x_i(t)$ can be represented as the sum of a time-locked brain signal $\varsigma(t)$ with variance σ_ς^2 and a noise $\eta_i(t)$ having variance σ_η^2:

$$x_i(t) = \varsigma(t) + \eta_i(t)\,, \tag{1}$$

where $\varsigma(t)$ is assumed to be trial invariant. The signal term $\varsigma(t)$ stands for a mixture of the time-locked brain responses corresponding to a channel. Correspondingly, the noise realization term $\eta_i(t)$ represents a mixture of the noise sources, such as spontaneous EEG and artifacts, at the same channel. In the case of the linear mixing model the mixture is a linear combination of the sources. The conventional averaging operation is performed for each channel separately and is described by the formula:

$$\widehat{\varsigma}(t) = \overline{x}_N(t) = \frac{1}{N} \cdot \sum_{i=1}^{N} x_i(t) = \varsigma(t) + \overline{\eta}_N(t), \tag{2}$$

where $\varsigma(t)$ is the time-locked ERP constituent (signal of interest) and $\overline{\eta}_N(t)$ is the noise constituent in the average. Term $\widehat{\varsigma}(t)$ denotes the estimate of the ERP signal. Let us define SNR in the average as the variance of ERP constituent divided by the variance of the noise constituent, that is

$$SNR_N = \frac{var\{\varsigma(t)\}}{var\{\overline{\eta}_N(t)\}} = \frac{\sigma_{\varsigma}^2}{\sigma_{\overline{\eta}}^2}, \tag{3}$$

where $\sigma_{\overline{\eta}}^2$ denotes the variance of the noise in the average. The resulting average in (2) is assumed to have higher SNR than a single trial. For this to be true, several assumptions must be met. First of all, it is required that the time-locked signal $\varsigma(t)$ is deterministic, i.e. that it does not change over trials. Second, noise $\eta_i(t)$ is supposed to be a zero mean ergodic process, and, as a consequence, we have

$$E[\eta_i(t)] = 0, \forall t = 1, \dots, T, \text{ and} \tag{4}$$

$$E[\eta_i(t)] = 0, \forall i = 1, \dots, N, \tag{5}$$

where $E[.]$ denotes the expectation operator. In equation (4) the averaging is carried out over realizations (trials) and equation (5) assumes averaging over time. Third, signal $\varsigma(t)$ and noise $\eta_i(t)$ must be uncorrelated:

$$E[\varsigma(t)\eta_i(t)] = E[\varsigma(t)]\, E[\eta_i(t)] = 0. \tag{6}$$

The expectation of the sum of two random variables equals to the sum of their expectations and, thus, the expectation of the data either across time or realizations should equal the expectation of the signal, provided that the expectation of noise is zero. Since the averaging over trials provides estimates of the true means of the data in every time point, it converges to signal values at these points, when the number of trials goes to infinity:

$$E\{\widehat{\varsigma}(t)\} = \varsigma(t) + \frac{1}{N}\sum_{i=1}^{N} E\{\eta_i(t)\} = \varsigma(t). \tag{7}$$

Therefore, if the above conditions are satisfied, the averaging operation reduces noise and estimates the signal with the level of accuracy depending on the number of averaged trials. Namely, the standard error (SE) accompanying the estimate (2) reads as

$$SE = \sqrt{E\{(\widehat{\varsigma}(t) - \varsigma(t))^2\}} = \frac{\sigma_\eta}{\sqrt{N}}. \tag{8}$$

The SE as given by this formula represents the amplitude of residual noise in ERP estimate that is proportional to the noise amplitude in the raw EEG data. SE diminishes proportionally to the square root of the number of trials N included in the averaged signal. The noise power $SE^2 = \sigma_\eta^2$ diminishes proportionally to N. Therefore, SNR as in (3) increases after averaging by N.

A realistic perspective, however, suggests that the above-mentioned assumptions are violated to some extent in practical situations. As a consequence, methods that are based on these assumptions and, in particular, plain averaging, may lose their accuracy.

2.6 Problem setting from the source separation perspective and model of the multichannel data

The general form of the non-stationary model of the physical process/system can be represented as

$$\mathbf{x}(t) = \mathbf{f}(\mathbf{s}(t), \boldsymbol{\theta}(t)), \tag{9}$$

where function $\mathbf{f}$ with parameters $\boldsymbol{\theta}(t) = [\theta_1(t)\ \theta_2(t)\ \ldots\ \theta_{K_\theta}(t)]^{\mathrm{T}}$ transforms the input sources $\mathbf{s}(t) = [s_1(t)\ s_2(t)\ \ldots\ s_{K_s}(t)]^{\mathrm{T}}$ at time t into the output vector $\mathbf{x}(t) = [x_1(t)\ x_2(t)\ \ldots\ x_{K_x}(t)]^{\mathrm{T}}$. Here K_x denotes the number of channels, K_s is the number of sources, and K_θ stands for the number of free parameters.

Regarding our specific domain, $\mathbf{s}(t)$ are realizations of the electromagnetic generators/sources located mainly in the human brain but also in other parts of the body and in the external environment. The components of $\mathbf{x}(t)$ represent the measurements of the mixed emissions of the sources captured at certain locations (electrodes) on the scalp. Finally, $\mathbf{f}$ denotes a function, which mixes the sources in places where electrodes are positioned.

For practical usage of the model (9), the form of the function $\mathbf{f}$ and free parameters $\boldsymbol{\theta}(t)$ should be defined explicitly together with the underlying constraints and assumptions imposed on input data. These specifications have an effect of adjusting the model to the problem being investigated.

We narrow the model (9) therefore for our problem area by specifying the form of the mixing function. Namely, we assume that the linear instantaneous mixing model with additive noise, which has acquired a considerable reputation in the research practice for being sufficiently accurate and reliable and yet simple, holds. [MBJS96, VO99, MP95]:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \boldsymbol{\eta}(t), \tag{10}$$

where $\mathbf{A}$ denotes the mixing matrix. Here $\boldsymbol{\eta}(t) = [\eta_1(t)\ \eta_2(t)\ \ldots\ \eta_{K_x}(t)]^T$ is the channel-wise additive noise usually assumed as having the covariance matrix of the following form $\mathbf{C}_{\boldsymbol{\eta}(t+\tau)\boldsymbol{\eta}(t)} = \sigma_\eta^2 \delta_\tau \mathbf{I}$ with the variance of the noise σ_η^2 and δ_τ denoting Kronecker delta, i.e. additive noise is modeled as zero mean, stationary, and spatio-temporally white [HKO01]. Note that sources $\mathbf{s}(t)$ include ERP and the other noise sources that follow the linear mixing model, and the term $\boldsymbol{\eta}(t)$ describes the noise that does not fit to the spatial model.

The generic non-stationary Blind Source Separation (BSS) problem defined on the model (9) consists in recovering parameters $\boldsymbol{\theta}(t)$ and sources $\mathbf{s}(t)$ based on the measurement vector $\mathbf{x}(t)$. In psychophysiology research it is common to consider stationary BSS problem that is defined on the model (10). In this case the mixing matrix $\mathbf{A}$ plays the role of unknown parameters. For simplicity let us assume now that the number of sources equals the number of simultaneous measurements, and realizations of both sources and their mixtures without additive noise form linearly independent subsets. These assumptions guarantee that the time course of any source can be obtained by the linear projection of the mixtures to a unique spatial direction associated with the source. Then the latter problem can be considered as solved successfully, if the inverse matrix $\mathbf{A}^{-1}$ is found.

Independent Component Analysis (ICA) is intended to solve the relaxed BSS problem. Provided that the sources are statistically independent, the purpose of ICA is to find the components, which can be scaled and negated versions of the original sources. Moreover, the order of the components or rows in the separation matrix $\mathbf{B}$ is an additional ambiguity. Thus, the ICA problem is considered to be successfully solved, if the estimated separation matrix in equation $\mathbf{y}(t) = \mathbf{B}\mathbf{x}(t)$ can be decomposed to $\mathbf{B} = \mathbf{QPA}^{-1}$, where $\mathbf{y}(t)$ denotes the vector of estimated components, $\mathbf{Q}$ is the diagonal sign/scaling matrix, and $\mathbf{P}$ stands for the permutation matrix. Note that each column of the inverse matrix $\mathbf{B}^{-1}$ reveals energy contributions of the corresponding component to the channels. Therefore the columns of the inverse of $\mathbf{B}$ are called spatial distributions, spatial maps, or topographies of the respective components.

To be able to solve the traditional ICA problem the linear instantaneous mixing model (10) is equipped with two additional assumptions:

Source independence: Sources are assumed to be statistically independent, i.e. their joint probability density function (pdf) is of the factorial form:

$$p(\mathbf{s}) = \prod_k p(s_k). \tag{11}$$

Non-Gaussianity of the sources: At most one source may have Gaussian distribution. The reason for this is that in the subspace spanned by Gaussian sources the mixtures of sources cannot be distinguished from the sources themselves, if only independence is assumed.

Usually there are additional constraints imposed on probability distribution functions of the sources (see Sect. 3.2). Since the target criteria used by ICA methods for estimating the separation matrix $\mathbf{B}$ are designed to obtain as much statistically independent components as possible, they are called independent components.

In the context of psychophysiology, our goal is ERP denoising in multichannel EEG data. As a fulcrum we intend to estimate such $\mathbf{B}$ that gives two sets of components spanning ERP and noise subspaces. The denoising is carried out by projecting back to the electrode field only the components related to ERP subspace and neglecting noise components. Therefore, compared to the relaxations made in ICA, we relax the BSS task even further by requiring only subspace separation. In this case the solution gets additional degrees of freedom. Namely, the angles between the components inside one subspace can be arbitrary. Therefore, the subspace separation problem is assumed to have been solved successfully, if the resulting separation matrix can be expressed as $\mathbf{B} = \mathbf{M}\mathbf{A}^{-1}$, where matrix $\mathbf{M}$ allows for mixing of the sources that belong to the same subspace.

In analogy to BSS problem, the equivalent term related to subspace separation problem can be termed Blind Subspace Separation. It is to be noted that neither source nor subspace separation can be totally blind. Rather sources and subspaces are distinguished based on (1) some subspace- or source-specific features or (2) mutual relationships, e.g. statistical independence. Similarly as ICA is related to BSS, the analogous Independent Subspace Analysis (ISA) concept is introduced for Blind Subspace Separation problem [Car98, HH00, HHI01]. ISA is a natural extension of ICA that additionally implies possible dependencies between the components inside one group (subspace), whereas the components from different groups are mutually independent. Other basic assumptions and indeterminacies in ISA are similar to those implied for ICA. Some additional assumptions may also be imposed by a specific methodology [The06]. The ISA concept has several closely related terms and interpretations appearing in literature that differ in methodology and constraints (see [The06] for a short review of ISA models). For example, the term Multidimensional ICA (MICA) was first introduced by Cardoso in [Car98]. Depending on the methodology, an additional restriction may be imposed on the group size or dimension of the subspaces as in [The04, HH00]. This is called k-ISA, where k is the dimension of the subspaces. The term Independent Feature Subspace Analysis was used in [HH00] to denote a technique that is a special case of MICA or k-ISA. In this method, the authors model the dependencies explicitly within a k-tuple.

From one side, the subspace separation problem in the psychophysiology context that is pursued here can be viewed as more specific compared to the standard ISA problem, since it includes additional problem-specific assumptions, which allow for discrimination between the signal and noise subspaces. On the other hand, some of the ISA assumptions, e.g., those related to probability distributions of the sources and independence (a weaker uncorrelatedness assumption is enough), are in general superfluous for our problem and can be excluded from consideration.

Formally, our goal is to find a transformation matrix $\mathbf{B}$ of size $K \times K$ on data $\mathbf{x}_i(t)$ so that the components of the vector $\mathbf{y}_i(t)$ obtained as $\mathbf{y}_i(t) = \mathbf{B}\mathbf{x}_i(t)$ can be unambiguously divided into the two groups related to ERP and noise subspaces. In practice, however, we expect that the subspaces are overlapped and can not be completely separated by a linear transformation, i.e. some components can only be mixtures of signal and noise sources. In this case we aspire to extract such components that for any dimensions of the signal (maximal SNR) and noise (minimal SNR) subspaces defined after the transformation there exists a division of the components to subspaces that provides optimal (in the sense of SNR) separation of the signal and noise. A specific proportion between the dimensions of the subspaces defines in this case the amount of ERP energy that is lost (passed to the noise subspace) after the dimension reduction. Therefore the smaller the size of the signal subspace is, the larger SNR and ERP energy loss are, and vice versa.

It is to be remembered that the overall data are of the form $x_{ik}^{pq}(t)$ (see Section 2.5). However, the most important dimensions that should be emphasized are channels k, time samples t, and trials i. Correspondingly the realizations of the sources contributing measurements $\mathbf{x}_i(t)$ are denoted as vector $\mathbf{s}_i(t) = [s_{i1}(t)\ s_{i2}(t)\ \dots s_{iK_s}(t)]^{\mathrm{T}}$. Additionally, we assume that the mixing matrix $\mathbf{A}$ in (10) is stationary, i.e. the same mixing model holds for all trials:

$$\mathbf{x}_i(t) = \mathbf{A}\mathbf{s}_i(t) + \boldsymbol{\eta}_i(t),\ \forall i = 1, \dots, N. \tag{12}$$

Basically, this assumption means that all sources are spatially stable. It is unlikely that this can be completely true in practical applications. Nevertheless we make this approximation that seems to be rather close to real circumstances and is supported by the theory of functional brain segregation, specialization, and integration discussed in Section 2.1. Thus, for example, the averaging procedure initially defined for one channel in (2) naturally transforms to multichannel averaging as:

$$\overline{\mathbf{x}}(t) = \frac{1}{N} \cdot \sum_{i=1}^{N} \mathbf{x}_i(t) = \mathbf{A}\overline{\mathbf{s}}(t) + \overline{\boldsymbol{\eta}}(t), \tag{13}$$

where $\overline{\mathbf{s}}(t) = \frac{1}{N} \cdot \sum_{i=1}^{N} \mathbf{s}_i(t)$ and $\overline{\boldsymbol{\eta}}(t) = \frac{1}{N} \cdot \sum_{i=1}^{N} \boldsymbol{\eta}_i(t)$.

Moreover, throughout the thesis we assume that $\mathbf{x}_i(t)$ have zero mean, i.e. they are centered:

$$\mathbf{x}_i(t) \ \leftarrow\ \mathbf{x}_i(t) - \frac{1}{T} \sum_{t=1}^{T} \mathbf{x}_i(t),\ \forall i = 1, \dots, N. \tag{14}$$

3 REVIEW OF RELATED METHODS

In this chapter we review the literature related to the development of one-channel denoising methods (Section 3.1) and source separation methods (Section 3.2) as well as their application to EEG data, (Section 3.3) and approaches to the validation of quality of data and results of ERP denoising (Section 3.4). This chapter does not pretend to be a complete review of the related methods. Therefore some approaches which might be expected by the reader may be missing. The main purpose of the chapter is to give to the reader an idea of the variety and breadth of the research efforts undertaken towards the considered problem and immerse the reader into the topic before we turn to reviewing the results of the thesis. Since the focus of the thesis is concentrated on spatial denoising approaches, of which ICA is probably the most prominent example, significant attention is paid to ICA-related methodologies. This disproportion is motivated by the intention to place the discussion within the context to which the proposed methods are most closely related.

3.1 Single-channel methods for ERP denoising

The single-channel ERP denoising methods are based on separating signal and noise in, e.g., time, frequency, or time-frequency domains.

A well-known example of the method that performs denoising in the temporal domain is traditional averaging. Other methods operating in the time domain are mainly targeted on modifying the data so that they more closely fulfil the declared properties, e.g., determinism of the signal and stationarity of the noise. Let us consider some of the temporal-domain methods in more detail.

In [LKS05] trimmed estimators for robust averaging were reviewed. In [DS00] an averaging algorithm based on orthogonal projection of the measurement vector to signal subspace was proposed. Also a version of the running noise-to-signal ratio (NSR) estimate for a single trial was considered in the same article. Weighted averaging is another important approach for improving conventional averaging [HRWL84, DM92, Łęs02, ŁG04]. In weighted averaging, each trial is multiplied by a weight inversely proportional to the noise estimated in the considered trial. The weighting allows for faster noise reduction during the averaging following. For more detailed discussion on this topic the reader can also refer to article [**PIII**].

Another large branch of methodologies that improve the performance of traditional averaging takes into account the variability of ERP amplitude and latency among trials. Since the assumption of unconditional determinism of ERP waveform across trials is unlikely to be fulfilled in practice, and variable amplitudes and latencies take place, some temporal synchronization of the trials to improve the quality of ERP estimate proves useful before averaging. Thus in [PMKG87] the assumed data model allows variable latencies over trials:

$$x_i(t) = \varsigma(t + \tau_i) + \eta_i(t). \tag{15}$$

The authors propose to find latency variations τ_i by using the maximum-likelihood approach and by employing the so-called iterative Fisher-scoring after transforming this model to the frequency domain. For the first time this model was proposed in [Woo67], where the method for estimating the latency of every single-trial event-related activity using a matched filter was also introduced. The latencies were determined there as time delays τ_i providing maximal cross-correlation between the averaged ERP template and a single trial. After that individual trials were adjusted to account for the latency variability prior to averaging. The problem of variable latencies was also addressed in [MKGP88], where the same method as in [PMKG87] for obtaining estimates of single latency shifts was discussed, and statistical test on the presence of latency jitter was proposed.

In [JV99] a more general model of the data was considered, where variable amplitudes were also taken into account:

$$x_i(t) = a_i\varsigma(t + \tau_i) + \eta_i(t). \tag{16}$$

The authors used an extension of the maximum-likelihood approach from [PMKG87] to estimate the parameters of ERP waveform variable across trials.

In [MTG84] two new statistical tests were proposed for testing (1) the variability of amplitude and (2) slowly changing signal model. Thus the first considered variability model is expressed as follows

$$x_i(t) = a_i\varsigma(t) + \eta_i. \tag{17}$$

The second model assumes that $\varsigma_i(t)$ is markedly different from $\varsigma_j(t)$ for large $|i - j|$ and close to $\varsigma_j(t)$ for $j = i + 1$. In [MGT90] the three statistical tests from [MTG84] and [PMKG87, MKGP88] for analyzing signal variation were summarized. See also [MGPK87] and [MGK88] for information on amplitude and latency jitter.

In [GMTS96] a method based on nonlinear alignment of pairs of local EP estimates and successive averaging of the aligned local averages was proposed to compensate for variable latencies. The method does not make prior assumptions about the properties of the EP waveform, while benefiting from accounting for the nonlinear character of latency distortions.

Methods related to denoising in the frequency domain consist in applying some filters either to the averaged response or to a single trial. Thus, in [Wal69] a Wiener filter based on estimated signal and noise spectra was proposed. An extensive discussion on the theory and practice of the application of *a posteriori* Wiener filter to averaged evoked potentials is presented in [dWM78]. A time domain *a posteriori* optimal filter was proposed for application to averaged responses in [BF88]. In [FB91] another optimal *a posteriori* time domain filter for average evoked potentials was derived based on the proposed estimators of signal and noise autocorrelations. In the study reported in [KGJ+07] optimal digital filtering is compared with the difference wave approach with both applied for mismatch negativity estimation within the paradigm from [PLL95]. In addition, authors estimate optimal parameters, i.e. cut-off frequencies, of the digital filter for mismatch negativity denoising within the specified paradigm.

Denoising in the time-frequency domain is often associated with wavelet decomposition methods and Gabor transform, which offer time-frequency representation of the signal [Mal99]. The denoising is performed in the following three steps: (1) the signal under consideration is decomposed into a wavelet domain consisting of coefficients that reflect contributions of the wavelet basis functions localized differently in the time-frequency plane, (2) the noise-associated coefficients are zeroed or shrunk, and (3) the time-frequency areas affiliated to the signal are transformed back to the temporal domain. The wavelet-based denoising methods can offer advantages over pure temporal or frequency domain methods, because various wavelet-like basis functions specialized in certain data patterns allow for better capturing of specific signal activities localized in the time-frequency plane. Specialization and double-domain localization properties of wavelet basis functions together result in better separability of the signal and noise. There are numerous studies related to multi-resolution wavelet analysis of EPs in the time-frequency domain, see, e.g., [TXrYCH93, AMTI97, ELG+00].

Another direction of research in filtering is related to nonlinear filtering methods [Jaz70, HLN94, Lip99]. The use of nonlinear approaches may be more appropriate compared to linear ones, if one assumes that nonlinear relationships underlie real processes. Then the nonlinear techniques allow for more accurate modeling of phenomena and capturing of data structures. The application of nonlinear procedures leads, in turn, to increased reliability of the results. There are numerous research studies completed and publications in this field [Ahm98]. For example, in [CCC+96] some of the nonlinear algorithms for processing biological signals are reviewed, which can be applied particularly to EEG processing and ERP denoising.

One of the earliest and the most straightforward approaches for improving the accuracy of EP estimate is related to data selection strategies and called *artifact rejection*. According to this method, measurements that are corrupted with artifacts are excluded from the averaging. Essentially the artifact rejection technique is based on artifact detection criteria. It is to be noted that when a trial is rejected, the relevant information may also be lost. Moreover, trial rejection decreases sample size for averaging and, therefore, also reduces the accuracy of the EP estimate. Hence when making a decision about artifact rejection one should always evaluate what is more appropriate - rejection or preservation of the trial. One possible solution for this evaluation problem is given by the method of sorted averaging proposed in [MvS99]. More details can be found, e.g., in [Bar86, ARS+99, BCM+03, KGK+08] and articles therein.

Summary: The developed methods for ERP denoising, which exploit data from each channel separately, allow for significant improvement of the reliability of ERP estimates. Nevertheless it seems that methods of this type have almost reached saturation point in their effectiveness. To further increase the performance of ERP denoising techniques additional information, like that encrypted in between-channel relationships, should be employed.

3.2 Source separation methods

In this section, we discuss the approaches and criteria developed to solve the relaxed BSS problem defined on the model (10). Since this is an extremely large and rapidly growing area, we consider only the main milestones of the work done in this direction. Moreover we further concentrate our attention on the issues related to the formalization of the separation criteria, rather than on the effective implementations of the algorithms based on these criteria.

First, we introduce the classification of the source separation methods. Then in Section 3.2.1 we discuss both the Principal Component Analysis (PCA) technique and whitening procedure based on eigenvalue decomposition. Afterwards, in Sections 3.2.2 and 3.2.3 we present a detailed overview of the BSS methods according to the previously defined classification.

The source separation methods can be classified into four basic categories by the *a priori* information or criteria they use to find a solution [CA02]:

Mutual independence, Non-Gaussianity, ICA: within this category the separation criteria employ some measure of mutual independence or non-Gaussianity of signals. In this case the higher-order statistics (HOS) are extensively exploited in solving the BSS problem. Only one Gaussian source is allowed, under the assumption of statistical independence, to be able to successfully solve BSS problem.

Temporal structure: the criteria of this category exploit the fact that, if sources have a temporal structure, then they also have non-zero autocorrelations. In this case, the assumption of uncorrelatedness of the sources, that is weaker than the independence assumption, can be used. As a consequence, second-order statistics (SOS) are enough to estimate mixing matrix and sources.

Diversities of the signals: these include different characteristic *a priori* known properties of the signals in temporal, spatial, and frequency domains or in combinations of these domains. The sources are thus initially divided into several categories by their discriminative properties in the chosen domains. The separation method, that must be aware of this division, automatically classifies found components into defined groups. The methods developed during our research belong to this class of source separation methods, that utilize diversities between groups of signals (see Chapter 4).

Non-stationarity: the second-order non-stationarity is usually considered within this category, which means that the variance of signals changes in time. The source separation problem under the variance non-stationarity assumption is solved within the framework of the second-order statistics, so that decorrelation is able to separate sources (see Section 3.2.2).

3.2.1 Principal component analysis and whitening

Principal Component Analysis [Jol86, Oja92, HKO01] is an orthogonal linear component decomposition technique that has two main interpretations of its optimized cost function.

From one side, PCA estimates components $y_k(t) = \mathbf{w}_k^{\mathrm{T}}\mathbf{x}(t)$, where $k = 1, \ldots, K$, $\mathbf{w}_k = [w_{k1}\ w_{k2}\ \ldots\ w_{kK}]^{\mathrm{T}}$, and $\mathbf{x}(t) = [x_1(t)\ x_2(t)\ \ldots\ x_K(t)]^{\mathrm{T}}$, one by one by maximizing the variance of each component

$$\mathcal{J}^{PCA}(\mathbf{w}_k) = E\{y_k^2\} = E\{(\mathbf{w}_k^{\mathrm{T}}\mathbf{x})^2\} = \mathbf{w}_k^{\mathrm{T}} E\{\mathbf{x}\mathbf{x}^{\mathrm{T}}\}\mathbf{w}_k = \mathbf{w}_k^{\mathrm{T}}\mathbf{C}_{\mathbf{x}}\mathbf{w}_k \tag{18}$$

subject to unit norm and orthogonality (with respect to previously found components) constraints $\mathbf{w}_{k_1}^{\mathrm{T}}\mathbf{w}_{k_2} = \delta_{k_1k_2}, \forall k_1, k_2 = 1, \ldots, K$, where $\delta_{k_1k_2}$ stands for the Kronecker delta, and $\mathbf{C}_{\mathbf{x}}$ is the covariance matrix of the data $\mathbf{x}$.

On the other hand, PCA components can be found by minimizing the function of mean-square error (MSE) between the original data and the data described by m-dimensional subspace

$$\mathcal{J}_{MSE}^{PCA}(\mathbf{w}_1, \ldots, \mathbf{w}_m) = E\{\|\mathbf{x} - \sum_{k=1}^{m}(\mathbf{w}_k^{\mathrm{T}}\mathbf{x})\mathbf{w}_k\|^2\}, \tag{19}$$

subject to orthogonality and unit norm constraints, where $\|.\|$ is the Euclidean norm.

In either case the optimization problems are solved in terms of eigenvectors and eigenvalues of the covariance matrix $\mathbf{C}_{\mathbf{x}}$. Namely, to find the eigenvalues one needs to solve the characteristic equation $\det(\mathbf{C}_{\mathbf{x}} - \lambda\mathbf{I}) = 0$ in variable λ. Here, $\mathbf{I}$ is the identity matrix, and $\det(\mathbf{B})$ is the determinant of the matrix $\mathbf{B}$. The solutions λ of the characteristic equation are called eigenvalues. For $K \times K$ square matrix $\mathbf{C}_{\mathbf{x}}$, there are K eigenvalues, which may have less than K distinct values in general. The column K-element eigenvector $\mathbf{e}_k$ associated with eigenvalue λ_k is found by solving the equation $\mathbf{C}_{\mathbf{x}}\mathbf{e}_k = \lambda_k\mathbf{e}_k$. Therefore all pairs of respective eigenvalues and eigenvectors satisfy $\mathbf{C}_{\mathbf{x}} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^{\mathrm{T}}$, where $\mathbf{E}$ is a matrix of K unit length pairwise orthogonal eigenvectors placed at columns and $\mathbf{\Lambda}$ is the diagonal matrix of corresponding eigenvalues ([HKO01]). Note that this form of decomposition is suitable for real symmetric matrices, which is what the covariance matrix is. In the general case transpose is replaced with matrix inverse.

One can see that PCA transformation $\mathbf{E}^{\mathrm{T}}\mathbf{x}$ transforms the data $\mathbf{x}$ so that the covariance matrix of the transformed data is diagonalized to $\mathbf{\Lambda}$. This means that all principal components are pair-wise uncorrelated. The eigenvalues are nothing else than the variances of the data projected to corresponding eigenvectors. The principal components are conventionally ordered according to decreasing eigenvalues, i.e. by reduction of their energy (variance) importance.

Often PCA is used as a preprocessing step for dimension reduction of the data by skipping the components explaining smallest data variations and assuming that the most important features of the data are described by the few components with the largest eigenvalues.

In the context of source separation, eigendecomposition plays an important role, since it is a composite part of another important preprocessing step called whitening or sphering. The data are said to be whitened if their covariance matrix is the identity matrix $\mathbf{C}_{\mathbf{z}} = E\{\mathbf{z}\mathbf{z}^{\mathrm{T}}\} = \mathbf{I}$, i.e. the individual spatial components of the data vector are uncorrelated and have unit variances. Since the PCA transformation diagonalizes the covariance matrix, all that is left to make data whitened is to scale the PCA components to unit variance. Thus, for example, the following transformation of the initial (centered) data $\mathbf{x}(t)$ leads to whitened data $\mathbf{z}(t)$:

$$\mathbf{z}(t) = \mathbf{V}\mathbf{x}(t) = \mathbf{E}\mathbf{\Lambda}^{-1/2}\mathbf{E}^{\mathrm{T}}\mathbf{x}(t) = \mathbf{V}\mathbf{A}\mathbf{s}(t), \tag{20}$$

where $\mathbf{V} = \mathbf{E}\mathbf{\Lambda}^{-1/2}\mathbf{E}^{\mathrm{T}}$ denotes the whitening matrix, and $\mathbf{\Lambda}^{-1/2}$ is a componentwise inverse square root operation that implements division of the PCA components by the corresponding standard deviations to enforce unit variance. In the above whitening procedure, final multiplication by $\mathbf{E}$ can be skipped, since it does only the backward rotation of the scaled PCA components.

The utility of the whitening consists in making independent sources orthogonal that reduces the subsequent independent component decomposition search to finding an orthogonal basis [HKO01]. Note that any orthogonal transformation applied to the whitened data results in whitened data again. Moreover, since whitening makes independent sources orthogonal, it allows orthogonalization to be used as a solution for preventing the convergence of the sequentially estimated components to the same point. Many ICA algorithms use whitening as a preprocessing step, e.g. FastICA [HO97, Hyv99a].

Basically the capabilities of PCA are limited with respect to the solution of the independent component separation problem, since it searches for spatially orthogonal and temporally uncorrelated components and focuses on the variance maximization / MSE minimization criteria that are unlikely to underlie real sources. In addition, even uncorrelated sources cannot be found by means of pure PCA - all that whitening does is just making uncorrelated sources orthogonal in a new coordinate system. To estimate the sources themselves the following orthogonal transformation (rotation) must accompany whitening. This further rotation is determined by using additional priors on sources such as temporal structure, non-stationarity, non-Gaussianity etc. Elaboration of this rotation is addressed by numerous algorithms, like AMUSE, SOBI, TDSEP, JADE, and FastICA that are described more thoroughly in the following sections.

3.2.2 Exploiting the temporal structure, non-stationarity, and diversity

The source separation methods based on the second-order statistical properties assume that the sources either exhibit temporal structure or they are non-stationary.

More precisely, if the temporal structure is implied, it is assumed that the delayed cross-covariance matrix of the initial sources reads as $\mathbf{C}_{\mathbf{s}(t+\tau)\mathbf{s}(t)} = E\{\mathbf{s}(t+\tau)\mathbf{s}(t)^{\mathrm{T}}\} = diag(\psi_1(\tau)\ \psi_2(\tau)\ \dots\ \psi_K(\tau))$ with auto-covariance of $s_k(t)$ denoted as $\psi_k(\tau) = E\{s_k(t+\tau)s_k(t)\}$, i.e. sources $s_k(t)$ are mutually uncorrelated, but non-zero inner-source auto-correlations exist. The assumptions underlying additive noise are taken as in Section 2.6. If pre-whitening is done as $\mathbf{z}(t) = \mathbf{V}\mathbf{x}(t)$ (see Section 3.2.1), sources become orthogonal in the whitened data space with the orthogonal separation matrix $\mathbf{B} = \mathbf{A}^{-1}\mathbf{V}^{-1}$. Since the delayed cross-covariance matrix of the sources is diagonal, to estimate sources one needs to find an orthogonal transformation on whitened data that diagonalizes $\mathbf{C}_{\mathbf{z}(t+\tau)\mathbf{z}(t)} = E\{\mathbf{z}(t+\tau)\mathbf{z}(t)^{\mathrm{T}}\}$. The diagonalization is carried out by eigendecomposition. This is the basic principle of AMUSE algorithm [TSHL90].

In SOBI (Second Order Blind Identification) [BAMCM97] and TDSEP (Temporal Decorrelation source SEParation) [ZM98] algorithms the same idea is exploited, but instead of one estimate of the time-delayed cross-covariance matrix for predefined τ, several cross-covariances for different τ are computed and jointly diagonalized approximately. This approach is considered to produce more robust results, and is, basically, a generalization of the AMUSE procedure. The overall cost function that is minimized by both methods has, therefore, the following form:

$$\mathcal{J}_{\mathrm{TDSEP}}(\mathbf{W}) = \sum_{i \neq j} \left(\frac{\sum_{t=1}^{T} y_i(t) y_j(t)}{T-1}\right)^2 + \sum_{k=1}^{N} \sum_{i \neq j} \left(\frac{\sum_{t=1}^{T} y_i(t) y_j(t+\tau_k)}{T-1}\right)^2. \tag{21}$$

Here, the first term minimizes the sum of off-diagonal elements of the correlation matrix with zero time lag; likewise, the second term does the same for the set of time-delayed second-order correlation matrices. The separation matrix $\mathbf{W}$ is included in (21) by taking $y_i(t) = \mathbf{w}_i^{\mathrm{T}}\mathbf{x}(t)$. Both methods take advantage of the whitening as the first step that explicitly minimizes the first term. Afterwards the approximate simultaneous diagonalization of several time-delayed second-order correlation matrices is carried out by using several Jacobi rotations [CS96].

The second-order statistics can also be used as discriminative indicators for source separation, when sources undergo variable variances. Thus, in [Man99] it was shown that in order to separate the instantaneous mixture of independent non-stationary sources the second-order statistics are sufficient. Namely, the diagonalization of the auto-correlation matrix separates the non-stationary signals.

SOS have limited capacity however for separating sources, and in many cases the distinguishability between different sources can only be defined in terms of higher-order statistical independence formalized by means of HOS and information theoretic concepts such as entropy and mutual information.

The concept of differential BSS assumes that the source signals obey two different features or values of some property, by which they can be distinguished. For example, the authors in [DBA04, TDH07] introduced the differential kurtosis measure that was intended for distinguishing the non-stationary sources from stationary ones. This approach also offers a partial solution to underdetermined problem, i.e. when the number of channels is less than the number of sources. The method is constructed so as to extract maximally clean non-stationary sources, while leaving the stationary sources out of interest, i.e., considering them as an inseparable subspace. This approach reveals an example of using the diversities between the two sorts of sources. The methods developed in the scope of this thesis belong to the same category: namely the subspace of the deterministic signal sources is contrasted with the subspace of the random noise sources (see Chapter 4). In general, the differential approaches can be considered as being optimal in the SNR sense, if the target class of the sources is considered as signal and other sources as noise.

3.2.3 Independent Component Analysis

The concept of Independent Component Analysis was introduced first in [Com94], where the search criteria related to the information theoretic viewpoint have also been proposed. Nevertheless the first pioneer work that yielded extensive explorations in the BSS area is considered to be the paper by Herault and Jutten [HJ86]. Their technique was based on a recurrent neural network model with a modified Hebbian learning rule and was able to separate mixtures of sub-Gaussian sources, that is to say those that have kurtosis less than that of the Gaussian distribution with the same mean and covariance. This approach was further advanced in [JH91, KJ94, CUR94].

Generally, there are two basic approaches to ICA: the projection pursuit and the information theoretic ([Hyv99c]). The slightly intuitive approach to ICA is the projection pursuit [Hub85, Fri87], where the components are estimated based on some pursued property they should possess. For example the central limit theorem says that the distribution of a sum of independent identically distributed random variables tends towards Gaussian distribution. Conversely, the most mutually independent components are obtained, when the sum of their dissimilarities to Gaussian distribution is maximized.

Several measures of non-Gaussianity can be used. One of them is the kurtosis - the fourth-order cumulant of a random variable. Kurtosis is zero for the Gaussian random variable and nonzero for most non-Gaussian variables. Negative kurtosis is an indicator that the distribution is sub-Gaussian. On the contrary, the super-Gaussian distributions have positive kurtosis. Therefore, the absolute value or square of kurtosis can be used for measuring the non-Gaussianity of the components and in the capacity of separation criterion. However, kurtosis as a statistic is highly sensitive to outliers when approximated by a finite sample. Another measure of non-Gaussianity is negentropy. It is based on differential entropy and is more robust to outliers, since good approximations of it are available. See [CT91, Pap91, HKO01, Hay09] for possible approximations of negentropy.

There are a number of algorithms that employ kurtosis and negentropy as projection pursuit indexes. Both the sequential and simultaneous component estimation versions are available. For example, FastICA algorithms are widely used [HKO01, HO97, Hyv99a].

Another way to approach the independence of the components related to information theoretic viewpoint is to maximize the log-likelihood function with respect to a linear transform **B** [GL90, PG97, PP96, Car97]:

$$\frac{1}{T}\log L(\mathbf{B}) = E\{\sum_{k=1}^{K}\log p_k(\mathbf{b}_k^T\mathbf{x})\} + T\log|\det\mathbf{B}|, \tag{22}$$

where $\mathbf{B} = \mathbf{QPA}^{-1} = [\mathbf{b}_1, \ldots, \mathbf{b}_K]^T$. There exist a number of different viewpoints that approach maximization of (22) per se, e.g., by maximizing the joint entropy of the outputs of a neural network [BS95, Car97], by minimizing the mutual information or maximizing the sum of negentropies/non-Gaussianities of individual components [HKO01, LGBS00, GF97], using the maximum-likelihood (ML) parametric density estimation and the Kullback-Leibler divergence [PP96], or starting from the negentropy perspective for projection pursuit [GF96].

While the basic algorithms are sharpened to separate specific types of sources (e.g., sub- or super-Gaussian), the extensions were proposed to deal with mixtures of sources having various distributions. For example, in [LGS99] an extended Infomax algorithm was proposed. It is suitable for separating mixtures of super-Gaussian and sub-Gaussian sources thanks to a parameterized probability distribution switching between sub- and super-Gaussian modes using the stability analysis proposed in [CL96]. The technique uses the natural gradient proposed by [Ama98] to speed up the convergence compared to the normal gradient search used in the original Infomax algorithm [BS95]. A similar method was also proposed by [GF96]. However, the authors used the sign of the kurtosis to switch between sub- and super-Gaussian regimes.

In [CS93] independence between components is addressed by neutralizing the correlations of the second and fourth orders. The algorithm called JADE (Joint Approximate Diagonalization of Eigenmatrices) consists of two steps: whitening of the multichannel data and the subsequent joint approximate diagonalization of a set of eigenmatrices of the fourth-order cumulant tensor. The fourth-order cumulant tensor is a four dimensional array with entries containing fourth-order cross-cumulants $\mathrm{cum}(x_i, x_j, x_k, x_l)$ [HKO01]. The overall cost function that is minimized by JADE has the following from

$$\mathcal{J}_{JADE}(\mathbf{W}) = \sum_{i\neq j}\left(\frac{\sum_{t=1}^{T} y_i(t)y_j(t)}{T-1}\right)^2 + \sum_{ijkl\neq iikl} cum(y_i, y_j, y_k, y_l)^2, \tag{23}$$

where $\mathbf{y} = \mathbf{W}\mathbf{x}$. Intrinsically this technique is similar to the maximization of the absolute value of kurtosis that is advocated by one of the implementations of the FastICA algorithm [HKO01]. Despite the theoretical validity, methods based on HOS exhibit inevitable problems. Namely, HOS are sensitive to outliers and require larger data samples for estimation.

An extension of the TDSEP algorithm to HOS called JADETD is a combination of TDSEP and JADE techniques [MPZ99]. This method uses the time structure and information carried by the higher order moments to separate latent sources. The resulting cost function minimized by JADETD can be obtained by combining basic criteria $\mathcal{J}_{\text{TDSEP}}$ (21) and $\mathcal{J}_{JADE}$ (23):

$$\begin{aligned}\mathcal{J}_{JADETD}(\mathbf{W}) = \sum_{k=0}^{N}\sum_{i\neq j}(\frac{\sum_{t=1}^{T} y_i(t)y_j(t+\tau_k)}{T-1})^2 + \\ \sum_{ijkl\neq iikl} cum(y_i, y_j, y_k, y_l)^2,\end{aligned} \tag{24}$$

where $\tau_0 = 0$. The delayed second-order correlations (first term) account for the information about temporal structure and allow for second-order correlations between components to be neutralised; whereas higher-order cumulants (second term) ensure the higher-order decorrelation.

Another approach to ICA is nonlinear PCA that provides type of nonlinear decorrelation [KJ94, KJ95]. The minimized cost function is adapted from (19) by changing linear terms to nonlinear:

$$\mathcal{J}_{MSE}^{NLPCA}(\mathbf{w}_1, \ldots, \mathbf{w}_m) = E\{\|\mathbf{x} - \sum_{k=1}^{m} g_k(\mathbf{w}_k^T\mathbf{x})\mathbf{w}_k\|^2\}, \tag{25}$$

where g_k are properly chosen nonlinear functions. It can be shown that after whitening and with certain choices of g_k this criterion becomes equivalent to ICA approaches based on, e.g., kurtosis and maximum likelihood [HKO01].

Moreover, a number of connected issues and models related to ICA have been addressed in numerous studies and publications, e.g. nonlinear mixtures, convolutive mixtures, noisy ICA [CC00, Hyv99b], and underdetermined mixtures [HKO01]. The problem of recovering sources under the underdetermined conditions, i.e. when the number of channels is less than the number of sources, is one of the most important and challenging ones addressed. This problem has a straightforward connection to the problem of finding the optimal data representations within overcomplete bases also called dictionaries [MZ93]. Although the representations of data are not unique within such bases, they may offer more efficient coding of data or more compact data representations. This is because the overcomplete number of basis functions designed to represent numerous specialized data patterns allows for better and more specific capturing of the data structure.

Referring to model (10) with the same number of channels as sources (mixing matrix $\mathbf{A}$ is a square matrix), the measurements $\mathbf{x}(t)$ can have a unique representation within the complete basis formed by sources $\mathbf{s}(t)$, and each element a_{ij}

in matrix $\mathbf{A}$ defines the coefficient or amount of contribution of a corresponding basis function $s_j(t)$ to the channel $x_i(t)$. However, in the case where more sources than observed mixtures are (mixing matrix $\mathbf{A}$ is not square, but rather has dimensions $d_1 \times d_2$, where $d_1 < d_2$) the data $\mathbf{x}(t)$ have multiple representations within the overcomplete basis $\mathbf{s}(t)$. In principle, using a basic ICA algorithm it is possible to find a square matrix $\mathbf{A}'$ such that the inverse of it will estimate the same amount of the most independent basis components as channels $\mathbf{y}(t) = (\mathbf{A}')^{-1}\mathbf{x}(t)$, but these components can only be mixtures of the original sources. In the normal mixture model finding the inverse of the mixing matrix also leads to estimating the sources. However, in the underdetermined case, one cannot estimate sources uniquely by inverting the rectangular mixing matrix $\mathbf{A}$, since it is only pseudo-invertible. Therefore in underdetermined problems one needs to perform two separate tasks: estimate the mixing matrix and sources. Thus, the source separation methods intended to solve underdetermined ICA problem are designed to have two separate steps each solving one part of the problem.

Usually the source separation methods within the underdetermined mixture model assume that sources also have sparse representation in some domain. In the time domain sources are unlikely to be sparse, so one needs to transform data first to a domain where this assumption may hold better, e.g. the time-frequency (TF) domain. For example, in [LAC+06] the mixing matrix is found based on ratios in the TF domain obtained by the wavelet transform. At the second step of the two-stage approach the authors estimate sources based on solving a linear programming problem. Also the recoverability analysis is presented to answer the question of which circumstances allow sources to be accurately estimated by solving the linear programming problem.

The probabilistic approach for solving the underdetermined problem was proposed in [LS00]. In the case of a square matrix $\mathbf{A}$ and zero additive noise the derivations of the authors converge to a standard ICA algorithm [BS95]. In the underdetermined case a different learning procedure is developed. In [LLGS99] empirical results from separating speech signals using this approach are presented.

An interesting approach to tackle the underdetermined problem was proposed in [KY04], where the underdetermined BSS problem was converted to the conventional problem by generating the missing channels. The algorithm is a two-step iterative procedure: (1) first, virtual channels are generated to maximize their conditional probability given the existing observations and unmixing matrix, then (2) based on the current estimates of the sources running parametric estimates of probability densities of the components are obtained, and unmixing matrix is updated using the extended Infomax algorithm.

Hence, there are many of approaches, viewpoints, and interpretations on ICA, but many of them are actually based on the same core principles and eventually converge to the same algorithms.

3.3 Application of Source Separation Methods to EEG

In this section, we briefly introduce the main milestones of the work done within the scope of application of the source separation methods to EEG data. The treatment is divided between ICA (Section 3.3.2) and other major approaches (Section 3.3.1).

3.3.1 Basic artifact removal approaches

The first attempts to remove eye activity artifacts from EEG recordings were based on regression between EEG and EOG channels either in the time [HG70, VGM82, GCD83] or frequency domain [WLM78, WVS83]. According to this approach, properly weighted data from EOG channels are subtracted from EEG channels. The weaknesses and problems of this method include: (1) dependence on having good reference EOG measurements and (2) potential loss of relevant EEG data. The EEG signal loss takes place because besides ocular activity EOG measurements contain EEG traces, and, thus, portions of EEG are also removed from EEG channels during subtraction. In order to reduce the cerebral activity in EOG channels used for computing regression coefficients, authors in [LPBS93] proposed low-pass filtering of EOG measurements.

Another set of techniques for removing EOG artifacts is related to fitting the spatiotemporal dipole models that explicitly model positions and orientations of dipoles representing sources [BS91b]. This approach, however, exhibits a number of bottlenecks. For any type of eye activity (blinks, movements, saccades) different dipole models need to be elaborated and fitted. Constructing the mixing model requires prior knowledge about the location and orientation of each dipole source or some method to estimate these from measurements. The estimation of the locations of sources based on measurements is related to finding the solution for the inverse problem [Cuf85, Sar87] that is difficult to address accurately, because of the simplified and roughly approximated models used in practice.

The method presented in [BS91a, BS94] avoids explicit estimation of the locations and orientations of the dipole sources by using PCA for direct computation of the topographies related to different EOG activities. Since the topographies are computed individually for each subject, this approach offers a good tailoring of the estimated model that takes into account features (head and brain topologies, and positions and orientations of the dipole sources) specific for a person. The EEG activity still needs to be modeled using the dipoles as before. For all that, PCA cannot completely separate eye activity from EEG, because it groundlessly restricts sources to be spatially orthogonal. In addition, it searches for temporally uncorrelated components that is a much weaker condition than the more physically-grounded independence assumption. All this makes estimates of ocular topographies inaccurate, although PCA offers clear advantages in performance over regression and pure spatio-temporal dipole models as shown in [LPBS93]. Moreover, PCA is effective mainly, when EOG is markedly bigger than EEG, and fails when they both have comparable variances [LSB97]. In addition, use of PCA - similarly to the pure spatio-temporal dipole models - requires a preliminary learning step, for which high-quality data samples containing different patterns (with and without eye artifacts) need to be obtained.

All the above mentioned methods share common inherent disadvantage related to the narrowness of the applicability area. For example, some artifacts, like those related to muscle activities, cannot be removed by these approaches, since no reference channels usually exist for them.

Another possible direction of artifact removal research is linked to spatial component decomposition methods. According to this approach, data are first decomposed to spatial components and artifactual components are identified. After that, EEG data are reconstructed back from the components, avoiding those related to artifacts.

For example, in [Har90] authors argue for the usefulness of PCA spatio-temporal components for EEG data analysis and interpretation. In their opinion, PCA components described by spatial distribution, temporal distribution, and amplitude represent spatio-temporal features that are physiologically meaningful and interpretable. Intuitively, PCA can be used, for example, for isolating ocular potentials. If the variance of a PCA component exceeds some threshold indicating abnormal concentration of energy in one component, this component is classified as artifactual and removed during inverse transformation. However, for other smaller artifacts and EEG activities the use of the maximal variance criterion for optimization is not grounded. Therefore bad sides of PCA in this application include: (1) orthogonality of the components that is unlikely natural to real sources, and (2) in general, motiveless maximal variance optimization criterion. Due to these reasons, PCA can not completely isolate ocular activity from other sources, and, thus, the loss of relevant EEG as well as incomplete removal of artifacts is possible during inverse transform.

3.3.2 Separation of EEG sources using ICA-based approaches

Recently developed ICA techniques are free from the numerous disadvantages inherent in the traditional specialized artifact removal methods. ICA algorithms can be applied to a variety of artifacts and not only eye-related ones. Moreover, the processing of different artifacts is unified within ICA, since the same model applies to all cases. ICA is capable of finding separation matrix without preliminary step, when calibration data are involved. In addition, ICA is an almost purely empirical and data-based solution and does not require modeling the locations and orientations of the generators in the head.

The results reported in [MBJS96] provide one of the first evidences that the application of ICA techniques to EEG data can be of benefit for the analysis of underlying processes. The basic Infomax algorithm from [BS95] was applied to both raw EEG data and averaged ERP responses. When applied to raw EEG, ICA was able to isolate different overlapping oscillatory activities and artifacts. Application to averaged ERP responses provided division of the complex

ERP waveforms to more psychophysiologically meaningful, temporally overlapping but spatially-separable subcomponents. In [MJB+97] the Infomax algorithm equipped with natural gradient [ACY96] and super-Gaussian prior was applied to decompose averaged auditory event-related brain responses into components with fixed scalp distributions and sparsely activated, maximally independent time courses. Each of these components were shown to appear at certain experimental conditions and, thus, associated with specific brain-processing mechanisms.

The extended Infomax algorithm [LS97] equipped with natural gradient, suitable for separating both super-Gaussian and sub-Gaussian sources, was applied to real EEG data for removal of different types of artifacts in [JHL+97]. In this work, authors compare performance of the ICA with the multiple-lag regression method [KMVS91], in order to show the advantages of ICA. Removal of electroencephalographic artifacts was also addressed in [JMH+00], where the extended Infomax was used. The successful results reflecting isolation and following removal of artifacts by ICA are presented in comparison to regression and PCA methods to demonstrate their drawbacks. The results reported in [JMM+01] prove one more time the usefulness of ICA for analyzing EEG, Magnetoencephalography (MEG), and functional Magnetic Resonance Imaging (fMRI) data. While the extended Infomax algorithm with natural gradient was used to decompose averaged ERP brain data into simpler physiologically meaningful subcomponents, authors also pay attention to separating sources of the raw EEG segments. Besides isolating artifacts, this allows patterns and phenomena that have varying temporal and spatial characteristics to be captured. Considering spatio-temporal dynamics may provide information about changes in subject's performance or state, and about complex interactions between different physiological processes and mechanisms. The "ERP-image" visualization tool was employed to show and analyze trial-to-trial variabilities. This visualization scheme enables the linkages between properties of EEG activity and specific tasks, events, and conditions to be found. In [JMW+01] ICA was also applied to raw EEG data, and similar technique for visualization of variations between trials was used. The results showed that ICA was able to separate distinct sources and activities, allowing: (1) removal of artifacts, (2) identification and segregation of stimulus- and response-locked EEG/ERP components on a single-trial base, (3) temporal synchronization of the response-locked activities from single trials before averaging to deal with the problem of latency jitter, (4) capturing and identifying context-sensitive and interpretable EEG dynamics by investigating the spatio-temporal variability between trials, and (5) segregation of spatially overlapping but temporally separable EEG activities that may be related to different task events and conditions.

In [Vig97] the FastICA algorithm developed in [HO97] was applied to separate ocular artifacts from other activity with promising results. In [VO99] the same method was applied to MEG data and resulted in successful removal of different types of artifacts. Moreover, another numerical study in the same paper was related to the application of ICA to averaged ERP responses. As a result, ICA allowed to separate complex ERP waveforms into simpler independent subcomponents, which were functionally distinct and reflected different psychophysiological processing mechanisms and phenomena. In [VHO96] and [KHV+97] the use of ICA for isolating and the subsequent removal of artifacts from EEG is also demonstrated by the example of the FastICA algorithm. In addition, authors attempt to provide clear psychophysiological interpretations of the obtained results. The report in [VSJ+00] reviews the results and shows the use of FastICA for both the identification and removal of artifacts from raw EEG/MEG data and for the decomposition and analysis of event-related brain signals.

In [WSST04] SOBI algorithm was successfully applied for ERP denoising on a single-trial level. This offered clear advantages in the following ERP classification step used in both basic psychophysiology and Brain Computer Interface (BCI) research. In [HYCK07] the results of SEP signals extraction from a single trial using SOBI were reported. The authors used artificially generated data, in which the SEP template obtained after ensemble averaging was mixed with different types of noises modeling EEG and power-line at various SNR levels. The correlation coefficient between the SEP template and the SOBI-extracted SEP estimate was used for the purposes of performance evaluation. In [TPZC00] authors have chosen SOBI for decomposing real MEG data from different subjects and tasks, because it provided sufficient accuracy of the solution at a reasonable computational cost compared to other methods. They demonstrate that, without using any specific domain knowledge and assumptions on the dipole/head model, SOBI was capable of separating (1) artifacts and EEG, (2) neuronal responses from different sensory modalities, and (3) neuronal responses from different processing stages within a given modality. Therefore, the obtained neuronal components received physiological and anatomical interpretations.

The ICA algorithm considered in [ASM03] provides a generalization of previously used ICA methodologies by considering EEG sources as spatio-temporal activity patterns and allows for better capturing of the dynamics of brain signals. With this approach sources are assumed to be spatially unstable and capable of changing their locations. The above assumptions lead to a convolutive mixture model in the temporal domain that is equivalent to multiplicative mixing of the complex signals in different bands of the frequency domain. The EEG data transformed into the spectral domain are decomposed to independent components using the complex or convolutive Infomax ICA algorithm. This approach applied to EEG data from the visual attention experiment has shown a better level of independence between sources compared to the standard ICA. Found sources explained spatio-temporal dynamics of brain data, had limited spectral extent, and were connected to the subject's behavior. In [ADSM06] a similar technique of complex ICA was successfully applied to fMRI recordings to extract patterns of spatio-temporal dynamics of brain activity.

In [MIFM07] the joint use of wavelet transform and ICA (WICA) for artifact removal from EEG is studied. In contrast to previous approaches combining wavelets and ICA, where wavelet denoising was only used on a pre- or post-ICA processing basis [ZG04], in the proposed multiresolution ICA method wavelet transform is on the other hand an integral part of ICA. According to the procedure under consideration, extended Infomax is applied only to wavelet coefficients representing frequency ranges, where artifact is present. This allows for possible loss of relevant information to be avoided, since the entire data are not processed by ICA. Indeed, if wavelet and ICA denoising methods are applied in a sequence, on every stage there is a danger of losing the relevant signal, because subspaces of signal and noise are partially overlapped in both the time-frequency and spatio-temporal domains. On the contrary, the combination of the two methods, when one method provides only a specialized noisy subset of its data representation domain to the second method, allows for localization of the area, where signal and noise overlap. The proposed technique takes into account the four major EEG rhythms (*alpha*, *beta*, *delta*, *theta*, see Section 2.1) by decomposing original data into the four wavelet ranges corresponding to each of the rhythms. The method thus allows for advantage to be taken of the a priori known characteristic features of EEG and artifacts in the frequency domain. WICA was compared with two other methods, where wavelet denoising was done only as a pre- or post-ICA processing step. The investigated data were obtained by mixing real EEG with synthesized artifacts of common types. Correlation between the artificially generated artifact and the artifactual component extracted by a method was used for quantitative assessment of the methods' performance. Results showed that WICA outperforms two other methods for all types of modeled artifacts. A similar idea of joint localization of the noisy data subset was employed in [CM06], where a different order of methods was used: wavelet thresholding was applied to independent components. The proposed technique (wICA) was compared with pure ICA on semi-simulated and real EEG data sets. The distortions of cerebral activity after ICA and wICA artifact suppressions were analyzed in time and frequency (amplitude and coherence/phase characteristics) domains. The results showed that wICA preserves characteristics of cerebral activity better than ICA alone.

Many studies have been done to compare different ICA algorithms between each other and with other techniques. For example, in [AIF+04] Infomax [BS95], extended Infomax [LS97], FastICA [Hyv99a], nonlinear ICA algorithms [YAC98, JK03], and MISEP [Alm03b, Alm03a] are reviewed, and the relative performances of these methods are compared against one another and, in particular, for their application to EEG data. Another comparative study is reported in [JHL+98], where an extended Infomax algorithm is compared with PCA as applied to artifact removal in EEG data. The results show the superiority of ICA. The drawbacks of PCA related to orthogonality of the components and an inconsistent maximal variance optimization criterion are discussed and explained. In [KJK+05] various ICA algorithms, MS-ICA [MS94], OGWE [MFBTGS03], JADE [Car99], SHIBBS [CBF96], Kernel-ICA [BJ02], and RADICAL [LMI03], are brought into comparison in the removal of ocular artifacts from EEG. The independence between components is assessed using a mutual information estimator, which is based on k-neighbor statistics and does not require prior knowledge about density functions [KSG04]. According to the results, the RADICAL algorithm has shown the best performance at separating the source signals from the observed

EEG measurements. In [IZ07] an iterative ICA technique (iICA) is set against the ensemble averaging and wavelet transform (WT) applied to EEG data for EP denoising. Methods were tested on both simulated and real EEG data sets. As results show, iICA outperforms other methods when applied to single trials and averaged responses. Performance of the methods was quantified using the average root-mean-square error and average correlation before and after processing.

The study reported in [CGAC08] focuses on removing muscle artifacts. The performance of four ICA algorithms, AMUSE, SOBI, Infomax, and JADE, was assessed at separating myogenic activity from EEG during sleep. The algorithms were applied to semi-simulated data, in which real sleep EEG measurements were combined with EMG of different powers. The efficiency of ICA algorithms was graded using three scores: computational time, Pearson's correlation coefficient, and mean-square error. According to the results, JADE showed the worst performance among used algorithms for temporal area. On the other hand, AMUSE was considerably faster than other methods and showed evenly good results independently of the power of muscle artifact for non-temporal regions. AMUSE was further applied to real EEG data and was successful at separating EMG from spontaneous EEG arousals at different sleep stages.

In the study presented in [GB01] the joint cumulant- and correlation-based (JCC) algorithm is developed. This algorithm is similar to that of JADETD and uses joint information obtained from higher order statistics and delayed cross-correlations to separate sources. The technique operates by simultaneous diagonalization of the set of time-delayed cross-correlation matrices used in SOBI and quadricovariance eigenmatrices as in JADE thereby combining the optimization criteria of both methods. The carefully designed experiment allowed for EEG/EOG data with properties that facilitated comparative evaluation of the performance of the three analyzed algorithms (SOBI, JADE, and JCC) to be obtained. The results and discussion showed the advantages and drawbacks of SOBI and JADE, and the usefulness of combining their beneficial properties within JCC. Thus, this study empirically proves that combining second- and higher-order information can offer better performance at separating EEG and artifacts than using these approaches separately.

In [GFF^{+}04] a framework for evaluating ICA methods as applied to eye blink artifact removal from multichannel EEG data was proposed. The performances of two algorithms, Infomax and FastICA, were compared at various synthetic configurations of data properties. The covariance between original blink-free and ICA-filtered EEG data was used as a quantitative metric for assessing the quality of the artifact extraction. Moreover, the averaging of filtered segments time-locked to blinks followed by visual inspection of the resulting time courses and topographies provided an additional qualitative characteristic of the cleanness of the filtered data from artifacts.

In [DSM07] the utility of ICA for improving the detection of artifacts is explored. Three methods, Infomax, SOBI, and FastICA, were used to decompose multidimensional EEG data into independent components revealing different sources contributing EEG measurements. Five techniques for detecting various types of artifacts were tested on EEG data with integrated artifacts of different kinds and levels of magnitude. As was shown, artifact detection techniques applied to ICA-decomposed data perform much better than if applied to raw data.

Several studies have focused on developing consistent and reliable criteria for automatic detection of artifactual, noisy, and study-relevant independent components. In [CV00] three classification criteria were explored with application to both simulated and real single trial EEG data: kurtosis, linear predictor, and Hurst exponent. In [VC02] the same criteria were also used to characterize independent components. The authors use modified CII algorithm [CACCR00] for component extraction. After useful components are selected using detection methods, they are filtered using the Wiener filter [Hay96] with parameters estimated using only signal subspace and projected back to electrodes. The approach was tested on both simulated and real EEG data and demonstrated advantage over ICA and filtering applied separately. The authors also show the principal necessity of performing ICA before applying signal detection mechanisms. In [DMS01] kurtosis and Shannon's entropy were used jointly as markers for component classification; whereas in [GMMV05] authors attempt to develop more efficient component classification criteria and empirically argue that the joint use of kurtosis and Renyi's entropy is more effective and allows some failures appearing when Shannon's entropy is used to be avoided. They use the extended Infomax algorithm in their computations on real EEG data.

The work reported in [BPZ$^+$04] presents a framework for identification and removal of artifacts from MEG consisting of three steps: (1) the application of the ICA algorithm, (2) the application of methods for automatic detection of artifactual activities based on statistical (kurtosis, entropy) and spectral (Power Spectrum Density correlation) characteristics of independent components, and (3) a control cycle on discrepancy between original and ICA-filtered data in order to avoid occasional loss of relevant data when rejecting 'artifactual' ICA components. The authors used the ICA algorithm called Cumulant-based Iterative Inversion in Signal Subspace (CIISS) that is an extension of the Cumulant-based Iterative Inversion (CII) algorithm [CACCR00, CACRC02] designed for noisy mixtures equipped with robust whitening procedure [BC00]. The approach was successfully tested on simulated data with sources modeling different EEG and artifactual activities as well as real MEG data.

3.4 Validation and assessing the quality of the results

In this chapter, we discuss how to evaluate and measure the quality of data or the results of source separation. We will consider the general-level suggestions to validation (Section 3.4.1) as well as the existing numerical techniques related to this issue (Section 3.4.2) and some practical limitations and advances (Section 3.4.3). Here, we examine only those approaches, which are applicable to real EEG data.

3.4.1 Quality evaluation measures: general overview

Properties of data known a priori from a problem area can be used either for performing the denoising (e.g., as contrasts for component extraction and selection procedures) or for validation purposes. For example, interesting characteristic features of the data can be considered as a specific content affiliated either with signal or noise in the temporal or the frequency domain (see, e.g., the temporal structure of ERP waveform in Figure 5). A specific probability distribution linked to a source is also an important property that can be used for distinguishing between sources. Basically, denoising or component extraction and validation criteria should form a complete set incorporating all knowledge available on data, in order to check all possible aspects of a method influencing the data.

The quality evaluation methods can be classified according to their nature as visual or manual and numerical or formal. All criteria can be evaluated either manually or numerically. In manual methods the qualitative validation is obtained based on an opinion of an expert or a pool of experts that possess a specific domain knowledge to exclude the subjective factor from the evaluation results. The expert knowledge is usually based on special data properties (e.g., waveform), whose quantification leads to numerical assessment.

For MMN studies, the following properties can be used for assessing the quality of data before and after a method application:

Polarity reversal. In data obtained using some specific experimental setups the polarity of the signals coming from the same source can be reversed in different channels. This may happen due to specific respective locations of the source, measuring sensors, and the reference sensors, and the specific orientation of the source dipole. For example, in our first data set from the study reported in [PLL95] MMN deflection is positive in mastoids M1 and M2 channels (see Figure 6). In contrary, MMN is negative in all other channels, except maybe Pz channel, where the polarity of MMN may vary. This spatial property of the data can be used (1) as a part of a component separation criterion, (2) to pick out correct components related to MMN after a transformation, and (3) for validation purposes.

Discriminating deviants. It is well-known in the theory and practice of electroencephalography research that a stronger deviant elicits a larger MMN peak with shorter latency. Some recent studies have indicated that the stronger deviant does not necessarily produce MMN with larger amplitude (see e.g., [HCJ+08]). We believe, however, that in many paradigms including the one from [PLL95] this relation holds, since it is supported by a substantial number of publications and observations [PLL95, HHKL07, KGJ+07]. Since 30 ms deviant is stronger (shorter in our case) in the mentioned paradigm, it should elicit MMN with larger amplitude and shorter latency than the deviant with 50 ms duration (see Figure 5). This criterion can be used mainly to verify the validity of a denoising technique that should strengthen the underlying relations between deviants. Moreover, if the difference between the two deviant responses is well formalized, it can be used as part of a separation criterion to regularize the solution and for component selection purposes.

SNR. SNR is a natural indicator of the quality of data and the success of denoising. If the SNR estimate increases after the method application, this serves as a partial confirmation of a successful denoising and the beneficial effect of the method. However the SNR considered alone presents an incomplete characteristic of the performance of denoising, since it does not provide information about possible loss of signal that may take place after denoising. The stronger requirement for the results of a denoising method to be counted successful is formulated as noise reduction, while retaining the signal power at the original level. Therefore when SNR is employed as a measure of quality of the denoising it should be accompanied by a properly chosen estimate of the signal loss (see Section 3.4.3). Moreover, target criteria of the component separation methods, which are optimal in the sense of SNR, can be used for validation purposes (see Section 4.3.2).

Group differences. This indicator is mainly used for validation purposes. If the two groups of participants (e.g., target and control) are known to deviate significantly at values of some ERP characteristic (e.g., amplitude or latency), the successful denoising should reveal the group difference in terms of the indicative characteristic.

Temporal structure. If the signal of interest has some known specific shape or regular structure in the temporal domain, this can be used for component separation as well as for validation (for example, see the temporal structure of the ERP waveform in Figure 5). For instance, in MMN studies the grand average possesses a known average pattern of the temporal structure across participants. Basically correlation between the subject average and the grand average can be used as a numerical criterion for component separation or validation. However, this criterion is overspecified, because besides the temporal pattern common for all participants it also contains and specifies the values of the amplitude and latency parameters that should, in fact, differentiate between the participants. As a result, subject specifics can be lost during the denoising based on this criterion and they can be disregarded during the validation process.

3.4.2 Estimating signal-to-noise ratio

The natural and most widely used measure of the quality of data is SNR [Pap91]. Therefore, in this section we consider methods for estimating SNR together with the powers of signal and noise as its composite parts.

Many techniques for estimating SNR in single trial/averaged responses have been developed [vdV00, Sch67, SRC74, MGT84, MTG84, RBH$^+$02, ÖD96, FB91], which essentially are based on the assumptions underlying the averaging (see Section 2.5). In [Sch67] the $(\pm)$-reference method was proposed for estimating the noise power in the averaged response. The $(\pm)$-reference is computed just as normal averaging, but half of the trials are taken with the negative sign:

$$\overline{x}'_N(t) = \frac{1}{N}\sum_{i=1}^{N}(-1)^i x_i(t) = \frac{1}{N}\sum_{i=1}^{N}(-1)^i(\varsigma_i(t) + \eta_i(t)) = \\ \frac{1}{N}\sum_{i=1}^{N}(-1)^i \eta_i(t) = \eta'(t). \tag{26}$$

If we have small odd number of trials N, it is desirable to use all of them in computing the estimates of signal and noise powers in the average for better accuracy. Thus, the following estimate of noise in the average can be used:

$$\eta'(t) = \frac{1}{2}\left(\frac{\sum_{i,i \bmod 2=1} x_i(t)}{N_1} - \frac{\sum_{i,i \bmod 2=0} x_i(t)}{N_2}\right), \tag{27}$$

where the two numerators in the above equation represent the summations of N_1 odd-numbered and N_2 even-numbered trials. N_1 and N_2 will, thus, differ at most by one trial.

In [SRC74] it has been shown that the variance of $\eta'(t)$ approximately equals to the variance of noise in averaged response of N trials:

$$var\{\eta'(t)\} \approx \sigma^2_{\overline{\eta}} \approx \frac{\sigma^2_{\eta}}{N}, \tag{28}$$

if the assumptions imposed on signal and noise properties in Section 2.5 are satisfied. The variance of noise in the averaged response should decrease proportionally to the number of averaged trials N. Therefore, the following formula can be used for signal-to-noise power ratio estimation in the averaged response:

$$\widehat{SNR}_N = \frac{\widehat{P}_{\varsigma}}{\widehat{P}_{\overline{\eta}}} = \frac{var\{\overline{x}_N(t)\}}{var\{\overline{x}'_N(t)\}} = \frac{var\{\varsigma(t) + \overline{\eta}_N(t)\}}{var\{\eta'(t)\}} \cong \\ \frac{var\{\varsigma(t)\}}{var\{\eta'(t)\}} \approx \frac{\sigma^2_{\varsigma}}{\sigma^2_{\eta}/N} = N\left(\frac{\sigma^2_{\varsigma}}{\sigma^2_{\eta}}\right), \tag{29}$$

where $\widehat{P}_{\varsigma}$ is the estimate of the signal power, and $\widehat{P}_{\overline{\eta}}$ stands for the estimate of the noise power in the average. The estimate of SNR measured in decibels (dB) can be computed using the result of formula (29) as:

$$(\widehat{SNR}_N)_{\mathrm{dB}} = 10\log_{10}\widehat{SNR}_N. \tag{30}$$

A similar procedure was proposed in [ED84], where the ratio of variances F_τ is calculated based on a single point reference in the evoked potential:

$$F_\tau = \left(\frac{\sigma^2_{\overline{x}}}{\sigma^2_{\tau}}\right), \tag{31}$$

where $\sigma_{\overline{x}}^2 = \frac{1}{T-1}\sum_{t=1}^{T}\overline{x}_N^2(t)$ is the variance of the averaged response $\overline{x}_N(t)$, and σ_τ^2 is the variance of the background noise, estimated at $t = \tau$:

$$\sigma_\tau^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i(\tau) - \frac{1}{N}\sum_{i=1}^{N}x_i(\tau))^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i(\tau) - \overline{x}_N(\tau))^2. \tag{32}$$

The time point τ can be arbitrarily chosen. It is, thereby, sensitive to time-locked artifacts. On the other hand, this allows for the investigation of the reproducibility of individual ERP features. F_τ reveals the estimate of average SNR in a single trial.

In [MGT84] the following set of formulae for estimating signal and noise powers, and SNR in a single trial, was proposed:

$$\widehat{\sigma}_\eta^2 = \widehat{P}_\eta = \frac{1}{N-1}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\left(x_i(t) - \overline{x}_N(t)\right)^2, \tag{33}$$

$$\widehat{\sigma}_\varsigma^2 = \widehat{P}_\varsigma = \frac{1}{T}\sum_{t=1}^{T}\overline{x}_N^2(t) - \frac{1}{N}\widehat{P}_\eta, \tag{34}$$

$$\widehat{\sigma}_x^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}x_i^2(t), \tag{35}$$

$$\widehat{SNR}_1 = \frac{\widehat{\sigma}_\varsigma^2}{\widehat{\sigma}_\eta^2} \approx \frac{\widehat{\sigma}_x^2 - \widehat{\sigma}_\eta^2}{\widehat{\sigma}_\eta^2}. \tag{36}$$

In [RTF88] the asymptotic distribution of the SNR estimator (36) was analyzed and method for constructing confidence intervals was proposed. In fact, quality measure F_τ is very closely related to $\widehat{SNR}_1$ proposed in [MGT84] (see [ÖD96]). The respective estimates of SNR in averaged responses $\widehat{SNR}_N$ can be obtained from single trial estimates (31) and (36) by multiplying $\widehat{SNR}_1$ and F_τ by N.

In [ÖD96] a running SNR estimator for averaged response optimized for on-line usage and based on formulas (33)-(36) was reported:

$$\widehat{SNR}_N = \frac{(1+\frac{1}{N})\sum_t(\sum_i x_i(t))^2 - \sum_t\sum_i x_i^2(t)}{\sum_t\sum_i x_i^2(t) - \frac{1}{N}\sum_t(\sum_i x_i(t))^2}. \tag{37}$$

In [RHB+97] this estimator was used for identification of the quality of ERP templates used for direct brain interface (DBI). However, this estimator lacks accuracy for small numbers of trials, because of some mathematical simplifications made in the derivation. Namely, it is biased towards underestimating the noise power and, thus, towards overestimating the signal power as shown in, e.g., [RBH+02]. In the latter work a more robust version of this running SNR estimator was used for quality assessment of ERP estimates:

$$\widehat{SNR}_N = \frac{(1+\frac{1}{N-1})(\frac{1}{N})\sum_t(\sum_i x_i(t))^2 - (\frac{1}{N-1})\sum_t\sum_i x_i^2(t)}{(\frac{1}{N-1})\sum_t\sum_i x_i^2(t) - (\frac{1}{N})(\frac{1}{N-1})\sum_t(\sum_i x_i(t))^2}. \tag{38}$$

3.4.3 Remarks about the practical limitations of SNR estimation approaches, signal loss estimation, and relation of the subspace separation criteria to the SNR estimation

Unfortunately, all of the considered methods start to output sufficiently accurate results only after a sufficiently large number of trials are collected [ÖD96]. This happens because the assumptions underlying these methods are close to being met only with large data samples. In particular, finite sample estimates of the signal and noise suffer from the influence of the present non-zero signal-noise and noise-noise cross-correlations, which are assumed to be zero by the methods and tend towards zero when the sample size increases.

To understand better the intimate reasons for the lack of accuracy, when the number of trials is small, let us decompose the estimate of the variance of the average into four additive components taking into account (1):

$$\begin{aligned} var\{\overline{x}_N(t)\} = \frac{1}{N^2} var\{\sum_{i=1}^{N} x_i(t)\} = \\ \frac{1}{N^2}\left(\sum_{i=1}^{N} var\{x_i(t)\} + 2\sum_{i<j} Cov_{x_i x_j}\right) = \\ \frac{1}{N^2}(\sum_{i=1}^{N}\left(\sigma_\varsigma^2 + 2Cov_{\varsigma\eta_i} + \sigma_{\eta_i}^2\right) + \\ 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\left(\sigma_\varsigma^2 + Cov_{\varsigma\eta_i} + Cov_{\varsigma\eta_j} + Cov_{\eta_i\eta_j}\right)) = \\ \sigma_\varsigma^2 + \frac{1}{N^2}\sum_{i=1}^{N}\sigma_{\eta_i}^2 + \frac{2}{N}\sum_{i=1}^{N} Cov_{\varsigma\eta_i} + \frac{2}{N^2}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} Cov_{\eta_i\eta_j}. \end{aligned} \tag{39}$$

Here, $Cov_{\xi_i\xi_j} = \frac{1}{T}\sum_{t=1}^{T}\xi_i(t)\xi_j(t)$ denotes the estimate of the covariance between the two quantities $\xi_i(t)$ and $\xi_j(t)$ that depend on time variable. One can see that the variance of the average consists of the following four terms: signal variance σ_ς^2, variance of the noise $\frac{1}{N^2}\sum_{i=1}^{N}\sigma_{\eta_i}^2$, cross-covariance between signal and noise $\frac{2}{N}\sum_{i=1}^{N} Cov_{\varsigma\eta_i}$, and cross-covariance between noises from different trials $\frac{2}{N^2}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} Cov_{\eta_i\eta_j}$. All methods assume that the sum of the two last terms equals to zero. This is approximately true when only large data sample is available. Otherwise, the cross-correlation terms are essentially large to degrade the accuracy of SNR estimate significantly. The influence of cross-correlation terms on SNR estimate manifests differently depending on whether they are counted as being signal or noise by the estimation procedure.

To visualize how the cross-correlation terms influence the estimates of SNR we did a numerical simulation illustrated in Figure 9. The simulated data consisted of 200 trials with 130 time samples each obtained by summing the randomly generated noise realizations $\eta_i(t)$ with variance $\sigma_\eta^2 = 10^3$ and the deterministic waveform $\varsigma(t)$ with variance $\sigma_\varsigma^2 = 8$. Hence, the SNR (SNR_1) was kept at $8 \cdot 10^{-3}$ on average over trials. Figure 9(a) contains the theoretical SNR in the

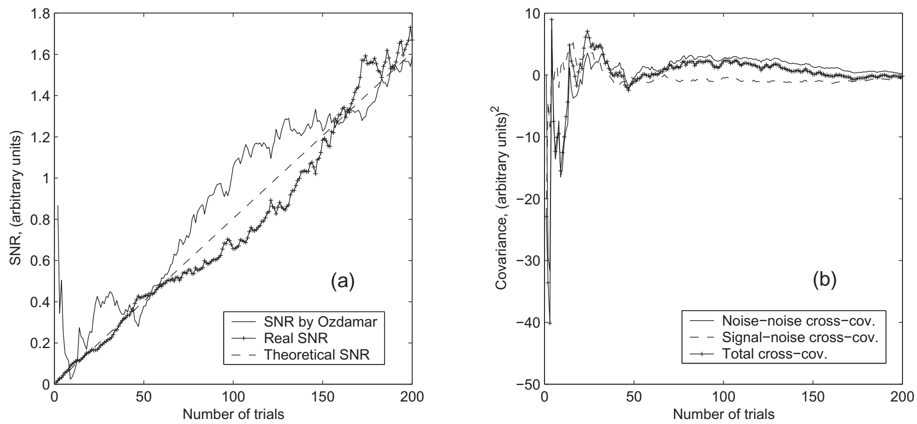

FIGURE 9 Illustration of the performance of SNR estimation: (a) SNR estimate by Özdamar formula (solid line), theoretical SNR (in dashed), real SNR (with crosses); (b) noise-noise cross-covariance term (solid line), signal-noise cross-covariance term (in dashed), sum of the noise-noise and signal-noise cross-covariance terms (with crosses).

average computed according to (28) and (29) as $SNR_i^{theoretical} = i \cdot SNR_1$, the estimate of SNR computed according to Özdamar's formula (38), and the real SNR computed by accounting for the cross-correlations as

$$SNR_i^{real} = \frac{\sigma_\varsigma^2}{var\{\overline{x}_i(t)\} - \sigma_\varsigma^2} = \frac{\sigma_\varsigma^2}{\frac{1}{N^2}\sum_{i=1}^N \sigma_{\eta_i}^2 + \frac{2}{N}\sum_{i=1}^N Cov_{\varsigma\eta_i} + \frac{2}{N^2}\sum_{i=1}^{N-1}\sum_{j=i+1}^N Cov_{\eta_i\eta_j}}. \tag{40}$$

In this formula all cross-correlation terms are designated as noise variance.

One can see that the real SNR and the estimate of SNR by Özdamar reveal approximately symmetric behavior around the theoretical SNR curve (Figure 9(a)). Moreover, real and Özdamar's SNR curves correlate with the cross-covariances shown in Figure 9(b). The positive correlation for Özdamar's SNR estimate is explained by fact that the energy of the cross-correlations is mainly assigned to the estimate of the signal energy by formulas (33) and (34), and, thus, this energy is placed to the numerator of SNR definition formula (3). On the contrary, all cross-covariances are treated as noise in the real SNR formula (40) and stand in the denominator of SNR definition formula (3). These observations explain the symmetry between Özdamar's and real SNR curves with respect to the theoretical SNR curve.

Based on the above discussion we draw a conclusion that the accuracy of SNR estimates can be improved, if one is able to segregate the powers of signal and noise from cross-covariances. This is, however, impossible in the framework of our assumptions and can be done only within an extended and augmented set of assumptions of a more specialized problem.

Nevertheless, often the main purpose of the quality estimates such as SNR is not assessing the absolute quality of data, but the comparison of different denoising methods and data sets. If the purely comparative goal is pursued, we do not need to find exact values of either SNR or signal and noise energies. Instead, if an SNR estimation method provides equally biased SNR estimates for any two data sets (e.g., before and after denoising), then the difference between these estimates will provide an unbiased estimate of the difference between the SNR of the two data sets.

Additionally, we would like to note that the optimization criteria of the methods presented in Chapter 4 can be used for SNR estimation, because they are targeted at discriminating between signal and noise components essentially in the same way as in the SNR estimation. The values of the optimized criteria affiliated to found spatial components are used for comparison between the quality of the components. These criteria can naturally be adapted for quality assessment in a single channel. This can be done by replacing the estimates of the variances along a spatial direction before and after data modification (e.g., averaging) by the estimates of the variances of single trial and modified response in the analyzed channel. For example, the criterion $\rho_2(\mathbf{w})$ (see equation (49)) can be adapted to estimate SNR individually for each channel as follows

$$\widehat{SNR}_1 = \frac{var\{\overline{x}_N(t)\}}{\frac{1}{N}\sum_{i=1}^{N} var\{x_i(t)\}}. \tag{41}$$

Here, in the numerator the estimate of variance of data after the averaging along a spatial direction is replaced by the estimate of variance of the averaged response at a single channel. Correspondingly, in the denominator the estimate of the single trial variance along a spatial direction is replaced by the estimate of the single trial variance in the channel. This estimator classifies data as of better quality, if the reduction of the variance after the averaging with respect to the variance in a single trial (raw data) is smaller. One can notice that the SNR estimate (41) is similar to one defined by formulas (33)-(36). This simple demonstration shows that principles of estimating SNR can be used to build criteria for spatial component separation and vice-versa. The problem with indeterminacy of cross-correlations remains, however, unsolved with the basic component separation criteria considered later (see Chapter 4 for further discussion).

As it was mentioned in Section 3.4.1, the SNR estimate cannot provide complete evaluation of the denoising results, since it does not tell us whether a signal loss takes place after the denoising or not. And if signal loss does take place, then to what extent. Therefore the issue of signal loss must be addressed separately. Let us consider possible approaches to validate a denoising method with respect to the signal loss.

If the number of trials available for one subject is so large that it allows for an almost noise-free ERP estimate, then the following strategy can be used. First, the set of trials is divided into the nonintersecting subsets, and the denoising is applied to each subset separately. Then the denoising results obtained for separate trial subsets are averaged. This averaged ERP estimate is then compared to the average of all trials to evaluate the signal loss.

However if the reliable ERP estimate cannot be constructed using all trials of one participant, the concept of grand average introduced in Section 2.3 can be employed to evaluate the signal loss tendency. First, a denoising method is applied to each subject's data separately. Then the signal loss is verified by comparing the two grand averages made before and after denoising. The discrepancy between the two grand averages signifies the disturbances introduced to the signal by the denoising procedure. This approach was approbated in article **[PI]**.

4 THESIS CONTRIBUTION

4.1 Introduction to the proposed methodology

The major idea encompassing the various ERP denoising methodologies discussed in this thesis consists in separating by a linear transformation ERP and noise subspaces in multichannel EEG data, assuming that data follow the linear instantaneous mixing model with electrode additive noise (12). By projecting back to the channel space only the components related to the signal subspace and neglecting noise components the denoising of the measurements is accomplished.

The main accent is put on developing consistent and reliable problem-specific criteria for distinguishing between the ERP and noise components in multidimensional EEG data space. The proposed approach to generate the criteria is to expose the data to some modification that changes signal and noise subspaces in different ways. Thus the signal and noise components are distinguished using these component-specific changes or the differential behavior they exhibit as a response to the induced perturbation.

We concentrate on those characteristic differences that are shown on the level of second-order statistical properties of data. Namely, the variances of data along a direction before and after performing the action on data are compared. Since the covariance matrix of the data contains all information regarding variances along different directions, we can also consider the process of comparing the two covariance matrices to identify the signal/noise discriminative features. Moreover, we focus attention on those modifications that exploit the data which have three dimensions of variation: channels, time samples, and trials. The differences between the two pre- and post-processing covariance matrices can be viewed on an additive or multiplicative level (see Figure 10). When referring to the additive differences we assume the subtraction of the variances of the original and modified data both projected to a direction $\mathbf{w}$:

$$\rho_{\Delta}(\mathbf{w}) = \mathbf{w}^{\mathrm{T}}\mathbf{C}_{\tilde{\mathbf{x}}}\mathbf{w} - \mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{x}}\mathbf{w} = \mathbf{w}^{\mathrm{T}}(\mathbf{C}_{\tilde{\mathbf{x}}} - \mathbf{C}_{\mathbf{x}})\mathbf{w} = \mathbf{w}^{\mathrm{T}}\Delta\mathbf{C}_{\mathbf{x}}\mathbf{w}, \tag{42}$$

where the matrix $\Delta\mathbf{C}_{\mathbf{x}}$ contains the additive difference between the covariance matrix of the original (initial) data $\mathbf{C_x}$ and the covariance matrix of the modified data $\mathbf{C_{\tilde{x}}}$. Here, the data $\mathbf{x}(t)$ is projected to a direction $\mathbf{w}$ as $y(t) = \mathbf{w}^{\mathrm{T}}\mathbf{x}(t)$, and the variance of the component $y(t)$ is obviously equal to $\mathbf{w}^{\mathrm{T}}\mathbf{C_x w}$. Therefore, the difference $\mathbf{w}^{\mathrm{T}}\Delta\mathbf{C_x w}$ between the two variances shows how the variance of the original data changed along the direction $\mathbf{w}$ after modification. We will see further (see Section 4.3) that variance changes after certain modifications depend on the SNR of data projected to $\mathbf{w}$. Therefore, this relation allows the directions explaining high SNR and those that describe low SNR to be distinguished.

Both the original $\mathbf{x}(t)$ and the modified $\tilde{\mathbf{x}}(t)$ data must be considered in an abstract sense that goes beyond the single trial or average (13) (see continuation of this section). As original data we can, for example, treat a single trial $\mathbf{x}(t) = \mathbf{x}_i(t)$ or all single trials considered together as one data sample or long sequence of concatenated channel-wise multidimensional trials $\mathbf{x}(t) = [\mathbf{x}_1(t)\ \mathbf{x}_2(t)\ \ldots\ \mathbf{x}_N(t)]$ etc. Correspondingly, the modified data in the simplest case can be the average $\tilde{\mathbf{x}}(t) = \overline{\mathbf{x}}(t)$ or many leave-one-out averages considered as one entity etc.

Note that additive differences of the variance as separation (source discrimination) criteria are valid only for whitened data, when source vectors form the orthogonal system (see Sections 3.2.1 and 4.3.1). When the multiplicative changes are analyzed, correspondingly, the ratio between the two variances is considered

$$\rho_{:}(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{C_{\tilde{x}}} \mathbf{w}}{\mathbf{w}^T \mathbf{C_x} \mathbf{w}}, \tag{43}$$

where $\rho_{:}(\mathbf{w})$ describes the multiplicative changes in variance of the data along the direction $\mathbf{w}$.

The principle of constructing the criteria $\rho_{\Delta}(\mathbf{w})$ and $\rho_{:}(\mathbf{w})$ is demonstrated in Figure 10, where the traditional averaging is used as data modification. The simulated data were obtained by randomly generated four realizations of the two sources $\mathbf{s}_i(t) = [s_{i1}(t)\ s_{i2}(t)]^{\mathrm{T}}$, $i = 1, \ldots, 4$ (Figure 10(a)). Then four trials $\mathbf{x}_i(t) = [x_{i1}(t)\ x_{i2}(t)]^{\mathrm{T}}$ of the simulated data were obtained by mixing the sources $\mathbf{x}_i(t) = \mathbf{A}\mathbf{s}_i(t)$ using randomly generated matrix $\mathbf{A}$ (Figure 10(b)). The contours of the scatter-plots of the original $\mathbf{x}_i(t)$ and averaged $\overline{\mathbf{x}}_4(t)$ data are indicated in Figure 10(c) together with the signal/noise directions and the variance components $\mathbf{w}^{\mathrm{T}}\mathbf{C_{\overline{x}}w}$, $\mathbf{w}^{\mathrm{T}}\mathbf{C_x w}$ of the criterion $\rho_{:}(\alpha)$ for an arbitrary direction $\mathbf{w}$. In Figure 10(d) the values of the criterion $\rho_{:}(\alpha)$ for different angles α corresponding to direction $\mathbf{w}$ are shown. Here, α_s denotes the angle corresponding to the signal direction, where $\rho_{:}(\alpha)$ is maximized, and α_n is the angle of the noise direction, where $\rho_{:}(\alpha)$ is minimized. In Figure 10(e) the same information is plotted as in Figure 10(c), but the initial and averaged data are whitened $\mathbf{z}_i(t) = \mathbf{V}\mathbf{x}_i(t)$, $\overline{\mathbf{z}}_4(t) = \mathbf{V}\overline{\mathbf{x}}_4(t)$. The variances of the initial and averaged whitened data are depicted as $\mathbf{w}^{\mathrm{T}}\mathbf{C_z w}$ and $\mathbf{w}^{\mathrm{T}}\mathbf{C_{\overline{z}}w}$, correspondingly. One can see that signal and noise directions become orthogonal in the whitened data space and, thus, the additive criterion $\rho_{\Delta}(\alpha)$ can be applied. It is again maximized for signal direction (at α_s) and minimized along noise direction (at α_n) as shown in Figure 10(f).

Here, the estimate of the covariance matrix of the original data is computed

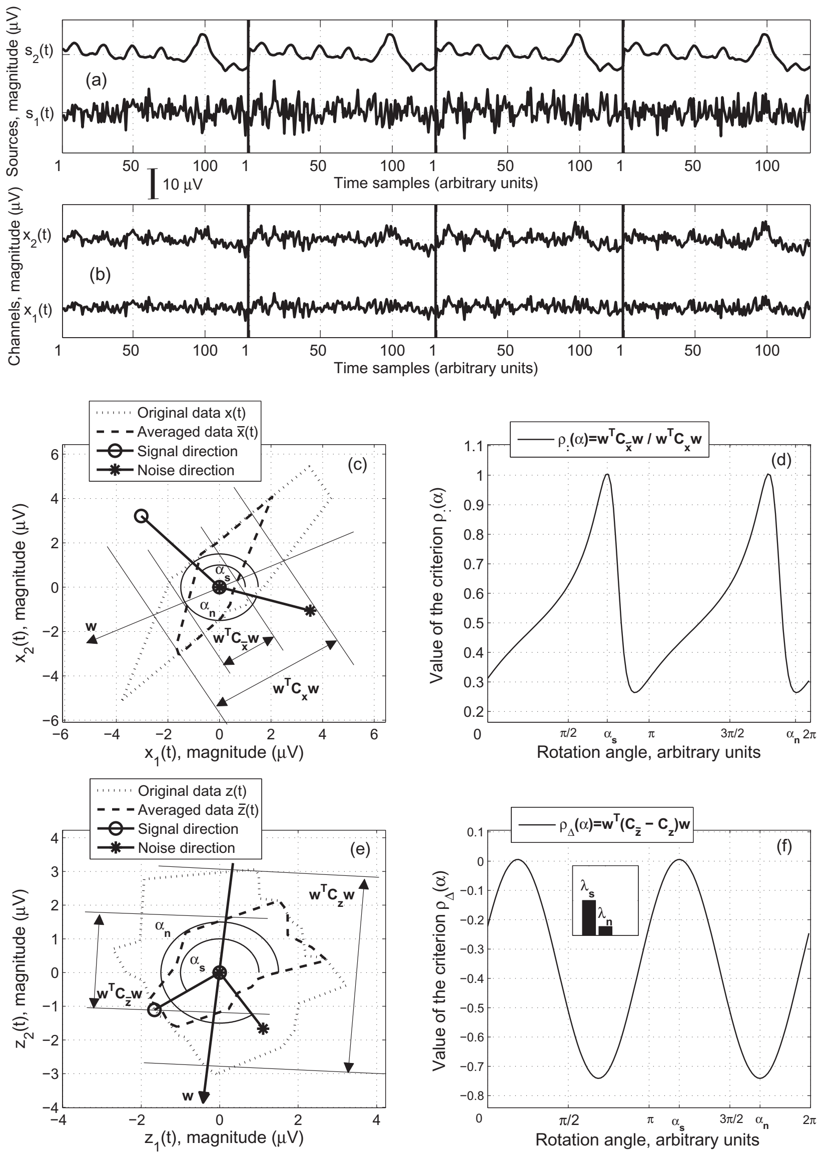

FIGURE 10 Schematic demonstration of the determination of signal/noise subspaces in EEG data by averaging: (a) sources $\mathbf{s}_i(t) = [s_{i1}(t)\ s_{i2}(t)]^{\mathrm{T}}$, $i = 1, \ldots, 4$; (b) trials $\mathbf{x}_i(t) = [x_{i1}(t)\ x_{i2}(t)]^{\mathrm{T}}$, $\mathbf{x}_i(t) = \mathbf{A}\mathbf{s}_i(t)$; (c) contours of the original $\mathbf{x}_i(t)$ and averaged $\overline{\mathbf{x}}_4(t)$ data, signal/noise directions, and indication of the criterion $\rho_{:}(\alpha)$ along $\mathbf{w}$; (d) behavior of the criterion $\rho_{:}(\alpha)$ depending on the rotation angle α, α_s denotes the angle corresponding to the signal direction and α_n is the angle of the noise direction; (e) contours of the whitened $\mathbf{z}_i(t) = \mathbf{V}\mathbf{x}_i(t)$ and whitened averaged data $\overline{\mathbf{z}}_4(t) = \mathbf{V}\overline{\mathbf{x}}_4(t)$ with indication of the signal/noise directions and criterion $\rho_\Delta(\alpha)$; (f) values of the criterion $\rho_\Delta(\alpha)$ for different angles α corresponding to direction $\mathbf{w}$; the eigenvalues of the PCA components (for $\mathbf{C}_{\overline{\mathbf{z}}}$) corresponding to the signal and noise sources are denoted as bars.

as

$$\widehat{\mathbf{C}}_{\mathbf{x}} = \frac{1}{M(T-1)} \sum_{j=1}^{M} \sum_{t=1}^{T} \underline{\mathbf{x}}_j(t) \underline{\mathbf{x}}_j(t)^{\mathrm{T}}, \tag{44}$$

where $\underline{\mathbf{x}}_j(t), j = 1, \ldots, M$, are some data samples obtained from the original single trials, which preserve the original mixing model. Basically, the procedure of constructing samples of original data should consist of linear operations applied to K-dimensional initial trials or K-dimensional time samples of the initial trials. For example, in the simplest case these data samples can be just all single trials $\underline{\mathbf{x}}_j(t) = \mathbf{x}_j(t), j = 1, \ldots, N$, and $M = N$. Similarly the N leave-one-out subaverages $\overline{\mathbf{x}}_{(j)}(t) = \frac{1}{N-1} \sum_{i=1, i \neq j}^{N} \mathbf{x}_i(t),\ j = 1, \ldots, N$, can be used as the data samples composing the original data $\underline{\mathbf{x}}_j(t) = \overline{\mathbf{x}}_{(j)}(t), j = 1, \ldots, N$, and $M = N$.

Thus, single trials are used to construct the samples of the original data. When the construction procedure is linear, the hypothetical sources that contribute these samples are obtained from the original sources in a similar way that samples of the original data are composed from the single trials. The important thing is that these samples should contain the information that is needed to obtain the separation matrix, i.e., preserve source-discriminative features over space. Therefore, the term *original data* is related to a reference point or origin from which the modification is started and to which the modified data are compared. The data samples of the original data, in turn, play the role of building blocks to obtain the modified data.

Correspondingly, the estimate of the covariance matrix of the modified data is obtained as follows

$$\widehat{\mathbf{C}}_{\tilde{\mathbf{x}}} = \frac{1}{L(T-1)} \sum_{j=1}^{L} \sum_{t=1}^{T} \tilde{\mathbf{x}}_j(t) \tilde{\mathbf{x}}_j(t)^{\mathrm{T}}, \tag{45}$$

where L denotes the number of different pieces of the modified data. For example, if single trials were our original data and if we use the traditional averaging as the data modification, then the modified data consist of only one piece $\tilde{\mathbf{x}}_1(t) = \overline{\mathbf{x}}(t)$, and $L = 1$, resulting in covariance matrix of the average $\hat{\mathbf{C}}_{\tilde{\mathbf{x}}} = \hat{\mathbf{C}}_{\overline{\mathbf{x}}}$. Similarly one can also compute the N leave-one-out subaverages. Then the modified data consist $L = N$ pieces $\tilde{\mathbf{x}}_j(t) = \overline{\mathbf{x}}_{(j)}(t),\ j = 1, \ldots, N$. Alternatively the average of the resulting separation matrices obtained on separate runs of a subspace separation algorithm for all $j = 1, ..., N$ can be used as a solution.

If the non-stationarity of the data properties is strong, then data should be segmented to smaller time periods satisfying the approximate stationarity. Specific solution corresponding to each of such segments can then be found.

4.2 Brief summary of the articles

ERPSUB algorithm presented in article **[PI]** was the first attempt to apply the problem-specific modification-based criterion to ERP and noise sources separation. ERPSUB exploits the additive differences found in whitened data before and after the averaging. The method appeared to be closely related to the Denoising Source Separation framework proposed by Särelä and Valpola [SV05]. Namely, ERPSUB can be considered as a particular derivative of the more general framework. ERPSUB was validated on EEG data obtained in the scope of the mismatch negativity study described in [PLL95, KGJ+07, HHKL07]. The results of tests have shown that the proposed technique essentially overcomes the traditional averaging procedure with regard to the effectiveness of ERP denoising.

From the algorithmic point of view, averaging is exploited twice in ERPSUB: first for a single channel denoising and then on multichannel level, where the spatial denoising is governed by the results obtained after the averaging. Namely, the two covariance matrices before and after the averaging are compared and the difference between them contains the information about the locations of the subspaces. From the methodological point of view, the main idea of ERPSUB is to utilize the information contained in all data dimensions: channels (spatial), time (temporal), and trials (synchronized temporal).

In [**PII**] we employed the classical SNR score as the cost function for the component extraction procedure instead of its traditional utilization as a validation utility. The results of the application of the obtained denoising procedure to the 9-dimensional EEG data have shown to be equivalent to the results of ERPSUB applied to the same data. This numerical study practically proved that both methods are optimal in the sense of SNR and are interchangeable. As a consequence of this conclusion the optimization criteria of both methods can be used not only for the component extraction, but also for the component selection and validation purposes, which are important components of the inter-methods comparison framework.

The goal of the research work presented in [**PIII**] was to study the possibilities of improving the performance of the averaging technique based on the problem-specific assumptions. The formalization of SNR concept was applied to solve explicitly the weighted averaging problem resulting in an averaging technique optimal in the SNR sense. Moreover, the weighted averaging and traditional averaging techniques, which belong to the class of methods for one-channel ERP denoising, were compared with the ERPSUB method operating on multi-channel level. The results of the comparison have shown that ERPSUB provides better quality of the ERP denoising, meaning that spatial information is an effective add-on that allows for more precise separability of the signal and noise.

In [**PIV**] the alternative subtraction data modification was used to design the subspace separation criterion used by the resulting DETSRC method. Similarly to ERPSUB it operates in whitened data space and employs the second-order changes in data appearing on the additive level. The application of the method to 9-dimensional EEG data set from mismatch negativity experiment [PLL95, KGJ+07, HHKL07] has shown that DETSRC outperforms the conventional averaging technique in the sense of ERP denoising.

The main purpose of the study reported in [**PV**] was to exclude the whitening step from ERPSUB algorithm. Since the estimation error is accumulated on each step of the algorithm containing approximations, it seems reasonable to reduce the number of computational steps to increase the final accuracy of the solution. For example, the assumption that the subspaces are orthogonal after whitening that is taken as fact for the subsequent component extraction procedure, cannot be absolutely valid, because of the unavoidable errors in estimating the covariance matrix of the data. The resulting method utilizes the differences observed on the multiplicative level and shows better performance than the traditional averaging with regard to the quality of ERP denoising. The method was tested on the same 9-dimensional EEG data as ERPSUB and DETSRC.

In [**PVI**] a generalized view on the previous approaches and the utilized criteria with and without whitening was presented. This allowed to systematically connect all previously developed methods in a form of a framework for methodology representation. The high-density 128-channel EEG data from ERP experiments reported in [HLGL07, HLGL08] were used for extensive analysis of the performance of the proposed methodology by example of DETSRC method.

Moreover, in [**PVI**] the benefits of switching between temporal and frequency domains are also illustrated. Namely, considering the problem in the frequency domain results in a dual effect. First, by making summation/subtraction data modifications (see Section 4.3.1) in the time domain and analyzing how different frequency bands are affected by these procedures, it should be possible to identify in which frequency bands signal is present. Specifically, by observing how different frequency bands lose or gain energy as a result of these operations and employing the ideas similar to those used for constructing the separation criteria in the spatio-temporal domain (see Section 4.3.1) one is able to determine the dominant location of the signal in the frequency domain. This, in turn, allows the optimal parameters (cut-off frequencies) of a band-pass filter, that reduces noise in the frequency domain, to be fixed. On the other hand, by performing an appropriate filtering using the filter parameters identified on the previous stage one can easily determine which spatial directions are affected by the filter and also the level of this affectedness. Since signal and noise directions should be affected differently in the sense of losing energy, they can be discriminated. Here band-pass and stop-band filtering are equivalent data modifications in the frequency domain to the proposed summation and subtraction data modifications in the synchronized time domain. This has a straightforward connection to the Denoising Source Separation framework proposed by Särelä and Valpola [SV05], where sources are distinguished based on their different behavior with respect to denoising.

4.3 Extended overview of the results

First, in Section 4.3.1 we introduce different forms of criteria for separating signal and noise subspaces in the framework of methodology sketched in Section 4.1. Then, in Section 4.3.2 we consider practical aspects of the exploitation of the proposed criteria and their properties. In Section 4.3.3 we analyze and discuss drawbacks of the proposed criteria and present possible extensions and improvements to the basic approaches.

4.3.1 Separation criteria

We propose the two basic data modifications that can be applied to three-dimensional data:

1. Trial summation: $\mathbf{x}_+(t) = \sum_{i=1}^{N} \mathbf{x}_i(t)$ or $\overline{\mathbf{x}}_+(t) = \frac{1}{N}\mathbf{x}_+(t) = \overline{\mathbf{x}}(t)$ (optional version with averaging).

2. Trial subtraction: $\mathbf{x}_-(t) = \sum_{i=1}^{N}(-1)^i \mathbf{x}_i(t)$ or $\overline{\mathbf{x}}_-(t) = \frac{1}{N}\mathbf{x}_-(t)$ (optional version with averaging).

Let us first consider the trial summation option. Furthermore, let us compare the two arbitrary directions $\mathbf{w}_1$ and $\mathbf{w}_2$, such that the SNR of the original data projected to $\mathbf{w}_1$ is larger than the SNR of the original data projected to $\mathbf{w}_2$. It is clear that the ratio between the variances of modified and initial data along the direction $\mathbf{w}_1$ should be larger than for the direction $\mathbf{w}_2$:

$$\frac{\mathbf{w}_1^{\mathrm{T}}\mathbf{C}_{\mathbf{x}_+}\mathbf{w}_1}{\mathbf{w}_1^{\mathrm{T}}\mathbf{C}_{\mathbf{x}}\mathbf{w}_1} > \frac{\mathbf{w}_2^{\mathrm{T}}\mathbf{C}_{\mathbf{x}_+}\mathbf{w}_2}{\mathbf{w}_2^{\mathrm{T}}\mathbf{C}_{\mathbf{x}}\mathbf{w}_2}, \tag{46}$$

$$\frac{\mathbf{w}_1^{\mathrm{T}}\mathbf{C}_{\overline{\mathbf{x}}_+}\mathbf{w}_1}{\mathbf{w}_1^{\mathrm{T}}\mathbf{C}_{\mathbf{x}}\mathbf{w}_1} > \frac{\mathbf{w}_2^{\mathrm{T}}\mathbf{C}_{\overline{\mathbf{x}}_+}\mathbf{w}_2}{\mathbf{w}_2^{\mathrm{T}}\mathbf{C}_{\mathbf{x}}\mathbf{w}_2}. \tag{47}$$

This takes place, because in the direction $\mathbf{w}_1$ the proportion of the deterministic signal is larger, and, therefore, variance increase after summation is larger also (variance reduction is smaller in case of averaging).

Conversely, larger SNR is obtained in data projected to a direction where a larger ratio between the variances of modified and original data is observed. Formalization of the above-mentioned ideas leads to the following criteria:

$$\rho_1(\mathbf{w}) = \frac{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{x}_+}\mathbf{w}}{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{x}}\mathbf{w}}, \tag{48}$$

$$\rho_2(\mathbf{w}) = \frac{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\overline{\mathbf{x}}_+}\mathbf{w}}{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{x}}\mathbf{w}}, \tag{49}$$

which should be maximized to find the maximum SNR components and minimized to locate the components describing minimum SNR (see Figure 10(c) and (d) for schematic illustration of $\rho_2(\mathbf{w})$ criterion). Here the criterion $\rho_1(\mathbf{w})$ stands for the summation data modification and the criterion $\rho_2(\mathbf{w})$ is used for averaging. Term $C_.$ denotes the covariance matrix of data denoted by the subscript (see Section 4.1).

Since the purpose is to extract subspaces and not sources, only the angle between the subspaces is important to determine. The angles between the components inside subspaces do not matter and, thus, can be restricted for example to be orthogonal thereby preventing convergence of consequently estimated components to the same point. The separation of the subspaces using the criteria (48) or (49) can be done without preliminary whitening of the data similarly as described in **[PV]** by maximizing/minimizing them with respect to $\mathbf{w}$ subject to the orthogonality and unit norm constraints of the resulting components.

For the subtraction-like data modifications the reasonings and explanations are similar, except that they are reversed. With respect to noise, subtraction and summation have equivalent impact: both result in increase (decrease when averaged) of the noise variance. However, on the contrary to the summation type of data modification, after the subtraction-like modification the relative to the original level variance increment in data is smaller (decrement is greater for averaging options) for those directions, where a larger deterministic signal is present, since its energy vanishes after the subtraction. Therefore, larger SNR is indicated by a smaller ratio between the variances of modified and original data along a direction.

The formal expressions for the criteria related to the subtraction type of data modification, which are equivalent to the summation-like criteria (48) and (49), are as follows:

$$\rho_3(\mathbf{w}) = \frac{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{x}_-}\mathbf{w}}{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{x}}\mathbf{w}}, \tag{50}$$

$$\rho_4(\mathbf{w}) = \frac{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\overline{\mathbf{x}}_-}\mathbf{w}}{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{x}}\mathbf{w}}. \tag{51}$$

The cost function $\rho_3(\mathbf{w})$ stands for the subtraction, and the criterion $\rho_4(\mathbf{w})$ is used for averaged subtraction. For both options, the respective criteria should be minimized to estimate maximum SNR components and maximized to find components with minimum SNR.

The criteria (48)-(51) can be further simplified by whitening the original data first, and performing the summation/subtraction modifications on these whitened data. Since whitening standardizes data to similar variances along all directions, the denominator in all criteria vanishes:

$$\rho_{1\mathbf{z}}(\mathbf{w}) = \frac{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}_+}\mathbf{w}}{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}\mathbf{w}} \equiv \mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}_+}\mathbf{w}, \tag{52}$$

$$\rho_{2\mathbf{z}}(\mathbf{w}) = \frac{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\overline{\mathbf{z}}_+}\mathbf{w}}{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}\mathbf{w}} \equiv \mathbf{w}^{\mathrm{T}}\mathbf{C}_{\overline{\mathbf{z}}_+}\mathbf{w}, \tag{53}$$

$$\rho_{3\mathbf{z}}(\mathbf{w}) = \frac{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}_-}\mathbf{w}}{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}\mathbf{w}} \equiv \mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}_-}\mathbf{w}, \tag{54}$$

$$\rho_{4\mathbf{z}}(\mathbf{w}) = \frac{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\overline{\mathbf{z}}_-}\mathbf{w}}{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}\mathbf{w}} \equiv \mathbf{w}^{\mathrm{T}}\mathbf{C}_{\overline{\mathbf{z}}_-}\mathbf{w}. \tag{55}$$

The equivalence relation $\equiv$ here means that $\arg\max_{\mathbf{w}}$ or $\arg\min_{\mathbf{w}}$ operator applied to criteria on the both sides produces equal results. In particular, the equivalence relation specifies to equality, if $\|\mathbf{w}\|_2 = 1$ or the projection is orthogonal. The criteria (52) and (53) are used in ERPSUB method in **[PI]**, and criteria (54) and (55) serve as a core of the DETSRC method in **[PIV]**.

When data are whitened, we do not need to compare the covariance matrix of the modified data with that of the original data in order to locate the signal/noise components, as it is implemented in criteria (48)-(51). Instead, to solve the separation problem we analyze the absolute values of variance of modified whitened data along different directions in the search space. Indeed, since the whitened data are standardized to similar variances along all directions, the modified whitened data become self-comparable or self-sufficient in the sense that they alone contain all the necessary information for separating the subspaces. This idea is similar to one used in Denoising Source Separation framework developed by Särelä and Valpola [SV05].

Note that the maximization vs. minimization of criteria (48)-(51) subject to orthogonality and unit norm constraints of the resulting vectors, as shown in [**PV**], lead to the extraction of different components and, thus, this approach finds non-orthogonal angle between the two subspaces. The situation is, however, different with the criteria (52)-(55): maximization vs. minimization of these criteria subject to orthogonality and unit norm constraints of the resulting vectors are nothing else than PCA and Minor Component Analysis (MCA) procedures. It is well known that they produce the same sets of pair-wise orthogonal components ordered differently: by a convention the components produced by PCA are ordered by their corresponding variances or eigenvalues in decreasing order, while the components of MCA are ordered in increasing order. In this case to find signal/noise directions we can use either PCA or MCA alone with only ordering difference. Thus, subspaces extracted by PCA and MCA using the criteria (52)-(55) are orthogonal. This shows that maximal/minimal SNR subspaces become orthogonal after whitening independently of whether they overlap (for example the problem is underdetermined) or not.

One possible variant of the resampling strategy for subtraction data modification may consist in subtracting from every multidimensional trial the average of all trials. The cumulative covariance matrix of such subtracted data is defined as:

$$\mathbf{C}_{\mathbf{x}-\bar{\mathbf{x}}} = \frac{1}{N(T-1)} \sum_{i=1}^{N} \sum_{t=1}^{T} (\mathbf{x}_i(t) - \bar{\mathbf{x}}(t))(\mathbf{x}_i(t) - \bar{\mathbf{x}}(t))^{\mathrm{T}}. \tag{56}$$

Then the resulting criterion for estimating the subspaces is a special case of (50):

$$\rho_3'(\mathbf{w}) = \frac{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{x}-\bar{\mathbf{x}}}\mathbf{w}}{\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{x}}\mathbf{w}}. \tag{57}$$

This criterion can be interpreted as clearing out the data variation due to the signal components by subtraction. Therefore, the ratio between the variances of subtracted and original data in the signal subspace should be minimal compared to the noise subspace. The simplified version for whitened data analogous to (54) is expressed as:

$$\rho_{3\mathbf{z}}'(\mathbf{w}) = \mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}-\bar{\mathbf{z}}}\mathbf{w}. \tag{58}$$

Using the latter criterion, signal subspace is extracted by minimizing the sum of N mean-square errors between the trials and their average, formally written $\arg\min_{\mathbf{w}} \sum_{i=1}^{N} var\{\mathbf{w}^{\mathrm{T}}(\mathbf{z}_i(t) - \bar{\mathbf{z}}(t))\}$. This is equivalent to minimizing the integral of variances (computed over trials) at all time points, formally expressed as

$\arg\min_{\mathbf{w}} \sum_{t=1}^{T} var\{\mathbf{w}^{\mathrm{T}}\mathbf{z}_i(t)\}$. When minimized, these criteria maximize the average similarity between the trials and their average or the average similarity between trials. In this definition each trial participates in forming the average pattern, to which it is then compared to test for similarity. To avoid the self-comparison one can subtract from each trial the leave-one-out subaverage computed without this trial $\arg\min_{\mathbf{w}} \sum_{i=1}^{N} var\{\mathbf{w}^{\mathrm{T}}(\mathbf{z}_i(t) - \overline{\mathbf{z}}_{(i)}(t))\}$. A similar criterion can be derived by approaching from an observation that adding a new trial to the average of sufficiently large number of trials should not change the average significantly, if this trial is not artifactual. Accordingly the signal directions are those where new trials introduce the least changes to the corresponding leave-one-out averages on the whole. To measure the change introduced by a new trial to the average, one can use MSE between the average of all trials and the average without the trial. Therefore the signal directions can be found by minimizing the sum of N mean-square errors between the leave-one-out subaverages and the average of all trials yielding to the cost function $\arg\min_{\mathbf{w}} \sum_{i=1}^{N} var\{\mathbf{w}^{\mathrm{T}}(\overline{\mathbf{z}}_{(i)}(t) - \overline{\mathbf{z}}(t))\}$. Moreover this idea can be used in testing for trial non-uniformity or variability (individually and as a whole). In other words the obtained individual MSE values measure the quality of each trial separately, and the mean of MSE values is a measure of average uniformity or stationarity of data.

The resampling option (56) is worthy of notice, in particular, because it has an interesting connection to the criterion of ERPSUB. Namely, consider the numerator of (57):

$$
\begin{aligned}
&\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{x}-\overline{\mathbf{x}}}\mathbf{w} = \frac{1}{N(T-1)}\sum_{t=1}^{T}\sum_{i=1}^{N}(\mathbf{w}^{\mathrm{T}}\mathbf{x}_i(t) - \mathbf{w}^{\mathrm{T}}\overline{\mathbf{x}}(t))^2 = \\
&\frac{1}{N(T-1)}\sum_{t=1}^{T}\sum_{i=1}^{N}((\mathbf{w}^{\mathrm{T}}\mathbf{x}_i(t))^2 - 2\mathbf{w}^{\mathrm{T}}\mathbf{x}_i(t)\overline{\mathbf{x}}^{\mathrm{T}}(t)\mathbf{w} + (\mathbf{w}^{\mathrm{T}}\overline{\mathbf{x}}(t))^2) = \\
&\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathbf{x}}\mathbf{w} - 2\mathbf{w}^{\mathrm{T}}(\frac{1}{T-1}\sum_{t=1}^{T}(\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i(t))\overline{\mathbf{x}}^{\mathrm{T}}(t))\mathbf{w} + \mathbf{w}^{\mathrm{T}}\mathbf{C}_{\overline{\mathbf{x}}}\mathbf{w} = \\
&\mathbf{w}^{\mathrm{T}}(\mathbf{C}_{\mathbf{x}} - \mathbf{C}_{\overline{\mathbf{x}}})\mathbf{w}. \qquad (59)
\end{aligned}
$$

Comparing (57) and (49) and taking into account (59) it follows that the criterion (57) is equivalent to the negated criterion (49) in the sense of the optimal argument. Therefore, the criterion (58) should produce the same set of the components as ERPSUB using the criterion (53), but the order of the components is reversed: the components of ERPSUB with the largest eigenvalues correspond to the components obtained using (58) with the lowest eigenvalues and vice versa.

4.3.2 Properties of the optimization criteria

The important benefit of using the suggested problem-specific criteria for optimization is their equivalence with estimates of quality measures similar to SNR (see e.g. comparison of formulae (33)-(36) and (41), (49) in Section 3.4.3). This happens because both the optimization criteria and methods for estimating quality measures like SNR are based on the same assumptions. In particular this means that the developed criteria are most effective in the sense of SNR under satisfied underlying data assumptions (i.e., (1)-(2), (4)-(6), (12)). Therefore, provided that the ERPSUB and DETSRC algorithms converge globally (the solution is explicit), they also provide a solution that is optimal in the sense of SNR. If, however, the assumptions are violated to some extent, then the criteria and the methods based on them are optimal rather in the sense of the estimate used to compute a quality measure. This property allows for exploiting the proposed criteria for optimization, validation, and classification purposes. In the following discussion, we explain this in detail.

The values of the proposed criteria in the directions of the found components correspond to the estimates of the proportion of signal or noise in raw data projected to these directions. From these estimates of proportions the estimates of SNR and NSR can be derived. Thus, for example, for the ERPSUB method the eigenvalues corresponding to the components reflect the estimates of proportion of the signal in data along the estimated directions. In other words, the rate of variance reduction after averaging of N trials can be used as a quality indicator. Similarly, for the DETSRC procedure the eigenvalues represent the estimates of the proportion of the noise in the data. Again the rate of variance reduction or increase after a subtraction indicates the quality of data. Therefore, the values of the criteria can be used for assessing the quality of the found components and denoised channels. Because of the interchangeability of the component separation and validation criteria, the traditional SNR estimates used in the psychophysiology context can also be used for separating the subspaces as it was practically tested in [**PII**].

In addition, since the nature of the problem-specific criteria used for component separation consists in discriminating between the properties of the signal and noise, the same criteria can also be used on the component selection or classification stage. Namely, the values of the target criterion corresponding to the estimated components naturally suggest the ordering of the components by the grade of their membership to either signal or noise class. For example, in the ERPSUB and DETSRC methods the output components are ordered by the corresponding eigenvalues or variances. In order to divide the extracted components into the two classes, a set of components from the beginning of the ordered sequence are assigned to the first class, and the rest to the second class. The specific meanings of the first and second class (i.e., signal or noise) depend on the context. For example, in ERPSUB the components with larger eigenvalues are more signal-like. On the contrary, in DETSRC these components have stronger noise affiliation.

If the components are extracted by a method (e.g., FastICA) that is unaware about the problem-specific categories *signal* and *noise*, external criteria that have knowledge about these concepts must be applied to solve the component classification problem. In this case the components can be sorted and classified using one of the proposed criteria (48)-(55).

The optimal performance of the proposed criteria in the validation and classification tasks allows for the creation of a framework for comparing the effectiveness of different methods in the sense of solving the subspace separation problem. Specifically, to compare methods the component classification procedure based on the designed criteria is applied first. The validation of the denoised channels using the same criteria is conducted afterwards. This scheme ensures uniform and fair treatment of the compared methods.

Furthermore, the values of the criteria corresponding to the found components can be used for estimating the dimensions of ERP and noise subspaces. In an ideal case there is a clear division of the components into the two groups - one having significantly larger respective values of the target criterion than the other. Each of these two groups corresponds to either ERP or noise subspace with the actual correspondence determined by the type of criterion used. In practical cases the clearness of the division of the components into the two sets is blurred and the border between the subspaces deteriorates. This situation may appear if the subspaces overlap or high signal-and-noise cross-correlations over sources or noise-and-noise cross-correlations over trials are present. The possible reasons that can lead to overlapped subspaces are as follows: the problem is underdetermined, an inaccurate mixing model, and non-stationarity of model properties. Therefore in the case of overlapped subspaces components become mixtures of both signal and noise sources. In this case analysis of the ordered sequence of values of a criterion allows one to consider a compromise between the amount of noise rejected and the amount of the signal loss. The smaller the number of the closest ERP-like components assigned to ERP subspace is, the larger the SNR obtained in the filtered channels, and the more noise is rejected, but also the more the signal is filtered out together with noise, and vice versa. Basically one needs to choose such proportion between the numbers of components designated to ERP and noise subspaces that provides a satisfactory trade-off between the SNR in the denoised channels and the amount of the signal loss.

Moreover the clearness of the difference between the two groups of components can be used for assessing the quality of the separation. On the other hand, taking into account the point that the problem-specific method performs optimally with regard to the set assumptions this clearness also reflects the quality of data, or the extent to which the underlying assumptions (e.g., (1)-(2), (4)-(6), (12)) are satisfied/violated.

Notice that when subspaces are overlapped, the proposed methods in general perform optimally in the sense of SNR estimate, but not in the sense of ERP energy loss. In particular the components ordered by using the suggested criteria may be ordered incorrectly in the sense of the ERP energy they contain. Imagine a situation where a component has poor SNR, but contains a lot of the signal energy. The correct ordering in the sense of ERP energy can be obtained by taking into account the original variances of the extracted components before whitening. On the other hand, if we additionally assume that the cleaner signal is (the larger its SNR) along a direction the larger also is its energy, then our algorithms become optimal in both SNR and ERP energy loss aspects. The latter assumption is rather reasonable, and is supported by practical studies. For example, in **[PI]** it was demonstrated that taking the first three maximal SNR components for reprojection results in insignificant signal loss.

The main problem with the proposed criteria remains the impossibility of segregating the signal-noise and noise-noise cross-correlations over space and trials (expected to be zero by the criteria) from the pure signal and noise energies without additional assumptions. During averaging and other manipulations with data these cross-correlations may occasionally exhibit behaviors similar to those expected from either signal or noise constituents, that makes them indistinguishable from the true signal and noise components. (See Sections 3.4.2 and 4.3.3 for more detailed explanations of this issue.)

4.3.3 Suggestions for avoiding the cross-correlations

As already mentioned, the main drawback of the proposed methods is their liability to the influence of empirically observed cross-correlations between sources in signal-noise and noise-noise categories. The reason for both types of cross-correlations to appear is the finite size of the time samples. Here we consider techniques that may partially overcome the harmful effect of these cross-correlations. The following considerations assume that ERPSUB is used with its averaging data modification [**PI**]. Moreover, we limit our attention only to those cross-correlations that weaken the contrast between subspaces and, thus, mislead separation methods based on second-order statistics in the case of the determined problem and zero additive noise.

Signal-noise correlations reveal their presence when applying a cost function to find a proper linear transformation. This includes the stages of (1) whitening and (2) seeking for signal and noise directions in the multidimensional data space, that is to say after the data are modified. For example, if a mixture of averaged noise sources has nonzero second-order correlation to signal sources, the weight of false direction is increased. To reduce the influence of cross-correlations of this type, the optimality of the solution can be assessed not only using the second-order but also the higher-order statistics. In other words when the assumptions on second-order statistical properties are violated, the similarity between the components of different trials needs to be estimated on the level of higher-order statistical information also. To incorporate the higher-order statistical similarity in the search procedure, the optimization criteria must take this information into account. The straightforward solution is related to the use of the mutual information concept. Following this idea, to maximize the similarity between different trials projected onto the same direction, one can maximize the sum of the mutual information between all pairs of trials $\sum_{i \neq j} I(\mathbf{w}^{\mathrm{T}}\mathbf{x}_i(t), \mathbf{w}^{\mathrm{T}}\mathbf{x}_j(t))$ or between all trials and their average $\sum_i I(\mathbf{w}^{\mathrm{T}}\mathbf{x}_i(t), \mathbf{w}^{\mathrm{T}}\overline{\mathbf{x}}(t))$ with respect to $\mathbf{w}$. The solution of this problem requires estimation of the probability densities of the sources. In fact, minimizing the average variance of between-trial differences used by the proposed methods is equivalent to the maximization of mutual information without priors on probability density functions of the sources (nothing more specific than the Gaussian distribution is intrinsically assumed). Indeed, if the variance of differences between trials is zero along a direction, the mutual information is maximized independently of what the distributions of the sources are. Exploiting the information on the probability densities has an effect of nonlinear weighting of the data points when estimating the similarity.

Noise-noise correlations appear when data are being modified, i.e., between trials. More specifically, this second type of undesirable cross-correlations is related to positive correlations between realizations of the noise sources from different trials. Considered together these noise sources may occasionally demonstrate behavior similar to that expected from signal sources, for example preserving relatively large variance while averaged. The difference to the previous case is that the resulting average of the noise sources may not necessarily correlate to the signal. We can say also that these are artifactual correlations over trials. In contrast, the signal-noise correlations are correlations over space. One solution to improve the performance of the separation in this case may be to exploit more robust estimates of the signal, for example weighted averaging and median instead of averaging [LKS05]. These functions are, however, non-commutative with respect to a linear transformation, and, thus, it is no longer possible to obtain simple solutions in terms of eigenvectors. In addition, median operation is non-differentiable, and the derivative of the cost function using weighted averaging seems complex and time consuming, if one tries to use the formal expressions of weighting coefficients directly.

Additional problems can be related to the phenomenon of additive noise that does not fit to the mixing model as other sources do. That is to say that the additive noise components may also exhibit correlations to the signal and noise sources either over time or trials.

Let us have a closer look at an example of a possible approach for reducing the harmful effect of undesirable noise-noise cross-correlations over trials. One way to accomplish this is to transform the original data to a domain where the influence of correlations on the results of the subspace separation is less strongly pronounced. The largest and, thus, the most harmful noise-noise cross-correlations are usually met in a lower-frequency band. Large higher-frequency noise-noise correlations are less probable in practice, and, hence, they are less important to address. To implement the domain transformation approach with focus on cancelation of lower-frequency noise-noise cross-correlations, we propose to use the first-order difference of the original signals instead of signals themselves during the computation of a separation matrix. Indeed, since differentiation highlights rapid changes (cf. image segmentation), that is to say it has an effect of a high-pass filter, trends and other lower-frequency noise behavior are reduced and, hence, have less opportunities to influence the results of a subspace separation procedure incorrectly. In contrast, systematic higher-frequency behavior is emphasized with respect to determining the solution of the subspace separation problem. Note that although differentiation increases the relative weight of higher-frequency noise components, they are unlikely to be synchronized over trials, and, hence, can be discriminated by a subsequently applied component extraction procedure. Therefore, in general, when combined with a deterministic component extraction technique, differentiation emphasizes higher-frequency deterministic components. This approach can hence be potentially useful for lowering the effect of noise and increasing the weight of signal in determining the solution in the cases when ERP waveform features also higher-frequencies compared to the suppressed lower-frequency noise, as is often the case, as in the example considered below.

Formally, the first-order difference is obtained from the initial data as

$$\Delta \mathbf{x}_i(t) = \mathbf{x}_i(t) - \mathbf{x}_i(t-1). \tag{60}$$

It is clear that the mixing model of the differentiated data does not change compared to the initial data, i.e.

$$\Delta \mathbf{x}_i(t) = \mathbf{A}\Delta \mathbf{s}_i(t) + \Delta \boldsymbol{\eta}_i(t), \tag{61}$$

where $\Delta \mathbf{s}_i(t) = \mathbf{s}_i(t) - \mathbf{s}_i(t-1)$ and $\Delta \boldsymbol{\eta}_i(t) = \boldsymbol{\eta}_i(t) - \boldsymbol{\eta}_i(t-1)$. In other words, finding the separation matrix in the domain of the first-order difference also provides the separation matrix for the initial data. Referring to adjusting, for example ERPSUB to include the differentiation step, the differentiation should be done first before all other steps. Then whitening, averaging, and PCA are carried out successively in a normal way.

Consider an example that demonstrates in which situations this procedure may help to improve the performance of the basic ERPSUB algorithm and in which fashion. For the purposes of demonstration we selected one participant's data collected using Pihko's paradigm [PLL95].

We compared the ERP estimates obtained by the three different procedures: (1) traditional averaging, (2) basic ERPSUB, and (3) ERPSUB with the differentiation step (ERPSUB-DIFF). Both estimates of ERPSUB were computed by projecting the first two components to the electrode field. The respective denoised channels for each of the three methods under discussion are shown in Figure 11. One can see that the traditional average comprises a very noisy estimate of ERP waveform. On the other hand, the ERPSUB operating in the domain of first-order difference provided estimates which are closer to our expectations and potentially more reliable than the estimates by two other approaches: (1) the standard parts tend to be approximately similar for both deviants as it should be, (2) on the other hand, the deviant responses differ in a way that also fits closer to the expected results. The respective components of ERPSUB for both cases and the corresponding eigenvalue distributions are illustrated in Figure 12. By analyzing components one notices that (1) for ERPSUB-DIFF most of ERP energy is concentrated in the first component for both deviants as we expect in an ideal case, (2) for basic ERPSUB the first components for both deviants contain mostly the contribution of a large trend and insignificant ERP energy, whereas most of ERP

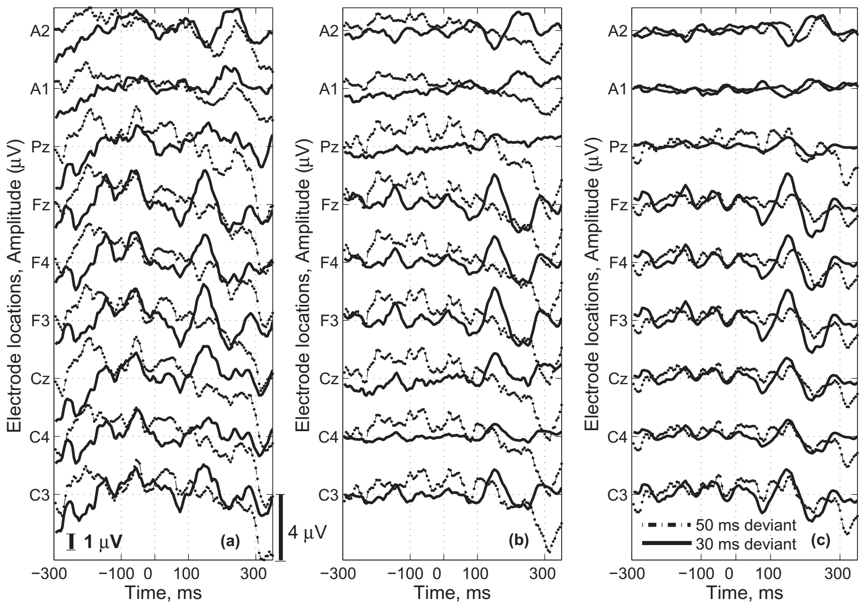

FIGURE 11 Comparison of ERP estimates obtained by (a) averaging, (b) basic ERPSUB, and (c) ERPSUB-DIFF for 50 ms deviant (dash-dotted) and 30 ms deviant (solid). One vertical tick marker is 4 μV. Negativity is up. The vertical scale bars and legend are related to all three subfigures. The data belong to one participant and were collected using Pihko's paradigm [PLL95]. The ERP estimates by ERPSUB were produced by projecting first two spatial components to the channel field. The respective components are shown in Figure 12.

energy is redistributed over the second and third components. The basic ERPSUB was able to extract a more or less clear ERP component and provide satisfactory denoising only for the 30 ms deviant, for which the raw responses should have theoretically larger SNR.

Based on this demonstration we conclude that large trial-to-trial noise correlations may get dominant priority in determining the solution. Even if the ERP component is almost correctly found, it might be ordered incorrectly, i.e. receive smaller weight compared to other noise-related components. As a result, the automatic component classification can fail. In general, large noise-noise correlations may reveal an attraction center for the search procedure and disorientate the optimization process. Therefore, the search procedure is less focused on the real signal, and ERP components extracted contain more noise.

Another simulation related to SNR estimation is demonstrated in Figure 13. We analyzed the behavior of the running SNR estimates in the averages of original and differentiated data in Fz channel. The SNR estimates were computed using the set of formulae (33)-(36) proposed by [MGT84]. The usefulness of these estimates consists not only in the factual quantitative estimation of the quality

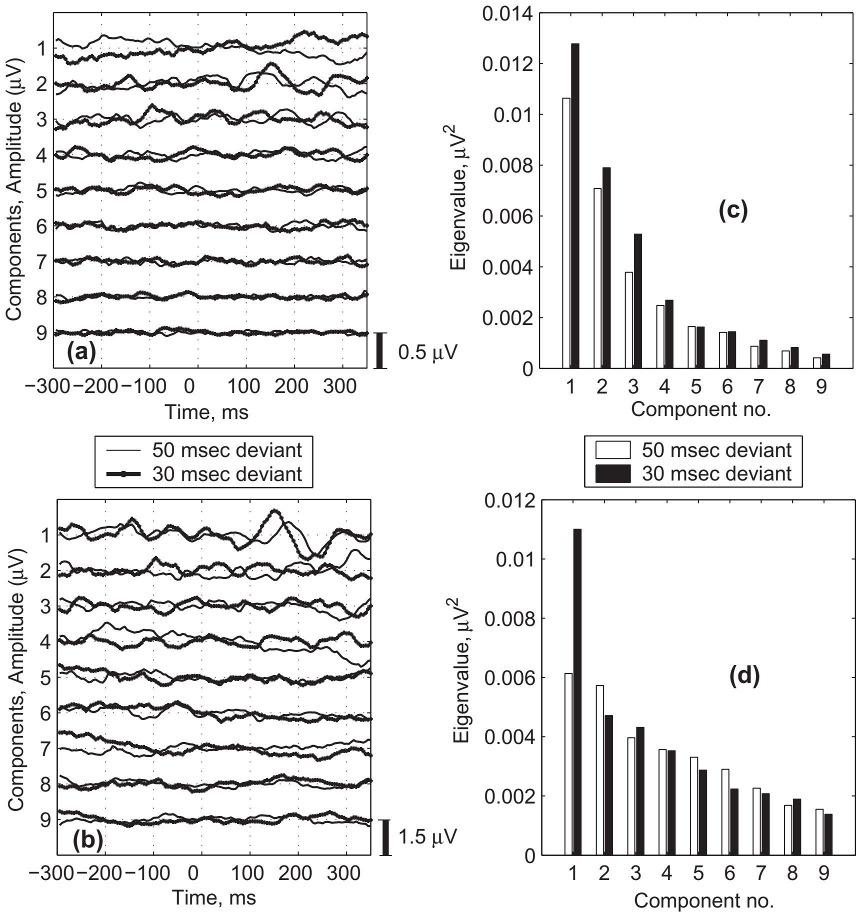

FIGURE 12 The components of (a) basic ERPSUB with the corresponding eigenvalue distribution (c) and (b) ERPSUB-DIFF with the corresponding eigenvalue distribution (d) found for 50 ms deviant (solid thin) and 30 ms deviant (solid thick). The components shown are in the original data domain for both methods, i.e., the original data are projected to component field using a separation matrix found either in original or differentiated data domain. The eigenvalues $diag(\overline{\mathbf{D}})$ are shown for the components in original (c) and differentiated (d) data domains, respectively.

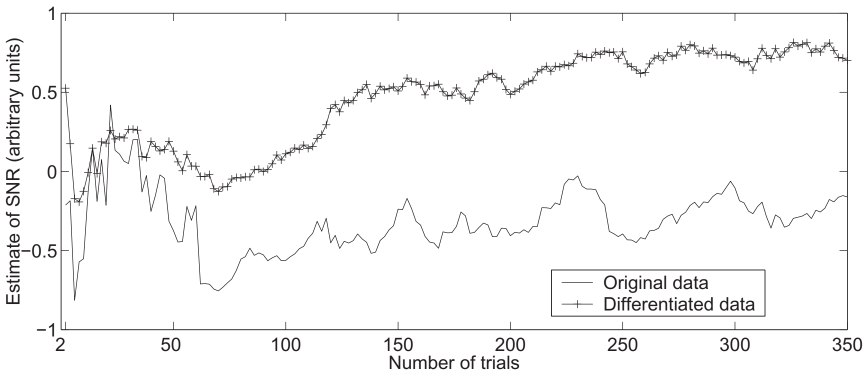

FIGURE 13 SNR in ERP estimates obtained by averaging original and differentiated data of a single participant from Fz channel and 30 msec deviant. The data were collected using Pihko's paradigm [PLL95]. The estimates of SNR were computed using the formulae (33)-(36).

of the averaged or raw data expressed in terms of SNR. Namely, analyzing the behavior of the running SNR estimate as the function of the number of trials provides an insight to the level, to which the assumptions regarding data are met, i.e., how uniform data are. In order averaging to work well, the running SNR estimate should increase linearly. The faster the running estimate of SNR in the average degenerates to a line-like behavior, the faster the underlying estimates of the signal and noise powers in raw data converge. From Figure 13 one can see that in the original data domain the SNR estimates are negative almost everywhere that means that whether the signal is underestimated or noise is overestimated. In contrast, for the differentiated data SNR estimates are correctly positive after a stabilization period in the beginning. Moreover, the stabilization period is more smooth in case of differentiated data. We may say that the whole running SNR estimate for the original data undergoes the stabilization period, because the local variations are too large and comparable to the main positive trend. In general, the running SNR estimate for the differentiated data reveals (1) a more steep positive trend that (2) faster becomes more regular, which means that (1) the SNR in the raw differentiated data is estimated higher than in the original raw data, and (2) the estimate of SNR in a single trial of differentiated data converges faster. All this indirectly supports our conjectures regarding the statement that in the domain of the first-order difference assumptions on data are met closer. This also means that more trials contribute to the increase of SNR while averaged. In other words, noise-noise correlations over trials have lesser chances to incorrectly affect the course of the optimization process and the final solution of the subspace separation problem.

We carried out also a preliminary validation of the proposed improvement using visual inspection and comparison of the results obtained for all subjects from the considered database. The results showed that the preprocessing by the first-order difference provides a systematic improvement over basic approaches. Nevertheless, the proposed approach needs to be verified more thoroughly. This is one of the directions in which we are going to develop in future research.

4.4 Author's contribution in joint publications

All major work reported in this thesis and related to (1) method development, (2) implementation, (3) analysis and interpretation of the results, and (4) documenting, has been done by the author of the thesis. The exceptions include (1) data collection and (2) conversion of the data to Matlab-readable format, which have been carried out primarily by colleagues and co-authors. The comments and suggestions of the colleagues and co-authors were useful in all stages of the preparation of the thesis. For an extended article-by-article description of the author's contribution see Appendix 1.

5 CONCLUSION

As a resume we emphasize the main achievements of the work reported in this thesis:

- A series of criteria for performing the separation of ERP components from other noise components in spatio-temporal domain have been proposed. These criteria are presented in a form of a framework reflecting the connections between different cost functions. Moreover, this framework and the history of its development highlight the ways that can be used to develop new criteria within the developed methodological and conceptual basis. The resulting methods based on the proposed criteria are (1) relatively simple, (2) fast, (3) robust with regard to different types of noise, and (4) possess a potential for further development and improvement.

- The application of the developed methods to real EEG data sets has shown significant gain in the sense of SNR in obtained ERP estimates as compared to the traditional averaging technique used as the reference. This improvement in ERP identification brings, in turn, benefits for psychophysiology research. From one side, the proposed methods provide more reliable estimates of ERP for a fixed number of trials. On the other hand, this allows the time of the experiment to be shortened, since a smaller number of trials is needed to obtain a reliable estimate of ERP waveform. The importance of shortening the time of the treatment stems from the need to keep the participant under approximately stationary conditions. Moreover some groups of probationers, like patients and infants, cannot tolerate long-lasting experiments.

- Practical aspects related to application of the developed techniques, validation, and interpretation of the results have also been considered. As a result, the obtained conclusions allowed a basis to be created for well-founded and sensible ways of application of the developed methods, as well as for elaboration of fair comparison framework for comparing performance of different methods and the quality of the results.

- A set of recommendations and guidelines for improving the reliability and robustness of the basic ERP extraction procedures was given based on practical and theoretical analysis, which discovered weak points peculiar to the discussed methods.

The second part of the conclusions is devoted to perspectives of the future work:

- Probably the main drawback of the proposed basic methods consists in the impossibility of segregating the non-zero estimates of signal-noise and noise-noise cross-correlations from the actual variance of the signal and noise components. These correlations may occasionally exhibit behavior similar to that expected from either signal or noise, and, thus, bias the results of the separation methods that may misinterpret these correlations as discriminative indicators. Therefore, further attempts must be made in order to avoid the erroneous influence of undesirable cross-correlations. One way to approach the solution of this issue is to consider the overall problem within some additional constraints or domain, where affiliation of different parts (variances and cross-correlations) of the total energy can be determined. In other words, our methods based on second-order statistics assume that these cross-correlations are zero, and we need to develop such approaches that are aware of the non-zero cross-correlations and are able to separate them from the rest of the energy of the data. This can also be viewed as a certain regularization or refinement of the problem that allows cross-correlations to be dealt with.

- Difficulties stipulated by non-zero signal-noise and noise-noise cross-correlations are also met during the validation stage. Since all SNR estimation methods assume that these correlations are zero, their presence also biases SNR estimates. Similarly to the problem discussed in the previous paragraph, to address this issue, data must be considered in some domain or equipped with additional assumptions, so that the cross-correlations can be separated from the variances of signal and noise components.

- Currently, in our implementations the component classification / selection process is carried out by taking for inverse transformation a predefined number of the first most ERP-like components determined from an ordered sequence of the eigenvalues. However, it is appealing to flexibly determine the most suitable number of ERP-affiliated components based on the quality of the separation explained by the eigenvalue distribution. This would provide a classification that is case-sensitive. To realize this, one obviously needs to construct a quantitative measure of the quality of the separation that is based on the distribution of the eigenvalues.

- Another possible advance, which we consider as being one of the potential perspective directions for further research, is the development of a fusion of the proposed ERP-tailored methods and general-purpose ICA algorithms. The motivation for this becomes clear, if we realize that (1) general-purpose

ICA techniques are able to separate sources of various kinds from each other with often satisfactory but not always optimal performance, (2) the proposed techniques, like ERPSUB, specialize in segregating the ERP from the other noise sources, but are not able to separate individual ERP subcomponents comprising ERP waveform. Therefore a combination of these two approaches, in which the advantages of both methodologies are combined for mutual benefits, would be interesting to study.

- Another perspective direction of the research is to consider other models of data than the linear instantaneous mixture model, which may be more natural and accurate for describing the considered phenomena. In particular, there is evidence that the mixture model is more close to being convolutive rather than linear (see, e.g., [ASM03]). The motivation for using convolutive models becomes apparent if we take into account the multiple reverberations of the electromagnetic waves from the skull tissue. This means that every sensor measures a sum of the delayed and attenuated versions of the same sources at each instance of time that grounds applications of convolutive models. To solve the ERP/noise separation problem for these models, the developed criteria must be properly adapted for frequency domain.

- In the future research we should consider a possibility to develop on-line versions of the proposed algorithms, which are more efficient implementations in the sense of saving computational resources.

YHTEENVETO (FINNISH SUMMARY)

Väitöskirjassa, suomenkieliseltä otsikoltaan "Herätepotentiaalien laskennallinen eristäminen EEG-havaintoaineistosta", tarkastellaan monikanavaiseen EEG (ElectroEcephaloGraphy) mittaukseen perustuvien ERP-signaalien (Event-Ralated Potential) kohinanpoistoa. Esitetty metodologia ERP-signaalin estimaatin luotettavuuden parantamiseksi perustuu kiinnostavan signaalin ja taustakohinan erotteluun eri aliavaruuksiin lineaarisen sekoittumamallin mukaisesti. Kiinnostavan signaalin sisältämän aliavaruuden projisointi takaisin alkuperäisen mittauksen elektronikentälle tuottaa luotettavamman ERP-estimaatin kuin mittausaineiston perinteinen keskiarvoistamiseen perustuva tarkastelu.

Työn keskeisenä tavoitteena oli etsiä johdonmukaisia ja luotettavia ongelmaspesifejä kriteerejä, joiden avulla kiinnostavan signaalin spatiaalinen vaihtelu ja taustakohina erotetaan toisistaan. Lisäksi työssä kehitettiin johdonmukainen ja objektiivinen vertailukehys eri ERP-laskentamenetelmien testaamiseksi. Vertailukehystutkimus liittyi kiinteästi myös aineiston luotettavuuden empiiriseen arviointiin. Kehitetyt kohinanpoistomenetelmät pystyvät harvoin saavuttamaan ideaalista lopputulosta, sillä niiden taustalla olevat perusolettamukset eivät ole sellaisenaan voimassa käytännön mittausaineistoille. Tarvitaankin lisää systemaattista ymmärrystä menetelmäkehityksen tueksi, ja tämäntyyppinen systemaattinen tarkastelu sekä suuntaviivoja mahdollisille ratkaisuille on työssä esitetty.

Kuten todellisille EEG-mittausaineistoille saadut laskennalliset tulokset osoittavat, päästään esitetyillä menetelmillä luotettavampaan lopputulokseen kuin perinteisesti psykofysiologiassa käytetyllä keskiarvoistamisella. Saaduilla tuloksilla voi olla merkittävä vaikutus kokeelliselle psykofysiologialle, sillä niiden avulla on mahdollista päästä lyhentämään mittaustilanteiden kestoa ja näin vähentämään irrelevanttien tekijöiden vaikutuksia saatuihin tuloksiin. Tällä on erityistä merkitystä lasten ja mm. neurologisista häiriöistä kärsivien testaukselle. ERP-estimaatin luotettavuuden paraneminen voi näin mahdollistaa yksittäisen koehenkilön diagnosoinnin aiempaa nopeammin. Toisaalta testiaikojen lyhentyminen mahdollistaa laajempien aineistojen keräämisen.

REFERENCES

[ACY96] S. Amari, A. Cichocki, and H. Yang. A new learning algorithm for blind signal separation. *Advances in Neural Information Processing Systems*, 8:757–763, 1996.

[ADSM06] Jörn Anemüller, Jeng-Ren Duann, Terrence J. Sejnowski, and Scott Makeig. Spatio-temporal dynamics in fMRI recordings revealed with complex independent component analysis. *Neurocomputing*, 69:1502–1512, 2006.

[Ahm98] N. U. Ahmed. *Linear and nonlinear filtering for scientists and engineers*. World Scientific Publishing Company, Singapore, 1998.

[AIF+04] B. Azzerboni, M. Ipsale, F. La Foresta, N. Mammone, and F.C. Morabito. A comparison of ICA algorithms in biomedical signal processing. In *15th Italian Workshop on Neural Nets, WIRN VIETRI 2004*, pages 313–320, Perugia, Italy, 2004.

[Alm03a] Luís B. Almeida. Faster training in nonlinear ICA using MISEP. In *4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA 2003)*, pages 313–320, Nara, Japan, 2003.

[Alm03b] Luís B. Almeida. MISEP - linear and nonlinear ICA based on mutual information. *Journal of Machine Learning Research*, 4:1297–1318, 2003.

[Ama98] S. Amari. Natural gradient works efficiently in learning. *Neural Computation*, 10:251–276, 1998.

[AMTI97] A. Ademoglu, E. Micheli-Tzanakou, and Y. Istefanopulos. Analysis of pattern reversal visual evoked potentials (prvep's) by spline wavelets. *IEEE Trans. Biomed. Eng.*, 44:881–890, 1997.

[ARS+99] Peter Anderer, Stephen Roberts, Alois Schlögl, Georg Gruber, Gerhard Klösch, Werner Herrmann, Peter Rappelsberger, Oliver Filz, Manel J. Barbanoj, Georg Dorffner, and Bernd Saletu. Artifact processing in computerized analysis of sleep EEG - a review. *Neuropsychobiology*, 40:150–157, 1999.

[ASM03] Jörn Anemüller, Terrence J. Sejnowski, and Scott Makeig. Complex independent component analysis of frequency-domain electroencephalographic data. *Neural Networks*, 16:1311–1323, 2003.

[BAMCM97] Adel Belouchrani, Karim Abed-Meraim, Jean-François Cardoso, and Eric Moulines. A blind source separation technique using second-order statistics. *IEEE Transactions on Signal Processing*, 45(2):434–444, 1997.

[Bar86] J.S. Barlow. Artifact processing (rejection and minimization) in EEG data processing. In F.H. Lopes da Silva, W. Storm van Leeuwen, and A. Remond, editors, *Handbook of Electroencephalography and Clinical Neurophysiology*, pages 15–62. Elsevier, Amsterdam, New York, 1986.

[Baş99a] E. Başar. *Brain function and oscillations. Vol. I: Principles and approaches*. Berlin: Springer Verlag, 1999.

[Baş99b] E. Başar. *Brain function and oscillations. Vol. II: Integrative brain function, neurophysiology and cognitive processes*. Berlin: Springer Verlag, 1999.

[BC00] Adel Belouchrani and Andrzej Cichocki. A robust whitening procedure in blind source separation context. *Electronics Letters*, 36:2050–2051, 2000.

[BCM+03] Colin Binnie, Ray Cooper, François Mauguière, John Osselton, Pamela Prior, and Brian Tedman. *Clinical Neurophysiology: EEG, Paediatric Neurophysiology, Special Techniques and Applications, Volume 2*. Elsevier Health Sciences, Amsterdam, London, 2003.

[BCP06] Mark F. Bear, Barry W. Connors, and Michael A. Paradiso. *Neuroscience*. Lippincott Williams & Wilkins, Philadelphia, Pennsylvania, 2006.

[BF88] Avi Blau and Miriam Furst. Time domain a posteriori optimal filter for average evoked potentials. In *IEEE Engineering in Medicine and Biology Society, 10th Annual International Conference*, volume 3, pages 1156–1157, 1988.

[BJ02] F. R. Bach and M. I. Jordan. Kernel independent component analysis. *Journal of Machine Learning Research*, 3:1–48, 2002.

[Bla90] Colin Blakemore, editor. *Vision: coding and efficiency*. Cambridge University Press, Cambridge, New York, 1990.

[BPZ+04] Giulia Barbati, Camillo Porcaro, Filippo Zappasodi, Paolo Maria Rossini, and Franca Tecchio. Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals. *Clinical Neurophysiology*, 115(5):1220–1232, 2004.

[BS91a] P. Berg and M. Scherg. Dipole models of eye activity and its application to the removal of eye artefacts from the EEG and MEG. *Clinical Physics and Physiological Measurements*, 12:49–54, 1991.

[BS91b] P. Berg and M. Scherg. Dipole models of eye movements and blinks. *Electroencephalogr Clin Neurophysiol*, 79(1):36–44, 1991.

[BS94] P. Berg and M. Scherg. A multiple source approach to the correction of eye artifacts. *Electroencephalogr Clin Neurophysiol*, 90(3):229–241, 1994.

[BS95] A. J. Bell and T. J. Sejnowski. An information-maximisation approach to blind separation and blind deconvolution. *Neural Computation*, 7(6):1129–1159, 1995.

[Buz06] György Buzsaki. *Rhythms of the brain*. Oxford University Press, Oxford, New York, USA, 2006.

[CA02] Andrzej Cichocki and Shun-Ichi Amari. *Adaptive Blind Signal and Image Processing*. John Wiley & Sons, Ltd., 2002.

[CACCR00] Sergio Cruces-Alvarez, Andrzej Cichocki, and Luis Castedo-Ribas. An iterative inversion approach to blind source separation. *IEEE Transactions on Neural Networks*, 11(6):1423–1437, 2000.

[CACRC02] Sergio Cruces-Alvarez, Luis Castedo-Ribas, and Andrzej Cichocki. Robust blind source separation algorithms using cumulants. *Neurocomputing*, 49(1):87–118, 2002.

[Cam90] Rainer Cammann. Is there no MMN in visual modality? *Behavioral and Brain Sciences*, 13(2):234–237, 1990.

[Car97] Jean-François Cardoso. Infomax and maximum likelihood for blind source separation. *IEEE Signal Processing Letters*, 4(4):112–114, 1997.

[Car98] J. F. Cardoso. Multidimensional independent component analysis. In *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'98)*, volume 4, pages 1941–1944, Seattle, WA, USA, 1998.

[Car99] Jean-François Cardoso. High-order contrasts for independent component analysis. *Neural Computation*, 11(1):157–192, 1999.

[CBF96] J.-F. Cardoso, S. Bose, and B. Friedlander. On optimal source separation based on second and fourth order cumulants. In *Proceedings of 8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing (SSAP 96)*, page 198, Corfu, Greece, 1996.

[CC00] S. Choi and Andrzej Cichocki. Blind separation of nonstationary sources in noisy mixtures. *Electronics Letters*, 36(9):848–849, 2000.

[CCC+96] S. Cerutti, G. Carrault, P.J.M. Cluitmans, A. Kinie, T. Lipping, N. Nikolaidis, I. Pitas, and M.G. Signorini. Non-linear algorithms for processing biological signals. *Computer Methods and Programs in Biomedicine*, 51(1-2):51–73, 1996.

[CGAC08] Maite Crespo-Garcia, Mercedes Atienza, and Jose L. Cantero. Muscle artifact removal from human sleep EEG by using independent component analysis. *Annals of Biomedical Engineering*, 36(3):467–475, 2008.

[CL96] J.-F. Cardoso and B. Laheld. Equivariant adaptive source separation. *IEEE Trans. on S.P.*, 45(2):434–444, 1996.

[CLN85] G. E. Chatrian, E. Lettich, and P. L. Nelson. Ten percent electrode system for topographic studies of spontaneous and evoked eeg activity. *Am. J. EEG Technol.*, 25:83–92, 1985.

[CM06] N. P. Castellanos and V. A. Makarov. Recovering EEG brain signals: artifact suppression with wavelet enhanced independent component analysis. *Journal of Neuroscience Methods*, 158(2):300–312, 2006.

[Com94] P. Comon. Independent component analysis - a new concept? *Signal Processing*, 36(3):287–314, 1994.

[CS93] J. Cardoso and A. Soloumiac. Blind beamforming for non-gaussian signals, 1993.

[CS96] Jean-François Cardoso and Antoine Souloumiac. Jacobi angles for simultaneous diagonalization. *SIAM Journal on Matrix Analysis and Applications*, 17(1):161–164, 1996.

[CT91] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, 1991.

[CTB07] John T. Cacioppo, Louis G. Tassinary, and Gary G. Berntson, editors. *Handbook of Psychophysiology*. Cambridge: Cambridge University Press, 3 edition, 2007.

[Cuf85] B. Neil Cuffin. A comparison of moving dipole inverse solutions using eeg's and meg's. *IEEE Transactions on Biomedical Engineering*, 32(11):905–910, 1985.

[CUR94] A. Cichocki, R. Unbehauen, and E. Rummert. Robust learning algorithm for blind separation of signals. *Electronics Letters*, 30(17):1386–1387, 1994.

[CV00] Andrzej Cichocki and Sergiy Vorobyov. Application of ICA for automatic noise and interference cancellation in multisensory biomedical signals. In *Proceedings of the Second International Workshop on Independent Component Analysis and Blind Signal Separation*, pages 621–626, Helsinki, Finland, 2000.

[Dau02] Jasper R. Daube, editor. *Clinical Neurophysiology*. Oxford University Press, USA, 2 edition, 2002.

[DBA04] Yannick Deville, Mohammed Benali, and Frédéric Abrard. Differential source separation for underdetermined instantaneous or convolutive mixtures: concept and algorithms. *Signal Processing*, 84:1759–1776, 2004.

[DM92] C. E. Davila and M. S. Mobin. Weighted averaging of evoked potentials. *IEEE Transactions on Biomedical Engineering*, 39(4):338–345, 1992.

[DMS01] A. Delorme, S. Makeig, and T. Sejnowski. Automatic artifact rejection for EEG data using high-order statistics and independent component analysis. In *Proceedings of the 3rd International Workshop on ICA*, pages 457–462, San Diego, 2001.

[DS00] Carlos E. Davila and Richard Srebro. Subspace averaging of steady-state visual evoked potentials. *IEEE Transactions on Biomedical Engineering*, 47(6):720–728, 2000.

[DSM07] Arnaud Delorme, Terrence Sejnowski, and Scott Makeig. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *NeuroImage*, 34:1443–1449, 2007.

[dSvLR87] F.H. Lopes da Silva, W. Storm van Leeuwen, and A. Ramond, editors. *Clinical applications of computer analysis of EEG and other neurophysiological signals*. Elsevier Science Ltd, Amsterdam, 1987.

[dWM78] J. P. C. de Weerd and W. L. J. Martens. Theory and practice of a posteriori wiener filtering of average evoked potentials. *Biological Cybernetics*, 30:81–94, 1978.

[EAWN98] Carles Escera, Kimmo Alho, István Winkler, and Risto Näätänen. Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of Cognitive Neuroscience*, 10(5):590–604, 1998.

[ED84] C. Elberling and M. Don. Quality estimation of averaged auditory brainstem responses. *Scand. Audiol.*, 13:187–197, 1984.

[ELG+00] Arndt Effern, Klaus Lehnertz, Thomas Grunwald, Guillén Fernández, Peter David, and Christian E. Elger. Time adaptive denoising of single trial event-related potentials in the wavelet domain. *Psychophysiology*, 37:859–865, 2000.

[FB91] Miriam Furst and Avi Blau. Optimal *a Posteriori* time domain filter for average evoked potentials. *IEEE Transactions On Biomedical Engineering*, 38(9):827–833, 1991.

[FDG+04] Juergen Fell, Thomas Dietl, Thomas Grunwald, Martin Kurthen, Peter Klaver, Peter Trautner, Carlo Schaller, Christian E. Elger, and Guillén Fernández. Neural bases of cognitive erps: more than phase reset. *Journal of Cognitive Neuroscience*, 16(9):1595–1604, 2004.

[FLA+06] Rain Ferenets, Tarmo Lipping, Andres Anier, Ville Jäntti, Sari Melto, and Seppo Hovilehto. Comparison of entropy and complexity measures for the assessment of depth of sedation. *IEEE Transactions on Biomedical Engineering*, 53(6):1067–1077, 2006.

[Fri87] J. H. Friedman. Exploratory projection pursuit. *Journal of the American Statistical Association*, 82(397):249–266, 1987.

[GB01] Irina F. Gorodnitsky and Adel Belouchrani. Joint cumulant and correlation based signal separation with application to EEG data analysis. In *Proceedings of 3-rd International Conference on Independent Component Analysis and Signal Separation*, pages 475–480, San Diego, California, USA, 2001.

[GCD83] G. Gratton, M. Coles, and E. Donchin. A new method for off-line removal of ocular artifact. *Electroenceph. clin. Neurophysiol.*, 55:468–484, 1983.

[GF96] M. Girolami and C. Fyfe. Negentropy and kurtosis as projection pursuit indices provide generalised ICA algorithms. In *Advances in Neural Information Processing Systems Workshop 9*, 1996.

[GF97] M. Girolami and C. Fyfe. Extraction of independent signal sources using a deflationary exploratory projection pursuit network with lateral inhibition. In *IEE Proceedings on Vision, Image and Signal Processing Journal*, volume 14, pages 299–306, 1997.

[GFF+04] Kevin A. Glass, Gwen A. Frishkoff, Robert M. Frank, Colin Davey, Joseph Dien, Allen D. Malony, and Don M. Tucker. A framework for evaluating ICA methods of artifact removal from multichannel EEG. In *ICA 2004*, pages 1033–1040, 2004.

[GL90] M. Gaeta and J.-L. Lacoume. Source separation without prior knowledge: The maximum likelihood solution. In *EUSIPO*, 1990.

[GMMV05] Antonino Greco, Nadia Mammone, Francesco Carlo Morabito, and Mario Versaci. Semi-automatic artifact rejection procedure based on kurtosis, Renyi's entropy and independent component scalp maps. In *Proceedings of International Conference on Signal Processing (ICSP 2005)*, volume 7, pages 22–26, Prague, Czech Republic, 2005.

[GMTS96] Lalit Gupta, Dennis L. Molfese, Ravi Tammana, and Panagiotis G. Simos. Nonlinear alignment and averaging for estimating the evoked potential. *IEEE Transactions on Biomedical Engineering*, 43(4):348–56, 1996.

[Han04] Todd C. Handy. *Event-Related Potentials : A Methods Handbook*. MIT Press (B&T), Camridge, Mass., 2004.

[Har90] Richard N. Harner. Singular value decomposition – a general linear model for analysis of multivariate structure in the electroencephalogram. *Brain Topography*, 3(1):43–47, 1990.

[Hay96] S. Haykin. *Adaptive filter theory*. Prentice Hall, Upper Saddle River, N.J., 3 edition, 1996.

[Hay09] Simon S. Haykin. *Neural Networks and Learning Machines*. Prentice Hall, New York, 3 edition, 2009.

[HCJ+08] J. Horváth, I. Czigler, T. Jacobsen, B. Maess, E. Schröger, and I. Winkler. MMN or no MMN: no magnitude of deviance effect on the MMN amplitude. *Psychophysiology*, 45(1):60–69, 2008.

[HG70] S. A. Hillyard and R. Galambos. Eye-movement artifact in the CNV. *Electroencephalogr Clin Neurophysiol*, 28:173–182, 1970.

[HH00] Aapo Hyvärinen and Patrik Hoyer. Emergence of phase and shift invariant features by decomposition of natural images into independent feature subspaces. *Neural Computation*, 12(7):1705–1720, 2000.

[HHI01] Aapo Hyvärinen, Patrik Hoyer, and M. Inki. Topographic independent component analysis. *Neural Computation*, 13(7):1525–1558, 2001.

[HHKL07] T. Huttunen, A. Halonen, J. Kaartinen, and H. Lyytinen. Does mismatch negativity show differences in reading disabled children as compared to normal children and children with attention deficit? *Developmental Neuropsychology*, 31(3):453–470, 2007.

[HJ86] J. Herault and C. Jutten. Space or time adaptive signal processing by neural network models. In J. S. Denker, editor, *Neural networks for computing: AIP conference proceedings*, volume 151, New York, 1986. American Institute for physics.

[HKO01] A. Hyvärinen, J. Karhunen, and E. Oja. *Independent component analysis*. John Wiley & Sons, Inc., 2001.

[HLGL07] J. Hämäläinen, P. H. T. Leppänen, T. Guttorm, and H. Lyytinen. N1 and p2 components of auditory event-related potentials in children with and without reading disabilities. *Clinical Neurophysiology*, 118:2263–2275, 2007.

[HLGL08] J. Hämäläinen, P. H. T. Leppänen, T. Guttorm, and H. Lyytinen. Event-related potentials to pitch and rise time change in children with reading disabilities and typically reading children. *Clinical Neurophysiology*, 119:100–115, 2008.

[HLN94] M. Hämäläinen, T. Lipping, and Y. Neuvo. Trends in nonlinear signal processing. *Methods of Information in Medicine*, 33(1):4–9, 1994.

[HM03] Tomas Hruby and Petr Marsalek. Event-related potentials - the p3 wave. *Acta Neurobiologiae Experimentalis*, 63:55–63, 2003.

[HO97] A. Hyvärinen and E. Oja. A fast fixed-point algorithm for independent component analysis. *Neural Computation*, 9(7):1483–1492, 1997.

[HRWL84] M. Hoke, B. Ross, R. Wickesberg, and B. Lütkenhöner. Weighted averaging - theory and application to electric response audiometry. *Electroenceph. Clin. Neurophysiol.*, 57:484–489, 1984.

[HS07] Erika Hoff and Marilyn Shatz, editors. *Blackwell Handbook of Language Development*. Blackwell Publishing, Oxford, UK, 2007.

[Hub85] P. J. Huber. Projection pursuit. *The Annals of Statistics*, 13(2):435–475, 1985.

[HYCK07] Liu Hongtao, Hu Yong, C.Q. Chang, and Luk D.K. Keith. Fast extraction of somatosensory evoked potential based on second order blind identification. In *29th Annual International Conference of the IEEE EMBS*, pages 5457–5460, Lyon, France, 2007.

[Hyv99a] A. Hyvärinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10(3):626–634, 1999.

[Hyv99b] Aapo Hyvärinen. Gaussian moments for noisy independent component analysis. *IEEE Signal Processing Letters*, 6(6):145–147, 1999.

[Hyv99c] Aapo Hyvärinen. Survey on independent component analysis. *Neural Computing Surveys*, 2:94–128, 1999.

[IZ07] Darshan Iyer and George Zouridakis. Single-trial evoked potential estimation: Comparison between independent component analysis and wavelet denoising. *Clinical Neurophysiology*, 118:495–504, 2007.

[Jas58] Herbert Henri Jasper. The ten-twenty electrode system of the international federation. report of the committee on methods of clinical examination in electroencephalography. *Electroenceph. Clin. Neurophysiol.*, 10:370–375, 1958.

[Jaz70] Andrew H. Jazwinski, editor. *Stochastic processes and filtering theory*. Academic Press, New York, 1970.

[JH91] C. Jutten and J. Herault. Blind separation of sources, part i: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, 24:1–10, 1991.

[JHL+97] Tzyy-Ping Jung, Colin Humphries, Te-Won Lee, Scott Makeig, Martin J. McKeown, Vicente Iragui, and Terrence J. Sejnowski. Extended ICA removes artifacts from electroencephalographic recordings. In *Advances in Neural Information Processing Systems 10*, pages 894–900, Denver, Colorado, United States, 1997.

[JHL+98] Tzyy-Ping Jung, Colin Humphries, Te-Won Lee, Scott Makeig, Martin J. McKeown, Vicente Iragui, and Terrence J. Sejnowski. Removing electroencephalographic artifacts: comparison between ica and pca. In *Neural Networks for Signal Processing VIII*, pages 63–72, 1998.

[JK03] Christian Jutten and Juha Karhunen. Advances in nonlinear blind source separation. In *Proceedings of the 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA 2003)*, pages 245–256, Nara, Japan, 2003.

[JMH+00] T.-P. Jung, S. Makeig, C. Humphries, T. W. Lee, M. J. McKeown, V. Iragui, and T.J. Sejnowski. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37:163–178, 2000.

[JMM+01] Tzyy-Ping Jung, Scott Makeig, Martin J. McKeown, Anthony J. Bell, Te-Won Lee, and Terrence J. Sejnowski. Imaging brain dynamics using independent component analysis. *Proceedings of the IEEE*, 89(7):1107–1122, 2001.

[JMW+01] Tzyy-Ping Jung, Scott Makeig, Marissa Westerfield, Jeanne Townsend, Eric Courchesne, and Terrence J. Sejnowski. Analysis and visualization of single-trial event-related potentials. *Human Brain Mapping*, 14:166–185, 2001.

[Jol86] I. Jolliffe. *Principal Component Analysis*. Springer, New York, 1986.

[JV99] Piotr Jaśkowski and Rolf Verleger. Amplitudes and latencies of single-trial ERP's estimated by a maximum-likelihood method. *IEEE Transactions On Biomedical Engineering*, 46(8):987–993, 1999.

[KGJ+07] I. Kalyakin, N. González, J. Joutsensalo, T. Huttunen, J. Kaartinen, and H. Lyytinen. Optimal digital filtering versus difference waves on the mismatch negativity in an uninterrupted sound paradigm. *Developmental Neuropsychology*, 31(3):429–452, 2007.

[KGK+08] Hyunseok Kook, Lalit Gupta, Srinivas Kota, Dennis L. Molfese, and Heikki Lyytinen. An offline/real-time artifact rejection strategy to improve the classification of multi-channel evoked potentials. *Pattern Recognition*, 41(6):1985–1996, 2008.

[KHV+97] Juha Karhunen, Aapo Hyvärinen, Ricardo Vigário, Jarmo Hurri, and Erkki Oja. Applications of neural blind separation to signal and image processing. In *Proceedings of the 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '97)*, volume 1, pages 131–134, Munich, Germany, 1997.

[KJ94] J. Karhunen and J. Joutsensalo. Representation and separation of signals using nonlinear PCA type learning. *Neural Networks*, 7(1):113–127, 1994.

[KJ95] J. Karhunen and J. Joutsensalo. Generalizations of principal component analysis, optimization problems, and neural networks. *Neural Networks*, 8(4):549–562, 1995.

[KJK+05] V. Krishnaveni, S. Jayaraman, P. M. Manoj Kumar, K. Shivakumar, and K. Ramadoss. Comparison of independent component analysis algorithms for removal of ocular artifacts from electroencephalogram. *Measurement Science Review*, 5(2):67–78, 2005.

[KMR06] Alexandra P. F. Key, Dennis L. Molfese, and Eva D. Ratajczak. ERP indicators of learning in adults. *Developmental neuropsychology*, 29(2):379–395, 2006.

[KMVS91] J. L. Kenemans, P. C. M. Molenaar, M. N. Verbaten, and J. L. Slangen. Removal of the ocular artifact from the eeg: a comparison of time and frequency domain methods with simulated and real data. *Psychophysiology*, 28(1):114–121, 1991.

[Kok97] A. Kok. Event-related potentials (ERP) reflections of mental resources: a rewiev and synthesis. *Biological Psychology*, 45:19–56, 1997.

[KSG04] Alexander Kraskov, Harald Stögbauer, and Peter Grassberger. Estimating mutual information. *Physical Review E*, 69(6):066138, 2004.

[KY04] Sang-Gyun Kim and Chang D. Yoo. Underdetermined independent component analysis by data generation. In *ICA 2004*, pages 445–452, 2004.

[LAC+06] Yuanqing Li, Shun-Ichi Amari, Andrzej Cichocki, Daniel W. C. Ho, and Shengli Xie. Underdetermined blind source separation based on sparse representation. *IEEE Transactions on Signal Processing*, 54(2):423–437, 2006.

[LAE+04a] H. Lyytinen, T. Ahonen, K. Eklund, T. Guttorm, P. Kulju, M.-L. Laakso, M. Leiwo, P. Leppänen, P. Lyytinen, A.-M. Poikkeus, U. Richardson, M. Torppa, and H. Viholainen. Early development of children at familial risk for dyslexia–follow-up from birth to school age. *Dyslexia: An International Journal of Research & Practice*, 10(3):146–178, 2004.

[LAE+04b] H. Lyytinen, M. Aro, K. Eklund, J. Erskine, T. Guttorm, M.-L. Laakso, P. Leppänen, P. Lyytinen, A.-M. Poikkeus, U. Richardson, and M. Torppa. The development of children at familial risk for dyslexia: birth to early school age. *Annals of Dyslexia*, 54(2):184–220, 2004.

[Łęs02] J. Łęski. Robust weighted averaging. *IEEE Transactions on Biomedical Engineering*, 49(8):796–804, 2002.

[ŁG04] J. Łęski and A. Gacek. Computationally effective algorithm for robust weighted averaging. *IEEE Transactions on Biomedical Engineering*, 51(7):1280–1284, 2004.

[LGBS00] Te-Won Lee, Mark Girolami, Anthony J. Bell, and Terrence J. Sejnowski. A unifying information-theoretic framework for independent component analysis. *International Journal Computers and Mathematics with Applications*, 39(11):1–21, 2000.

[LGH+05] Heikki Lyytinen, Tomi K. Guttorm, Tiina Huttunen, Jarmo Hämäläinen, Paavo H.T. Leppänen, and Manu Vesterinen. Psychophysiology of developmental dyslexia: a review of findings including studies of children at risk for dyslexia. *Journal of Neurolinguistics*, 18:167–195, 2005.

[LGS99] Te-Won Lee, Mark Girolami, and Terrence J. Sejnowski. Independent component analysis using an extended infomax algorithm for mixed sub-gaussian and super-gaussian sources. *Neural Computation*, 11(2):417–441, 1999.

[Lip99] T. Lipping. Nonlinear digital filtering of physiological signals. *Medical & Biological Engineering & Computing*, 37(suppl. 1):73–76, 1999.

[Lip01] Tarmo Lipping. *Processing EEG during anaesthesia and cardiac surgery with non-linear order statistics based methods*. PhD thesis, Tampere University of Technology, Pori School of Technology and Economics, 2001.

[LJ04] T. Lipping and V. Jäntti. EEG signal in monitoring brain function in anesthesia and intensive care: a review. *Proceedings of the Estonian Academy of Sciences: Engineering*, 10(2):95–109, 2004.

[LJYHH95] T. Lipping, V. Jäntti, A. Yli-Hankala, and K. Hartikainen. Adaptive segmentation of burst-suppression pattern in isoflurane and enflurane anesthesia. *International Journal of Clinical Monitoring and Computing*, 12:161–167, 1995.

[LKS05] Zbigniew Leonowicz, Juha Karvanen, and Sergei L. Shishkin. Trimmed estimators for robust averaging of event-related potentials. *Journal of Neuroscience Methods*, 142(1):17–26, 2005.

[LLGS99] T. W. Lee, M. S. Lewicki, M. Girolami, and T. J. Sejnowski. Blind source separation of more sources than mixtures using overcomplete representation. *IEEE Signal Processing Letters*, 6(4):87–90, 1999.

[LMI03] Erik G. Learned-Miller and John W. Fisher III. Independent components analysis by direct entropy minimization. Technical Report UCB/CSD-03-1221, University of California at Berkeley, 2003.

[LPBS93] O. G. Lins, T. W. Picton, P. Berg, and M. Scherg. Ocular artifacts in recording EEG and event-related potentials II: source dipoles and source components. *Brain Topography*, 6(1):65–78, 1993.

[LS97] Te-Won Lee and Terrence J. Sejnowski. Independent component analysis for sub-gaussian and super-gaussian mixtures. In *4th Joint Symposium on Neural Computation*, volume 7, pages 132–139, La Jolla, CA: University of California, San Diego, 1997.

[LS00] Michael S. Lewicki and Terrence J. Sejnowski. Learning overcomplete representations. *Neural Computation*, 12(2):337–365, 2000.

[LSB97] T. D. Lagerlund, F. W. Sharbrough, and N. E. Busacker. Spatial filtering of multichannel electroencephalographic recordings through principal component analysis by singular value decomposition. *Journal of Clinical Neurophysiology*, 14(1):73–82, 1997.

[Luc05] Steven J. Luck. *An Introduction to the Event-Related Potential Technique*. The MIT Press, Cambridge, Mass., 2005.

[Mal99] Stéphane Mallat. *A wavelet tour of signal processing*. Academic Press, 2 edition, 1999.

[Man99] Ali Mansour. The blind separation of non stationary signals by only using the second order statistics. In *Fifth International Symposium on Signal Processing and its Applications (ISSPA'99)*, pages 235–238, Brisbane, Australia, 1999.

[MBJS96] S. Makeig, A. J. Bell, T. Jung, and T. J. Sejnowski. Independent component analysis of electroencephalographic data. *Advances in neural information processing systems*, 8:145–151, 1996.

[MFBTGS03] Juan J. Murillo-Fuentes, Rafael Boloix-Tortosa, and Francisco J. González-Serrano. Initialized jacobi optimization in independent component analysis. In *4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA 2003)*, pages 1053–1058, Nara, Japan, 2003.

[MGK88] J. Möcks, T. Gasser, and W. Köhler. Basic statistical parameters of event-related potentials. *J. Psychophysiol.*, 2(1):61–70, 1988.

[MGPK87] J. Möcks, T. Gasser, D. T. Pham, and W. Köhler. Trial-to-trial variability of single potentials: Methodological concepts and results. *Int. J. Neurosci.*, 33:25–32, 1987.

[MGT84] J. Möcks, T. Gasser, and P. D. Tuan. Variability of single visual evoked potentials evaluated by two new statistical tests. *Electroencephalography and Clinical Neurophysiology*, 57:571–580, 1984.

[MGT90] Joachim Möcks, Theo Gasser, and Pham Dihn Tuan. Testing and estimation for variable single evoked brain potentials. *IEEE Engineering In Medicine And Biology Magazine*, pages 40–43, 1990.

[MIFM07] Nadia Mammone, Giuseppina Inuso, Fabio La Foresta, and Francesco Carlo Morabito. Multiresolution ICA for artifact identification from electroencephalographic recordings. In *11th International Conference KES 2007, XVII Italian Workshop on Neural Networks*, volume I, pages 680–687, Vietri sul Mare, Italy, 2007.

[MJB+97] Scott Makeig, Tzyy-Ping Jung, Anthony J. Bell, Dara Ghahremani, and Terrence J. Sejnowski. Blind separation of auditory event-related brain responses into independent components. *Proceedings of the National Academy of Sciences USA*, 94:10979–10984, 1997.

[MKGP88] J. Möcks, W. Köhler, T. Gasser, and D. T. Pham. Novel approaches to the problem of latency jitter. *Psychophysiology*, 25:217–226, 1988.

[MMB+08] Dennis L. Molfese, Victoria J. Molfese, Jennifer Beswick, Jill Jacobi-Vessels, Peter J. Molfese, Alexandra P. F. Key, and Gillian Starkey. Dynamic links between emerging cognitive skills and brain processes. *Developmental neuropsychology*, 33(6):682–706, 2008.

[Mol00] D. L. Molfese. Predicting dyslexia at 8 years of age using neonatal brain responses. *Brain and language*, 72(3):238–245, 2000.

[MP95] Jaakko Malmivuo and Robert Plonsey. *Bioelectromagnetism - Principles and Applications of Bioelectric and Biomagnetic Fields*. Oxford University Press, New York, 1995.

[MPZ99] Klaus-Robert Müller, Petra Philips, and Andreas Ziehe. JADETD: Combining higher-order statistics and temporal information for blind source separation (with noise). In *International Workshop on ICA*, pages 87–92, 1999.

[MS94] L. Molgedey and H. Schuster. Separation of independent signals using time-delayed correlations. *Physical Review Letters*, 72(23):3634–3637, 1994.

[MTG84] Joachim Möcks, Pham Dinh Tuan, and Theo Gasser. Testing for homogeneity of noisy signals evoked by repeated stimuli. *The Annals of Statistics*, 12(1):193–209, 1984.

[MVMZ04] Klaus-Robert Müller, Ricardo Vigário, Frank Meinecke, and Andreas Ziehe. Blind source separation techniques for decomposing event-related brain signals. *International Journal of Bifurcation and Chaos*, 14(2):773–791, 2004.

[MvS99] R. Mühler and H. von Specht. Sorted averaging - principle and application to auditory brainstem responses. *Scandinavian Audiology*, 28(3):145–9, 1999.

[MWT+99] S. Makeig, M. Westerfield, J. Townsend, T.-P. Jung, E. Courchesne, and T.J. Sejnowski. Functionally independent components of the early event-related potentials in a visual spatial attention task. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 354:1135–44, 1999.

[MZ93] S. G. Mallat and Z. F. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12):3397–3415, 1993.

[Nää92] Risto Näätänen. *Attention and brain function*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1992.

[NdS05] Ernst Niedermeyer and Fernando Lopes da Silva, editors. *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, Philadelphia, USA, 5 edition, 2005.

[NGM78] R. Näätänen, A. W. K. Gaillard, and S. Mäntysalo. Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42:313–329, 1978.

[NJW05] R. Näätänen, T. Jacobsen, and I. Winkler. Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. *Psychophysiology*, 42(1):25–32, 2005.

[Nol01] John Nolte. *The Human Brain: An Introduction to Its Functional Anatomy*. Mosby Inc., Orlando, Florida, U.S.A., 5 edition, 2001.

[NP87] R. Näätänen and T. W. Picton. The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, 24:375–425, 1987.

[NPR89] R. Näätänen, P. Paavilainen, and K. Reinikainen. Do event-related potentials to infrequent decrements in duration of auditory stimuli demonstrate a memory trace in man? *Neuroscience Letters*, 107:347–352, 1989.

[NPRA07] R. Näätänen, P. Paavilainen, T. Rinne, and K. Alho. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, 118(12):2544–2590, 2007.

[NW99] R. Näätänen and I. Winkler. The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, 125(6):826–859, 1999.

[ÖD96] 0. Özdamar and R. E. Delgado. Measurement of signal and noise characteristics in ongoing auditory brainstem response averaging. *Annals of Biomedical Engineering*, 24(6):702–715, 1996.

[Oja92] E. Oja. Principal components, minor components, and linear neural networks. *Neural Networks*, 5:927–935, 1992.

[OP01] R. Oostenveld and P. Praamstra. The five percent electrode system for high-resolution EEG and ERP measurements. *Clin. Neurophysiol.*, 112:713–719, 2001.

[PAO+00] T.W. Picton, C. Alain, L. Otten, W. Ritter, and A. Achim. Mismatch negativity: Different water in the same river. *Audiology and Neuro-Otology*, 5:111–139, 2000.

[Pap91] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 3 edition, 1991.

[PG97] D.-T. Pham and P. Garrat. Blind separation of mixture of independent sources through a quasi-maximum likelihood approach. *IEEE Trans. on Signal Proc.*, 45(7):1712–1725, 1997.

[PK95] J. Polich and A. Kok. Cognitive and biological determinants of p300: An integrative review. *Biological Psychology*, 41:103–146, 1995.

[PLL95] E. Pihko, T. Leppäsaari, and H. Lyytinen. Brain reacts to occasional changes in duration of elements in continues sound. *NeuroReport*, 6:1215–18, 1995.

[PMKG87] D. T. Pham, J. Möcks, W. Köhler, and T. Gasser. Variable latencies of noisy signals: Estimation and testing in brain potential data. *Biometrika*, 74(3):525–533, 1987.

[PP96] B. Pearlmutter and L. Parra. A context-sensitive generalization of ICA. In *International Conference on Neural Information Processing*, pages 151–157, 1996.

[RBH+02] Mitchell M. Rohde, Spencer L. BeMent, Jane E. Huggins, Simon P. Levine, Ramesh K. Kushwaha, and Lori A. Schuh. Quality estimation of subdurally recorded, event-related potentials based on signal-to-noise ratio. *IEEE Transactions On Biomedical Engineering*, 49(1):31–40, 2002.

[RC96] Michael D. Rugg and Michael G. H. Coles, editors. *Electrophysiology of mind: event-related brain potentials and cognition*. Oxford University Press, Oxford, U.K., 1996.

[Res01] Richard M.D. Restak. *The Secret Life of the Brain*. Joseph Henry Press, Washington, DC, 2001.

[RHB$^+$97] M. M. Rohde, J. E. Huggins, S. L. BeMent, S. P. Levine, R. K. Kushwaha, L. A. Schuh, and D.A. Ross. Use of a running signal to noise ratio to identify voluntary event related cortical potentials. In *19th International Conference - IEEE/EMBS*, 1997.

[RTF88] Jonathan Raz, Bruce Turetsky, and George Fein. Confidence intervals for the signal-to-noise ratio when a signal embedded in noise is observed over repeated trials. *IEEE Transactions On Biomedical Engineering*, 35(8):646–649, 1988.

[Sar87] Jukka Sarvas. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Physics in Medicine and Biology*, 32(1):11–22, 1987.

[Sch67] H. Schimmel. The ($\pm$)reference: accuracy of estimated mean components in average response studies. *Science*, 157:92–94, 1967.

[Sch96] E. Schröger. A neural mechanism for involuntary attention shifts to changes in auditory stimulation. *Journal of Cognitive Neuroscience*, 8:527–539, 1996.

[SJ05] John M. Stern and Jerome Engel Jr. *Atlas of EEG Patterns*. Lippincott Williams & Wilkins, Philadelphia, 2005.

[SRC74] H. Schimmel, I. Rapin, and M. M. Cohen. Improving evoked response audiometry with special reference to the use of machine scoring. *Audiology*, 13:33–65, 1974.

[SS08] Louis A. Schmidt and Sidney J. Segalowitz, editors. *Developmental psychophysiology: theory, systems, and methods*. Cambridge University Press, New York, USA, 2008.

[SV05] Jaakko Särelä and Harri Valpola. Denoising source separation. *Journal of Machine Learning Research*, 6:233–272, 2005.

[TDH07] Johan Thomas, Yannick Deville, and Shahram Hosseini. Differential fast fixed-point algorithms for underdetermined instantaneous and convolutive partial blind source separation. *IEEE Transactions on Signal Processing*, 55(7):434–444, 2007.

[The04] F.J. Theis. Uniqueness of complex and multidimensional independent component analysis. *Signal Processing*, 84:951–956, 2004.

[The06] Fabian J. Theis. Towards a general independent subspace analysis. In *Proceedings of Neural Information Processing Systems (NIPS 2006)*, 2006.

[TPZC00] Akaysha C. Tang, Barak A. Pearlmutter, Michael Zibulevsky, and Scott A. Carter. Blind source separation of multichannel neuromagnetic responses. *Neurocomputing*, 32(33):1115–1120, 2000.

[TSHL90] L. Tong, V. C. Soon, Y. F. Huang, and R. Liu. AMUSE: A new blind identification algorithm. In *ISCAS*, pages 1784–1787, New Orleans, LA, 1990.

[TT09] Shanbao Tong and Nitish V. Thankor, editors. *Quantitative EEG analysis methods and clinical applications*. Artech House Publishers, 2009.

[TXrYCH93] N. V. Thakor, G. Xin-rong, S. Yi-Chun, and D. F. Hanley. Multiresolution wavelet analysis of evoked potentials. *IEEE Trans. Biomed. Eng.*, 40:1085–1094, 1993.

[VC02] Sergiy Vorobyov and Andrzej Cichocki. Blind noise reduction for multisensory signals using ICA and subspace filtering, with application to EEG analysis. *Biological Cybernetics*, 86(4):293–303, 2002.

[vdV00] Maarten van de Velde. *Signal validation in electroencephalography research*. PhD thesis, Eindhoven University of Technology, 2000.

[VGM82] R. Verleger, T. Gasser, and J. Möcks. Correction of EOG artifacts in event-related potentials of EEG: Aspects of reliability and validity. *Psychophysiology*, 19(4):472–480, 1982.

[VHO96] Ricardo Vigário, Aapo Hyvärinen, and Erkki Oja. ICA fixed-point algorithm in extraction of artifacts from EEG. In *Proceedings of 1996 IEEE Nordic Signal Processing Symposium*, pages 383–386, Espoo, Finland, 1996.

[Vig97] R. Vigário. Extraction of ocular artifacts from EEG using independent component analysis. *Electroencephalography and Clinical Neurophysiology*, 103:395–404, 1997.

[VO99] Ricardo Vigário and Erkki Oja. Independent component analysis of human brain waves. In *International Work-Conference on Artificial and Natural Neural Networks (IWANN'99)*, volume II, pages 238–247, Alicante, Spain, 1999.

[VSJ+00] Ricardo Vigário, Jaakko Särelä, Veikko Jousmäki, Matti Hämäläinen, and Erkki Oja. Independent component approach to the analysis of EEG and MEG recordings. *IEEE Transactions On Biomedical Engineering*, 47(5):589–593, 2000.

[Wal69] D. O. Walter. *A posteriori* wiener filtering of average evoked responses. *Electroencephalography and Clinical Neurophysiology*, 27:61–70, 1969.

[WC98] I. Winkler and I. Czigler. Mismatch negativity: deviance detection or the maintenance of the 'standard'. *NeuroReport*, 9(17):3809–3813, 1998.

[Win07] István Winkler. Interpreting the mismatch negativity (MMN). *Journal of Psychophysiology*, 21(3-4):147–163, 2007.

[WLM78] J. L. Whitton, F. Lue, and H. A. Moldofsky. Spectral method for removing eye-movement artifacts rom the EEG. *Electroencephalogr Clin Neurophysiol*, 44:735–741, 1978.

[Woo67] C. D. Woody. Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. *Med. biol. Eng.*, 5:539–553, 1967.

[WS95] I. Winkler and E. Schröger. Neural representation for the temporal structure of sound patterns. *NeuroReport*, 6(4):690–694, 1995.

[WSST04] Yan Wang, Matthew T. Sutherland, Lori L. Sanfratello, and Akaysha C. Tang. Single-trial classification of ERPs using second-order blind identification (SOBI). In *Third International Conference on Machine Leaming and Cybernetics*, pages 4246–4251, Shanghai, China, 2004.

[WVS83] J. C. Woestenburg, M. N. Verbaten, and J. L. Slangen. The removal of the eye-movement artifact from the EEG by regression analysis in the frequency domain. *Biological Psychology*, 16:127–147, 1983.

[YAC98] Howard Hua Yang, S. Amari, and Andrzej Cichocki. Information-theoretic approach to blind separation of sources in nonlinear mixture. *Signal Processing*, 64(3):291–300, 1998.

[You90] Robert M. Young. *Mind, Brain and Adaptation in the Nineteenth Century: Cerebral Localization and Its Biological Context from Gall to Ferrier*. Oxford University Press, USA, 1990.

[Zek78] S. M. Zeki. Functional specialisation in the visual cortex of the rhesus monkey. *Nature*, 274:423–428, 1978.

[ZG04] W. Zhou and J. Gotman. Removal of EMG and ECG artifacts from EEG based on wavelet transform and ICA. In *26th Annual International Conference of the IEEE EMBS (EMBC)*, pages 392–395, 2004.

[ZM98] A. Ziehe and K. Müller. TDSEP – an efficient algorithm for blind separation using time structure. In *ICANN'98*, pages 675–680, 1998.

APPENDIX 1 AUTHOR'S CONTRIBUTION IN JOINT PUBLICATIONS: AN EXTENDED DESCRIPTION

Article PI - The author of the thesis is the principal author of this article. All major ideas related to elaboration of the method were proposed by the first author, whereas the co-authors made an important contribution to the preparation of the article during discussions and at the stage of scientific writing.

Article PII - All major work related to the preparation of the article including method development, implementation, testing, and scientific writing was conducted by the first author. At all stages of the development of the article the first author received important suggestions and recommendations from the co-authors.

Article PIII - While the contribution of the first author was dominant with respect to method development, implementation, validation, and scientific reporting, the comments and suggestion of the co-authors contributed significantly to the preparation of the article.

Article PIV - The article was written primarily by the first author in close cooperation with the other authors. At all stages of the preparation of the article, from the proposal of the idea and method development to the implementation and scientific reporting, the co-authors provided important comments.

Article PV - The contribution of the first author was dominant in all aspects of the preparation of the article except the aforementioned data collection and preprocessing.

Article PVI - The article has been written by the first author in close collaboration with the other authors. The design of the statistical analysis was proposed by Igor Kalyakin. Moreover, Section 2.2 of the article related to experimental data description has been primarily provided and supervised by the co-authors from the Department of Psychology.