Aus dem Bereich Systemische Neurowissenschaften und Neurotechnologie,
Theoretische Medizin und Biowissenschaften
der Medizinischen Fakultät der Universität des Saarlandes
Homburg/Saar

EFFICIENT ACQUISITION AND DENOISING OF FULL–RANGE EVENT–RELATED POTENTIALS FOLLOWING TRANSIENT STIMULATION OF THE AUDITORY PATHWAY

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
der medizinischen Fakultät
der Universität des Saarlandes

2019

vorgelegt von

Manuel Christoph Kohl, M.Sc.
geboren am 29. Mai 1989 in Saarbrücken

Datum der Promotion : 17. Juli 2020

Dekan : Prof. Dr. med. Michael D. Menger

Referenten : Prof. Dr. rer. nat. Dr. rer. med. habil. Daniel J. Strauss,

Prof. Dr. med. Karsten Schwerdtfeger

„To truly explore is to exhaust countless possibilities.
For an inch of progress, we're willing to walk miles."

– LEI JUN –

Dedicated to those who make a difference.

Zusammenfassung

Diese Arbeit steht im Zusammenhang mit aktuellen Entwicklungen auf dem Gebiet der ereigniskorrelierten Potentiale (EKP) des humanen auditorischen Systems, insbesondere der schnellen, entfaltungsbasierten EKP-Aufzeichnung sowie einzelantwortbasierten Vorverarbeitungs–, Entrauschungs– und nachgelagerten Analysemethoden. Ziel ist die Bereitstellung eines vollständigen Methodensatzes, der eine schnelle, zuverlässige Erfassung der gesamten elektrophysiologischen Aktivität entlang der Hörbahn vom Hirnstamm bis zum Cortex ermöglicht, die als Folge transienter akustischer Stimulation auftritt. Das vorliegende Manuskript gliedert sich in drei aufeinander aufbauende Untersuchungsbereiche :

Zunächst wird die generelle Machbarkeit der gleichzeitigen Aufzeichnung von Einzelantworten der auditorischen Hirnstammpotentiale zusammen mit mittelspäten und späten EKP anhand von Referenzmessungen an 15 normalhörenden Probanden demonstriert. Es werden hierzu geeignete Erfassungsparameter (Abtastrate, Bandpassfiltereinstellungen und Interstimulusintervalle) ermittelt, gefolgt von einer Signalanalyse der resultierenden EKP im Hinblick auf deren dominante intrinsische Skalen, um auf dieser Grundlage die Eigenschaften einer optimalen Signaldarstellung mit maximal reduzierter Anzahl an Abtastpunkten zu bestimmen, die durch nichtlineare Neuabtastung auf eine logarithmische Zeitbasis realisiert wird. Hierbei wird ein Kompressionsverhältnis von 16.59 erzielt. Zeit–Skalen–Analysen der uniform und logarithmisch abgetasteten EKP–Einzelantworten zeigen, dass bei der kompressiven Neuabtastung keine relevante Information verloren geht, was durch eine vergleichende Auswertung der resultierenden, gemittelten Wellenformen zusätzlich gestützt wird – alle prominenten Wellen bleiben sichtbar und sind hinsichtlich ihrer charakteristischen Latenzen und Amplituden von der Neuabtastung weitgehend unbeeinflusst. Die uniforme und logarithmische Signalrepräsentation werden hinsichtlich ihrer Anfälligkeit für die üblicherweise bei der EKP–Aufzeichnung auftretenden physiologischen und technischen Störquellen vergleichend untersucht.

Obwohl bereits eine Fülle von gut etablierten Ansätzen für die Entrauschung von EKP–Einzelantwortdarstellungen zur Verbesserung der Signalqualität und/oder zur Reduktion der benötigten Erfassungszeiten existiert, erfordern die wesentlich veränderten Störeigenschaften der vorliegenden, logarithmisch abgetasteten Einzelantwortdarstellungen im Gegensatz zu ihrem uniformen Äquivalent eine Neubewertung der verfügbaren Methoden für diese Art von Daten. Darüber hinaus werden zwei neuartige, effiziente Entrauschungsalgorithmen geboten, die auf der Koeffizientenmanipulation einer Sinogramm–Repräsentation bzw. einer analytischen, diskreten Wavelet–Zerlegung der Einzelantworten basieren und gemeinsam mit zwei etablierten Entrauschungsmethoden einer vergleichenden Leistungsbewertung unterzogen werden. Um einen umfassenden Vergleich zu ermöglichen, werden der im ersten Teil dieser Arbeit erhaltene EKP–Messdatensatz sowie synthetischen Daten eingesetzt, die mithilfe eines phänomenologischen EKP-Modells bei verschiedenen Signal–Rausch–Abständen (SRA) erzeugt wurden, wobei die individuellen Anstiege in mehreren Zielmetriken zur objektiven Bewertung der Performanz herangezogen werden. Die erhaltenen Ergebnisse deuten darauf hin, dass die vorgeschlagenen Entrauschungsalgorithmen die etablierten Methoden sowohl in den eingesetzten Zielmetriken als auch mit Blick auf die Laufzeiten deutlich übertreffen.

Weiterhin wird ein effizientes Reizsequenzoptimierungsverfahren für den Einsatz mit entfaltungsbasierten EKP–Aufzeichnungsmethoden vorgestellt, das eine konsistente Rauschunterdrückung innerhalb eines breiten Frequenzbands erreicht. Ein neuartiges Stimulus–Präsentationsparadigma für die schnelle, verschachtelte Erfassung auditorischer Hirnstammpotentiale, mittlelspäter und später Antworten durch alternierende Darbietung von optimierten, dichter Stimulussequenzen und nachgelagerter, langsamer Einzelstimulation wird eingeführt und in 20 normalhörenden Probanden evaluiert. Entfaltete Sequenzantworten, die frühe und mittlere EKP enthalten, werden mit den nachfolgenden späten Antworten fusioniert, wobei eine Zeit–Frequenz-aufgelöste, gewichtete Mittelung unter Berücksichtigung von Regularität über Einzelantworten hinweg zum Einsatz kommt. Diese erreicht einheitliche SRA der resultierenden EKP–Signale über alle untersuchten Zeitskalen hinweg. Die erhaltenen, gemittelten EKP–Wellenformen weisen Morphologien auf, die sowohl mit einschlägigen Literaturwerten als auch mit den im ersten Teil dieses Manuskripts erhaltenen Referenzaufnahmen konsistent sind, wobei alle markanten Wellen deutlich in den Gesamtmittelwerten sichtbar sind. Das neuartige Stimulationsparadigma verkürzt die Erfassungszeit um den Faktor 3.4 und vergrößert gleichzeitig den erreichten SRA erheblich. Die Ergebnisse deuten darauf hin, dass die vorgeschlagene verschachtelte Stimuluspräsentation und die nachgelagerte EKP–Verarbeitungsmethodik zur schnellen, zuverlässigen Extraktion neuronaler Korrelate der gesamten auditorischen Verarbeitung im Rahmen zukünftiger Studien geeignet sind.

Abstract

This body of work relates to recent advances in the field of human auditory event–related potentials (ERP), specifically the fast, deconvolution–based ERP acquisition as well as single–response based preprocessing, denoising and subsequent analysis methods. Its goal is the contribution of a cohesive set of methods facilitating the fast, reliable acquisition of the whole electrophysiological response generated by the auditory pathway from the brainstem to the cortex following transient acoustical stimulation. The present manuscript is divided into three sequential areas of investigation :

First, the general feasibility of simultaneously acquiring auditory brainstem, middle–latency and late ERP single responses is demonstrated using recordings from 15 normal hearing subjects. Favourable acquisition parameters (i.e., sampling rate, bandpass filter settings and interstimulus intervals) are established, followed by signal analysis of the resulting ERP in terms of their dominant intrinsic scales to determine the properties of an optimal signal representation with maximally reduced sample count by means of nonlinear resampling on a logarithmic timebase. This way, a compression ratio of 16.59 is achieved. Time–scale analysis of the linear–time and logarithmic–time ERP single responses is employed to demonstrate that no important information is lost during compressive resampling, which is additionally supported by a comparative evaluation of the resulting average waveforms – here, all prominent waves remain visible, with their characteristic latencies and amplitudes remaining essentially unaffected by the resampling process. The linear–time and resampled logarithmic–time signal representations are comparatively investigated regarding their susceptibility to the types of physiological and technical noise frequently contaminating ERP recordings.

While in principle there already exists a plethora of well–investigated approaches towards the denoising of ERP single–response representations to improve signal quality and/or reduce necessary aquisition times, the substantially altered noise characteristics of the obtained, resampled logarithmic–time single response representations as opposed to their linear–time equivalent necessitates a reevaluation of the available methods on this type of data. Additionally, two novel, efficient denoising algorithms based on transform coefficient manipulation in the sinogram domain and on an analytic, discrete wavelet filterbank are proposed and subjected to a comparative performance evaluation together with two established denoising methods. To facilitate a thorough comparison, the real–world

ERP dataset obtained in the first part of this work is employed alongside synthetic data generated using a phenomenological ERP model evaluated at different signal–to–noise ratios (SNR), with individual gains in multiple outcome metrics being used to objectively assess algorithm performances. Results suggest the proposed denoising algorithms to substantially outperform the state–of–the–art methods in terms of the employed outcome metrics as well as their respective processing times.

Furthermore, an efficient stimulus sequence optimization method for use with deconvolution–based ERP acquisition methods is introduced, which achieves consistent noise attenuation within a broad designated frequency range. A novel stimulus presentation paradigm for the fast, interleaved acquisition of auditory brainstem, middle–latency and late responses featuring alternating periods of optimized, high–rate deconvolution sequences and subsequent low–rate stimulation is proposed and investigated in 20 normal hearing subjects. Deconvolved sequence responses containing early and middle–latency ERP components are fused with subsequent late responses using a time–frequency resolved weighted averaging method based on cross–trial regularity, yielding a uniform SNR of the full–range auditory ERP across investigated timescales. Obtained average ERP waveforms exhibit morphologies consistent with both literature values and the reference recordings obtained in the first part of this manuscript, with all prominent waves being visible in the grand average waveforms. The novel stimulation approach cuts acquisition time by a factor of 3.4 while at the same time yielding a substantial gain in the SNR of obtained ERP data. Results suggest the proposed interleaved stimulus presentation and associated postprocessing methodology to be suitable for the fast, reliable extraction of full–range neural correlates of auditory processing in future studies.

Parts of the contents of the present dissertation manuscript have already been published in Kohl and Strauss (2016), Kohl *et al.* (2019a) and Kohl *et al.* (2019b).

Keywords :

Human Auditory Event–Related Potentials (ERP), Auditory Full–Range Response (AFRR), Nonlinear ERP Resampling, ERP Single Response Denoising, Fast Deconvolution–Based ERP Acquisition

List of Symbols

$\mathbb{N}$: The set of natural numbers.

$\mathbb{Z}$: The set of integer numbers.

$\mathbb{R}$: The body of real numbers.

$\mathbb{C}$: The body of complex numbers.

$\mathbb{X}_+$: The strictly positive subset of a body or set $\mathbb{X}$.

$\mathscr{X}$: Denotation of arbitrary sets.

L^p : The Lebesgue space of p–integrable functions.

ℓ^p : The Lebesgue space of p–summable sequences.

$\mathbb{X}^n$: The space of $n-$dimensional vectors x over $\mathbb{X}$.

$\mathbb{X}^{m\times n}$: The space of $m\times n-$dimensional matrices $\boldsymbol{X}$ over $\mathbb{X}$.

$||x||_p$: The p–norm of a sequence or vector x.

$||\boldsymbol{X}||_F$: The Frobemius norm of a matrix $\boldsymbol{X}$.

$\overline{x}$: The arithmetic mean value of a sequence or vector x.

$\varnothing\overline{x} \pm \sigma_x$: Compound denotation of mean and standard deviation.

$\propto$: Denotation of a proportional relationship.

δ : The Dirac unit impulse.

Δ_T : The Dirac comb with period T.

$\top$: Matrix transposition.

$\cdot$: The matrix product.

$\langle \cdot, \cdot \rangle$: The inner product.

$(\,\cdot\,)$: The continuous index operator.

$[\,\cdot\,]$: The discrete index operator.

$\lceil \cdot \rceil$: The ceiling function.

$\lfloor \cdot \rfloor$: The floor function.

$|\cdot|$: The absolute value function.

$*$: The convolution operator / complex conjugate (superscript).

$\mathfrak{F}\{\cdot\}$: The Fourier transform operator.

$\mathfrak{M}\{\cdot\}$: The Mellin transform operator.

$\mathfrak{S}\{\cdot\}$: The scale transform operator.

$\mathfrak{R}\{\cdot\}$: The Radon transform operator.

$\mathfrak{W}_{\Psi}\{\cdot\}$: The continuous wavelet transform operator with mother wavelet Ψ.

$\mathfrak{X}^{-1}$: The inverse of operator $\mathfrak{X}$.

List of Abbreviations

ABR Auditory Brainstem Response. 19, 23–26, 28, 29, 32, 34, 49, 50, 54, 56–62, 79, 82–85, 87, 88, 93, 96, 101–103

AC Auditory Cortex. 11, 14–16, 19, 21

ADJAR Adjacent Response Removal. 29, 49, 55

AFRR Auditory Full–Range Response. 24, 26, 27, 29, 31–36, 38, 42, 44, 45, 47–50, 57–66, 68–80, 82–91, 93–97, 103, 104

ALR Auditory Late Response. 19, 21, 22, 24–26, 29, 34, 49, 50, 54–56, 58–60, 62, 79, 93, 96, 101–103

AMLR Auditory Middle–Latency Response. 19, 21, 24–26, 34, 49, 50, 54, 56, 58–62, 79, 82–85, 93, 96, 101–103

AN Auditory Nerve. 11, 13, 19, 97

AP Auditory Pathway. 11, 13–17, 19, 21, 22, 96, 97

CLAD Continuous Loop Averaging Deconvolution. 50, 51, 56, 96

CLD Continuous Loop Deconvolution. 51, 53, 54, 56, 58–62, 79, 81–85, 92, 93, 96, 101

CN Cochlear Nucleus. 11, 13, 14, 17, 19

CWI Correlation Waveform Index. 47, 48, 91

CWT Continuous Wavelet Transform. 31, 33, 36, 38, 63, 64, 66, 87, 88, 102

DE Differential Evolution. 50, 53, 54, 92, 97

DWT Discrete Wavelet Transform. 41, 45

EEG Electroencephalogram. 17, 18, 33, 36, 39, 40, 51, 54, 55, 59, 64, 67, 68, 88–90

ERP Event–Related Potential. 17–19, 21–32, 34–36, 39–56, 58–66, 68–79, 82–85, 87–97, 101–104

FDE Frequency–Domain Error. 53, 92

FFR Frequency–Following Response. 25, 55, 96

FFT Fast Fourier Transform. 33, 42, 50, 58, 91

FIR Finite Impulse Response. 51

FT Fourier Transform. 35, 97

IC Inferior Colliculus. 14, 17, 19

IIR Infinite Impulse Response. 30, 58

ISI Interstimulus Interval. 24, 26, 27, 29, 39, 49, 52–57, 79, 81, 88, 92, 93, 95–97, 101

JR Jitter Ratio. 53, 92

LL Lateral Lemniscus. 14, 19

LT Laplace Transform. 34, 52

LTI Linear Time–Invariant. 51, 97

MGN Medial Geniculate Nucleus. 14, 15, 17

MT Mellin Transform. 34, 42

NAF Noise Amplification Factor. 54, 55, 79, 80, 92

NLD Nonlinear Diffusion. 41

NLM Nonlocal Means. 41, 42, 47, 72, 91, 95

PSD Power Spectral Density. 55, 79, 80

RANDWT Rational Analytic Discrete Wavelet Transform. 45, 46, 59–61, 79, 82

RMS Root Mean Square. 18, 39, 40, 55, 79

RN Residual Noise. 39, 40, 48, 64, 67, 68, 75, 77, 88, 89, 91

RT Radon Transform. 42–44, 90, 95

SDR Sinogram Domain Reweighting. 44, 73, 90–92, 95

SNR Signal–to–Noise Ratio. 24, 26, 29, 34, 39, 40, 47–50, 54–56, 59, 62, 68, 76, 79, 85, 90–93, 95, 96, 103

SOC Superior Olivary Complex. 13, 14, 19

SPL Sound Pressure Level. 11, 29

ST Scale Transform. 35, 64, 65, 87, 88, 97

SWAPS Split Wavelet Amplitude & Phase Smoothing. 46, 47, 74, 91, 92, 95

TV Total Variation. 41

UGM Unidirectional Gaussian Means. 40–42, 46, 47, 71, 91, 92, 95

WPS Wavelet Phase Stability. 36, 38, 64–66, 87, 88

Contents

Zusammenfassung **1**

Abstract **3**

List of Symbols **5**

List of Abbreviations **7**

Contents **11**

1 Introduction **15**

1.1 Background 15

1.1.1 The Human Auditory Pathway 15

1.1.2 Transient Auditory ERP 21

1.2 Motivation and Objectives of This Work 28

2 Materials and Methods **31**

2.1 Reference AFRR Acquisition 31

2.1.1 Preliminary Considerations 31

2.1.2 Subjects 31

2.1.3 Stimulus Presentation and ERP Recording 32

2.1.4 ERP Processing and Analysis 34

2.2 Phenomenological AFRR Model 36

2.3 Optimization of AFRR Signal Representation 38
  2.3.1 Preliminary Considerations 38
  2.3.2 Nonuniform Resampling 40
  2.3.3 Noise Characteristics 43
2.4 Denoising of AFRR Single Responses 44
  2.4.1 Preliminary Considerations 44
  2.4.2 Denoising Using the Radon Transform 47
  2.4.3 Denoising Using Analytic Wavelet Filterbanks 49
  2.4.4 Comparative Performance Evaluation 51
2.5 Fast Deconvolution–Based AFRR Acquisition 53
  2.5.1 Preliminary Considerations 53
  2.5.2 Stimulus Sequence Optimization 55
  2.5.3 Interleaved Stimulus Presentation Paradigm 59
  2.5.4 Subjects 61
  2.5.5 Stimulus Presentation and ERP Recording 61
  2.5.6 ERP Processing 62
  2.5.7 Comparative Statistical Analysis 66

3 Results 67

3.1 Reference AFRR Acquisition 67

3.2 Optimization of AFRR Signal Representation 68

3.3 Denoising of AFRR Single Responses 72

3.4 Fast Deconvolution–Based AFRR Acquisition 83

4 Discussion 91

4.1 Reference AFRR Acquisition 91

4.2 Optimization of AFRR Signal Representation 91

4.3 Denoising of AFRR Single Responses 93

4.4 Fast Deconvolution–Based AFRR Acquisition 96

5 Conclusion and Future Work 99

A Appendix 105

List of Figures 109

List of Tables 113

Bibliography 115

List of Publications 137

Acknowledgments 139

1 Introduction

1.1 Background

1.1.1 The Human Auditory Pathway

The Auditory Pathway (AP) is a neural macrostructure of the human brain, which emanates from the inner hair cells of the cochlea, located within the inner ear, and terminates with projections to the temporal regions of the cerebral cortex, i.e., the Auditory Cortex (AC). It is responsible for mediating the sensory modality of hearing and is composed out of several neural clusters (nuclei) interconnected by ascending and, albeit to a lesser extent, descending fibres (Møller, 2011). A schematic illustration of the various anatomic structures constituting the auditory pathway and their respective interconnections is given in Fig. 1.1.

Peripheral Level

Following an acoustical stimulation of the outer ear and its subsequent mechanical transduction through the middle ear into the oval window of the cochlea, the approximately 3500 inner hair cells in the organ of Corti, which is located between the *scala media* and the *scala tympani*, are selectively excited by the formation of a pressure wave traversing the cochlea from the oval window to the round window (Schmidt *et al.*, 2017). Due to a gradual increase in longitudinal stiffness of the basilar membrane, which separates the *scala vestibuli* from the other cochlear compartments, a frequency–resolved, selective neural input to the afferent fibres of the 8$^{\text{th}}$ cranial nerve, the Auditory Nerve (AN), is achieved, where single neurons are tuned to single frequencies. As a result of the cochlear geometry and the mechanical properties of the basilar membrane, acoustical frequencies audible to the average normal hearing human reside within the range of 20 Hz – 20 kHz (Møller, 2011).

The AN transmits the neural excitation patterns to the first processing stage along the AP, the Cochlear Nuclei (CN) located within the rostral medulla (Celesia and Hickok,

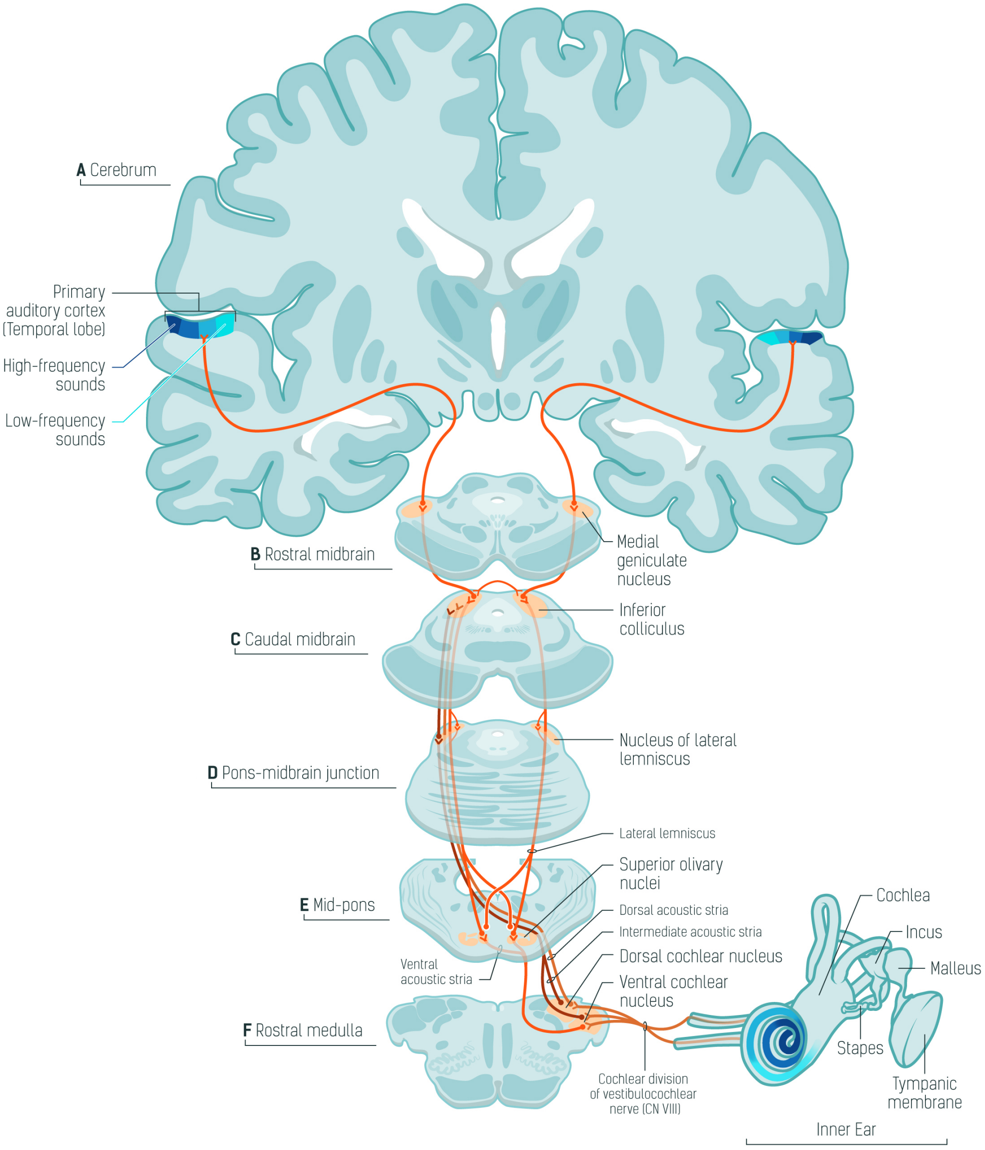
A Cerebrum
Primary auditory cortex (Temporal lobe)
High-frequency sounds
Low-frequency sounds
B Rostral midbrain
Medial geniculate nucleus
C Caudal midbrain
Inferior colliculus
D Pons-midbrain junction
Nucleus of lateral lemniscus
Lateral lemniscus
Superior olivary nuclei
E Mid-pons
Dorsal acoustic stria
Intermediate acoustic stria
Dorsal cochlear nucleus
Ventral cochlear nucleus
Ventral acoustic stria
F Rostral medulla
Cochlear division of vestibulocochlear nerve (CN VIII)
Cochlea
Incus
Malleus
Stapes
Tympanic membrane
Inner Ear

Figure 1.1: Illustration of slices through anatomic structures of the human auditory pathway. Reproduced with slight modifications from (Stanishevskaya, 2016).

2015). Within an acoustical dynamic range of 20 – 40 dB Sound Pressure Level (SPL), the neural response rates of AN fibres corresponding to the characteristic frequencies dominantly excited by the acoustic stimulus increase approximately proportional to the stimulus intensity, bottoming out at subthreshold levels. By contrast, sound intensities largely exceeding the above dynamic range will only lead to neglilible rate increases, but result in the recruitment of fibres with neighbouring characteristic frequencies as their respective thresholds are increasingly exceeded (Roederer, 2009).

Brainstem Level

As is the case with the AN, all subsequent processing stages of the ascending human AP from the brainstem level onwards consequently realize a tonotopic organization, mostly in the form of an internal spatial encoding with sharp, level–tolerant frequency tuning curves of single neurons in case of the brainstem–level nuclei (McLaughlin *et al.*, 2010). Additionally, the dorsal and ventral CN spatially separate high and low frequency content even further by predominantly processing a disproportionally large amount of the associated neural responses (Celesia and Hickok, 2015). The primary role of the CN in auditory processing is thought to reside in an enhancement of spectral component separation by means of a contrast–increasing, multilayered lateral inhibition structure (Kral and Majernik, 1996).

The CN project to the Superior Olivary Complex (SOC), which is located within the mid–pons. It consists of the medial and lateral superior olive as well as the trapezoid body, the latter two being strongly interconnected (Moore, 1987). As is the case with the CN, the lateral and medial superior olive largely separate processing of neural information corresponding to high and low frequency sound components (Biacabe *et al.*, 2001). Since the SOC nuclei are the first neural structure receiving fibres ascending from the *contralateral* auditory path (Grothe, 2000), they play a key role in the perception of binaural spatial cues, a vital prerequisite of directional hearing (i.e., sound source localization and separation). The medial superior olive is assumed to be responsible for the processing of interaural latency and phase differences present in lower–frequency sounds (below 3—4 kHz), presumably using a temporal cross–correlation structure in the form of bipolar coincidence neurons (Joris and Yin, 2007; Campbell and King, 2004), while the lateral superior olive appears to be predominantly sensitive to interaural level differences present in higher frequency sounds, integrating them by means of inhibitory contralateral projections (Rees and Palmer, 2010).

Many afferent fibres emanating from the CN and SOC join to provide input to the Lateral Lemniscus (LL) (Glendenning *et al.*, 1981), which is located within the pons–midbrain junction and divided into a dorsal and a ventral substructure. With the dorsal partition receiving input from both the ipsi– *and* contralateral auditory paths and the ventral substructure exclusively receiving contralateral input, the LL further contributes to the interlinkage of neurally encoded binaural auditory information, and projects to the Inferior Colliculus (IC) (Møller, 2011).

Midbrain Level

The IC is the largest of all brainstem– and midbrain–level auditory nuclei and is located within the caudal midbrain (Schmidt *et al.*, 2017). Receiving input from lower–level nuclei of both the ipsi– and contralateral ascending paths, it represents the terminal point of contact for the auditory brainstem, integrating information from all inferior nuclei as the vast majority of their ascending fibres synapse in it. The IC is divided into a purely auditory, central nucleus and a laterally surrounding, external subdivision composed of auditory and somatosensory fibres (Celesia and Hickok, 2015). Given their distinct tonotopic organization and rich interconnection of fibres from both ascending auditory paths as well as the lateral interconnection of the left and right central nuclei (Binns *et al.*, 1992), the IC presumably plays an important role in frequency discrimination (Rees and Palmer, 2010) as well as the perception of interaural time differences (Fitzpatrick *et al.*, 2002).

Being interlinked between the IC and the AC, the Medial Geniculate Nucleus (MGN) constitutes the primary sensory relay mechanism of the AP. As one of the thalamic nuclei, it is located within the rostral midbrain and comprises of a ventral, medial and dorsal subdivision. The ventral MGN receives input from the brainstem– and midbrain–level nuclei of the lemniscal pathway and projects directly to the primary AC, hence it forms the main ascending information route of the AP. By contrast, the medial and dorsal MGN receive and integrate multimodal sensory input and dominantly project to the nonprimary AC areas (Bizley, 2017). The roles of the MGN besides sensory gating are ranging from sound recognition and sound source localization over emotional responses to sounds to the direction of auditory attention (Rees, 2009).

Cortical Level

The AC of both cerebral hemispheres represent the terminal structures of the ascending human AP and are located on and around the superior temporal gyri (i.e., Heschl's gyri) within the *fissura sylvii* (Ramachandran, 2002). Being a part of the cerebral cortex makes them the phylogenetically newest auditory structure in the mammalian brain (Kaas, 2008). The AC are by far the largest and most complex neural systems participating in the modality of hearing and can be coarsely subdivided into three tightly interlinked regions based on their cytoarchitectonic and physiological properties as well as their dominant directions of information flow (Scott *et al.*, 2017) : Each primary AC constitutes the respective core area and mainly receives input from the ventral (and, albeit to a much lesser extent, the medial) MGN of the thalamus. It is surrounded by and strongly connected with the belt region (secondary AC), which receives additional input from the dorsal MGN and is itself laterally interconnected with the parabelt region (tertiary AC), located dorsally on the superior temporal gyrus (Kaas and Hackett, 2000).

In contrast to the inner workings of the small, rather separated and less complex nuclei that are characteristic of auditory brainstem anatomy, the intricate connectome of the thalamocortical auditory system and its manifold functional modalities, collectively termed *„auditory thalamocortical transformation"* by Winer *et al.* (2005), are much less well understood to date (Eggermont, 2010). Even more so than the preceding structures of the AP, the AC plays a key role in the fusion of binaural auditory information (Pantev *et al.*, 1996) and also features a distinct tonotopy (Pantev *et al.*, 1988), albeit more in terms of musical pitch rather than the underlying acoustical frequencies represented by the cochleotopy of the early brainstem–level nuclei. This cortical pitch representation is a veritable example of neuroplasticity as it is subject to incremental, functional reorganization due to, e.g., cochlear damage or the continued pursuit of a musical skill (Pantev *et al.*, 1998).

Findings of Näätänen *et al.* (2001) suggest a broad variety of pertinent perceptual achievements (construction of a coherent auditory scene, recognition of familiar spectrotemporal patterns, detection of novel or deviant sounds and anticipation of future auditory events) to be largely due to a primitive *„sensory intelligence"* and subsequent higher cognitive processes well localized to the AC, which organize the auditory input and arguably perform some form of invariant extraction. Furthermore, the AC has been shown to play a paramount role in both preattentive and attentive filtering of sounds by gating the important ones to awareness based on their degree of dissimilarity with the preceding auditory input (Jääskeläinen *et al.*, 2004).

While it is well–accepted since the pioneering work of Hubel *et al.* (1959) that the neural behaviour in the AC is strongly modulated by attention, the intricate interplay of preattentive and attentive processes and their respective roles in, among others, the above–mentioned perceptual achievements are only beginning to be unraveled. A notable consensus is the subdivision of attentional processes into a conscious, task–dependent and voluntary top–down variety and input–dependent, involuntary bottom–up or *„pop–out"* attention largely based on the salience of the auditory input (Fritz *et al.*, 2007). Top–down attentional processes in particular are closely linked to the functions of the corticofugal system, which is presented in the next subsection. Another well–backed work hypothesis is the presence of two distinct neural representations of auditory processing streams in the primate brain (Arnott *et al.*, 2004), the first of them projecting from the caudal part of the superior temporal gyrus to the parietal regions and primarily processing spatial cues of the auditory scene. By contrast, the second stream is thought to emanate from the anterior lateral AC belt and dominantly process information related to the patterns or *„objects"* within the auditory scene (Rauschecker and Tian, 2000).

Being a subdivision of the richly interconnected sensory cortex, the AC also integrates multimodal sensory information (predominantly from the visual modality) into the construction of coherent auditory scenes especially when presented with incomplete or noisy acoustical input (Sekiyama *et al.*, 2003), which leads to, e.g., considerable improvements in overall speech intelligibility for hearing–impaired listeners provided that they can see the lip movements of the speaker. Albeit midbrain–level structures like the superior colliculus also contribute to multisensory integration (Stein *et al.*, 2009), substantial visual modulation of auditory perception, as demonstrable via the McGurk illusion, has been shown to be accompanied by distinct neural correlates of the visual stimuli on the human AC very early into auditory processing (Smith *et al.*, 2013), hence auditory multisensory integration is probably mostly dominated by processes on the cortical level.

Descending Auditory Pathway

The descending auditory pathway or corticofugal auditory system forms multiple top–down feedback paths capable of propagating cortical–level information back to preceding processing stages of the ascending AP (Suga *et al.*, 2011), with its most peripheral effectors being the hair cells of the cochlea (Xiao and Suga, 2001). The many outputs of the corticofugal system constitute one of the largest pathways in the human brain with many auditory and non–auditory targets on the presynaptic cortical, thalamic, midbrain and brainstem levels (Winer, 2005). The descending auditory efferents reach the CN, the IC as well as the MGN (Winer and Lee, 2007). Consequently, pertinent experiments involving macroscopic electrophysiological correlates of the associated neural activity (e.g., Woldorff and Hillyard, 1991; Müller *et al.*, 2009) show a modulation of *early* auditory processing by higher–level cortical processes such as selective attention.

The influence of corticofugal efferents on the perception of auditory input is exercised by a continuous alteration of bottom–up auditory information processing, ranging from subtle reshaping of response properties for single neurons along the central AP (entailing among others a mechanism for dynamic gain control) to potentially substantial and lasting reorganization of existing tonotopic and computational neural maps (Suga *et al.*, 2000).

1.1.2 Transient Auditory ERP

Event–Related Potentials (ERP) can be defined in a very general way as the totality of measurable neural response patterns which reproducibly occur time–locked to a specific stimulus[1]. ERP obtained from the human scalp EEG have played and continue to play a paramount role in neuroscience research and clinical neurodiagnostics as large–scale neural correlates of many sensory and cognitive processing modalities (Oken and Phillips, 2009).

---

[1] Note that the conceptual delimination between the *evoked* (exogeneously dependent, i.e., stimulus–driven) and *event–related* (primarily influenced by endogeneous factors) parts of stimulus–locked neural activity (Goodin, 2012) – while potentially very important in certain contexts – is omitted for the sake of terminological simplicity within the scope of this work, as the presented signal processing methods solely consider the ERP from an epiphenomenological viewpoint.

Given that scalp EEG recordings, depending on the electrode geometries and positions, effectively yield a spatially weighted integral over the infintesimal voltage elements due to microscopic, subcutaneous current flows caused by the superposition of neural electric dipoles (predominantly those of the apical dendrites of cortical pyramidal neurons), ERP acquired from scalp EEG leads likewise enjoy an excellent temporal resolution, whereas their spatial resolution is greatly reduced (Malmivuo and Plonsey, 1995). The former property is a crucial prerequisite to properly relating these large–scale electrophysiological responses to their respective neural generators in detail, particularly when considering that, e.g., the earliest prominent wave of the auditory ERP to transient stimuli already appears within one millisecond post–stimulus (Picton *et al.*, 1974).

Since the pioneering demonstrational work of Dawson (1954), it is an established consensus that the acquisition of transient ERP has to rely on repeated measurements in some way to obtain meaningful results. This is primarily owed to the fact that the sought–after prominent waves are generally orders of magnitudes smaller than the competing background noise floor of oscillatory EEG activity, which typically resides in the range between 10 – 100 μV Root Mean Square (RMS) (Malmivuo and Plonsey, 1995). Hence, the underlying rationale of ERP acquisition is the implicit assumption of a linear signal–plus–noise superposition model, where the stimulus–locked ERP activity constitutes a deterministic signal and the spontaneous oscillatory EEG – despite being substantially larger in amplitude – merely forms an additive, stochastic fluctuation thereof (Dawson, 1951). Under this model assumption, the observational error should successively approach zero the higher the number of repeated measurements being averaged, as the stochastic signal components will tend to cancel out on average (Taylor, 1997). Note that this behaviour of course nicely extends to the generally nondeterministic physiological artifacts (e.g., sporadic muscle activity) frequently contaminating the acquired ERP signals.

Stimulus Waveforms

Many different stimulus waveforms have been proposed for the acquisition of transient auditory ERP, ranging from short clicks and tonebursts (Hall, 2007) to the more recently introduced chirp stimuli (e.g., Fobel and Dau, 2004; Cargnelutti *et al.*, 2017), which, by carefully compensating for the frequency–dependent stimulus travel time along the basilar membrane, are intended to achieve a more synchronized discharge of the cochlear inner hair cells, thereby significantly improving the resulting average ERP waveform morphology compared to the traditional click (Dau *et al.*, 2000). Consequently, frequency–specific chirp stimuli have been applied with good results in brainstem evoked response audiometry (Bell *et al.*, 2002; Corona-Strauss *et al.*, 2012).

Signal Morphology and Neural Generators

Fig. 1.3 shows the signal morphology[2] of the human transient auditory ERP as first reported by Picton *et al.* (1974), which is characterized by a temporal succession of alternating polarity fluctuations composed out of up to 16 prominent waves with characteristic latencies, all of them occuring within less than one second post–stimulus. Specific subsets of the prominent waves are terminologically aggregated into the early, middle–latency and late auditory ERP based on their neural generators. The earliest waves $I - VII$, ranging from around 1 to 10 ms, are associated with neural activity ascending from the AN via the brainstem–level nuclei of the AP up to the IC, thus collectively being termed Auditory Brainstem Response (ABR) (Picton, 2010). At the brainstem level, each prominent wave in the ERP corresponds to neural activity well localized to a specific auditory nucleus : While waves I and II can be attributed to the activity of the distal and proximal ipsilateral AN, wave III originates mainly from the CN with contributions from the ipsilateral SOC. Waves IV and V are compound potentials with multiple ipsi- and contralateral sources such as the LL termination at the IC as well as the direct pathways from the CN to the IC. Wave V is typically followed by a negative peak arising from dendritic activity in the contralateral IC (Hall, 2007).

By contrast, for the Auditory Middle–Latency Response (AMLR) (ranging from $10-50$ ms) and subsequent Auditory Late Response (ALR) ($50-300$ ms), the prominent waves are much less well localized to specific neural structures as they arise from large–scale thalamocortical projections and subsequent cortical activity (Di and Barth, 1992). Source topography analysis of the neural activity have revealed the AMLR components from 30 ms onwards (waves N_a, P_a and N_b) to be originating from sources

---

[2] Note that the displayed ERP average waveforms were recorded using a vertex–negative electrode montage, a practice quite widespread in clinical applications. By contrast, all other displays of ERP data appearing throughout this manuscript consistently feature a vertex–positive electrode montage.

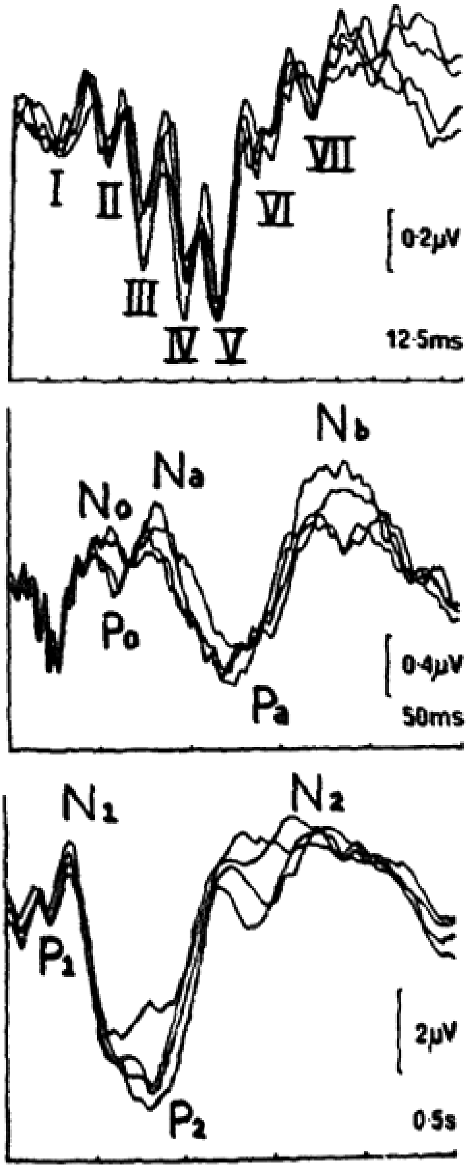

Figure 1.2: Auditory Brainstem–, Middle–Latency and Late Responses. Reproduced with slight modifications from Picton *et al.* (1974).

distributed medio–laterally along Heschl's gyri, with the earlier waves being generated in the postero–medial part (i.e., the primary AC) and the later waves forming within the more lateral, secondary AC (Liégeois-Chauvel *et al.*, 1994). Likewise, the formation of the most prominent part of the ALR, the N_1–P_2 complex, presumably involves three dominant cerebral generators in the form of dipole sources, two of which are oriented vertically on the supratemporal plane with the third being oriented laterally in the magnopyramidal temporal field (Scherg *et al.*, 1989).

Utility in Clinical Medicine and Neuroscience Research

The transient auditory ERP have many clinical applications in assessing functions of the human AP. Contemporary medical uses include the objective estimation of hearing thresholds, prominently applied in the auditory screening of, e.g., neonates, infants or uncooperative patients (Glasscock *et al.*, 1979), diagnosis of cochlear pathologies as well as tumors of the 8th cranial nerve (Stockard and Rossiter, 1977), intraoperative monitoring purposes and diagnosis of cerebral death (Burkard *et al.*, 2006). The considerable diagnostic value of transient auditory ERP is essentially due to the following beneficial properties :

- Simple, well–established measurement setups, stimulation paradigms and signal processing methods

- Reproducible neural correlates which usually generalize well across subjects within homogeneous groups

- High sensitivity for exogeneous (i.e., stimulus characteristics) and endogeneous factors (e.g., attention effects)

- High sensitivity for a broad variety of pathological conditions

One minor drawback of human auditory ERP resides in the fact that the AMLR and ALR components in particular undergo gradual age–induced changes in their signal morphology due to an alteration of the underlying source topology and associated dipole orientations, with the majority of ERP patterns usually being established as of age six and several other components slowly maturing at rates between $10 - 50\,\%$ per year (Ponton *et al.*, 2002). While this behaviour aggravates the quantitative comparison of auditory ERP acquired from groups with considerably different age distributions, the reported intra–subject test–retest–repeatability of both the brainstem–level and thalamocortical responses is remarkable even when employing more complex stimulus waveforms such as speech fragments (Bidelman *et al.*, 2018).

As already pointed out in Sect. 1.1.1, cortical auditory processing is strongly modulated by attention. Consequently, transient ALR have been at the center of interest in assessing these attentional mechanisms for decades since the pioneering demonstration of Picton and Hillyard (1974). However, recent findings of Reichenbach *et al.* (2016) and Forte *et al.* (2017) impressively demonstrate the feasibility of acquiring reliable *subcortical* correlates of attentional processing all the way down to the fragile brainstem–level components using running speech stimuli. This novel form of neural correlates will very likely improve the overall understanding of processing stages along the AP in future research as it enables additional quantitative insight into the role of corticofugal processing in auditory attention.

An attentional effect of potentially huge clinical relevance is *cortical inhibition*, which can be indirectly quantified by the amount of peak amplitude reduction in the N_1–P_2 complex of the subsequent ALR responses to paired stimulus presentations. A substantial lack thereof has been proposed as an objective indicator of attention deficit disorders in adult patients by Schubert *et al.* (2014). Moreso, reduced cortical inhibition correlates with reduced performance in a speech intelligibility task (González–Trejo *et al.*, 2013) as well as a reduced ability to perform an auditory selective attention task in a simulated driving environment (González–Trejo *et al.*, 2015), additionally making it a possible predictor for driving performance in modern driving environments with increasing levels of distraction due to the advent of in–car infotainment systems.

Another attentional longterm effect measurable in the cortical auditory ERP to repetitive, subthreshold stimulus presentation is *habituation* (Prosser *et al.*, 1981; Mariam *et al.*, 2009; Kern *et al.*, 2010). It must be clearly distinguished from stimulus–specific adaptation, which primarily occurs in the auditory thalamic nuclei (Anderson *et al.*, 2009) and leads to a significant amplitude reduction in the obtained ERP compared to the response to the first few stimuli (Zhang *et al.*, 2009). Despite several refinements of the available methodology for assessment of habituation using ALR single trials (e.g., Mariam *et al.*, 2012; Mortezapouraghdam *et al.*, 2015), the quantification of stimulus habituation has not yet seen noteworthy application in clinical diagnostics to the present date, most likely due to the substantial acquisition times involved. Nevertheless, recent research indicates dysfunctional long–term habituation of matched tone stimuli to be a potential objective measure for classification of the decompensation degree in sufferers of tonal *tinnitus aureum* (Lehser *et al.*, 2015) with a promising therapeutic approach involving spectral notching of the individual patients' tinnitus frequencies using commercially available hearing aids (Haab *et al.*, 2019).

The last application of transient auditory ERP to be presented in this overview[3] lies with the growing body of research on *binaural interaction*. Following the initial demonstrations of binaural difference potentials in the transient–evoked human ABR by Levine (1981) and subsequent investigations of analogous middle–latency and late binaural difference potentials (McPherson and Starr, 1993), this field has seen an increasing number of scientific contributions during recent years (e.g., Henkin *et al.*, 2015; Francart *et al.*, 2018). The substantial susceptibility of binaural brainstem–level (Riedel and Kollmeier, 2006) and subsequent cortical (Ungan *et al.*, 2001) difference potentials with respect to interaural timing and intensity differences makes them a valuable tool in quantifying the degree of binaural interaction along the auditory pathway (ideally leading to the two slightly deviating sounds being fused into a single percept), which is of paramount interest in, e.g., evaluating and improving the fitting quality of the increasingly ubiquitous hearing devices. Consequently, successful recent work investigating the impact of nonlinear frequency compression on binaural interaction in hearing aid users (Klauke *et al.*, 2015) as well as electrode paring in bilateral cochlear implantees (Hu *et al.*, 2016) emphasizes the suitability of the employed measures. Building upon the work of Schebsdat *et al.* (2018) on artifact reduction in ERP acquired from cochlea implant users and the fast ERP acquisition methods developed in Kohl *et al.* (2019a), our upcoming research demonstrates the benefits of carefully compensating the interaural delay in asymmetric hearing treatments involving one cochlea implant and one hearing aid with respect to brainstem–level binaural interaction (Schebsdat *et al.*, 2019), which is additionally backed by recent psychoacoustic investigations by Zirn *et al.* (2019).

---

[3] Please note that providing a more exhaustive list of contemporary and possible future applications for auditory ERP in clinical neurodiagnostics and neuroscience research is beyond the scope of this work, hence the inclined reader is refered to Picton (2010) for further reference within this regard.

1.2 Motivation and Objectives of This Work

The foundational work of (Picton *et al.*, 1974) effectively systematized the acquisition of ABR, AMLR and ALR as *separate* components of the auditory ERP, with the subsequently established recording parameters such as Interstimulus Intervals (ISI), acquisition time windows and bandpass filter settings favourable for each of these components (e.g., Hall, 2007) essentially being the result of careful, empirical optimization. As a consequence, the vast majority of studies involving transient auditory ERP have adhered and continue to adhere to these established acquisition methods and associated recording parameters given that their normative character enables comparability of the obtained signals across different studies, which is a very reasonable approach in and of itself.

However, sticking with the traditional recording *„recipes"*, if only for the sole reason of them being tried and tested, is not without drawbacks, as a substantial amount of information potentially valuable in answering research questions is irrecoverably lost due to the signal conditioning involved in obtaining either of the three ERP components. Very much by contrast, Michelini *et al.* (1982) have demonstrated the general feasibility of simultaneously acquiring *all* auditory ERP components in a single measurement by means of a parametric, progressive decimation to warp the obtained ERP signals on a non–uniform timebase. Fig. 1.3 displays an example average ERP waveform obtained using this approach. Even more impressively, the employed methods barely stretched the technological limits of mainstream computational hardware available back at the time. An in–depth analysis of the waveform morphology exhibited by the obtained signals was subsequently given in Arslan *et al.* (1984). Despite this substantial pioneering contribution to ERP recording methodology, simultaneous acquisition of transient ABR, AMLR and ALR – henceforth referred to as Auditory Full–Range Response (AFRR) throughout this manuscript – while potentially offering improvements in the understanding of different auditory processing stages and their dynamic interaction, has not seen any noteworthy application in neuroscience research to the present date. This is most likely due to the unacceptably large recording times necessary to acquire all three auditory ERP components at reasonable Signal–to–Noise Ratios (SNR) when employing the approach of Michelini *et al.* (1982) as the proposed method has to rely on very low ISI to prevent waveform disturbance due to overlap of subsequent single responses.

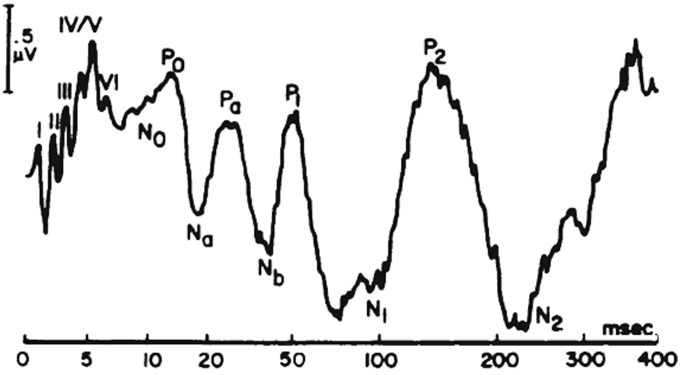

Figure 1.3: Full–Range Auditory ERP. Reproduced with slight modifications from Michelini *et al.* (1982).

Even though this promising approach has neither found its way into contemporary ERP recording methodology nor seen any notable refinements during recent years aside of our own contributions (Kohl and Strauss, 2016; Kohl *et al.*, 2019b), a recent surge of research on alternative methods for the simultaneous acquisition of brainstem–level and cortical ERP (e.g., Bidelmann, 2015; Slugocki *et al.*, 2017) strongly emphasizes the general desirability of the line of thought that motivated the work of Michelini *et al.* (1982) to begin with. As will be analyzed in greater detail during the literature review of Sect. 2.5.1, the above–mentioned more recent approaches accurately reflect the need to substantially reduce the large acquisition times inherent to this endeavour, which is achieved by means of interleaving very dense stimulation periods with longer pauses to more or less simultaneously yield an auditory brainstem Frequency–Following Response (FFR) and a cortical response. While being laudable efforts in principle, the proposed approaches effectively sacrifice valuable information in exchange for their notably fast acquisition speed, as the employed brainstem FFR lacks the temporal resolution necessary to relate the resulting response to their respective neural generators in detail. By contrast, the work of Holt and Özdamar (2014) demonstrated the feasibility of simultaneously recording high–quality time–resolved ABR and AMLR components at high stimulation rates using a deconvolution–based acquisition approach to separate the overlapping responses, which in turn is incapable of obtaining a well–defined ALR due to the detrimental effects of increased stimulus rates on their response amplitudes (Holt and Özdamar, 2016).

The aspiration of the present dissertation resides in answering the following research questions raised during the process of reviewing the prior art by Michelini *et al.* (1982), Holt and Özdamar (2014) and Holt and Özdamar (2016) :

(1) As the AFRR obviously spans three full timescales ($10^{-3} - 10^{0}$ s), a compact signal representation for the obtained ERP single responses is desirable to facilitate memory–efficient storage as well as fast, computationally inexpencive denoising. Hence, if simultaneously acquired ABR, AMLR and ALR arguably are best displayed on a non–uniform timebase, what is the optimum warping function ? Can sufficient empirical support for its optimality be gathered and, more importantly, is it possible to demonstrate that no relevant signal features are lost during the associated resampling process ? How does the resampling of AFRR single responses affect their susceptibility to the noise sources frequently contaminating ERP recordings compared to uniformly sampled single responses ?

(2) Are the state–of–the–art ERP single response denoising methods suitable for use with the resampled signal representation, and are there any more efficient ways to denoise the obtained AFRR single responses ?

(3) How can the acquisition of AFRR single responses be realized in a more time–efficient way by reducing the employed ISI, without at the same time introducing a loss of signal quality in the obtained ERP ?

Sect. 2.3 addresses questions (1), with empirical support for the suitability of the introduced resampling method being based on the analysis of a reference AFRR dataset acquired from 15 normal hearing subjects in Sect. 2.1. Starting from a short review of established methods for ERP single response denoising, two novel, efficient denoising approaches are subsequently proposed in Sect. 2.4, followed by a comparative performance analysis in order to answer questions (2). To facilitate a thorough evaluation, the resampled real–world ERP dataset gathered in Sect. 2.1 is employed alongside synthetic data generated at different SNR using the AFRR model developed in Sect. 2.2. Concluding this body of work, question (3) is addressed in Sect. 2.5 by introducing a novel interleaved stimulus presentation paradigm inspired by Bidelmann (2015) and subsequent deconvolution–based ERP processing methods centered around the work of Holt and Özdamar (2014), which are successfully evaluated in 20 normal hearing subjects.

2 Materials and Methods

2.1 Reference AFRR Acquisition

2.1.1 Preliminary Considerations

Answering the research questions raised in the previous section obviously necessitates acquiring a reference dataset of transient–evoked full–range auditory ERP. The dataset needs to comprise a number of samples large enough to facilitate statistical analyses in order to be suitable as a base for the subsequent investigation and comparative evaluation of the proposed signal processing methods. Hence, a first study was conducted to obtain the necessary data. The ERP acquisition parameters applied in this study were chosen based on the following considerations :

- The time window used to epoch the acquired ERP should be large enough to fully encompass all three timescales of interest.

- The sampling rate and the spectral bandwidth of the obtained responses after filtering should be suitable to reliably retain all relevant signal details across investigated timescales. A beneficial side effect of using larger bandwidths resides in the minimization of potential filtering artifacts (i.e., pre– and post–ringing) perturbing the waveform morphology of the acquired ERP (see Rousselet (2012) for a review).

- The employed ISI should reflect a reasonable compromise between acquisition speed and the quality of the obtained cortical ERP components, which strongly benefit from larger ISI (Davis *et al.*, 1966).

2.1.2 Subjects

The study was conducted at Saarland University of Applied Sciences (Saarbrücken, Germany), with 15 subjects (10 males, 5 females, ages ⌀26 ± 3 years) being recruited from the social environment of the author to participate in the reference AFRR recording.

None of the subjects exhibited any known history of audiological or neurological disorders. All of them were classified as normal hearing using puretone audiometry with hearing thresholds consistently below 15 dB hearing level within the frequency range of 0.5 – 8 kHz (Gelfand, 2009). The subjects were informed about the experimental procedure, which was designed in accordance with the Declaration of Helsinki and approved by the institutional review board and the ethics comittee of the *Ärztekammer des Saarlandes*, and signed a consent form.

2.1.3 Stimulus Presentation and ERP Recording

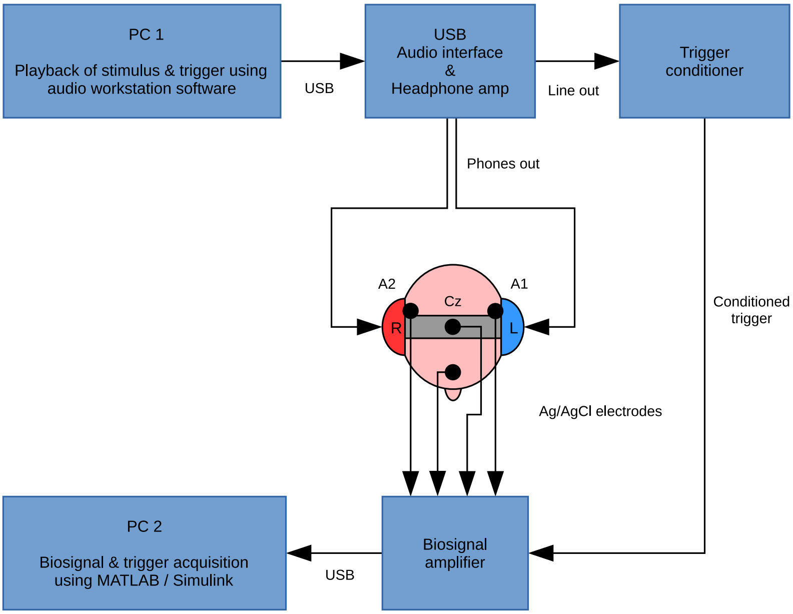
PC 1
Playback of stimulus & trigger using audio workstation software
USB
USB
Audio interface
&
Headphone amp
Line out
Trigger conditioner
Phones out
A2
A1
Cz
R
L
Conditioned trigger
Ag/AgCl electrodes
PC 2
Biosignal & trigger acquisition using MATLAB / Simulink
USB
Biosignal amplifier

Figure 2.1: Stimulus Presentation and ERP Acquisition Setup.

A schematic overview of the employed stimulus presentation and ERP acquisition setup is given in Fig. 2.1. The broadband chirp stimulus optimized for eliciting the ABR in humans as proposed by Fobel and Dau (2004) was used for stimulation (edge frequencies 0.1 – 10 kHz, 9.9 ms duration). Stimuli were presented binaurally to the subjects at 70 dB peak–to–peak equivalent SPL using a desktop computer connected to a USB audio interface (Scarlett 2i4, Focusrite, UK) with its integrated headphone amplifier driving a pair of standard audiometry headphones (HDA–200, Sennheiser, Germany). The chirp stimulus was calibrated according to IEC 60645–3 (International Electrotechnical Commission, 2007) using an artificial ear and a sound level meter (N° 4153 and 2250, Brüel & Kjaer, Denmark). Stimulus waveforms and corresponding acquisition triggers were generated offline as 24 bit PCM wave files at 48 kHz sampling rate using MATLAB™ R2016a (The MathWorks, USA) and played back using multichannel audio workstation software (Studio One 2, PreSonus, USA). The stimulus wavefiles comprised a total amount of 10^3 stimuli to reliably acquire the ABR components at a reasonable SNR (Hall, 2007), presented at an ISI of 1 s[1] ±10 % equally distributed jitter to obtain the additional SNR advantage due to averaging as pointed out by Woldorff (1993), which resulted in 16:50 min overall acquisition time. Stimulus polarity was alternated for each successive chirp to equally distribute potential electromagnetic influences from the headphones on the acquired ERP, eventually cancelling them out during averaging.

Subjects were instructed to sit on a lounger in a darkened, acoustically controlled environment, relax with their eyes closed, but not fall asleep and avoid any movement during ERP acquisition. The auditory ERP were recorded using four passive Ag/AgCl cup electrodes placed on the left and right mastoids (A2/A1, active leads), the vertex (Cz, reference lead) and the forehead (ground) of the subjects. Electrode impedances were kept below 5 kΩ. The raw auditory ERP leads and corresponding trigger signals were recorded sample–synchronously without analog prefiltering at 19.2 kHz sampling rate using a second desktop computer connected to a medical–grade biosignal amplifier with an external trigger conditioner (g.USBamp and g.Trigbox, g.tec Medical Engineering, Austria).

---

[1] Note that this particular ISI value is the result of empirical optimization based on variational repeated measurements in a single subject. As reported by Davis *et al.* (1966), ALR components benefit progressively from ISI increases up to (obviously highly impractical) presentation rates as low as one 6th of Hz. Preliminary tests involving rates well below 1 Hz are consistent with this finding, with ALR morphology in the obtained AFRR averages appearing progressively distorted. Attempts at mitigating this detrimental effect by subsequently employing the Adjacent Response Removal (ADJAR) overlap correction method introduced by Woldorff (1993) – and promoted by none less than Picton *et al.* (2000) as a part of their contemporary recording standards for higher–rate ALR acquisition – notably proved unsuccessful, hence the ISI value of 1 s was established before conducting this study as a reasonable compromise.

2.1.4 ERP Processing and Analysis

Signal conditioning, epoching and ERP analysis were carried out in MATLAB$^{\mathrm{TM}}$. The raw ERP signals were zero–phase filtered using a 4th order Butterworth Infinite Impulse Response (IIR) bandpass filter with lower and upper 3 dB cutoff frequencies of 1 Hz and 1 kHz to eliminate baseline wandering and high–frequency noise. In order to reliably remove potential powerline interference and its harmonics from the signals, a set of 4th order Butterworth IIR notch filters was employed, centered at all integer multiples of 50 Hz within the pass band with 3 dB bandwidths of 5 millioctaves.

By epoching the conditioned discrete–time ERP raw signals $u \in \ell^2(\mathbb{R})$ using the acquisition time window from 1 ms – 1 s after each rising edge of the associated trigger signals (compensating for the 9.9 ms delay introduced by the chirp stimulus), a set $\mathcal{U} = \{ u_m \in \mathbb{R}^N \mid m = 1\,..\,M,\ M,N \in \mathbb{N}\}$ of ERP single responses was obtained for each subject and electrode lead, respectively. This sampled single–response representation was first proposed by Jung *et al.* (1999), coining the term *„ERP image"* in the context of data visualization. Note that it has also been interpreted as a *„sweep matrix"* $\boldsymbol{U} \in \mathbb{R}^{M \times N}$ containing the M single responses as row vectors, particularly within the context of single–trial ERP denoising algorithms (e.g., Strauss *et al.*, 2013). Hence, both interpretations and notations are used as deemed appropriate throughout this manuscript as they can be considered interchangeable.

In order to assess the obtained ERP and compare their waveform morphology to the previous results of Michelini *et al.* (1982), average waveforms u_a were calculated for each single response set. Prior to averaging, single responses were rendered offset–free by sample–wise subtraction of their respective arithmetic mean values $\overline{u}_a$. Responses exceeding a bipolar amplitude threshold of 70 μV were rejected as artifacts. To ensure a high data quality of the grand average, single–subject average waveforms u_a were visually screened, with 4 of them being excluded from the grand average either due to exhibiting no prominent waveform morphology or containing postauricular muscle responses (O'Beirne and Patuzzi, 1999) exceeding 5 μV zero–to–peak. Fig. 3.1 displays the resulting grand average waveform of the remaining 11 single–subject ERP averages.

As a prerequisite to the subsequent AFRR model development and optimization of the AFRR signal representation, a time–frequency analysis of the obtained single responses was performed. Wavelet transforms in particular have been repeatedly employed for this purpose with great success in multiple ERP studies (e.g., Samar *et al.*, 1995; Demiralp and Ademoglu, 2001; Quian Quiroga *et al.*, 2001; Murali and Kulish, 2007; Wang *et al.*, 2007) due to numerous desirable properties. By applying the Continuous Wavelet Transform (CWT)

$$\mathfrak{W}_{\Psi}\{f\}(a,b) = |a|^{-\frac{1}{2}} \left\langle f\,,\,\psi_{(a,b)} \right\rangle \tag{2.1}$$

where $\psi_{(a,b)} = \Psi(\frac{\cdot - b}{a})$ denotes a set of translated and dilated functions derived from a normalized, admissible mother wavelet Ψ (Aldroubi and Benedetto, 1996), time–scale representations of any signal $f \in L^2$ can be obtained. In the case of the present sampled, discrete–time single responses $u_m \in \mathbb{R}^N$, the above inner product conveniently reduces to

$$\mathfrak{W}_{\Psi}\{u_m\}(a,b) = \frac{1}{\sqrt{a}} \sum_{n=1}^{N} u_m[n]\, \Psi^{*}\left(\frac{n-b}{a}\right), \; a \in \mathscr{A}, b \in \mathscr{B} \tag{2.2}$$

where $\mathscr{B} = \{0\,..\,N-1\}$ denotes the set of $N \in \mathbb{N}$ discrete–time translations and $\mathscr{A} \subset \mathbb{R}_{+}^{L}$ denotes the uniquely sampled set of $L \in \mathbb{N}$ continuous dilations used for analysis. Given the dimensionless center frequency f_c of Ψ, the sampling rate dependent dilations a can be associated with pseudofrequencies $f_p = f_c\, f_s\, a^{-1}$. This way, 200 analysis pseudofrequencies were spaced logarithmically within the frequency range from 1 Hz to 1 kHz. Likewise, the discrete–time translations b can be associated with temporal shifts $\tau = b\, f_s^{-1}$. This effectively yields a time–frequency representation of u_m. The complex Morlet with a shape parameter of 2π as first introduced by Goupillaud *et al.* (1984) was employed as mother wavelet Ψ, hence the obtained transform coefficients $\mathfrak{W}_{\Psi}\{u_m\} \in \mathbb{C}^{L \times N}$ constitute analytic signals with maximum time–frequency resolution (Gabor, 1946). Fig. 3.2 displays the real part of the grand average CWT coefficients obtained from the AFRR dataset.

2.2 Phenomenological AFRR Model

During the past decades, systems neuroscience has made significant progress at providing computational models of different parts of the auditory system (Meddis *et al.*, 2010). Out of these models, those concerned with the brainstem–level neural structures are capable of predicting the transient ABR signal morphology in great detail (e.g., Harte *et al.*, 2010; Verhulst *et al.*, 2015). There have been similarly motivated attempts at providing physiology–based models of the cortical auditory ERP (e.g., Kerr *et al.*, 2008), with a multitude of other available models operating on a rather coarse level of detail by considering macroscopic dipole sources and their spatiotemporal interaction (e.g., Scherg and von Cramon, 1986; Verkindt *et al.*, 1995). To the knowledge of the author, there exists no cohesive modeling framework for the electrophysiological responses of the *whole* auditory pathway to the present date. Since the development of such a multiscale model is well beyond the scope of this work, a simple, phenomenological model based on the AFRR dataset obtained in the previous section is proposed in the following for the subsequent comparative evaluation of ERP single–response denoising methods.

As there currently does not seem to be an established consensus on whether the evoked or the phase–reset hypothesis (or a mixture of both) holds in terms of ERP formation (Yeung *et al.*, 2004), the proposed model is based on the signal–plus–noise model established by Dawson (1954). The system response of any sufficiently large ensemble of interconnected, weakly nonlinear processes due to a transient excitation can be reasonably well approximated by a Gaussian distribution due to the central limit theorem (Pólya, 1920). Hence, the recorded AFRR electrode lead signals $r(t)$ due to excitation with a stimulus train

$$s(t) = \sum_{m=1}^{M} \delta\left(t - t_s[m]\right) \tag{2.3}$$

comprising $M \in \mathbb{N}$ stimulus onsets in the form of Dirac unit impulses are modeled by a linear superposition of Gaussian functions representing the subsequent prominent waves of the ERP as also proposed by Mohseni *et al.* (2007) and an additive, independent noise process $n(t)$:

$$r(t) = s(t) * \left(\sum_{w=1}^{W} a_p[w] \exp\left(\frac{t - \tau_p[w]}{\sigma_p[w]} \right) \right) + g_n\, n(t) \tag{2.4}$$

In Eqn. 2.4, $a_p \in \mathbb{R}^W$ denotes the peak amplitudes of the $W \in \mathbb{N}$ prominent waves, $\tau_p \in \mathbb{R}_+^W$ denotes their respective latencies post–stimulus and $\sigma_p \in \mathbb{R}_+^W$ their temporal widths. $g_n \in \mathbb{R}_+$ is a variable factor adjusting the gain of noise superimposed on the deterministic Gaussian mixture.

To determine the respective parameters for the Gaussian functions from the reference AFRR dataset acquired in the previous section, the prominent peaks were located in the grand average real part of CWT coefficients as displayed in Fig. 3.2 using a gradient method initialized with approximate values obtained by point & click. Tab. A.3 shows the respective values for peak translation τ_p, pseudofrequency f_p and amplitude u_p. The temporal widths σ_p were calculated from the respective pseudofrequencies as

$$\sigma_p[\cdot] = \frac{1}{4\pi f_p[\cdot]} \tag{2.5}$$

The additive, stochastic noise process $n(t)$ was realized by means of spectral resynthesis as proposed by Yeung *et al.* (2004). 28 available standard vertex–mastoid recordings of ongoing EEG acquired at 19.2 kHz sampling rate without analog prefiltering in adult subjects using the same setup and recording evironment as the reference AFRR acquisition study were averaged in the frequency domain. To generate a random noise signal of arbitrary length as part of the model output, the averaged half–side magnitude spectrum was first interpolated to the necessary sample count, multiplied with unit phasors drawn from a random, circular uniform distribution and transformed into a time–domain signal via inverse Fast Fourier Transform (FFT). The approach of Yeung *et al.* (2004) has been shown to retain the signal characteristics of ongoing EEG including deterministic disturbances such as powerline interference to a high degree (making them virtually indistinguishable from a real EEG recording), which is of paramount interest for the intended purpose of the proposed model.

To ensure the comparability of real–world and synthetic AFRR data, the model output was conditioned using the same processing parameters as described in the previous section, yielding bandlimited ERP images $\boldsymbol{U} \in \mathbb{R}^{M \times N}$ with $M = 10^3$. The gain factor g_n was chosen as appropriate to realize any specified SNR in terms of the Frobemius norms $||\boldsymbol{U}||_F$ of split ERP images containing only the deterministic and the stochastic part of the epoched model output, respectively.

2.3 Optimization of AFRR Signal Representation

2.3.1 Preliminary Considerations

Michelini *et al.* (1982) have argued that simultaneously acquired ABR, AMLR and ALR are probably best displayed on a nonuniform timebase, with several more or less progressive warpings being proposed to emphasize different temporal regions of the ERP. However, no empirical support for their respective optimality was given apart from visual inspection. The measured latencies of prominent waves as reported by Picton *et al.* (1974) (see Tab. A.1), time–frequency windowing values conventionally used in auditory ERP acquisition (Hall (2007), see Fig. 2.2) and the time–frequency peak translations and pseudofrequencies obtained in the previous section (see Tab. A.3) are visibly in support of Michelini *et al.* (1982) as they all suggest a roughly proportional time dependence of both the interpeak intervals and the corresponding pseudoperiods of the response.

An appropriate framework to quantify the degree of signal self–similarity across timescales hinted at by the above observation is given by the Mellin Transform (MT) (Butzer and Jansche, 1997) :

$$\mathfrak{M}\{f\}(p) = \int_{\mathbb{R}_+} f(t)\, t^{p-1}\, dt \ , \ p \in \mathbb{C} \tag{2.6}$$

Being closely related to the Laplace Transform (LT), which expands a time–domain signal using dampened harmonic oscillations of different frequencies as basis functions, the MT

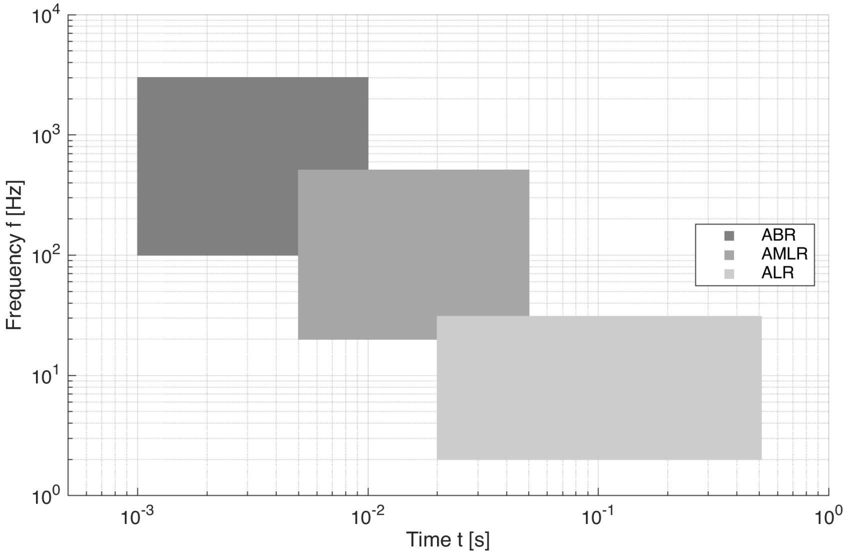

Figure 2.2: Conventional time–frequency windowing of auditory ERP components according to Hall (2007)

decomposes it into dampened oscillations with instantaneous frequencies temporally increasing (or decreasing) at different rates. The MT, particularly its various discrete–time implementations, have seen numerous successful applications in analyzing signals containing modes with rising or falling frequencies as they are remarkably robust to stationary noise (e.g., Bertrand *et al.*, 1990; Zwicke and Kiss, 1983). By substituting $p = -is + \frac{1}{2}$, $s \in \mathbb{R}_+$, the Scale Transform (ST) can be derived, whose transform magnitude enjoys the property of invariance with respect to temporal scaling of f (De Sena and Rocchesso, 2007; Cohen, 1993) much in the same way that the Fourier Transform (FT) magnitude is invariant with respect to temporal shifts. Hence, the single–subject AFRR averages u_a obtained in Sect. 2.1 were analyzed in terms of their intrinsic scales by applying the ST. As the Parseval theorem holds for the scale transform coefficients $\mathfrak{S}\{u_a\}(s)$, a sparse distribution of coefficient energy across few scales is linked to a high degree of self–similarity across timescales within the corresponding time–domain AFRR averages. This property in turn is a necessary prerequisite for a compact, information–preserving signal representation on a progressively decimated timebase as proposed by Michelini *et al.* (1982). Fig. 3.3 shows the obtained grand average ST coefficient magnitudes.

Another key aspect to be examined within this context is the degree of signal synchronization across single responses, as it is directly linked to the prominence of ERP components over the competing noisefloor of oscillatory EEG activity. Multiple amplitude–independent, time–frequency resolved measures to quantify this synchronization degree have been proposed independently by different research groups (e.g., Tallon-Baudry *et al.*, 1996; Corona-Strauss *et al.*, 2009; Klimesch *et al.*, 2014), all of them essentially being varieties of circular resultant vector length (Fisher, 1993) in the complex–valued coefficient domain of a suitable transform. Here, the Wavelet Phase Stability (WPS) (Low and Strauss, 2011)

$$\Gamma_\varphi(f_p,\tau) = \frac{1}{M}\left|\sum_{m=1}^{M} \exp\bigl(i\,\varphi_m(f_p,\tau)\bigr)\right| \tag{2.7}$$

is employed, where $\varphi_m(f_p,\tau) = \arg(\mathfrak{W}_\Psi\{u_m\}(f_p,\tau))$ denotes the time–frequency resolved instantaneous phases of the single–response CWT coefficients obtained in Sect. 2.1. Note that Γ_φ assumes values from the bounded range $[0,1]$ where 1 corresponds to perfect phase synchronization across single responses, effectively indicating which areas of the time–frequency plane carry the information of event–related signal components and which can be considered redundant based on their lack of such. Fig. 3.4 shows the grand average of WPS calculated separately for all ERP images.

2.3.2 Nonuniform Resampling

Results of the preliminary investigations (see Sect. 3.2) strongly suggest a *logarithmic* timebase to be the most adequate for a compact representation of the AFRR. Hence, a suitable resampling operator is proposed in the following. Note that the interpeak intervals and pseudoperiods of successive prominent waves in the AFRR being roughly proportional to their time of appearance will result in an almost equidistant distribution of peaks with uniform sample resolution after data transformation. The resampling of an AFRR single response $u_m \in \mathbb{R}^N$ sampled at the fixed rate F_s can be realized by first introducing a variable sampling rate $f_s \in \mathbb{R}_+^N$ logarithmically decreasing over time.

Considering the linear timebase

$$t[n] = t_l + (n-1)\,\frac{t_u - t_l}{N-1}\ ,\ n = 1\,..\,N \tag{2.8}$$

of the uniformly sampled single response with $N \in \mathbb{N}$ sampling points equidistantly spaced within the time window $[\,t_l\,,\,t_u\,] \subset \mathbb{R}_+$, the time–dependent, logarithmically decreasing 3 dB cutoff frequency

$$f_c[n] = \exp\left(\log(f_u) - (n-1)\,\frac{\log(f_u) - \log(f_l)}{N-1}\right)\ ,\ n = 1\,..\,N \tag{2.9}$$

of a bandlimiting lowpass filter is defined. In Eqn. 2.9, the range $[\,f_l\,,\,f_u\,] \subset \mathbb{R}_+$ denotes the frequency window, with its edges temporally coinciding to the end and the beginning of the response time window, respectively. The minimum sample count necessary to represent the progressively bandlimited, resampled response is then given by

$$N' = \left\lceil\ \frac{2}{F_s}\sum_{n=1}^{N} f_c[n]\ \right\rceil \tag{2.10}$$

due to the Nyquist theorem. Using this critically reduced sample count, the logarithmically spaced timebase

$$t'[n] = \exp\left(\log(t_l) + (n-1)\,\frac{\log(t_u) - \log(t_l)}{N'-1}\right)\ ,\ n = 1\,..\,N' \tag{2.11}$$

is defined for the resampled response. By assembling a matrix containing N' normalized Gaussian windows as row vectors, with their respective temporal shifts and 3 dB cutoff frequencies given by the successive values of t' and f_c', the resampling operator

$$\boldsymbol{R}[n,\cdot] = \frac{\sqrt{2\pi}\, f_c'[n]}{F_s} \exp\left(-2\left(\pi f_c'[n]\left(t[\cdot] - t'[n]\right)\right)^2\right), \; n = 1\,..\,N' \tag{2.12}$$

is obtained, where

$$f_c'[n] = f_u - (n-1)\,\frac{f_u - f_l}{N' - 1}, \; n = 1\,..\,N' \tag{2.13}$$

denotes the time–dependent cutoff frequency on the logarithmic timebase. Applying the operator $\boldsymbol{R} : \mathbb{R}^N \rightarrow \mathbb{R}^{N'}$ to resample an AFRR single response u_m then reduces to calculating the matrix product

$$u_m{}' = \left(\boldsymbol{R} \cdot \left(u_m{}^\top\right)\right)^\top \tag{2.14}$$

Note that in contrast to the critical exponential resampling investigated by De Sena and Rocchesso (2007), the presented resampling operator is of course inherently lossy in nature as it effectively reduces the dimensionality of the input data. Using the time window $[10^{-3}, 1]$ s and frequency window $[1, 10^3]$ Hz at a sampling rate of $F_s = 19.2$ kHz, the compression ratio exhibited by the operator amounts to $N/N' = 16.59$. Hence, it is necessary to demonstrate that this substantial dimensionality reduction – while being the very reason for the endeavour in the first place – does not lead to a loss of relevant information in the case of AFRR single responses.

To facilitate this demonstration, the CWT of the resampled single responses $u_m{}'$ was calculated for the reference AFRR dataset obtained in Sect. 3.1 according to Eqn. 2.2. While the translations b can still be associated with temporal translations τ by virtue of Eqn. 2.11 in this case, dilations a cannot be related anymore to specific constant pseudofrequencies f_p due to the progressively increasing sampling intervals. Fig. 3.5 displays the real part of the grand average CWT coefficients obtained in this way and Fig. 3.6 shows the associated WPS according to Eqn. 2.7. Note that the CWT coefficients of the resampled single responses can also be interpreted as a sparse subset of a chirplet transform (Mann and Haykin, 1991) of the linear–time responses as the employed transform atoms correspond to tilted time–scale atoms in the uniformly sampled time domain.

2.3.3 Noise Characteristics

It has previously been shown by Özdamar and Delgado (1996) that the RMS value u_{RMS} of the Residual Noise (RN) present within an ERP waveform after repeated averaging of ERP single responses closely satisfies the power law

$$u_{RMS}(m) \propto \frac{1}{\sqrt{m}} \tag{2.15}$$

where $m \in \mathbb{N}$ denotes the number of averaged responses. Note that this particular relationship only holds necessarily for a white noise under the signal–plus–noise model assumption (Dawson, 1954), with other types of additive noise featuring deviating spectral compositions potentially yielding different rates of decay. For example, as has been pointed out by Woldorff (1993), the amount of jitter in the employed ISI is linked to a significant, albeit frequency–specific SNR advantage, hence additionally influencing this decay rate.

To empirically quantify the above relationship, facilitating an objective comparison of the uniformly sampled and resampled response representations regarding potential SNR benefits of the latter, a $1/n^{\alpha}$ model was fitted to the RMS progressions of cumulative averages by least–squares regression. The averages were obtained by epoching 10^3 *„single responses"* from an ongoing noise signal using ISI of 1 s ±10 % equally distributed jitter (see Sect. 2.1.3). By repeating this process 25 times, averaged progressions of RMS values estimating the RN as a function of trial count were obtained. This process was conducted for the uniformly sampled and resampled response representations, with the random EEG noise signals being generated using the spectral resynthesis method and bandlimited as described in Sect. 2.2 to ensure signal properties close to physiological conditions. Additionally, the RMS decay profiles were calculated and regressed for both representations using a stationary 50 Hz sine as the ongoing noise signal to investigate the behaviour of both representations with respect to the powerline interferences commonly contaminating ERP recordings. Figs. 3.7 and 3.8 show the resulting RMS progressions and their associated regression statistics for both types of noise, respectively.

As the logarithmic resampling operator involves a progressive averaging of samples in the linear time domain due to the employed Gaussian lowpass filter, the resulting resampled responses will necessarily exhibit a temporal decrease in RN. While the effect of this skew will most likely be limited by the fact that EEG activity closely follows an inverse power law with less energy being allotted to the critical higher frequencies, a substantial temporal skew in the *effective* SNR profile remains very likely as the amplitudes of prominent waves in the auditory ERP follow the opposite temporal relationship. Note that this undesirable departure from SNR uniformity across timescales will be subsequently alleviated in Sect. 2.5. To quantify the degree of temporal skew in RN for the resampling method in question, the sample–wise standard deviations across single responses were investigated in the 25 ERP images and their resampled counterparts for the synthetic EEG noise, as they have previously been shown to be a suitable estimate for the RN (Riedel *et al.*, 2001). Hence, least squares regressions of a $1/t^{\alpha}$ model were performed on the obtained standard deviation profiles, yielding their respective temporal rates of decay. Fig. 3.9 shows the time–dependent RN profiles and their associated regression statistics.

2.4 Denoising of AFRR Single Responses

2.4.1 Preliminary Considerations

During the past decade, various approaches to improve the signal quality of single–response ERP images have emerged independently[2]. The merit of applying such methods is essentially twofold, as the primary effect of SNR reduction in a single–response representation will under certain circumstances also lead to a reduced amount of responses necessary to achieve an acceptable average waveform, hence lessening acquisition times. This second potential benefit can obviously not be obtained by algorithms which perfectly preserve the average waveform of the unfiltered ERP image (i.e., linear operators on a per–sample basis such as partial averaging or Unidirectional Gaussian Means (UGM) filters). These methods definitely have their place in elevating the average waveform morphology and its potential gradual changes across the experiment (e.g., during habituation (Prosser *et al.*, 1981)) over the competing noise floor within the ERP image representation – in fact, they can be safely considered as the most *„honest"* approaches towards two–dimensional ERP denoising for this very property given the lack of information about the underlying response. However, this comes at the price of substantial averaging blur, potentially obscuring very localized events of interest within the responses. Hence, the focus of this investigation shall reside on nonlinear operators due to the second potential benefit outlined above.

In principle, nonlinear two–dimensional ERP denoising can draw from a rich body of work on algorithms originating from the field of computer vision. For example, Mustaffa *et al.* (2010) have successfully applied Nonlinear Diffusion (NLD) filtering for the purpose in question. Another well–established algorithm involving the constrained minimization of the Total Variation (TV) norm (Rudin *et al.*, 1992) has later been shown to be outperformed by contemporary wavelet–based methods (Figueiredo *et al.*, 2006), hence it will not be investigated here. The Nonlocal Means (NLM) algorithm (Buades *et al.*, 2005), which enjoys broad popularity, has been applied with very good results to ERP image denoising by Strauss *et al.* (2013) and Schubert *et al.* (2014), significantly outperforming the above–mentioned NLD approach. Despite its good performance, the underlying rationale is partially questionable due to possible nonlocal contributions of signal components originating from different neural generators (this can be alleviated in principle by restricting them to a suitably narrow search window with the obvious downside being decreased filter performance). A different, promising framework for the purpose in question is given by the wavelet–based denoising approaches due to numerous desirable properties outlined in Aldroubi and Benedetto (1996). Consequently, Quian Quiroga and Garcia (2003) and Ahmadi and Quian Quiroga (2013) proposed variations of wavelet thresholding to denoise ERP single responses. However, their approach suffers from a considerable drawback by design : since it relies on the critically sampled Discrete Wavelet Transform (DWT) (Mallat, 1989) for its low computational complexity, the inherent lack of shift invariance effectively prevents any meaningful integration of transform coefficients corresponding to localized signal components in adjacent single

---

[2] Note that the established one–dimensional ERP denoising methods operating on the acquired lead signals prior to epoching such as Kalman filtering (e.g., Mohseni *et al.*, 2007) are not investigated within the scope of this work for the sake of brevity, given that two–dimensional denoising approaches enjoy a far greater potential by virtue of integrating signal information both temporally *and* from adjacent responses.

responses into the coefficient thresholding rule and additionally renders the implementation particularly prone to jitter. In an attempt to alleviate these shortcomings, we proposed an alternative algorithm based on an overcomplete analytic wavelet filterbank (Kohl *et al.*, 2019b) which showed significantly increased performance over the UGM filter in a quantitative comparison using real–world ERP recordings. However, preliminary investigations with the present *resampled* single response representations did not reveal a comparable benefit, emphasizing the necessity to reevaluate the available approaches regarding their suitability for this type of data.

A commonality of many successful denoising algorithms is the projection of signals into a transform domain using a suitable linear operator and subsequent modification of the transform coefficients instead of directly working with the original time–domain signals. The filtered signals are then obtained by applying the inverse of the employed transform operator to the modified coefficients, with the transform pair preferably enjoying the perfect reconstruction property. One of the most recognized examples of such a denoising scheme is probably given by the work of Donoho (1995) on wavelet shrinkage. A suitable choice of transform pair effectively concentrates the energy of the sought–after signal components within few coefficients in the transform domain, leaving the remaining coefficients to predominantly encode noise by virtue of the Parseval theorem. Within this framework, the process of shrinking the energy of coefficients most likely representing unwanted signal components – either determined by the properties of the employed transform (e.g., Kohl and Strauss, 2016) or a suitable measure derived from the coefficients themselves (e.g., Aldroubi and Benedetto, 1996; Kohl *et al.*, 2019b) – has an additional potential advantage as it reduces the likelihood of filtering artifacts compared to *additive* filtering approaches such as UGM or NLM, selectively *removing* signal energy as opposed to achieving noise suppression by weighted recombination of different signal components.

Preceding development of the two novel denoising approaches outlined in the following sections, several linear transforms were investigated regarding their capability of separating signal and noise in the transform domain using the resampled AFRR single–responses in question. While orthogonal expansions using Laguerre polynomials (Mandyam and Ahmed, 1996) and Bessel functions of the first kind (Cree and Bones, 1993) as bases, having been successfully employed for a variety of signal compression applications, notably failed to produce a sparse coefficient distribution of the ERP signal components, the desired behaviour is displayed to a remarkable degree by the Radon Transform (RT) (e.g., Fig. 3.10), putting it at the center of the first presented denoising algorithm (Sect. 2.4.2). Given the successful track record of wavelet–based denoising methods, the second proposed algorithm (Sect. 2.4.3) is based on an overcomplete, perfectly reconstructing discrete wavelet filterbank. Note that calculating the FFT of the resampled single responses yields their discrete MT as proposed by De Sena and Rocchesso (2007) in a computationally efficient way. However, judging by preliminary experiments, this signal representation – while being a straightforward approach following the reasoning of Sect. 2.3.1 – turned out to provide no significant benefit over the above wavelet representation for the intended application despite its considerably sparse coefficient distribution.

2.4.2 Denoising Using the Radon Transform

The RT of a compactly supported, two–dimensional distribution $f(x, y) \in L^1(\mathbb{R}^2)$ constitutes a set of line integrals or *„projections"* of f along all straight lines or *„beams"* with radial distances $r \in \mathbb{R}$ from origin and projection angles $\phi \in [-\frac{\pi}{2}, \frac{\pi}{2}[$:

$$\mathfrak{R}\{f\}(r, \phi) = \int_{\mathbb{R}} f\left(r \cos\phi + \tau \sin\phi, r \sin\phi - \tau \cos\phi\right) d\tau \qquad (2.16)$$

This set of projections, collectively termed the *„sinogram"* of f, can be used to reconstruct the underlying distribution f by virtue of the Fourier slice theorem (Radon, 1917). In practice, i.e., considering discrete distributions $\boldsymbol{F} \in \mathbb{R}^{M \times N}$, the inverse RT is usually numerically approximated either by filtered backprojection (e.g., using the ideal, bandlimited *Ram–Lak* filter) or by iterative reconstruction methods (Kak *et al.*, 2002).

In the discrete sinogram $\boldsymbol{S} \in \mathbb{R}^{P \times P}$ obtained by applying the RT on an ERP image $\boldsymbol{U} \in \mathbb{R}^{M \times N}$, where $P = \lceil \sqrt{M^2 + N^2} \rceil$ denotes its diagonal dimensions, large magnitudes of single projections $\boldsymbol{S}[\cdot, p]$, $p = 1 \mathinner{..} P$ correspond to signal components with a high degree of similarity across the ERP image under the respective projection angles

$$\phi[p] = -\frac{\pi}{2} + (p - 1)\,\frac{\pi}{P - 1} \qquad (2.17)$$

thus accumulating in the associated directional sums. By contrast, signal components which exhibit large amounts of amplitude fluctuation under the projection angle will tend to cancel out on average. As the energy of projections under flat angles $|\phi| \approx 0$ (with $\phi = 0$ denoting vertical direction) directly corresponds to signal components strongly time–locked to the stimulus (i.e., the ERP), the projections under angles ϕ close to $\pm\frac{\pi}{2}$ will predominantly represent the noisy signal components both due to the transform's inherent directionality and the Parseval theorem. Hence, to attenuate the noisy signal components within the ERP image, the individual projections of $\boldsymbol{S}$ need to be reweighted accordingly using a normalized weight vector w' prior to obtaining the filtered ERP image by filtered backprojection of $\boldsymbol{S}'$:

$$\boldsymbol{S}'[\cdot,p] = \boldsymbol{S}[\cdot,p]\, w'[p] \tag{2.18}$$

The weight vector w was calculated according to Eqn. 2.19 in order to both penalize reduced signal energy of the individual projections (i.e., instability across the ERP image under the projection angle) and to progressively suppress the contribution of projections with increasing absolute angle $|\phi|$ (i.e., attenuating noisy signal components). This is achieved by multiplying the angular energy distribution $e_\phi[p] = ||\boldsymbol{S}[\cdot,p]||_2^2$ of the discrete sinogram with a suitable angular support term. To preserve overall signal energy, the weight vector was normalized according to Eqn. 2.20.

$$w[p] = e_\phi[p]\cos(\phi[p])^2 \tag{2.19}$$

$$w'[p] = w[p]\,\frac{||e_\phi[\cdot]||_2^2}{||w[\cdot]||_2^2} \tag{2.20}$$

Fig. 3.10 shows the sinogram $\boldsymbol{S}$ for a single ERP image and 3.11 displays the angular energy distributions e_ϕ and corresponding normalized weights w' for the resampled reference AFRR ERP images. Throughout this manuscript, the algorithm outlined above will be referred to by the term Sinogram Domain Reweighting (SDR). Note that Beylkin

(1987) have independently proposed a method quite similar in nature within the context of geosignal analysis, which is based on their implementation of a discrete RT. In principle, the above algorithm is not restricted to the intended application at all. For example, by replacing w' with a normalized von Mises distribution $d_{vM}(\mu, \varkappa)$ (Fisher, 1993), a versatile, global isotropic denoising operator applicable under arbitrary angles $\mu \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ with adjustable filter strength $\varkappa \in [0, +\infty[$ can be trivially obtained.

2.4.3 Denoising Using Analytic Wavelet Filterbanks

The second proposed ERP image denoising algorithm was based on the Rational Analytic Discrete Wavelet Transform (RANDWT) and its inverse as introduced by Bayram (2013), which features a flexible time–frequency covering with arbitrary redundancy hence approximate shift invariance, thereby circumvening the major drawback of DWT–based approaches (e.g., Ahmadi and Quian Quiroga, 2013; Quian Quiroga and Garcia, 2003). This particular choice of transform was further motivated by its analytic time–scale atoms, enabling split processing of time–frequency resolved instantaneous amplitudes and phases, as well as its computationally efficient implementation as an iterated filterbank which enjoys the perfect reconstruction property.

A schematic representation of the proposed denoising algorithm is depicted in Fig. 2.3. Given a sampled ERP image $\boldsymbol{U} \in \mathbb{R}^{M \times N}$ containing the single responses as row vectors, multiscale representations were derived by separate decomposition of each single response into $J \in \mathbb{N}$ subbands with analytic detail coefficients $\boldsymbol{U_j}$, $j = 1 .. J$ and one real–valued approximation $\boldsymbol{U_a}$ using a RANDWT filterbank (Bayram, 2013). By choosing sampling factors $p/q = 7/9$, $r/s = 1/2$ and Q-factor 3, a decomposition with a moderate degree of redundancy ($R = 2.25$) was achieved using the maximum amount of $J = \lfloor (\ln(\frac{4s}{rN})) / (\ln(\frac{r}{s})) \rfloor = 19$ available subbands for the resampled AFRR single responses.

The analytic coefficients of each subband ERP image were then split up into instantaneous amplitudes $|\boldsymbol{U_j}|$ and complex phases $\boldsymbol{e}^{\boldsymbol{i}\varphi_j}$, which were separately processed prior to recombination. A normalized, unidirectional Gaussian smoothing operator operating along the single–response dimension with a standard deviation of $\sigma = 5$ was employed to separately regularize the instantaneous amplitude and phase structures of each subband[3].

Approximation coefficients $\boldsymbol{U_a}$ were independently smoothed across responses using the same operator. Finally, the inverse RANDWT was applied to the processed coefficients, yielding the filtered ERP image $\boldsymbol{U'}$. The underlying motivation for performing an *independent* regularization of the coefficients' amplitudes and phases is rather straightforward : As the event–related signal components are characterized by a substantial degree of morphological regularity across responses, manifesting itself by magnified phase stabilities in the transform domain (see, e.g., Corona-Strauss *et al.*, 2009; Low and Strauss, 2011; Klimesch *et al.*, 2014), it takes a considerably higher amount of additive noise to distort the instantaneous phases to the point of reliably destroying their characteristic structure than to deform the time–domain amplitudes beyond recoverability. Consequently, an amplitude–independent phase regularization preserves the characteristic temporal locations of the prominent waves to a higher degree than the simple expedient of smoothing the time–domain signals across single–responses (e.g., by UGM filtering). Throughout this manuscript, the algorithm outlined in this section will be referred to by the term Split Wavelet Amplitude & Phase Smoothing (SWAPS).

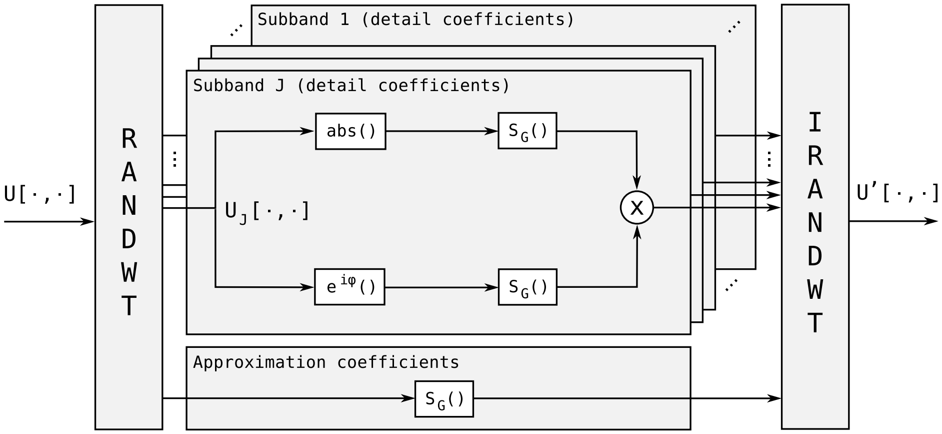

Figure 2.3: Schematic illustration of the SWAPS denoising algorithm.

---

[3] Note that preliminary experiments employing the iterative, isotropic phase denoising algorithm introduced by Villa *et al.* (2010) to regularize the instantaneous phase structures, while having been applied in Kohl *et al.* (2019b) with good results using an identical RANDWT signal decomposition of *uniformly* sampled ERP images, did not reveal any significant benefit over the faster, convolution–based Gaussian smoothing on the present data, hence the approach was abandoned within the scope of this work.

2.4.4 Comparative Performance Evaluation

Alongside the two proposed novel ERP image denoising algorithms, two of the established methods were selected for comparative performance evaluation : The UGM filter with a standard deviation of $\sigma = 5$ (to establish a baseline performance and to facilitate a direct comparison with the SWAPS algorithm using the same value of σ) and the NLM filter with a patch size of 10×10 (Strauss *et al.*, 2013), which has been previously shown to considerably outperform several alternative approaches outlined in Sect. 2.4.1.

All of the algorithms were separately used on the 15 resampled single–subject reference AFRR ERP images $\boldsymbol{U}$ obtained in Sect. 2.1.4, yielding filtered ERP images $\boldsymbol{U}'$. Fig. 3.12 shows the grand average of unfiltered reference ERP images, Fig. 3.13 the corresponding grand average waveform and Figs. 3.14 – 3.21 analogously show the grand average ERP images and waveforms after filtering with the four algorithms in question. To further investigate the dependence of algorithm performances on the SNR of the provided data in detail, a set of 25 synthetic ERP images each was generated for every SNR ranging from −18 dB to 12 dB in steps of 6 dB using the phenomenological AFRR model developed in Sect. 2.2 and processed analogously. Fig. 3.24 shows the grand average of synthetic ERP images at an SNR of –6 dB and Fig. 3.25 the corresponding grand average waveform, respectively.

In order to objectively quantify and compare the performance of the investigated algorithms, four different outcome metrics were calculated for each pair of unfiltered and filtered ERP images. The overall waveform reproducibility (i.e., robustness of average signal morphology) was quantified by the Correlation Waveform Index (CWI), i.e., the Pearson correlation (Upton and Cook, 2008)

$$\rho(u_1, u_2) = \frac{\sum_{n=1}^{N} (u_1[n] - \overline{u}_1)(u_2[n] - \overline{u}_2)}{\sqrt{\sum_{n=1}^{N} (u_1[n] - \overline{u}_1)^2} \sqrt{\sum_{n=1}^{N} (u_2[n] - \overline{u}_2)^2}} \tag{2.21}$$

calculated between the separately obtained average waveforms of odd and even single responses (Bernarding *et al.*, 2010) for each ERP image. Following the approach of Riedel *et al.* (2001), the RN of unfiltered and filtered ERP images under the signal–plus noise model (Dawson, 1954) was estimated by the sample–wise standard deviations across single responses

$$\sigma_u[\cdot] = \sqrt{\frac{\sum_{m=1}^{M} (\boldsymbol{U}[m,\cdot] - u_a[\cdot])^2}{M-1}} \tag{2.22}$$

which were further reduced into a single scalar value by means of the euclidian norm $RN_{est} = ||\sigma_u[\cdot]||_2$ as proposed by Strauss *et al.* (2013). Since the underlying waveform is known exactly for the synthetic AFRR in contrast to the real–world data, the similarity between the average waveforms and the underlying waveform (i.e., the degree of structure preservation) was additionally quantified for the unfiltered and filtered synthetic ERP images by means of their respective CWI. For every pair of unfiltered and filtered ERP images, the individual gains in each outcome metric due to the filtering process were calculated to facilitate objective comparison. Fig. 3.22 shows boxplots for the individual gains g_{CWI} in waveform reproducibility and Fig. 3.23 for the individual gains g_{RN} in the residual noise estimate for the reference AFRR. Likewise, Figs. 3.26 and 3.27 show these gains for the synthetic AFRR in dependence of their respective SNR. Boxplots of the individual gains g_{SIM} in average waveform similarity to the underlying waveform for the synthetic ERP data are given in Fig. 3.28. Lastly, the average execution times t_E exhibited by each of the investigated algorithms[4], which obviously are quite interesting to compare since they closely correlate with the involved computational costs, are displayed in Fig. 3.29. Given the rather unambiguous results of this evaluation (see Sect. 3.3), an additional analysis with regards to statistical significance and effect sizes was omitted for the sake of brevity.

---

[4] Note that the four investigated algorithms were implemented in MATLAB™ following established optimization practices (Altman, 2014) wherever possible. Execution times were measured on an available desktop workstation featuring a Sandy Bridge Intel Core i7 quad–core processor with 3.4 GHz fixed clock speed and 16 GB of RAM, running MATLAB™ R2019a on Microsoft Windows 10.

2.5 Fast Deconvolution–Based AFRR Acquisition

2.5.1 Preliminary Considerations

The results displayed in Sects. 3.1 and 3.2 confirm the general feasibility of simultaneously acquiring ABR, AMLR and ALR single responses to transient stimuli in one measurement using high sampling rates as first proposed by Michelini *et al.* (1982) and subsequently refined in Kohl and Strauss (2016). Despite the efforts outlined in the previous parts of this manuscript, reducing the computational load and memory demand associated with processing these signals by introducing a suitable resampling method, the rather large acquisition time still limits the applicability of the AFRR as a viable diagnostic tool. Currently, to obtain an acceptable SNR down to the fragile brainstem–level activity, at least 10^3 single trials must be acquired, requiring acquisition times of 15 min per condition owing to the large ISI of around 1 s, which in turn are necessary to reliably obtain cortical components. Therefore, AFRR acquisition remains potentially prone to signal degradation due to, e.g., electrode drifting. Nevertheless, the approach promises to improve the understanding of different processing stages along the auditory pathway and their dynamic interaction. Within this context, one key advantage of single–response AFRR recordings over the simple expedient of sequentially recording ABR, AMLR and ALR to transient stimuli at their respective optimum rates, which yields average ERP waveforms of good quality within smaller acquisition times (around 5 min per condition), resides in the fact that the obtained responses reflect the *undisturbed* propagation of neural activity along the entire auditory pathway following the solitary stimuli as close as possible, which is of paramount interest if the influence of slow, cortical–level phenomena such as inhibitional or habituational effects are to be studied together with earlier auditory processing stages.

Some past and more recent efforts have successfully demonstrated simultaneous acquisition of parts Özdamar and Kraus (1983) or the full range Holt and Özdamar (2014) of auditory ERP using considerably smaller ISI. Consequently, various alternative methods to eventually reduce acquisition time by disentangling overlapping ERP obtained at higher stimulation rates have been proposed by independent researchers. A notable early effort is given by the work of Woldorff (1993) on iterative estimation and subtraction of the overlap caused by adjacent responses (ADJAR), which has subsequently been advocated by Picton *et al.* (2000). The majority of contemporary efforts fall within the category of deconvolution–based methods such as least squares deconvolution (Bardy *et al.*, 2014a,b), maximum length sequence deconvolution (Bohórquez and Özdamar, 2006; Peng *et al.*, 2017a), Wiener filter deconvolution (Wang *et al.*, 2006) and Continuous Loop Averaging Deconvolution (CLAD) (Özdamar *et al.*, 2003; Delgado and Özdamar, 2004). Out of these, the computationally efficient, FFT–based CLAD method has proven particularly useful in simultaneously obtaining high–quality ABR and AMLR to transient stimuli at high presentation rates, although the obtained ALR components suffer from substantial rate–induced detriment (Holt and Özdamar, 2014, 2016). However, the loss of temporal resolution across the duration of the experiment inherent to deconvolution approaches as opposed to the prevalent single–trial based analysis poses a considerable drawback since many state–of–the–art ERP denoising methods (see Sect. 2.4.1 for a review) rely on single response representations in the form of ERP images (Jung *et al.*, 1999). Furthermore, measures quantifying the robustness of signal morphology across recording channels (Giroldini *et al.*, 2016) and (predominantly) across single responses, having been repeatedly proposed by independent researchers under the designations phase–locking factor (Tallon-Baudry *et al.*, 1996), phase–locking index (Brockhaus-Dumke *et al.*, 2008; Klimesch *et al.*, 2014) and phase (synchronization) stability (Corona-Strauss *et al.*, 2009; Low and Strauss, 2011), consistently outperformed amplitude average based measures on a regular basis in the respective ERP studies, thus emphasizing the additional diagnostic value of ERP single responses from a signal processing perspective.

In this part of the manuscript, a novel approach towards the fast acquisition of AFRR single–trial recordings using an interleaved, deconvolution–based stimulus presentation paradigm (Sect. 2.5.3) and associated signal processing methodology (Sect. 2.5.6) is introduced. Additionally, an efficient method for designing optimized stimulus sequences for ERP deconvolution based on Differential Evolution (DE) is proposed and evaluated (Sect. 2.5.2). In order to assess the novel AFRR acquisition approach, chirp–evoked auditory ERP (Fobel and Dau, 2004; Holt and Özdamar, 2014; Corona-Strauss *et al.*, 2009) were recorded from a group of 20 normal hearing subjects (Sect. 2.5.4, 2.5.5) and processed using the novel method. The reference AFRR dataset obtained in Sect. 2.1 was employed to facilitate an objective comparison in terms of waveform morphology and SNR advantage (Sect. 2.5.7). While the fast, interleaved acquisition method may not obtain *undisturbed* auditory activity to solitary transient stimuli like the considerably slower prior art recording paradigm (Sect. 2.1.3) does, the obtained results (Sect. 3.4) strongly suggest it to be a desirable alternative in obtaining quality full–range neural correlates of auditory activity within reasonable time wherever this strict requirement can be relaxed (e.g., in fast auditory screening and functional assessment applications).

2.5.2 Stimulus Sequence Optimization

Although brain dynamics can generally be considered highly nonlinear and time–variant, the model implicitly underlying deconvolution approaches in ERP acquisition is the linear superposition of single responses to subsequent stimulus presentations (Dawson, 1954; Kristensen *et al.*, 2017), which has been shown to hold sufficiently for the brainstem and middle–latency components of the auditory ERP (Holt and Özdamar, 2014; Holt, 2015; Bohórquez and Özdamar, 2008). As depicted in Fig. 2.4, the recorded ERP activity $r_n(t)$ due to stimulation with a stimulus onset train $s(t)$ is modeled as the output $r(t)$ of a causal Linear Time–Invariant (LTI) system with a compactly supported impulse response $h(t)$ (i.e., a Finite Impulse Response (FIR) filter) superimposed with additive noise $n(t)$ comprising spontaneous EEG background activity, physiological and technical artifacts.

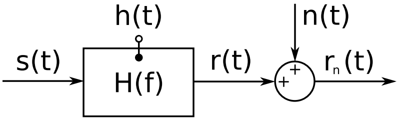

Figure 2.4: Linear superposition ERP model. The measured event–related response $r_n(t)$ of the auditory system to a train of transient stimulations $s(t)$ is composed of the causal, linear, time–invariant system response of H superimposed with an additive noise process $n(t)$.

Thus, the recorded ERP activity can be expressed as

$$r_n(t) = \underbrace{s(t) * h(t)}_{r(t)} + n(t), \quad h \in L^2(\mathbb{R}). \tag{2.23}$$

In the case of Continuous Loop Deconvolution (CLD), and subsequently CLAD (Delgado and Özdamar, 2004), $s(t)$ represents the periodic extension of a finite–length stimulus onset pattern and can thus be expressed as

$$s(t) = \sum_{l=1}^{k} \Delta_T \left(t - \sum_{m=1}^{l} t_{ISI}[m] \right), \tag{2.24}$$

where $t_{ISI} \in \mathscr{T}^k$ and $k \in \mathbb{N}$. $\mathscr{T} \subset \mathbb{R}_+$ denotes the range of ISI with admissible jitter, t_{ISI} denotes the sequence of ISI employed between the k subsequent stimulus presentations and $\Delta_T(t)$ denotes the Dirac comb with period $T = \sum t_{ISI}[\cdot]$. Consequently, a noisy estimate $h_e(t)$ of the underlying system response can be obtained by deconvolving the recorded ERP activity with the stimulus onset train :

$$h_e(t) = \underbrace{r(t) *^{-1} s(t)}_{h(t)} + \underbrace{n(t) *^{-1} s(t)}_{n_r(t)} \tag{2.25}$$

Eqn. (2.25) shows that the application of the deconvolution operator $*^{-1}s(t)$ additionally filters the noise superimposed on the recorded ERP activity, yielding a superposition of the underlying system response and a noise residual $n_r(t)$. It is straightforward to see from the frequency domain[5] equivalent

$$H_e(f) = H(f) + \frac{N(f)}{S(f)} \tag{2.26}$$

of Eqn. (2.25) that the magnitude of the reciprocal transfer function $S^{-1}(f)$ corresponding to the stimulus onset train quantifies the spectral amplification of additive noise introduced by the deconvolution operator (Özdamar and Bohórquez, 2006). Hence, using the ϵ–insensitive, quadratic error functional

---

[5] Note that strictly speaking, Eqn. 2.25 would have to be expressed in terms of the LT as $s(t)$ and $r(t)$ are not L^1 functions per definition. However, the frequency–domain interpretation can be employed without loss of generality as any real–world excitation will be of finite duration.

$$E(t_{ISI}) = \max\left(\frac{1}{|S(f)|^2} - 1, 0\right) \tag{2.27}$$

it is possible to optimize stimulus onset sequences t_{ISI} for noise attenuation in a designated frequency range of interest by solving the minimization problem

$$\min_{t_{ISI} \in \mathscr{T}^k} \left\| E\left(t_{ISI}\right) \right\|^2_{L^2(\mathscr{F})} \tag{2.28}$$

where $\mathscr{F} \subset \mathbb{R}_+$ denotes the range of frequencies considered for optimization. Note that depending on the actual choice of k, $\mathscr{T}$ and $\mathscr{F}$, this problem does not necessarily have one unique solution. Due to the morphologic dependence of auditory ERP on the ISI (e.g., Nelson *et al.*, 1997), particularly at high presentation rates (Holt and Özdamar, 2016), the range $\mathscr{T}$ generally has to be chosen as narrow as possible to allow for the applicability of a linear ERP activity superposition model at all. Meanwhile, the range $\mathscr{F}$ is typically desired to be as large as possible to capture all relevant details of the auditory ERP. As a result of these contradicting requirements, the error functional E becomes highly convoluted. In practice, $\mathscr{T}$ is additionally quantized by the greatest common divisor of the sampling rates used for stimulus presentation and ERP acquisition, further aggravating the search for satisfactory solutions t_{ISI}.

DE was introduced by Storn and Price (1997) as an iterative, global optimization metaheuristic over continuous spaces, which is based on vector populations of candidate solutions, mutation and recombination. It has since been successfully applied to a broad range of complex optimization problems in signal processing (e.g., Storn, 2007) and showed exceptional performance in benchmarks involving both highly convoluted and quantized functionals (Price *et al.*, 2005), thereby representing a natural choice for the minimization problem denoted in Eqn. (2.28). Consequently, Huang *et al.* (2014) proposed the optimization of CLD stimulus sequences by DE, albeit using an additive functional composed of the ISI Jitter Ratio (JR) and a Frequency–Domain Error (FDE) similar to Eqn. (2.27). This is intrinsically suboptimal due to the mutual exclusiveness of minimal JR and FDE, effectively hampering any choice of optimization algorithm by design. Subsequently, a solution space contraction variant of DE had to be employed for minimization to ensure convergence, involving a regularization factor to dynamically weight the contribution of JR and FDE across iterations. By contrast, the classical *DE/rand/1/bin* strategy as recommended by Price *et al.* (2005) was applied in this work with good results and fast convergence, minimizing the ϵ–insensitive FDE functional (2.27) over the designated frequency range $\mathscr{F}$ subject to the solution space constraint $t_{ISI} \in \mathscr{T}^k$ to limit the JR to an acceptable value, as denoted in Eqn. (2.28).

The parameters used for sequence optimization were chosen to obtain ABR and AMLR components at an SNR comparable to that of ALR, which is a desirable prerequisite for the subsequent signal processing (Sect. 2.5.6), particularly given the considerable temporal skew in SNR due to logarithmic resampling as demonstrated in Sect. 2.3.3. Note that it takes thousands of single responses to obtain reliable ABR averages, while ALR typically take less than 100 responses for the same data quality (Hall, 2007). Hence, this ratio is represented by the optimized sequence (see Fig. 3.31, top and Tab. A.2) comprising $k = 16$ stimuli. The admissible ISI range was set to $\mathscr{T} = [25, 43]$ ms, resulting in a mean presentation rate of $\overline{R} = 31.6$ Hz (Fig. 3.31, center) since previous work suggests rates around and below 40 Hz to be favorable in eliciting large amplitude AMLR components in the deconvolved responses due to a resonance effect (Galambos *et al.*, 1981; Özdamar *et al.*, 2007). A frequency range $\mathscr{F} = [20, 750]$ Hz corresponding to the main signal energy of waves V to N_b was designated for sequence optimization. The lower bound of $\mathscr{F}$ seems reasonable as waves past N_b were shown to be very ineffeciently (if at all) elicited by the CLD method at comparably high rates (Delgado and Özdamar, 2004). By contrast, its rather low upper bound constitutes a tradeoff, as waves I to III will be lost due to filtering, which is made up for by the strong and uniform gain in noise attenuation unachievable by broader ranges. The optimized sequence was obtained by applying 10^4 iterations of the *DE/rand/1/bin* algorithm to the minimization problem denoted in Eqn. 2.28 using a population of 10^3 candidate vectors, scaling factor bounds of 0.1 and 0.9, a crossover probability of 0.25 and the solution space constraint $\mathscr{T} = [25, 45]$ ms. The lower boundary of $\mathscr{T}$ corresponds to the target presentation rate of 40 Hz. Empirical experiments have shown DE to reliably converge towards solutions with noise attenuation (i.e., $|S^{-1}(f)| < 1$) across $\mathscr{F}$ if the upper boundary of $\mathscr{T}$ is chosen to be at least 1.8 times the lower boundary. Hence, the value was set to 45 ms. Consequently, the obtained sequence exhibits consistent noise attenuation (Fig. 3.31, bottom).

A major criterion for assessing ERP acquisition methods is the SNR of the resulting ERP, with it being common in the context of deconvolution methods to quantify the SNR advantage over conventional averaging by a suitable metric. Özdamar and Bohórquez (2006) estimated the SNR advantage due to deconvolution using the Noise Amplification Factor (NAF) c_{dec}, which was calculated using the ISI sequence. Peng *et al.* (2017b) recently proposed an improved NAF metric g_{dec} under the well justified assumption of an $1/f^{\alpha}$ model with $\alpha \approx 1$ underlying the additive EEG noise rather than a white process, which is well supported by previous findings of Ferree and Hwa (2003). In order to objectively quantify the SNR advantage of the obtained sequence, 28 available standard vertex–mastoid recordings of ongoing EEG acquired at 19.2 kHz in adult subjects were analyzed. 200 single responses bandlimited to the frequency range $\mathscr{F}$ were epoched and averaged for each single subject. Least–squares regressions of the $1/f^{\alpha}$ noise model to the individual average Power Spectral Densities (PSD) were performed to allow an accurate estimation of g_{dec} (see Fig. 3.30 for an example). Additionally, the empirical NAF g_{emp} was calculated based on the RMS ratio of raw and deconvolved averages.

2.5.3 Interleaved Stimulus Presentation Paradigm

Bidelmann (2015) recently proposed the interleaved acquisition of the Auditory Brainstem FFR and the ALR in a single measurement. In order to separately elicit both the FFR and ALR optimally, a stimulus presentation paradigm featuring alternating periods of clustered presentation at high rates and single presentations at low rates was developed. Slugocki *et al.* (2017) presented a similar method using short bursts of clicks as stimuli, yielding both a brainstem FFR and a cortical response to the stimulus cluster, which were separately analyzed. While providing simultaneous insight into auditory processing on the brainstem and cortical levels with short acquisition times and good SNR, the lack of temporal resolution inherent to the FFR–based approaches severely limits the possibilities of relating the brainstem level activity to the respective generators in detail. Presacco *et al.* (2010) evaluated a deconvolution method using bursts of clicks followed by long pauses as stimuli, incorporating the 200 ms post–stimulus EEG in the deconvolution process to improve the quality of middle–latency and late ERP components. The disadvantages of this approach reside in the lack of an analytical equation for the NAF (rendering it impractical for efficient sequence design using optimization algorithms) and the lack of temporal resolution across the experiment. Preliminary experiments of our own involving repeated presentation of a sequence with logarithmically decreasing successive ISI and conventional response averaging, while substantially increasing acquisition time, turned out to yield progressively distorted responses with no distinct cortical components, which could not be alleviated by subsequent application of the ADJAR algorithm (Woldorff, 1993) hence probably being largely due to detrimental rate effects.

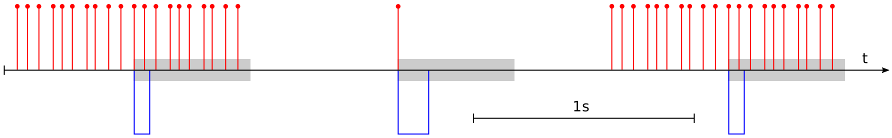

Figure 2.5: Schematic illustration of the interleaved stimulus presentation paradigm. Red stems : stimulus presentation times, blue lines : acquisition triggers, gray areas : acquisition time windows.

Here, a novel stimulus presentation paradigm following the reasoning of Bidelmann (2015) is introduced, interleaving high–rate and low–rate stimulation periods to alleviate their respective drawbacks. As depicted in Fig. 2.5, the high–rate periods feature two repetitions of the low–jitter CLD sequence optimized in Sect. 2.5.2 to elicit large amplitude ABR and AMLR components. Each first repetition elicits the overlap of preceding single responses necessary for the circular deconvolution of the succeeding sequence response with the stimulus onset train. Hence, only the responses to every second sequence presentation were acquired, separately deconvolved and arranged as an ERP image. This approach simultaneously yields the ABR and AMLR components at the SNR advantage of around 7 dB as determined in Sect. 2.5.2, while at the same time retaining the desired resolution across the duration of the experiment as opposed to the deconvolution of *averaged* responses to repeated stimulus sequence presentations (i.e., CLAD). However, as the CLD method has been previously shown to be rather ineffective at obtaining ALR components when using high stimulation rates (Holt and Özdamar, 2016) with the effective frequency range of the employed sequence being optimized accordingly, it is necessary to subsequently acquire the ALR using a separate stimulus presentation detached from the CLD sequence. Since a reduction of stimulation rate has been previously reported to progressively benefit the amplitudes of obtained ALR components down to 1/6 Hz (Davis *et al.*, 1966), ISI of 2 s were employed between the high–rate periods of

CLD sequence presentation and the single stimulus presentations to optimally elicit the ALR. While Budd and Michie (1994) reported the ISI range around 100 ms to be optimal in eliciting large–amplitude N_1 waves, potential detrimental effects on the other ALR components were not investigated, hence the large ISI were given preference in this study to reliably obtain undisturbed ALR. A uniformly distributed jitter of ± 10 % was added to the ISI to minimize the influence of habituational and refractory effects as well as to obtain the additional SNR advantage during averaging pointed out by Woldorff (1993). Acquisition triggers of different durations were associated with the alternating stimulation periods for response segmentation during postprocessing.

2.5.4 Subjects

The study was carried out at Saarland University of Applied Sciences (Saarbrücken, Germany). 20 subjects (11 males, 9 females, ages ∅33 ± 14 years) were recruited from the social environment of the authors to participate in the evaluation of the novel AFRR acquisition paradigm, none of which had a known history of audiological or neurological conditions. All of them were classified as normal hearing using puretone audiometry, with hearing thresholds below 15 dB hearing level within 0.5 – 8 kHz (Gelfand, 2009). Subjects were informed about the experimental procedures, which were designed in accordance with the Declaration of Helsinki and approved by the responsible authorities, and signed a consent form.

2.5.5 Stimulus Presentation and ERP Recording

To ensure comparability of the obtained results, the setup described in Sect. 2.1.3 (see Fig. 2.1) was reused for this study. The chirp stimulus optimized for eliciting the ABR in humans, as proposed by Fobel and Dau (2004), was used for stimulation (edge frequencies 0.1 – 10 kHz, 9.9 ms duration). Stimuli were presented binaurally to the subjects at 70 dB peak–to–peak equivalent sound pressure level using a USB audio interface (Scarlett 2i4, Focusrite, UK) and standard audiometry headphones (HDA–200, Sennheiser, Germany). The chirp stimulus was calibrated according to IEC 60645–3 (International Electrotechnical Commission, 2007). Stimulus waveforms and corresponding acquisition triggers were generated offline at 48 kHz sampling rate using MATLAB^{TM} R2016a and played back using multichannel audio workstation software (Studio One 2, PreSonus, USA). Stimulus waveforms comprised 100 repetitions of the interleaved presentation paradigm outlined in Sect. 2.5.3 for evaluation of the novel method. This amounts to 1700 net stimulus presentations and an overall acquisition time of 8:30 min, resulting in a relative speed improvement of 3.4 over the prior art reference AFRR recording featuring 10^3 stimuli presented at an ISI of 1 s ±10 % equally distributed jitter (16:50 min overall acquisition time). Stimulus polarity was alternated for each successive chirp to equally distribute electromagnetic influence from the headphones on the acquired ERP.

Subjects were instructed to sit on a lounger, relax with their eyes closed, but not fall asleep and avoid any movement during ERP acquisition. The auditory ERP were recorded using four passive Ag/AgCl electrodes placed on the left and right mastoids (A2/A1, active channels), the vertex (Cz, reference channel) and the forehead (ground) of the subjects. Impedances were kept below 5 kΩ. The raw auditory ERP and corresponding trigger signals were recorded without analog prefiltering at 19.2 kHz sampling rate using a medical grade biosignal amplifier and trigger conditioner (g.USBamp and g.Trigbox, g.tec Medical Engineering, Austria).

2.5.6 ERP Processing

ERP preprocessing and analysis were carried out in MATLAB^{TM} . ERP signals were zero–phase filtered using 4th order Butterworth IIR bandpass and notch filters (5 millioctaves notch bandwidth, centered at all integer multiples of 50 Hz within $\mathscr{F}$) to eliminate baseline wandering, high–frequency noise and powerline interference. Signals acquired for evaluation of the novel paradigm were conditioned twice to account for the necessary different bandpass filter settings ($\mathscr{F} = [20, 750]$ Hz for the CLD sequence responses, $\mathscr{F} = [1, 750]$ Hz for the subsequent single AFRR to preserve low–frequency ALR components). Reference AFRR recordings were bandlimited to $\mathscr{F} = [1, 750]$ Hz to facilitate objective comparison.

Separate single–trial ERP images (Jung *et al.*, 1999) were generated for the responses to CLD sequence and subsequent single AFRR stimulus presentations (evaluation of the novel paradigm) as well as for the reference AFRR recordings by epoching the conditioned signals according to the respective acquisition triggers, compensating the 9.9 ms delay introduced by the chirp stimulus. Responses exceeding a zero–to–peak amplitude threshold of 40 μV (CLD sequence responses) and 70 μV (subsequent single AFRR and reference AFRR) in any electrode channel were rejected as artifacts. The transient ABR/AMLR to each CLD sequence presentation were recovered from the sequence responses by FFT–based circular deconvolution with the stimulus onset train. ERP images containing more than 25 % artifact trials, no prominent average waveform morphology or postauricular muscle responses (O'Beirne and Patuzzi, 1999) exceeding 5 μV zero–to–peak were excluded from further analysis, leaving 15 (CLD ABR/AMLR and subsequent single AFRR) and 10 (reference AFRR) single–subject datasets. CLD ABR/AMLR and subsequent single AFRR ERP images were truncated to contain the greatest common subset of 85 artifact–free single trials as a prerequisite for the fusion algorithm proposed in the following paragraph. Reference AFRR ERP images were truncated separately to encompass the same amount of 863 single trials. ERP images were resampled on a logarithmic timebase with the range $[10^{-3}, 5 \cdot 10^{-1}]$ s using the Gaussian resampling operator (Sect. 2.3.2) with edge frequencies of 1.5 kHz and 10 Hz.

The above preprocessing steps effectively yield two resampled responses to each repetition of the interleaved high–rate/low–rate paradigm, with the components of interest (ABR/AMLR and ALR) being separately elicited close to optimally and roughly exhibiting the same SNR. To obtain the full–range auditory evoked response to each repetition of the paradigm, it is therefore necessary to fuse the signal content of both partial responses in a meaningful way. A straightforward approach to achieve this would be the temporal concatenation of the first response up to a certain time point (e.g., 20 ms) and the second response from that point onwards. However, this will result in sharp discontinuities, which could be alleviated in principle either by smoothing or by temporal crossfading (bearing the question how to choose the associated parameters optimally). To circumvent this problem, an alternative fusion algorithm is proposed, which will always produce smooth waveforms without the need to impose additional assumptions about the transition as it solely relies on information present within the responses themselves. The method is based on the reasoning that the event–related EEG activity constitutes a regularity of specific signal parts localized in time and scale across adjacent trials, which is backed quite well by previous findings involving measures quantifying the associated phase stability (e.g., Corona-Strauss *et al.*, 2009; Low and Strauss, 2011; Tallon-Baudry *et al.*, 1996; Klimesch *et al.*, 2014). Hence, it makes sense to quantify the *„event–relatedness"* of all signal components by their cross–trial regularity and obtain weighted averages based on this measure given its high correlation with the desired average waveform reproducibility. To achieve this, a multiresolution signal processing scheme based on the RANDWT and its inverse as introduced by Bayram (2013) was employed. This particular choice of transform pair was motivated by its flexible, constant–Q time–frequency covering with adjustable redundancy, which is necessary for the obtained analytic transform coefficients to be approximately shift invariant (i.e, largely unaffected by signal jitter) as well as robustly invertible after substantial manipulation. Furthermore, a computationally efficient implementation is given in the form of an iterated filterbank with perfect reconstruction. Starting from each ERP image $\boldsymbol{U} \in \mathbb{R}^{M \times N}$ containing $M \in \mathbb{N}$ single trials as row vectors, multiresolution representations of $\boldsymbol{U}$ were derived by decomposing the single trials into $J \in \mathbb{N}$ subbands with analytic detail coefficients $\boldsymbol{U_j} \in \mathbb{C}^{M \times N_j}$, $j = 1..J$ and one real–valued approximation $\boldsymbol{U_a} \in \mathbb{R}^{M \times N_a}$ using a RANDWT filterbank. By choosing sampling factors $p/q = 9/10$, $r/s = 1$ and Q–factor 1.5, a decomposition with substantial redundancy $R = 10$ was achieved using the maximum amount of $J = \lfloor \ln\left(\frac{4s}{rN}\right) / \ln\left(\frac{r}{s}\right) \rfloor = 40$ available subbands. The analytic coefficients of each subband ERP image $\boldsymbol{U_j}$ were split into amplitudes $|\boldsymbol{U_j}|$ and phases $\varphi_{\boldsymbol{j}}$. Examining the cross–trial phase stability $\Gamma(\varphi_{\boldsymbol{j}})$, exemplarily shown for the resampled CLD ABR/AMLR and subsequent single AFRR ERP images of a single subject in Fig. 3.32, it is obvious to see by comparing the time–scale distributions that Γ indeed assumes its highest values around the temporal occurences of either ABR/AMLR (Fig. 3.32, top) or ALR (Fig. 3.32, bottom) components in support of the initial assumption.

Since the deviant key areas of Γ in both distributions indicate a rather clear separability of ABR/AMLR and ALR components by cross–trial phase synchronization, the localized phase stabilities $\Gamma_\sigma(\varphi_{\boldsymbol{j}})$ across small, adjacent neighbourhoods of trials (i.e., restricted by a Gaussian window with a standard deviation of $\sigma = 5$) is proposed as a suitable weight factor to form a partition of unity for the recombination of subband detail coefficients $\boldsymbol{U_j}$:

$$\boldsymbol{U}'_{\boldsymbol{j}}[m,\cdot] = \frac{\boldsymbol{U}^{\mathbf{1}}_{\boldsymbol{j}}[m,\cdot]\,\Gamma_\sigma(\varphi^{\mathbf{1}}_{\boldsymbol{j}})[m,\cdot] + \boldsymbol{U}^{\mathbf{2}}_{\boldsymbol{j}}[m,\cdot]\,\Gamma_\sigma(\varphi^{\mathbf{2}}_{\boldsymbol{j}})[m,\cdot]}{\Gamma_\sigma(\varphi^{\mathbf{1}}_{\boldsymbol{j}})[m,\cdot] + \Gamma_\sigma(\varphi^{\mathbf{2}}_{\boldsymbol{j}})[m,\cdot]} \tag{2.29}$$

In Eqn. (2.29), $\boldsymbol{U}_j^{\mathbf{1}}$ and $\boldsymbol{U}_j^{\mathbf{2}}$ denote the subband detail coefficients of CLD ABR/AMLR and subsequent single AFRR ERP images, $\Gamma_\sigma(\varphi_j^{\mathbf{1}})$ and $\Gamma_\sigma(\varphi_j^{\mathbf{2}})$ the associated localized phase stabilities and $\boldsymbol{U}_j'$ the recombined detail coefficients. Because the associated approximation coefficients $\boldsymbol{U}_a^{\mathbf{1}}$ and $\boldsymbol{U}_a^{\mathbf{2}}$ lack a cross–trial stability metric comparable to Γ due to being real–valued, a simple linear transition

$$\boldsymbol{U}_a'[m,\cdot] = \frac{(N_a - \cdot)\, \boldsymbol{U}_a^{\mathbf{1}}[m,\cdot] + (\cdot - 1)\, \boldsymbol{U}_a^{\mathbf{2}}[m,\cdot]}{N_a - 1} \tag{2.30}$$

was performed in order to smoothly interpolate approximation coefficients from the CLD ABR/AMLR to the subsequent single AFRR. Recombined approximation and detail coefficients were normalized (the latter separately for each subband) to preserve overall signal energy :

$$\boldsymbol{U}_a''[m,\cdot] = \boldsymbol{U}_a'[m,\cdot]\, \frac{||\boldsymbol{U}_a^{\mathbf{1}}[m,\cdot]||_2^2 + ||\boldsymbol{U}_a^{\mathbf{2}}[m,\cdot]||_2^2}{||\boldsymbol{U}_a'[m,\cdot]||_2^2} \tag{2.31}$$

$$\boldsymbol{U}_j''[m,\cdot] = \boldsymbol{U}_j'[m,\cdot]\, \frac{||\boldsymbol{U}_j^{\mathbf{1}}[m,\cdot]||_2^2 + ||\boldsymbol{U}_j^{\mathbf{2}}[m,\cdot]||_2^2}{||\boldsymbol{U}_j'[m,\cdot]||_2^2} \tag{2.32}$$

The fused single–trial AFRR ERP image $\boldsymbol{U}'$ was then obtained by applying an inverse RANDWT filterbank to project the recombined, normalized coefficients $\boldsymbol{U}_j''$ and $\boldsymbol{U}_a''$ back into time domain.

2.5.7 Comparative Statistical Analysis

For each single–subject ERP image (i.e., CLD ABR/AMLR, subsequent single AFRR, fused AFRR and reference AFRR), time–domain amplitude averages u_a and the sample–wise cross–trial standard deviations σ_u (Eqn. 2.22) were calculated, with all amplitude averages being additionally detrended by subtracting least–squares fit 2nd order polynomials to reliably eliminate potential slow drifts across investigated timescales. Likewise, the grand average of ERP images, corresponding detrended amplitude averages and sample–wise standard deviations were calculated based on the single–subject datasets. In order to objectively compare the data quality of ERP obtained using the novel and prior art acquisition approaches, an empirical SNR estimate under the signal–plus–noise model (Dawson, 1951) given by

$$SNR_{est} = \frac{|| u_a[\cdot] ||_2}{|| \sigma_u[\cdot] ||_2} \tag{2.33}$$

was calculated for each single–subject ERP image, separately for the entire signal length and temporally localized to the occurences of ABR ($[1, 10]$ ms), AMLR ($[10, 50]$ ms) and ALR ($[50, 300]$ ms) (Hall, 2007). Note that due to u_a and σ_u both being normalized by M, SNR_{est} is independent of the number of trials present in the investigated ERP image, hence its values are comparable between fused and reference AFRR. To quantify the consistency of SNR improvement due to the proposed fusion of CLD ABR/AMLR with subsequent single AFRR trials, individual SNR gains were calculated and the statistical significance of SNR increase was evaluated using a one–sided Wilcoxon–Mann–Whitney U–test. Likewise, SNR estimates calculated from the fused AFRR ERP images yielded by the novel acquisition approach and ERP images obtained from the reference AFRR recordings were assessed with a one–sided U–test to evaluate the significance of SNR improvement of the novel method over the prior art. Exact values of prominent wave peak latencies t_p and amplitudes u_p were derived semi–automatically from the single–subject ERP average waveforms u_a using a gradient method initialized with approximate values obtained by point & click. Peak latencies and amplitudes of the fused AFRR and reference AFRR averages were analyzed for significant differences by means of a two–sided U–test.

3 Results

3.1 Reference AFRR Acquisition

Fig. 3.1 shows the grand average ERP waveform of the 11 screened single–subject reference AFRR recordings. Gray areas denote the standard deviation across subjects within the dataset. Fig. 3.2 displays the time–frequency resolved real part of grand average CWT coefficients obtained from the reference AFRR single responses. Note that despite being displayed using a logarithmic time axis to better reveal the signal morphology for the reasons outlined in Sect. 2.3.1, the underlying signals are uniformly sampled at 19.2 kHz sampling rate.

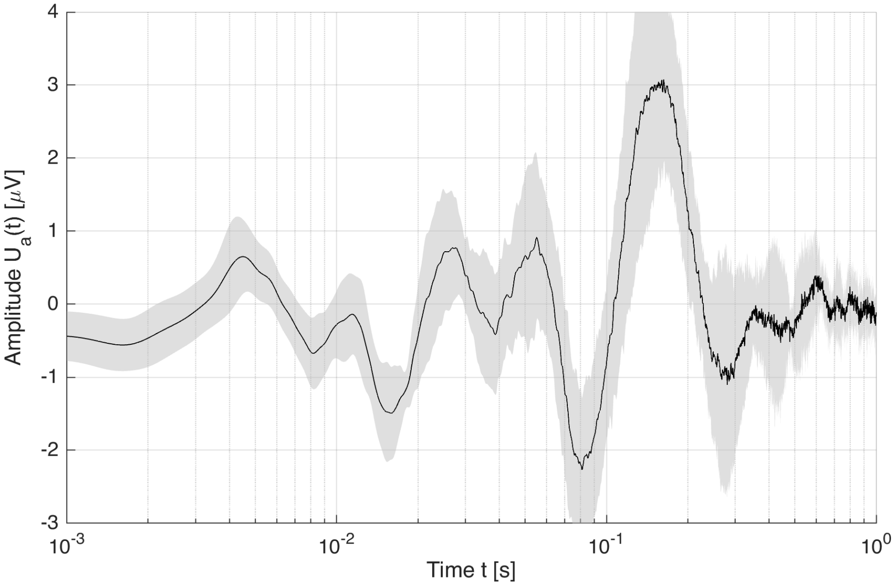

Figure 3.1: Grand average ERP waveform of reference AFRR recordings.

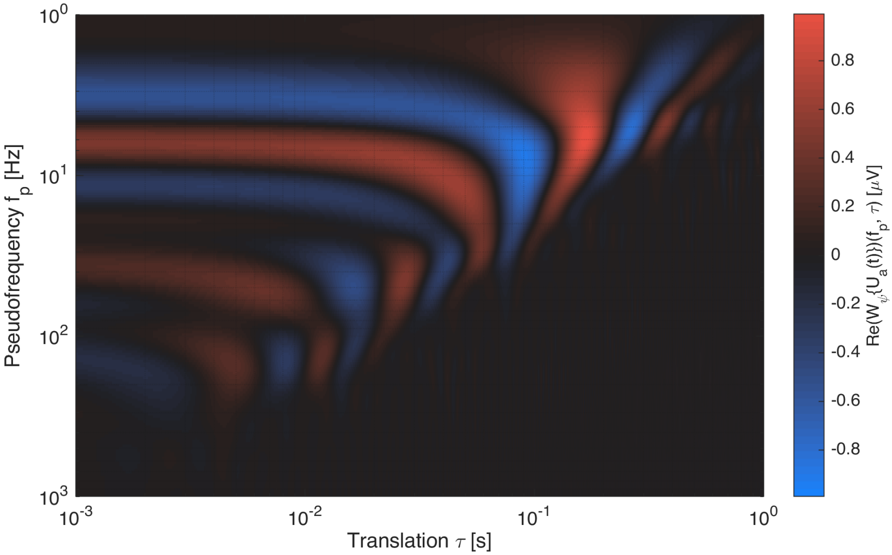

Figure 3.2: Real part of grand average reference AFRR CWT coefficients.

3.2 Optimization of AFRR Signal Representation

Fig. 3.3 displays the magnitudes of the grand average ST coefficients for the uniformly sampled single–subject reference AFRR averages. Fig. 3.4 shows the grand average of time–frequency resolved WPS obtained from the associated single–subject ERP images, respectively. Figs. 3.5 and 3.6 display the time–scale resolved grand average CWT real part and corresponding WPS for the single–response representations resampled on a logarithmic timebase as described in Sect. 2.3.2. The decay profiles of estimate RN for the uniformly sampled and resampled single–response representations and their associated regression statistics are given in Fig. 3.7 (RN as a function of trial count for the synthetic EEG noise), Fig. 3.8 (RN as a function of trial count for the stationary 50 Hz noise) and Fig. 3.9 (RN as a function of time post–stimulus for the synthetic EEG noise), respectively.

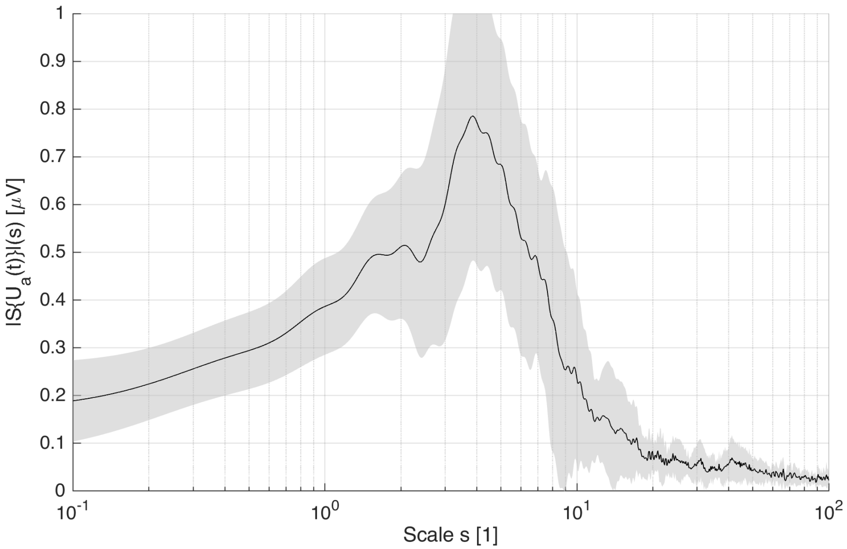

Figure 3.3: Magnitude of grand average AFRR ST coefficients.

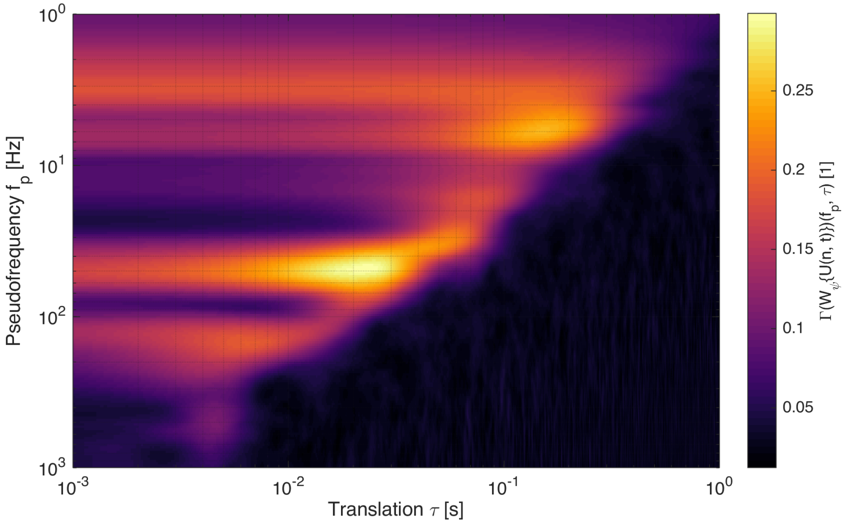

Figure 3.4: Grand average of WPS for the uniformly sampled AFRR ERP images.

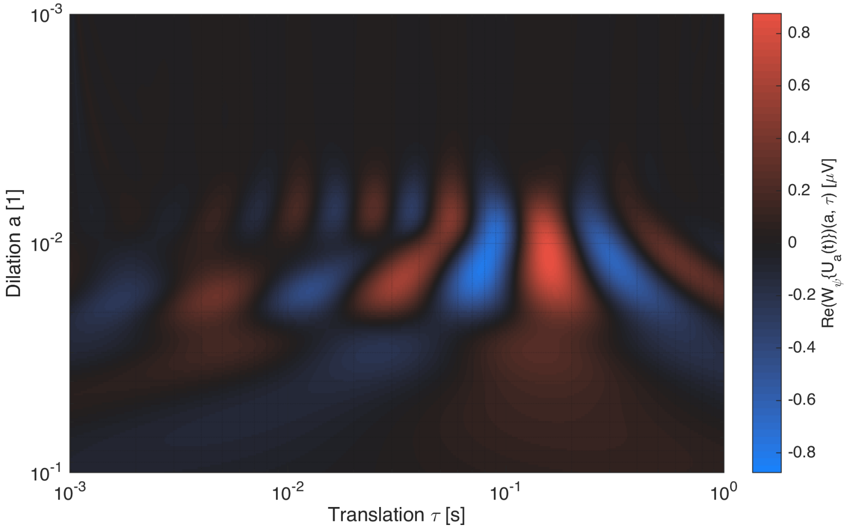

Figure 3.5: Real part of grand average resampled AFRR CWT coefficients.

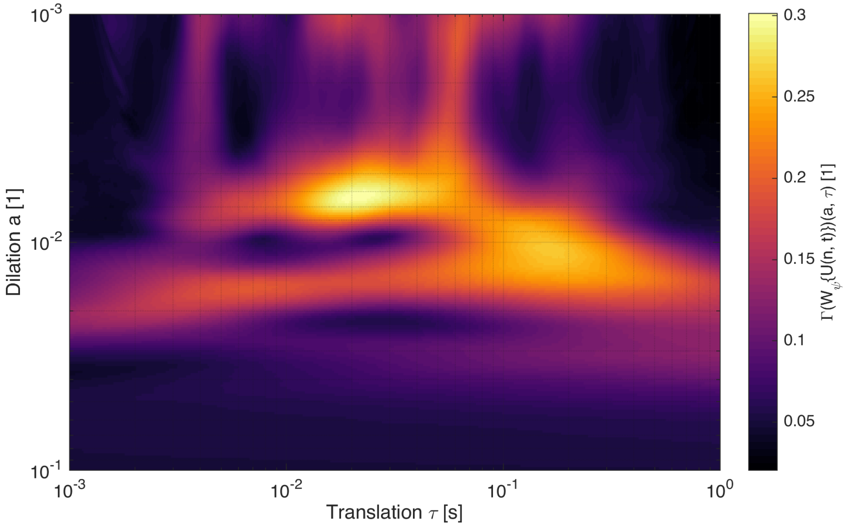

Figure 3.6: Grand average of WPS for the resampled AFRR ERP images.

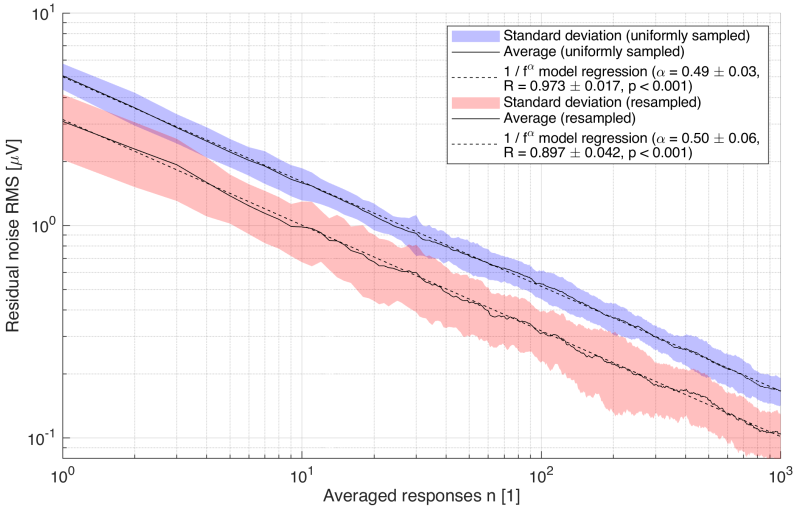

Figure 3.7: Decay profiles and model regressions of RN as a function of trial count (synthetic EEG noise).

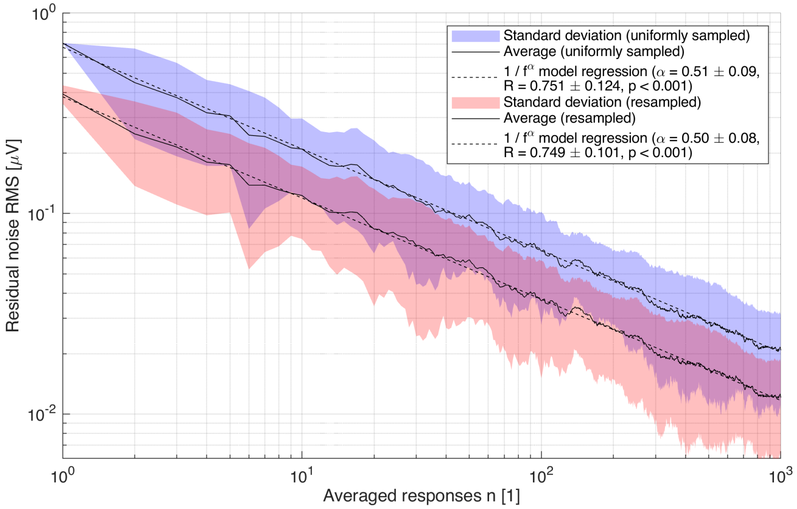

Figure 3.8: Decay profiles and model regressions of RN as a function of trial count (stationary 50 Hz noise).

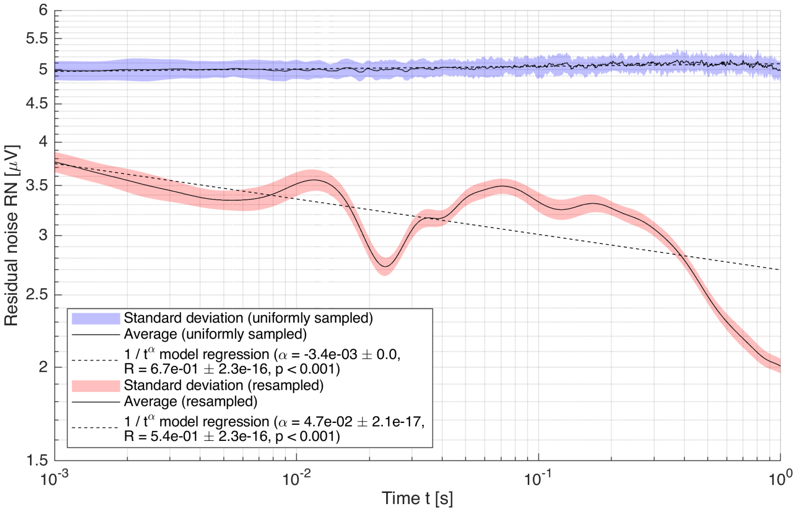

Figure 3.9: Decay profiles and model regressions of RN over time post–stimulus (synthetic EEG noise).

3.3 Denoising of AFRR Single Responses

Fig. 3.10 exemplarily shows the sinogram of a resampled single–subject ERP image. Fig. 3.11 displays the angular energy distributions e_ϕ (blue) and normalized weight vectors w' (red) for the resampled reference AFRR ERP images. Coloured areas denote the standard deviation across subjects. Fig. 3.12 shows the grand average of unfiltered reference ERP images, Fig. 3.13 the corresponding grand average waveform and Figs. 3.14 – 3.21 analogously show the grand average ERP images and waveforms after filtering with the four algorithms in question. Likewise, Fig. 3.24 displays the grand average of synthetic ERP images (SNR = -6 dB) and Fig. 3.25 the corresponding grand average ERP waveform. Note that here, gray areas denote the standard deviation *across single responses*. Fig. 3.22 shows boxplots for individual gains g_{CWI} in waveform reproducibility and Fig. 3.23 for individual gains g_{RN} in the residual noise estimate. Likewise, Figs. 3.26 and 3.27 show these gains for the synthetic AFRR in dependence of their respective SNR. Boxplots of individual gains g_{SIM} in average waveform similarity to the underlying waveform for the synthetic ERP data are given in Fig. 3.28. In all boxplots, circles denote the arithmetic mean values. The average algorithm execution times t_E are displayed in Fig. 3.29.

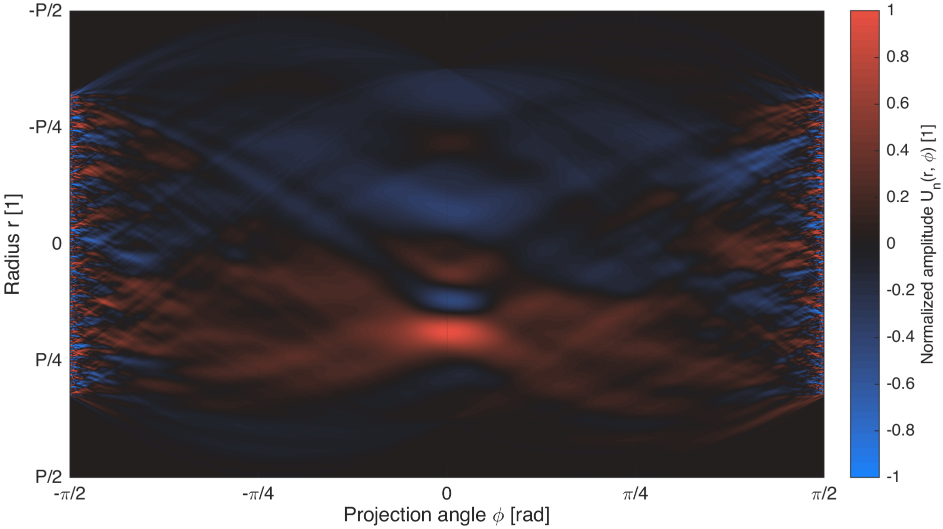

Figure 3.10: Example sinogram of resampled single–subject ERP image.

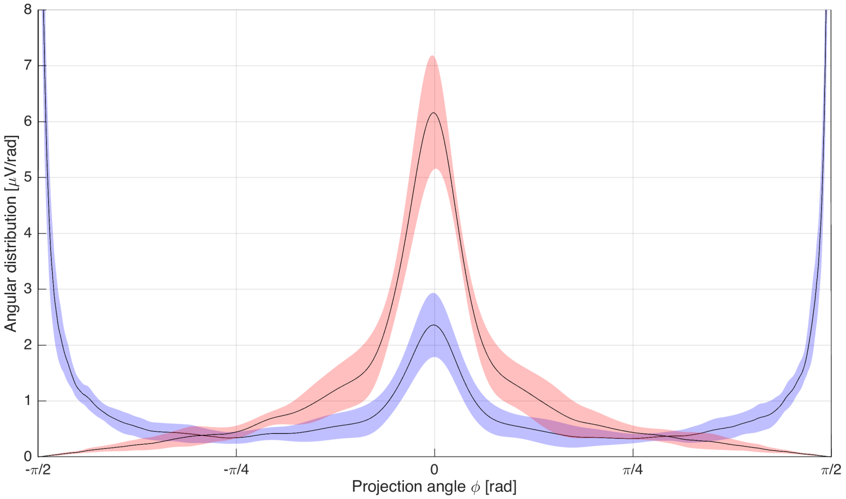

Figure 3.11: Angular energy distributions e_ϕ (blue) and normalized weight vectors w' (red) of resampled reference AFRR dataset.

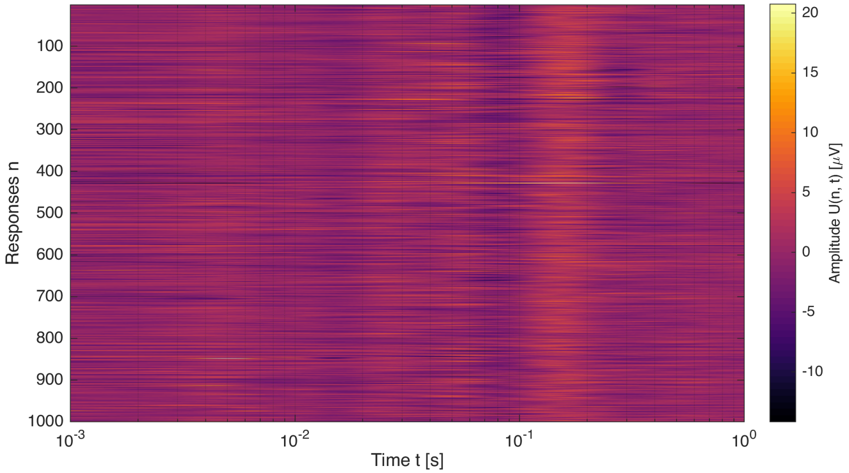

Figure 3.12: Grand average ERP image of resampled reference AFRR dataset (unfiltered).

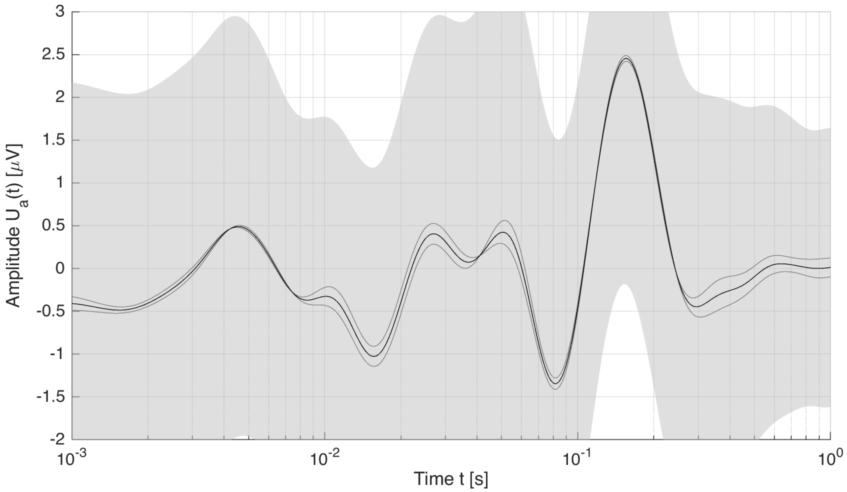

Figure 3.13: Grand average ERP waveform of resampled reference AFRR dataset (unfiltered).

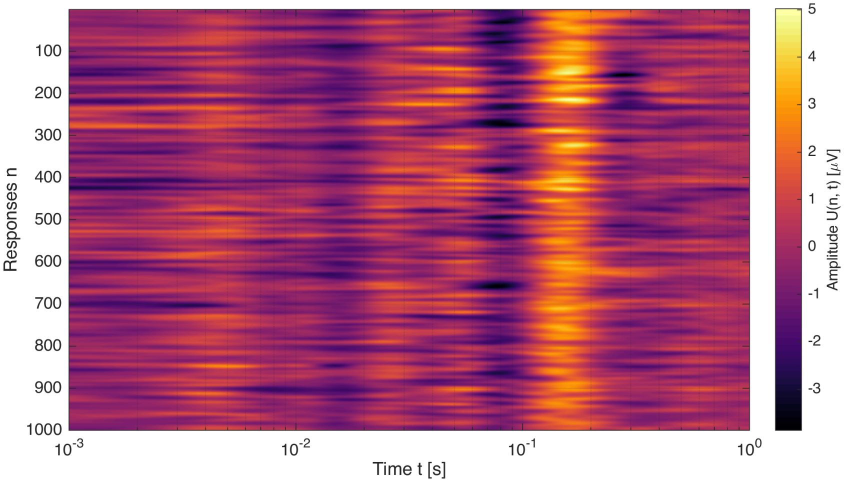

Figure 3.14: Grand average ERP image of resampled reference AFRR dataset (UGM filtered).

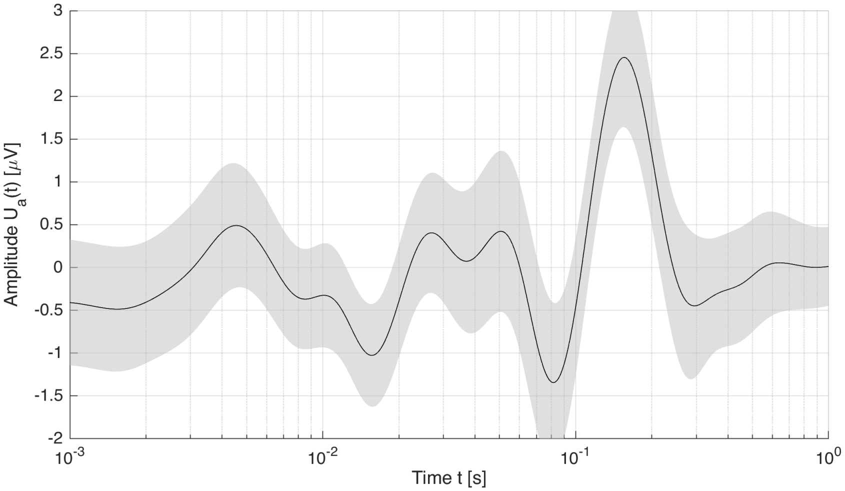

Figure 3.15: Grand average ERP waveform of resampled reference AFRR dataset (UGM filtered).

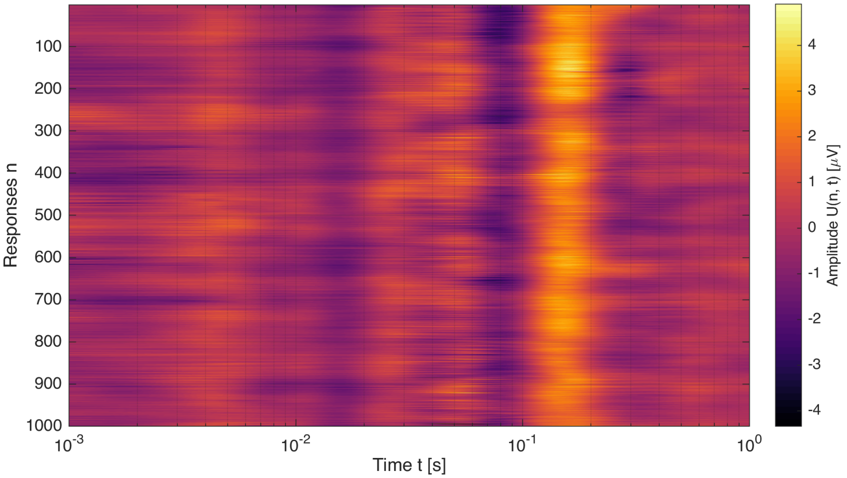

Figure 3.16: Grand average ERP image of resampled reference AFRR dataset (NLM filtered).

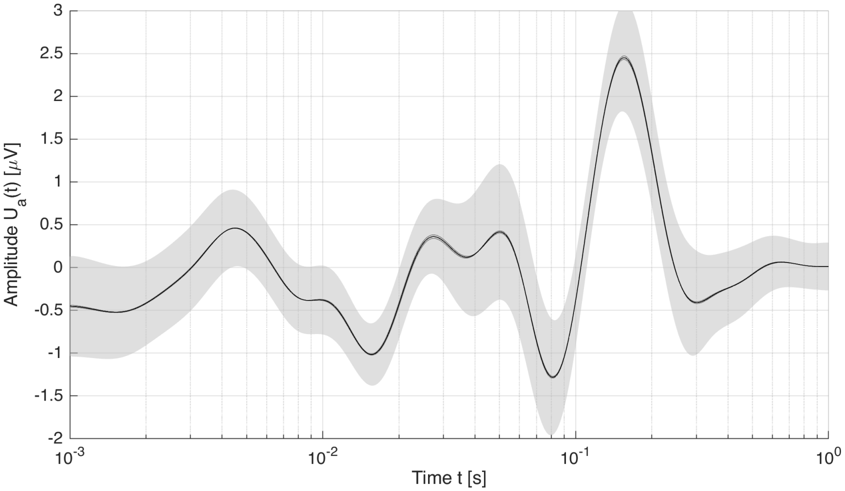

Figure 3.17: Grand average ERP waveform of resampled reference AFRR dataset (NLM filtered).

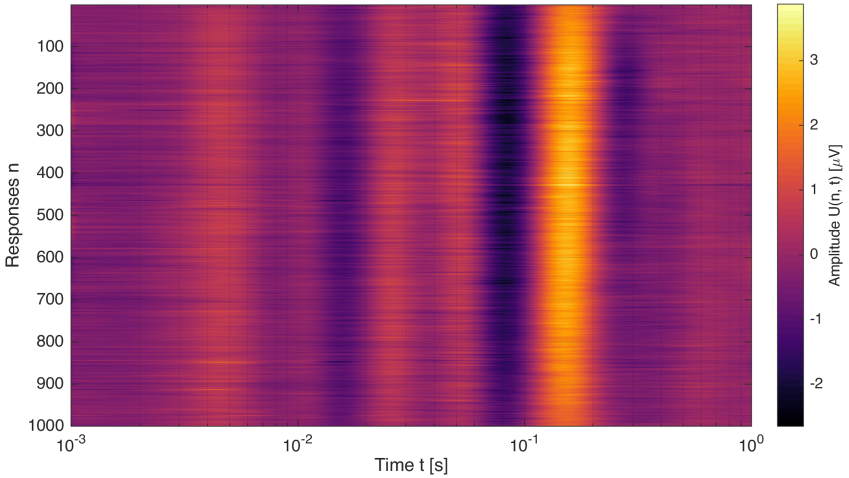

Figure 3.18: Grand average ERP image of resampled reference AFRR dataset (SDR filtered).

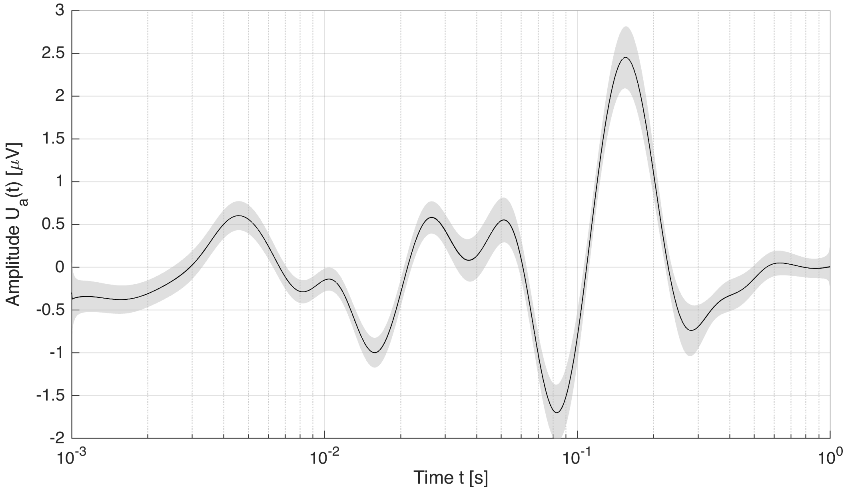

Figure 3.19: Grand average ERP waveform of resampled reference AFRR dataset (SDR filtered).

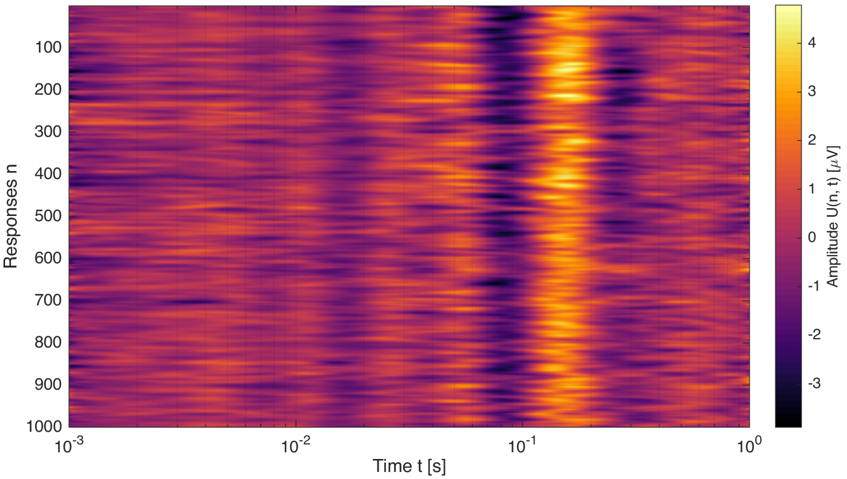

Figure 3.20: Grand average ERP image of resampled reference AFRR dataset (SWAPS filtered).

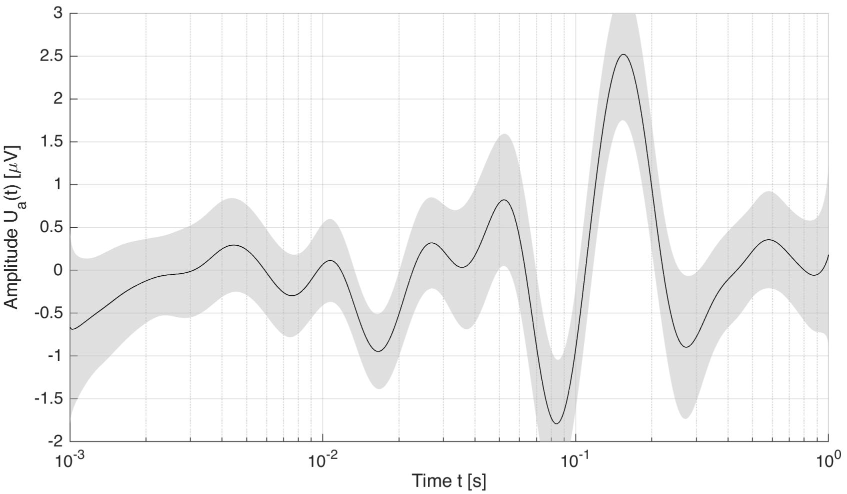

Figure 3.21: Grand average ERP waveform of resampled reference AFRR dataset (SWAPS filtered).

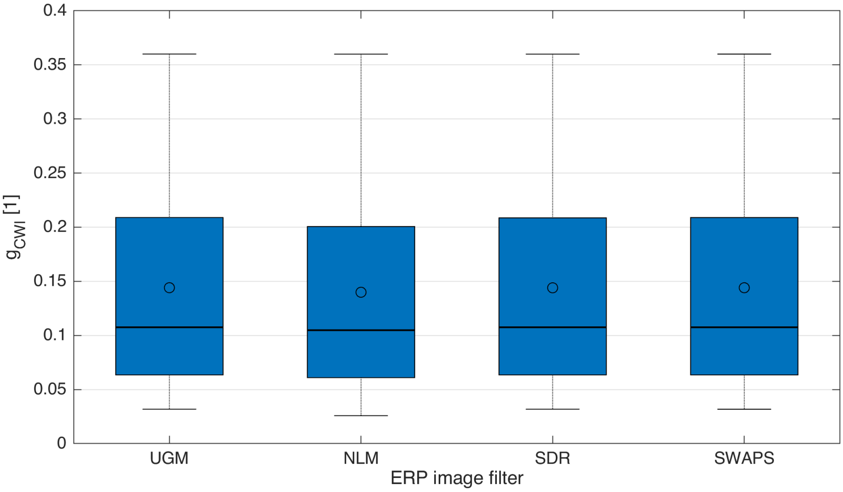

Figure 3.22: Boxplots of individual gains g_{CWI} in waveform reproducibility for the investigated ERP image filters (reference AFRR).

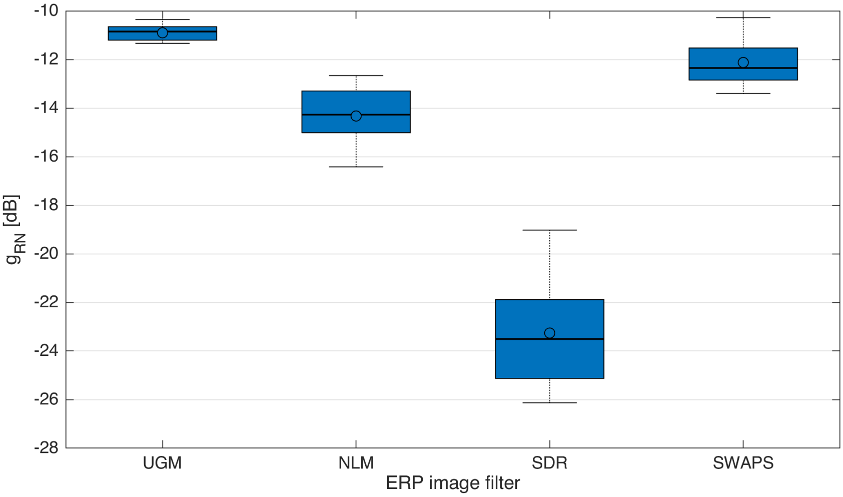

Figure 3.23: Boxplots of individual gains g_{RN} in RN for the investigated ERP image filters (reference AFRR).

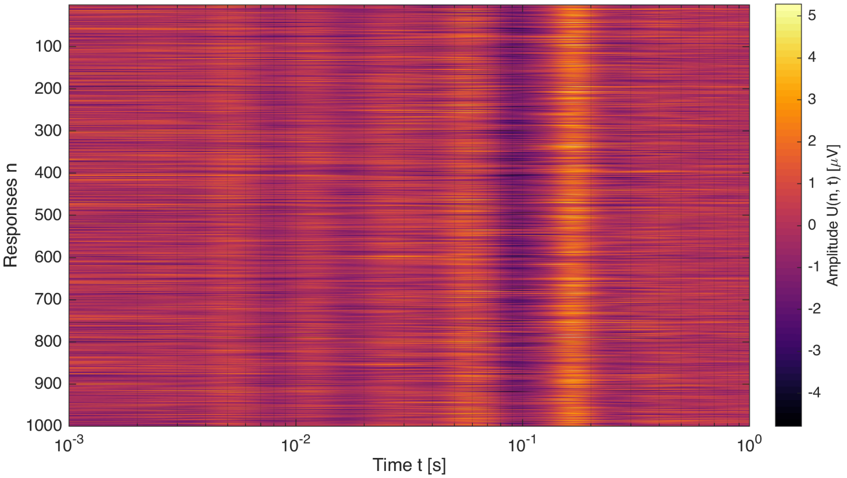

Figure 3.24: Grand average ERP image of resampled synthetic AFRR dataset at −6 dB SNR (unfiltered).

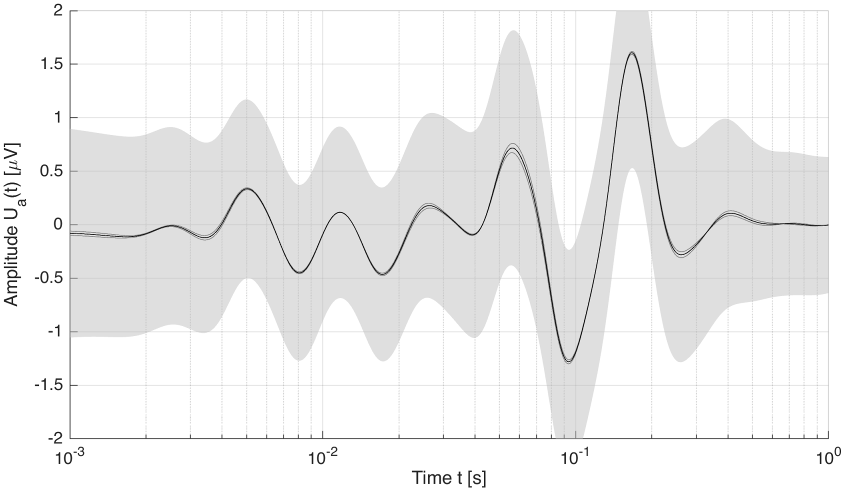

Figure 3.25: Grand average ERP waveform of resampled synthetic AFRR dataset at −6 dB SNR (unfiltered).

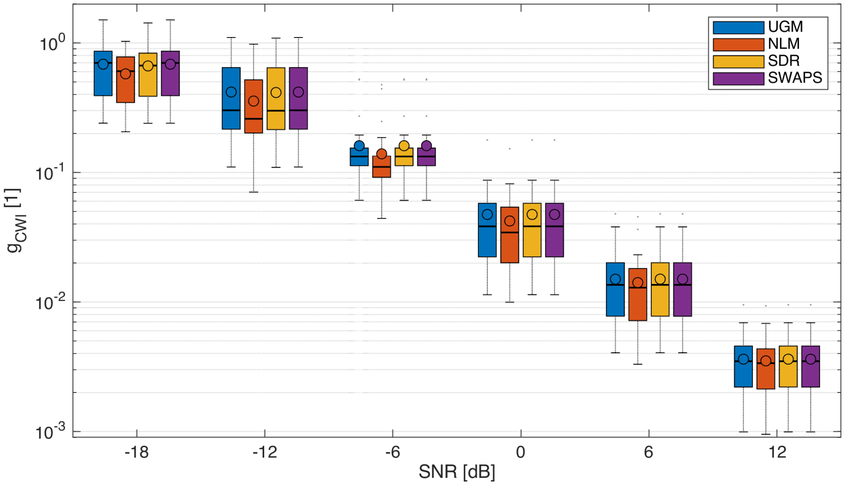

Figure 3.26: Boxplots of individual gains g_{CWI} in waveform reproducibility for the investigated ERP image filters (synthetic AFRR).

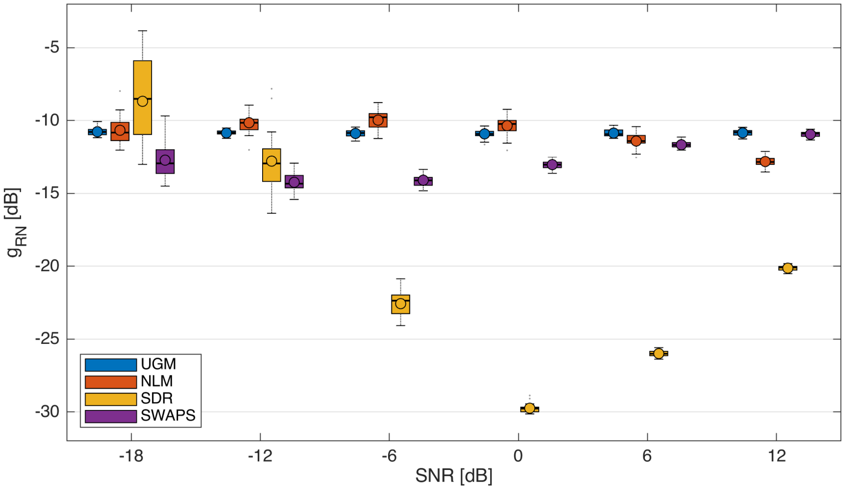

Figure 3.27: Boxplots of individual gains g_{RN} in RN for the investigated ERP image filters (synthetic AFRR).

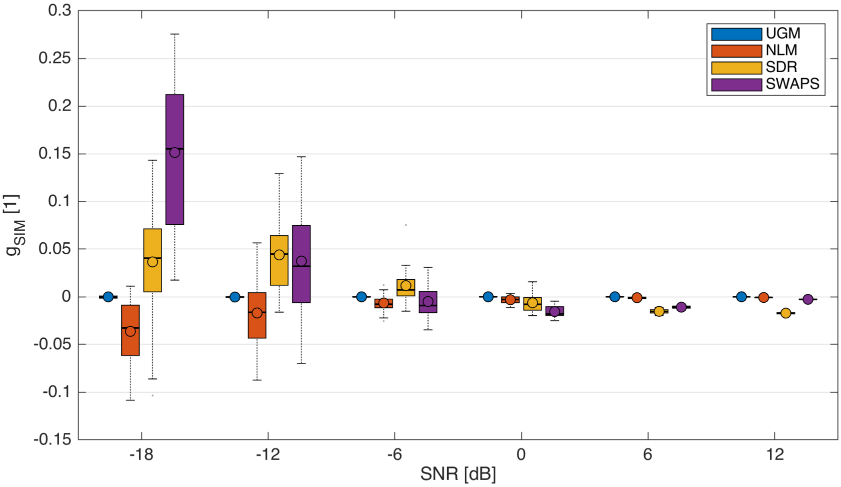

Figure 3.28: Boxplots of individual gains g_{SIM} in structural similarity for the investigated ERP image filters (synthetic AFRR).

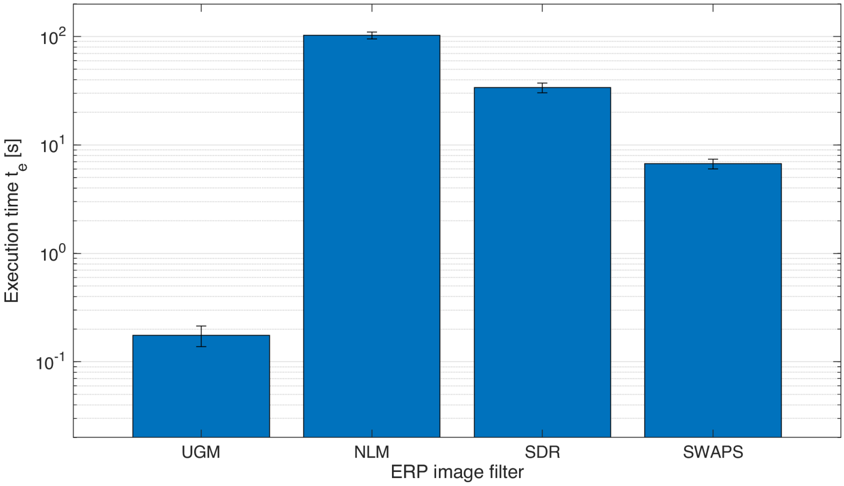

Figure 3.29: Average execution times of the investigated ERP image filters.

3.4 Fast Deconvolution–Based AFRR Acquisition

Fig. 3.31 shows the properties of the stimulus ISI sequence optimized for CLD. The estimate SNR advantage due to deconvolution as proposed by Özdamar and Bohórquez (2006) assumes a value of $c_{dec} = 7.87$ dB for the optimized sequence. Fig. 3.30 displays an example single–subject NAF estimation and the associated regression statistics. Least squares regressions of the $1/f^{\alpha}$ noise model to the individual average PSD resulted in $\alpha = \varnothing 1.01 \pm 0.27$ in support of Peng *et al.* (2017b) and $g_{est} = \varnothing 7.51 \pm 0.12$ dB. The empirical NAF based on the RMS ratio of raw and deconvolved averages was calculated to be $g_{emp} = \varnothing 7.46 \pm 0.52$ dB. Hence, both theoretical and empirical results strongly suggest a consistent SNR advantage of no less than 7 dB for the obtained CLD sequence.

A single–subject example of the cross–trial RANDWT phase stability Γ is given in Fig. 3.32 for the deconvolved CLD ABR/AMLR (Fig. 3.32, top) and subsequent single AFRR (Fig. 3.32, bottom) ERP images, respectively. Fig. 3.33 shows the grand average of single–trial CLD ABR/AMLR, subsequent single AFRR, fused AFRR and reference AFRR ERP images obtained from A2 electrode recordings. Fig. 3.34 displays the corresponding average amplitude waveforms (black lines) together with their respective cross–trial standard deviations (gray areas). Separate averages of split datasets containing the odd and even single trials (dark gray lines) were additionally included for qualitative assessment of waveform reproducibility.

Boxplots of the employed SNR estimate (Eqn. 2.33) calculated separately for the four temporal ranges (entire signal, ABR, AMLR and ALR) in each single–subject AFRR ERP image (recordings at A2) are displayed in Fig. 3.35. Boxplots of individual gain in estimate SNR due to the proposed fusion of CLD ABR/AMLR with subsequent single AFRR trials are given in Fig. 3.36. The p–values and effect sizes r corresponding to the one–sided U–test for significant increase in individual SNR estimates between subsequent single AFRR and fused AFRR ERP images are given in Tab. A.4. Likewise, values of p and r for the one–sided U–test between individual SNR estimates of fused AFRR and reference AFRR ERP images are displayed in Tab. A.5. Fig. 3.37 shows boxplots of the peak latencies t_p and amplitudes u_p obtained from fused AFRR and reference AFRR average waveforms (A2 recordings). In all given boxplots, arithmetic mean values are additionally denoted by circles.

According to the conducted two–sided U–tests, no significant differences between peak latencies of fused AFRR and reference AFRR average waveforms were found ($p \gg 0.05$). Peak amplitudes exhibit significant differences for waves N_a, P_a, N_1 and P_3. Hence, additional one–sided U–tests were performed individually for each of the waves in question (Tab. A.6). Note that effect sizes for wave P_3 cannot be provided due to the small number of 6 overall observations in the individual reference AFRR averages.

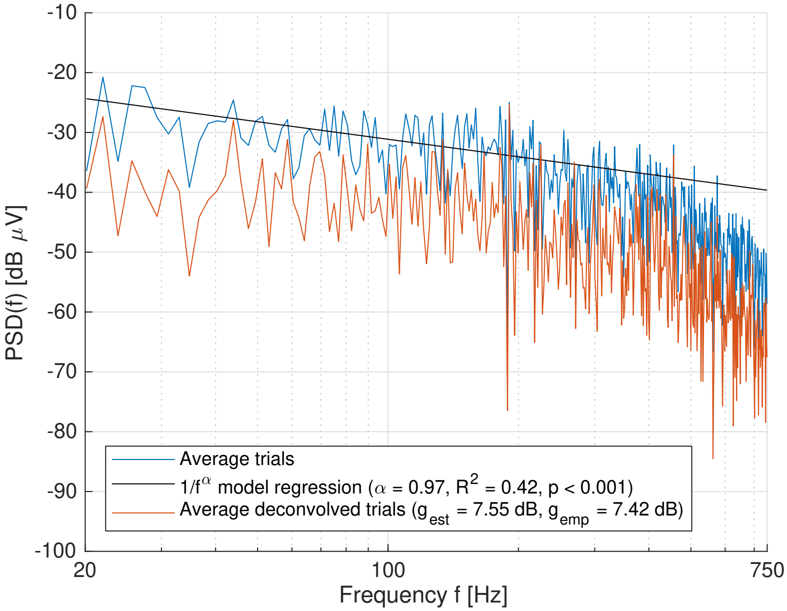

Figure 3.30: Example single–subject NAF estimation. Blue line : PSD of the raw average, black line : $1/f^{\alpha}$ noise model regression, red line : PSD of the deconvolved average.

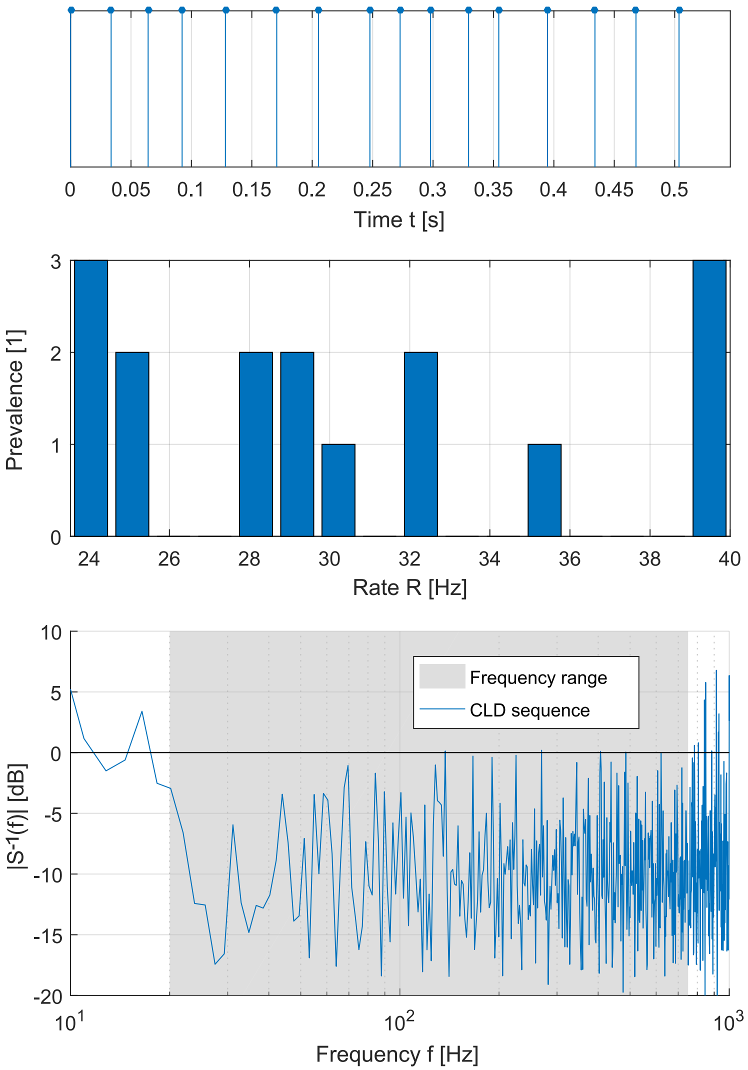

Figure 3.31: Properties of the stimulus ISI sequence optimized for CLD. From top to bottom : Stimulus onset train, presentation rate histogram, reciprocal transfer function (i.e., spectral noise amplification).

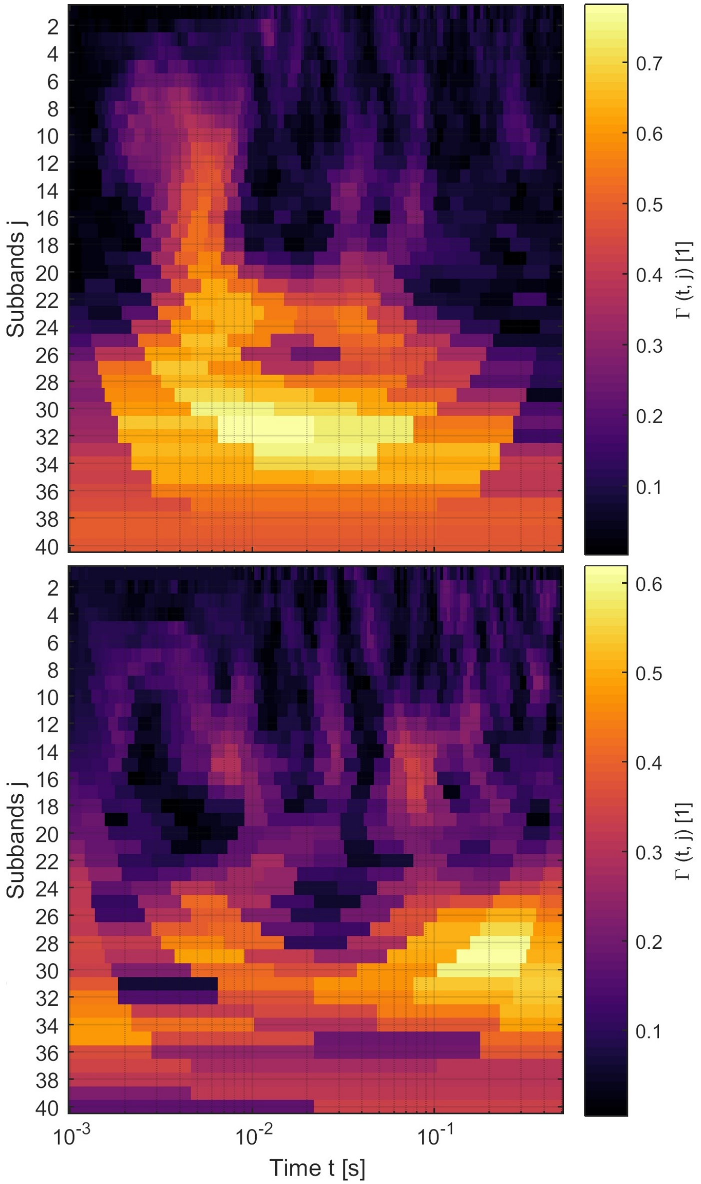

Figure 3.32: Example single–subject cross–trial RANDWT phase stability Γ. Top : CLD ABR/AMLR ERP image, bottom : subsequent single AFRR ERP image.

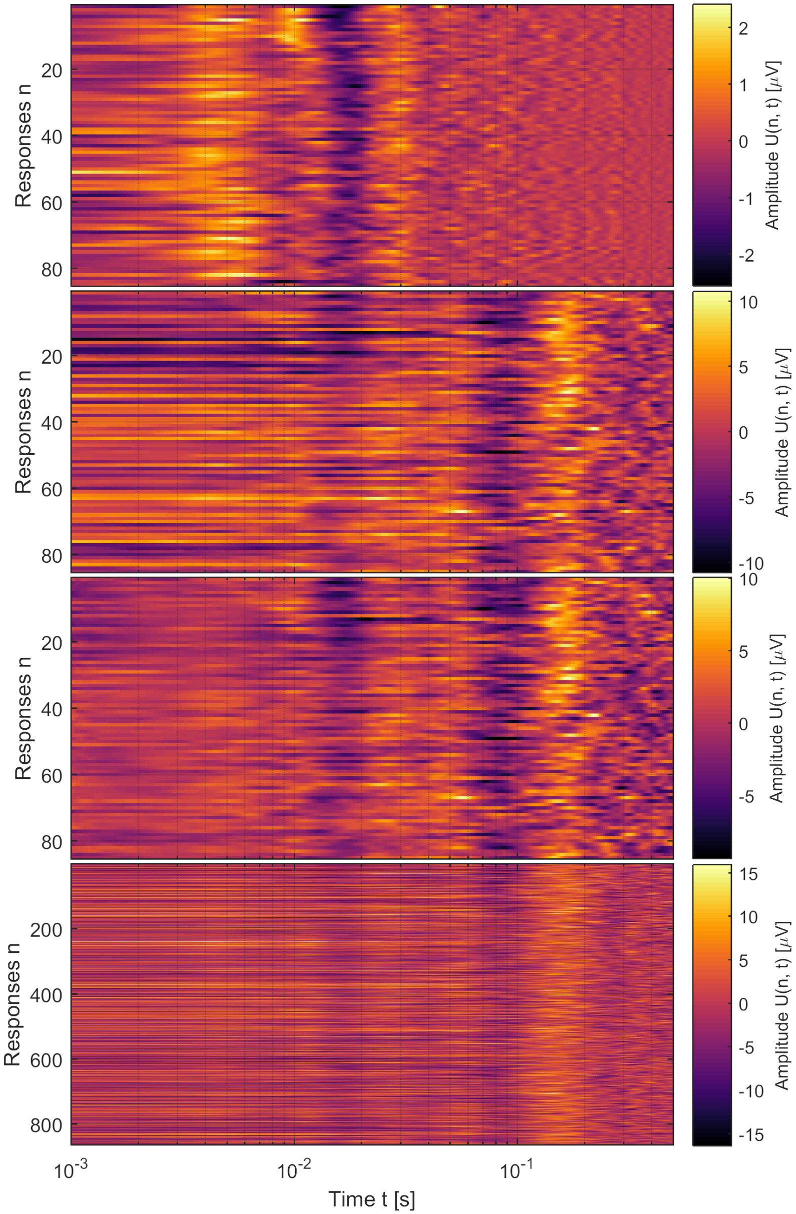

Figure 3.33: Grand average ERP images of recordings at A2. From top to bottom : CLD ABR/AMLR, subsequent single AFRR, fused AFRR and reference AFRR.

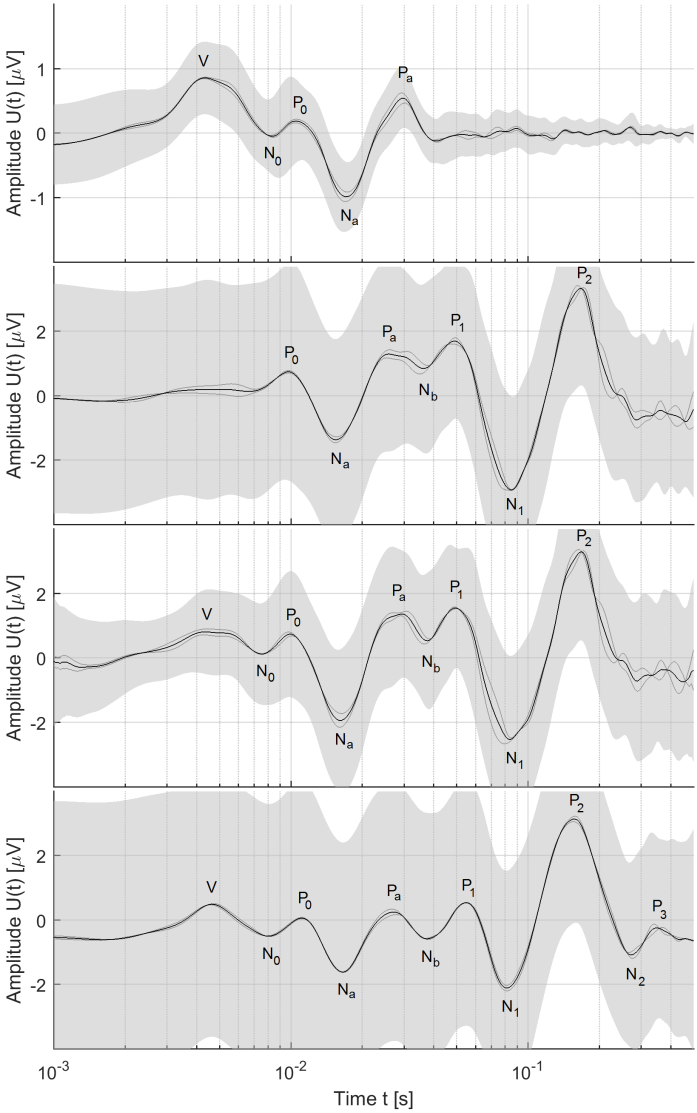

Figure 3.34: Grand average ERP waveforms of recordings at A2. From top to bottom : CLD ABR/AMLR, subsequent single AFRR, fused AFRR and reference AFRR. Black lines denote the average of all trials, dark gray lines the separate averages of odd/even trials and gray areas the sample–wise cross–trial standard deviation.

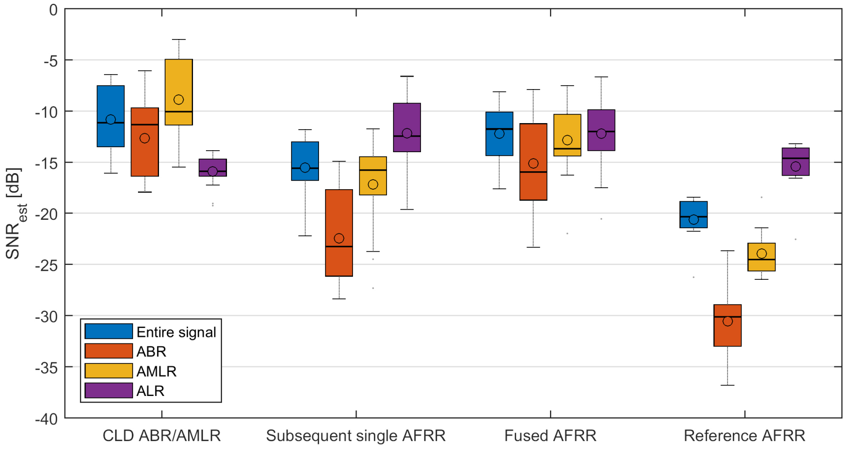

Figure 3.35: Boxplots of individual SNR estimates calculated for A2 recordings of CLD ABR/AMLR, subsequent single AFRR, fused AFRR and reference AFRR ERP images.

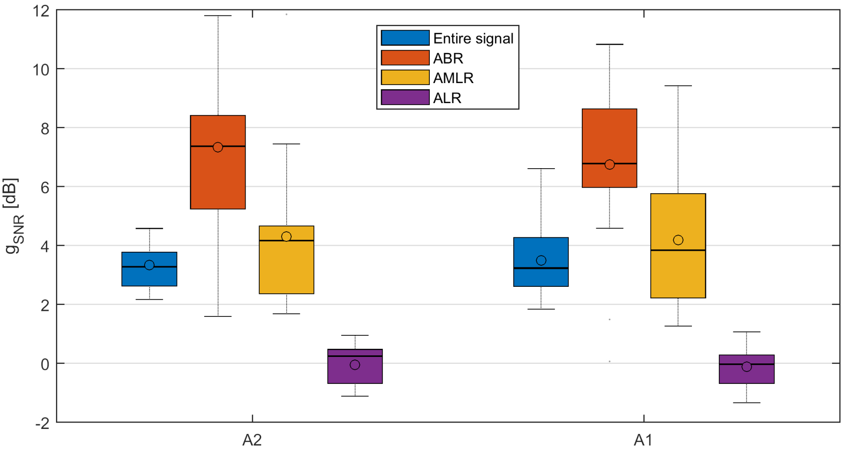

Figure 3.36: Boxplots of individual SNR gains of fused AFRR over subsequent single AFRR ERP images calculated separately for A2 and A1 recordings.

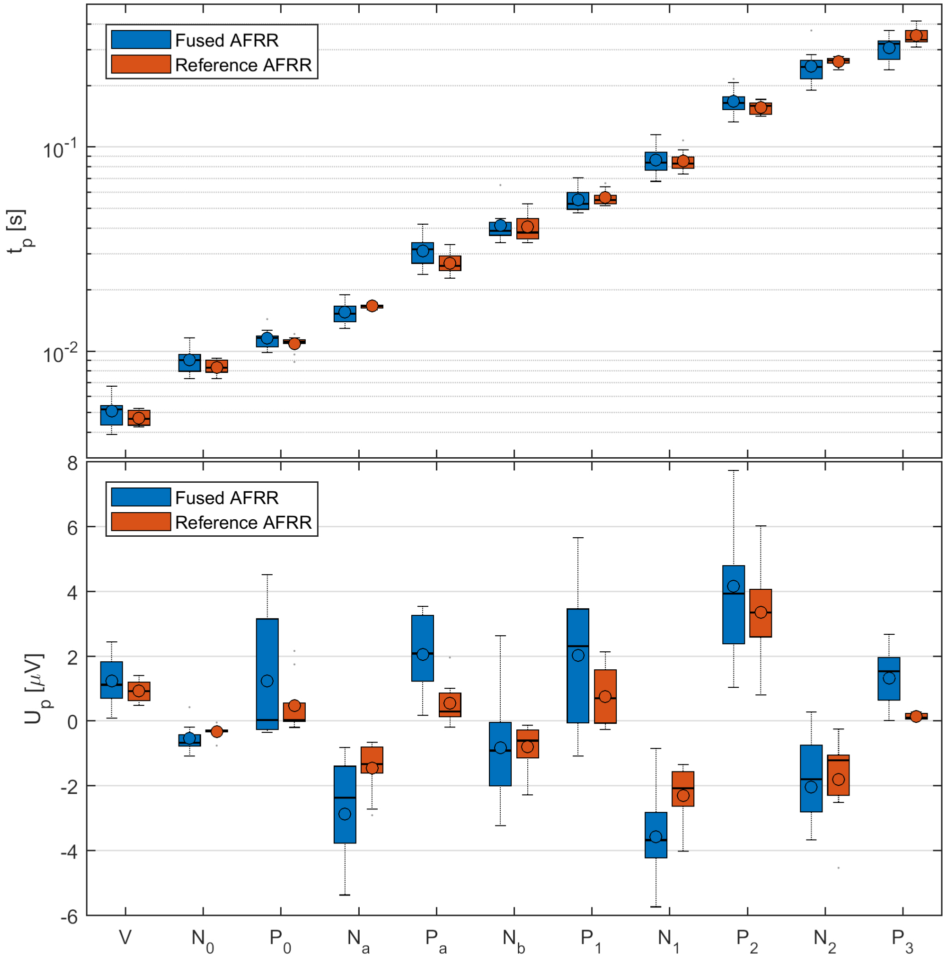

Figure 3.37: Boxplots of individual prominent wave peak latencies t_p and amplitudes u_p obtained from fused AFRR and reference AFRR average waveforms recorded at A2.

4 Discussion

4.1 Reference AFRR Acquisition

On a first note, the obtained single–subject and grand average AFRR waveforms (Fig. 3.1) exhibit a distinct ERP waveform morphology. Amplitudes and latencies of the prominent waves (encompassing V, N_0, P_0, N_a, P_a, N_b, P_1, N_1, P_2 and N_2 with the notable absence of the earliest ABR waves $I - III$, which is most likely due to their general fragility and the potentially slightly too low value for the upper cutoff frequency of 1 kHz used during signal conditioning) are very similar to the results reported by Michelini *et al.* (1982) and appear to be well in accordance with the pertinent literature values Picton *et al.* (1974) (see Tab. A.1). Comparing the average ERP waveforms to the real part of grand average CWT coefficients (Fig. 3.2), the time–frequency resolved waveforms appear to be considerably smoother than their time–domain counterparts due to the high redundancy of the CWT signal representation. This desirable property enables the robust determination of prominent waves' peak translations τ_p and their corresponding amplitudes u_p as well as their respective spectral emphasis f_p, which notably facilitates the additional detection of wave III (see Tab. A.3) despite its very small amplitude, which plausibly explains its absence from the time–domain average waveforms.

4.2 Optimization of AFRR Signal Representation

Judging by the magnitudes of grand average ST coefficients obtained from the reference AFRR average waveforms (Fig. 3.3), the time–domain signals indeed exhibit a substantial degree of self–similarity across investigated timescales as the majority of energy within the ST coefficient distribution is confined within a narrow interval around a single dominant scale and its first integer multiple. Hence, the first necessary property for a successful lossless resampling on a nonuniform timebase is exhibited by the signals as anticipated.

Further examining the grand average of WPS across single responses within the uniformly sampled AFRR ERP images (Fig. 3.4), it is evident from the sharp discontinuity and subsequent complete lack of phase syncronization within the lower right triangular part of the time–frequency plane that the underlying signals carry a temporally decreasing amount of event–related frequency content, with the temporal decrease of spectral emphasis very closely following a $1/f$ relationship as it appears as a straight line in the double–logarithmic visualization. As a result, a *logarithmic* warping function can be safely considered optimal for the underlying data, reproducibly resulting in a uniform sampling resolution of the resampled signals across all investigated timescales, hence arguably achieving the most compact, lossless AFRR signal representation possible. This finding is further backed up by the time–scale resolved grand average CWT real parts and WPS obtained from the actual resampled single response representations (Figs. 3.5 and 3.6), as they exhibit a very sparse distribution around two dominant, *close to constant* dilations[1] as expected, with all prominent signal features of their linear–time equivalents being mapped into this representation. On another note, the bandpass frequency window of $[1, 10^3]$ Hz used to condition (and subsequently resample) the reference AFRR dataset additionally proves to be close to the presumable optimum since the secondary diagonal in Fig. 3.4 effectively separates the time–frequency plane into ERP signal components and the lack thereof as aimed for. Note that the very high–frequent early ABR components (i.e., waves $I - III$) partly cross this diagonal in the time–frequency plane, which may further explain their complete absence from the resampled signals (see, e.g., Fig. 3.13). However, the negative impact of this localized anomaly can be easily alleviated by increasing the upper cutoff frequency accordingly at the minor expense of a diminished compression ratio when using the resampling operator, which will most probably still reside within the same order of magnitude (factor 16.95 for the chosen parameters). Based on the obtained results, a less aggressive cutoff frequency of around 2 – 3 kHz appears reasonable and is recommended for further applications.

The obtained decay profiles of estimate RN as a function of trial count, evaluated for the additive EEG noise signals (Fig. 3.7), are well in accordance with the theoretical prediction of a $1/\sqrt{n}$ relationship (Özdamar and Delgado, 1996) with α assuming an average regressed value of ≈ 0.5. However, for the stationary 50 Hz noise (Fig. 3.8), the profile enjoys a substantially faster decay rate of $\alpha \approx 0.75$, which can be safely ascribed to the reduced Wiener entropy of the stationary sine as opposed to the broadband EEG noise, allowing for a more effective attenuation due to averaging under the employed ISI jitter (Woldorff, 1993). Furthermore, the resampled response representation shows a distinct relative lead over its uniformly sampled counterpart of 38.3 % (EEG noise) and 41.9 % (stationary 50 Hz noise) on average, which is consistent across the whole duration of the experiment, hinting at an equally high amount of noise suppression when using less than half the number of single responses as for the uniformly sampled representation. This is probably due to the beneficial effects of temporal Gaussian smoothing during resampling as every form of weighted averaging (be it within time or across responses) will necessarily result in a suppression of RN to some degree. Consequently, the time–resolved RN profiles (Fig. 3.9) of the resampled signal representation consistently subside below their uniformly sampled counterparts by about 2 μV on average. While the uniform representation exhibits a flat temporal RN profile, the resampled representation expectably shows a coarse $1/t^{\alpha}$ trend under the progressive temporal averaging of the Gaussian filter, which is additionally modulated by the spectral composition of the additive EEG noise : Note the 50 Hz notch filter employed during signal conditioning corresponding to a distinct dip at around 20 ms with the dominant α– and β–band signal energy and their gradual rolloff towards lower frequencies likewise being faithfully mapped into the temporal regions corresponding to their dominant periods. This rather plausible finding of a temporally resolved frequency–selectivity of the proposed resampling operator strongly emphasizes the necessity to meticulously remove all stationary noise components from the raw ERP data by notch filtering *prior* to epoching and resampling, as the resampled response representation is obviously potentially prone to aggregation of *temporally localized* artifacts in contrast to the uniformly sampled representation if signals are not conditioned appropriately.

---

[1] Note that the coefficient profile of these dilations is closely related to the ST of the underlying, uniformly sampled signals (Fig. 3.3) for reasons outlined by De Sena and Rocchesso (2007).

4.3 Denoising of AFRR Single Responses

On a first note, the example sinogram of a resampled single–subject AFRR ERP image (Fig. 3.10) shows a remarkable degree of separation between the coefficients corresponding to ERP components and additive noise. Note the large smooth, connected areas of similar values around flat projection angles ϕ visibly corresponding to the time–domain amplitude morphology (in fact, the central projection at $\phi = 0$ is by definition *identical* to the average amplitude waveform). This necessarily leaves the coefficients under steeper angles to encode the more *irregular* signal components both due to the transform's directionality and the Parseval theorem, which is supported by their considerably more disjoint structure. Consequently, the associated angular distributions e_ϕ (Fig. 3.11, blue) exhibit two dominant areas of accumulating energy with remarkably low standard deviation across the single–subject dataset, the central one of which can be safely considered as predominantly representing the sought–after ERP components. The normalized weight vectors w' (Fig. 3.11, red) attenuate the coefficient contributions around the orthogonal incident angles $|\phi| \approx \pi/4$ as expected. Hence, a reliable reduction of additive noise is anticipated for the SDR filter based on these findings.

It is already evident from visual comparison of the filtered ERP images and the associated average waveforms (Figs. 3.14 – 3.21) to their unfiltered counterparts (Figs. 3.12 and 3.13) that each of the four investigated ERP image denoising algorithms considerably reduces the amount of additive EEG noise. As anticipated above, this holds particularly true for the SDR filter, which produces remarkably smooth results (Fig. 3.18) with the by far lowest residual cross–trial standard deviations (Fig. 3.19) of all employed filters. However, this presumable benefit is actually double–edged, as the SDR filter, due to being based on the RT, effectively constitutes a *global* operator (which is *not* the case for the other three contenders), putting it at a considerable disadvantage should the underlying waveform morphology gradually undergo substantial changes across the duration of the experiment as is the case with the *nonstationary* ERP traces encountered in, among others, habituation assessment (e.g., Mariam *et al.*, 2009). As this property is intrinsic to the employed RT hence not alleviable, potential applications of the SDR filter must involve a preliminary check on whether the criterion of a sufficiently stable ERP morphology across trials is met at least approximately to avoid potential filtering artifacts and loss of relevant information.

Further comparison of the objective outcome metrics provides a more detailed insight into the respective filter performances. Judging by the individual gains g_{CWI} in waveform reproducibility, all filters improve the resampled ERP images about the same amount of ≈ 0.1 for the real–world AFRR recordings (Fig. 3.22). The same holds for the synthetic AFRR dataset (Fig. 3.26) at each investigated SNR, with a monotonous gain increase towards lower values. This behaviour is expectable as there isn't much to be improved with respect to this measure at higher SNR values to begin with.

Considering the individual gains g_{RN} in RN, which is arguably the most important metric to be investigated within this context, both results for the real–world dataset (Fig. 3.23) and the synthetic AFRR (Fig. 3.27) indicate a clear advantage of the SDR filter over its contenders. However, the synthetic data reveals this lead to be maximal at rather high SNR values and to disproportionally decline towards the more realistic range of -12 dB and less, where the other filters equal or exceed its performance. Note that the UGM filter conveniently provides a constant baseline performance of ≈ -10 dB for comparison across all investigated SNR as its attenuation of additive noise is virtually independent of the SNR given that the filter constitutes a linear operator. The NLM filter provides a performance comparable to UGM with respect to this measure, albeit it excels slightly at higher SNR values. By contrast, the rather simple SWAPS algorithm outperforms both by about $2 - 5$ dB on average within the more relevant lower SNR range. Note that this finding interestingly appears reversed within the real–world data, where the NLM and SWAPS filters lead by ≈ 3 dB and ≈ 1 dB over UGM, respectively.

Individual gains g_{SIM} in similarity of the filtered average waveforms to the underlying ERP as obtained from the synthetic dataset (Fig. 3.28) expectably show all investigated filters to exhibit close to no increase at the higher SNR for the same reasoning as mentioned above regarding the gains in CWI. The UGM filter again provides a baseline value of 0 for comparison as it perfectly preserves the average ERP morphology for reasons outlined in Sect. 2.4.1. Notably, both the SDR and the SWAPS filter start to exhibit progressively increasing benefits with respect to this outcome metric within the critical SNR range below -6 dB, while the NLM filter exhibits the opposite behaviour (i.e., slightly, but steadily departing from preservance of the underlying waveform).

Lastly, to put the above performances of each investigated filter into perspective, their computational costs as quantified by the algorithm execution times (Fig. 3.29) remains to be compared. Here, the UGM filter expectably outperforms all of the alternative approaches substantially given that it boils down to a single, FFT–accelerated convolution. It is followed by the SWAPS algorithm with an approximately 35–fold computational footprint, the SDR filter (factor 175) and the NLM filter (factor 500), which appears to be rather inefficient in comparison. In the light of these figures, the SWAPS filter can be safely considered the most promising of the compared algorithms for use with the resampled single–response AFRR representation, as it follows a sound theoretical rationale, is reasonably fast, provides a consistent performance improvement over the baseline (i.e. UGM) and alternative approaches within the most relevant SNR range and – most importantly – doesn't suffer from the potential drawbacks of an underlying global operator as is the case with SDR. Hence, it can be potentially recommended for a broader range of related *a posteriori* ERP denoising tasks, even if reliable ground truth regarding the underlying ERP morphology may be lacking. However, it needs to be pointed out that the implicit assumption of globally (SDR) or at least locally (SWAPS) stable, *vertical* ERP traces is – intentionally – hardwired into both proposed methods, which is one of the key reasons for their good performance on the present data as it fulfills this assumption. It is worth noting that these tailored approaches will very likely fall short of successfully denoising the later waves of the cortical ERP (i.e., from 300 ms onwards) as these potentials generally exhibit large inter–trial variability both in their respective peak amplitudes and latencies due to being modulated by higher cognitive processes.

4.4 Fast Deconvolution–Based AFRR Acquisition

The properties of the obtained ISI sequence optimized for use with the CLD method by means of DE (Fig. 3.31 and Tab. A.2) can be considered exceptional as it achieves consistent noise attenuation of no less than 7 dB within a broad frequency range of [20, 750] Hz while simultaneously exhibiting rather low JR compared to the prior art of Özdamar and Bohórquez (2006), which involved time–consuming *manual* optimization of the individual ISI, yielding 0.82 dB of noise attenuation as the best result. While the related work of Huang *et al.* (2014) arguably demonstrates the superior convergence behaviour of the employed solution–space contraction DE over *DE / rand / 1 / bin* (reaching convergence within hundreds of iterations as opposed to $\approx 10^4$), this particular advantage bears the question of practical relevance and may be at least partially attributable to the substantially more narrow frequency ranges with bandwidths of 120 Hz and less being designated for optimization of the FDE within their work. Unfortunately, no objective NAF metrics were provided by Huang *et al.* (2014) to further validate their results, aggravating a comparative evaluation of both optimization approaches which was omitted within the scope of the present work for this reason.

It is clearly visible from the grand average ERP images (Fig. 3.33) and their corresponding average waveforms (Fig. 3.34) that the proposed novel ERP acquisition and processing methodology yields quality simultaneous recordings of the early, middle–latency and late auditory ERP components. The visible prominent waves include V, N_0, P_0, N_a and P_a for the CLD ABR/AMLR and P_0, N_a, P_a, N_b, P_1, N_1 and P_2 for the subsequent single AFRR. Subsequent single AFRR recordings expectably exhibit large amplitude fluctuations, resembled by high amounts of cross–trial standard deviation, and small net contributions to the average waveforms particularly within the temporal range of the brainstem–level components, which is probably due to their generally rather low SNR and the small number of 85 single responses being available for averaging. By contrast, the CLD ABR/AMLR responses exhibit visibly lower cross–trial standard deviations and well–defined ABR and AMLR components with noteworthy amounts of late auditory ERP components being absent from the averages as anticipated (Holt and Özdamar, 2014).

Further examination of the full–range auditory evoked responses obtained by the proposed fusion of CLD ABR/ AMLR with subsequent single AFRR trials reveals the responses to encompass all prominent waves within the resulting average waveforms while exhibiting lower and more uniform cross–trial standard deviations across all investigated temporal ranges judging by direct visual comparison, as was aimed for when optimizing the CLD sequence. These findings suggest a substantial increase in signal quality due to single–trial fusion especially for the early and middle-latency components, supported by the respective distributions of single–subject estimate SNR (Fig. 3.35, columns 2 and 3) and their individual gains (Fig. 3.36). Accordingly, the corresponding one–sided U–tests (Tab. A.4) show that increases in estimate SNR are significant with large effect sizes within the temporal ranges of ABR and AMLR (and subsequently for the entire signal). No significant increase in SNR was observed for the ALR, owing to the above–mentioned lack of noteworthy signal contribution from the CLD ABR/AMLR during fusion within this temporal range. Comparing the grand average fused AFRR and reference AFRR, it is evident that the proposed acquisition approach outperforms the prior art acquisition method not only by a substantial reduction of recording time (factor 3.4), but also in terms of the magnified SNR and its notable increase in temporal uniformity (Fig. 3.35, columns 3 and 4), supported by significant p–values of the respective one–sided U–test (Tab. A.5) with large (ABR/AMLR) and moderate (ALR) effect sizes. The latter finding is probably attributable to an addup of the expected SNR advantages due to the optimized deconvolution sequence of around 7 dB and the averaging of obtained responses with a larger ISI jitter compared to the reference recording (Woldorff, 1993).

The grand average waveform morphology of fused AFRR enjoys consistency with the obtained reference AFRR averages across investigated timescales. Accordingly, all prominent waves visible in the reference AFRR averages are also present in the fused AFRR grand average waveforms, with the exception of waves N_2 and P_3. Waves I to III are notably absent from both the fused and the reference AFRR waveforms as expected, owing to the rather aggressive upper cutoff frequency resulting from the tradeoff made during sequence optimization. Peak latencies and amplitudes of detected waves in the fused AFRR averages (Fig. 3.37) are consistent with reference averages and literature values (Picton *et al.*, 1974; Hall, 2007). The detected peak latencies do not differ significantly between fused and reference AFRR ($p \gg 0.05$), whereas amplitudes differ significantly for waves N_a, P_a, N_1 and P_3. However, the results of the corresponding one–sided U–tests (Tab. A.6) revealed the latter differences to be consistently in favor of the novel acquisition approach (significant, albeit with moderate effect size). Hence, no *adverse* effects on peak latencies and amplitudes of prominent waves in the obtained auditory ERP were discoverable during the comparative evaluation of the novel method and the prior art acquisition approach.

5 Conclusion and Future Work

The present work successfully subjected the motivation and reasoning underlying previous efforts by Michelini *et al.* (1982) towards the simultaneous acquisition of auditory brainstem–, middle–latency and late ERP and their proposed representation on a non-linear timebase to an in–depth investigation, particularly with regards to the optimal warping function. Obtained results of AFRR signal analysis regarding their dominant intrinsic scales and their time–frequency resolved cross–trial phase stability (effectively determining the degree of stimulus–locking within signal components) strongly support the optimality of a logarithmic warping to achieve a compact AFRR signal representation without loss of important signal features. The lack of such detrimental information loss was subsequently demonstrated for the proposed logarithmic resampling method adapted to the temporally decreasing spectral bandwidth of the signals in question, which achieves a substantial compression ratio of 16.95 while additionally exhibiting a distinct SNR advantage progressively increasing with time post–stimulus due to beneficial effects of the employed Gaussian averaging.

In an attempt to systematically sound out the available room for further improvements in its signal quality by means of denoising, the proposed resampled AFRR signal representation was employed to compare two established (i.e., UGM and NLM) and two newly proposed (SDR and SWAPS) two–dimensional ERP image denoising algorithms on the basis of several objective quality metrics using a real–world reference dataset of AFRR ERP images and synthetic data obtained from a phenomenological AFRR model at different SNR. Both proposed denoising approaches substantially surpass the prior art methods in the comparative performance evaluation both in terms of the employed quality metrics and their respective computational costs. Future work on improving the newly introduced SDR filter may involve its adaptation to a discrete RT transform pair, e.g., employing the fast implementation proposed by Press (2006) and its exact, iterative inverse to further reduce the involved computational cost.

Finally, a novel method for the fast acquisition of auditory full–range evoked potentials (AFRR) using an interleaved, deconvolution–based ERP acquisition approach was introduced and evaluated, which is accompanied by an efficient method for the automated optimization of ISI sequences suitable for deconvolution, reliably maximizing the spectral attenuation of additive noise within the deconvolved responses subject to a designated time–frequency constraint. As a sidenote, the proposed sequence optimization method has already been successfully applied to the design of stimulus sequences suitable for the fast, simultaneous acquisition of brainstem–level and subsequent binaural difference potentials in bimodal cochlea implantees by the time of this writing (Schebsdat *et al.*, 2019), enabling the objective evaluation of such asymmetric hearing treatments with respect to their binaural benefit. More general, the results obtained in the present work suggest the proposed methodology to be a promising toolset for simultaneous acquisition of neural correlates originating fom all subsequent processing stages along the AP. The interleaved acquisition approach employs successive, alternating periods of high–rate CLD stimulation and low–rate stimulation to separately elicit both the ABR/AMLR and ALR at their respective optimum rates, fusing the resulting ERP single trials based on their time–frequency resolved cross–trial regularity. It improves upon pure CLAD aquisition (Holt and Özdamar, 2014), simultaneously yielding quality ABR, AMLR *and* ALR components with additional resolution across the duration of the experiment. Furthermore, it improves two–fold upon the similarly motivated interleaved acquisition approach of Bidelmann (2015) by yielding *temporally resolved* ABR and AMLR components and by introducing a distinct SNR advantage due to the preceding stimulus sequence optimization, as both of these beneficial properties are not available from the brainstem FFR. Lastly, the proposed method improves upon the prior art reference AFRR acquisition at larger ISI, cutting acquisition time by a factor of 3.4 while at the same time exhibiting a substantially improved, temporally uniform SNR of the obtained ERP.

Since all of the pertinent research questions initially raised in Sect. 1.2 could be answered satisfactorily, the conjoint use of the proposed body of methods for AFRR acquisition, denoising and analysis can be safely recommended for a variety of applications in neuroscience research. Potential use cases essentially fall into two categories : First, as the involved necessary high trial counts of undisturbed ALR activity acquired at large ISI meet the necessary preconditions of the first acquisition approach as employed for the reference AFRR recording, future studies involving habituation assessment (which are currently exclusively based on cortical responses) may benefit from the additional information provided by the AFRR at virtually no additional expense, arguably enabling improved insight into preceding contributions of subcortical, precognitive processing to the commonly investigated cortical activity. The same holds in principle for the analysis of other cortical–level phenomena such as inhibition currently based on the ALR, which may well be accompanied by thalamocortical and/or corticofugal processes potentially additionally visible within AFRR recordings. The second category of possible applications is more geared towards clinical auditory diagnostics. Since the fast, interleaved deconvolution approach demonstrably yields quality full–range auditory activity within reasonable acquisition times, it arguably constitutes a valuable tool for auditory screening applications. To facilitate a potential adoption of this method in clinical settings, which obviously requires its translation into medically certified hard– and software, a reference implementation on an established medical–grade ERP acquisition system is currently under development as a collaborative effort of the *Systems Neuroscience & Neurotechnology Unit* and *Pilot Blankenfelde GmbH*.

Note that from a research viewpoint, the validity of the LTI signal–plus–noise ERP model has been repeatedly – and rightfully – questioned despite its ubiquity. In particular, the estimation of higher–order Volterra kernels, which naturally extend the concepts of impulse response and convolution to *nonlinear* systems, have been successfully employed for the quantification of nonlinear behaviour and signal self–interaction within the auditory system both at the levels of the AN (Eggermont, 1993) and the cortical spectrotemporal receptive fields (Klein *et al.*, 2000). As the reliable identification of higher–order Volterra kernels in the time domain has proven exceptionally difficult in the presence of additive noise, most estimation methods evade to the FT domain to obtain more robust results. Within the context of the logarithmic signal mapping and its beneficial properties in compactly representing the AFRR, it remains to be investigated whether a suitable stimulus sequence optimization minimizing the influence of additive noise *within the ST domain* may allow for a robust estimation of the 2$^{\text{nd}}$ order Volterra kernel in addition to the 1$^{\text{st}}$ order kernel given by the ERP waveform to capture the nonlinear behaviour exhibited by the AP. A quantification of these quadratic AFRR components obviously is of substantial interest in further improving the understanding of binaural fusion, which by its very nature constitutes the *interacting* proportion of auditory neural activity hence arguably eludes a detailed quantification within a linear analysis framework. However, this challenging endeavour shall be the subject of future work.

To conclude, all of the signal processing methods developed within the present work (i.e., ISI sequence optimization by means of DE, stimulus waveform generation, ERP preprocessing, nonlinear resampling, ERP image filters and single–trial fusion) are publicly made available under the MIT license, split across multiple Git repositories hosted at GitLab to facilitate adaption and verification by other researchers in future studies.

https://gitlab.com/manuelchristophkohl/

A Appendix

Table A.1: Peak latencies t_p of prominent waves for the ABR, AMLR and ALR as reported by Picton *et al.* (1974).

	ERP component	prominent waves					
t_p [ms]	ABR	I	II	III	IV	V	VI
		1.5	2.6	3.8	5.0	5.8	7.4
	AMLR	N_0	P_0	N_a	P_a	N_b	
		8.9	12	16	25	36	
	ALR	P_1	N_1	P_2	N_2		
		50	83	161	290		

Table A.2: Successive ISI of the optimized CLD stimulus sequence.

	stimulus # [1]			
ISI [ms]	01 – 33.54	02 – 30.63	03 – 28.13	04 – 35.83
	05 – 42.50	06 – 34.58	07 – 42.50	08 – 25.00
	09 – 25.21	10 – 31.46	11 – 25.00	12 – 40.31
	13 – 39.27	14 – 33.96	15 – 35.63	16 – 42.50

Table A.3: Peak translations τ_p, pseudofrequencies f_p and amplitudes u_p of prominent waves as obtained from the grand average real part of CWT coefficients.

ERP component	wave	τ_p [ms]	f_p [Hz]	u_p [μV]
ABR	III	2.67	499.5	0.17
	IV	5.17	164.5	0.49
AMLR	N_0	8.34	153.4	-0.55
	P_0	11.73	143.1	0.56
	N_a	17.62	64.4	-0.57
	P_a	25.95	60.1	0.47
	N_b	45.32	30.0	-0.36
ALR	P_1	56.37	13.5	1.07
	N_1	95.38	12.6	-1.05
	P_2	170.90	7.2	0.92
	N_2	242.60	6.9	-0.71
	P_3	311.40	7.1	0.45

Table A.4: p–values and effect sizes r of the one–sided U–test between SNR estimates SNR_{est} of individual subsequent single AFRR and fused AFRR ERP images (recordings at A2 and A1).

		Entire signal	ABR	AMLR	ALR
A2	p	$3.09 \cdot 10^{-3}$	$6.08 \cdot 10^{-4}$	$3.09 \cdot 10^{-3}$	$4.83 \cdot 10^{-1}$
	r	0.71	0.84	0.71	0.01
A1	p	$5.06 \cdot 10^{-3}$	$9.33 \cdot 10^{-4}$	$1.13 \cdot 10^{-2}$	$5.00 \cdot 10^{-1}$
	r	0.66	0.80	0.59	0.00

Table A.5: p–values and effect sizes r of the one–sided U–test between SNR estimates SNR_{est} of individual fused AFRR and reference AFRR ERP images (recordings at A2 and A1).

		Entire signal	ABR	AMLR	ALR
A2	p	$1.79 \cdot 10^{-5}$	$1.79 \cdot 10^{-5}$	$2.89 \cdot 10^{-5}$	$1.23 \cdot 10^{-2}$
	r	0.83	0.83	0.80	0.45
A1	p	$1.13 \cdot 10^{-4}$	$1.79 \cdot 10^{-5}$	$7.24 \cdot 10^{-5}$	$3.16 \cdot 10^{-2}$
	r	0.74	0.83	0.76	0.37

Table A.6: p–values and effect sizes r of the one–sided U–test between peak amplitudes u_p of fused AFRR and reference AFRR ERP averages (recordings at A2 and A1) for waves N_a, P_a, N_1 and P_3[1].

		N_a	P_a	N_1	P_3
A2	p	$2.15 \cdot 10^{-2}$	$1.55 \cdot 10^{-3}$	$2.45 \cdot 10^{-2}$	$4.35 \cdot 10^{-3}$
	r	0.40	0.60	0.39	–
A1	p	$1.23 \cdot 10^{-2}$	$3.37 \cdot 10^{-2}$	$2.85 \cdot 10^{-2}$	$3.45 \cdot 10^{-3}$
	r	0.45	0.38	0.39	–

---

[1] Note that due to the small number of overall observations for waves P_3, effect sizes r cannot be reported.

List of Figures

1.1 Illustration of slices through anatomic structures of the human auditory pathway . . . 16

1.2 Auditory Brainstem–, Middle–Latency and Late Responses . . . 24

1.3 Full–Range Auditory ERP . . . 29

2.1 Stimulus Presentation and ERP Acquisition Setup . . . 32

2.2 Conventional time–frequency windowing of auditory ERP components . . . 39

2.3 Schematic illustration of the SWAPS denoising algorithm . . . 50

2.4 Linear superposition ERP model . . . 55

2.5 Schematic illustration of the interleaved stimulus presentation paradigm . . . 60

3.1 Grand average ERP waveform of reference AFRR recordings . . . 67

3.2 Real part of grand average reference AFRR CWT coefficients . . . 68

3.3 Magnitude of grand average AFRR ST coefficients . . . 69

3.4 Grand average of WPS for the uniformly sampled AFRR ERP images . . . 69

3.5 Real part of grand average resampled AFRR CWT coefficients . . . 70

3.6 Grand average of WPS for the resampled AFRR ERP images . . . 70

3.7 Decay profiles and model regressions of RN as a function of trial count (synthetic EEG noise) . . . 71

3.8 Decay profiles and model regressions of RN as a function of trial count (stationary 50 Hz noise) 71

3.9 Decay profiles and model regressions of RN over time post–stimulus (synthetic EEG noise) 72

3.10 Example sinogram of resampled single–subject ERP image 73

3.11 Angular energy distributions e_ϕ and normalized weight vectors w' of resampled reference AFRR dataset 73

3.12 Grand average ERP image of resampled reference AFRR dataset (unfiltered) 74

3.13 Grand average ERP waveform of resampled reference AFRR dataset (unfiltered) 74

3.14 Grand average ERP image of resampled reference AFRR dataset (UGM filtered) 75

3.15 Grand average ERP waveform of resampled reference AFRR dataset (UGM filtered) 75

3.16 Grand average ERP image of resampled reference AFRR dataset (NLM filtered) 76

3.17 Grand average ERP waveform of resampled reference AFRR dataset (NLM filtered) 76

3.18 Grand average ERP image of resampled reference AFRR dataset (SDR filtered) 77

3.19 Grand average ERP waveform of resampled reference AFRR dataset (SDR filtered) 77

3.20 Grand average ERP image of resampled reference AFRR dataset (SWAPS filtered) 78

3.21 Grand average ERP waveform of resampled reference AFRR dataset (SWAPS filtered) 78

3.22 Boxplots of individual gains g_{CWI} in waveform reproducibility for the investigated ERP image filters (reference AFRR) 79

3.23 Boxplots of individual gains g_{RN} in RN for the investigated ERP image filters (reference AFRR) 79

3.24 Grand average ERP image of resampled synthetic AFRR dataset at –6 dB SNR (unfiltered) 80

3.25 Grand average ERP waveform of resampled synthetic AFRR dataset at –6 dB SNR (unfiltered) 80

3.26 Boxplots of individual gains g_{CWI} in waveform reproducibility for the investigated ERP image filters (synthetic AFRR) 81

3.27 Boxplots of individual gains g_{RN} in RN for the investigated ERP image filters (synthetic AFRR) 81

3.28 Boxplots of individual gains g_{SIM} in structural similarity for the investigated ERP image filters (synthetic AFRR) 82

3.29 Average execution times of the investigated ERP image filters 82

3.30 Example single–subject NAF estimation 84

3.31 Properties of the stimulus ISI sequence optimized for CLD 85

3.32 Example single–subject cross–trial RANDWT phase stability Γ 86

3.33 Grand average ERP images of recordings at A2 87

3.34 Grand average ERP waveforms of recordings at A2 88

3.35 Boxplots of individual SNR estimates (CLD ABR/AMLR, subsequent single AFRR, fused AFRR and reference AFRR) 89

3.36 Boxplots of individual SNR gains (fused AFRR over subsequent single AFRR) 89

3.37 Boxplots of individual prominent wave peak latencies t_p and amplitudes u_p 90

List of Tables

A.1 Peak latencies t_p of prominent waves for the ABR, AMLR and ALR 105

A.2 Successive ISI of the optimized CLD stimulus sequence 105

A.3 Peak translations τ_p, pseudofrequencies f_p and amplitudes u_p of prominent waves 106

A.4 p–values and effect sizes r (one–sided U–test between SNR estimates of individual subsequent single AFRR and fused AFRR) 107

A.5 p–values and effect sizes r (one–sided U–test between SNR estimates of individual fused AFRR and reference AFRR) 107

A.6 p–values and effect sizes r (one–sided U–test between peak amplitudes u_p of fused AFRR and reference AFRR) for waves N_a, P_a, N_1 and P_3. 108

Bibliography

Ahmadi, M., and Quian Quiroga, R. (**2013**). "Automatic Denoising of Single–Trial Evoked Potentials," NeuroImage **66**, 672–680, doi: `10.1016/j.neuroimage.2012.10.062`.

Aldroubi, A., and Benedetto, J. (**1996**). *Wavelets in Medicine and Biology*, 1 ed. (Taylor & Francis, New York, USA), isbn: `978-0849394836`.

Altman (**2014**). *Accelerating MATLAB Performance: 1001 Tips to Speed Up MATLAB Programs*, 1 ed. (Apple Academic Press, Palm Bay, USA), isbn: `978-1482211290`.

American Bible Society (**1995**). *Holy Bible: Contemporary English Version*, 1 ed. (American Bible Society, Philadelphia, USA), isbn: `978-1585160211`.

Anderson, L. A., Christianson, G. B., and Linden, J. F. (**2009**). "Stimulus–Specific Adaptation Occurs in the Auditory Thalamus," J. Neurosci. **29**(22), 7359–7363, doi: `10.1523/jneurosci.0793-09.2009`.

Arnott, S. R., Binns, M. A., Grady, C. L., and Alain, C. (**2004**). "Assessing the Auditory Dual–Pathway Model in Humans," Neur. Imag. **22**(1), 401–408, doi: `10.1016/j.neuroimage.2004.01.014`.

Arslan, E., Prosser, S., and Michelini, S. (**1984**). "Simultaneous Recording of Auditory Evoked Potentials. Relationships Among the Fast, Middle and Long Latency Components," Scand. Audiol. **13**(2), 75–81, doi: `10.3109/01050398409043043`.

Bardy, F., Dillon, H., and Van Dun, B. (**2014**a). "Least–Squares Deconvolution of Evoked Potentials and Sequence Optimization for Multiple Stimuli Under Low–Jitter Conditions," Clin. Neurophysiol. **125**(4), 727–737, doi: `10.1016/j.clinph.2013.09.030`.

Bardy, F., Van Dun, B., Dillon, H., and McMahon, C. M. (**2014**b). “Deconvolution of Overlapping Cortical Auditory Evoked Potentials Recorded Using Short Stimulus Onset–Asynchrony Ranges,” Clin. Neurophysiol. **125**(4), 814–826, doi: `10.1016/j.clinph.2013.09.031`.

Bayram, I. (**2013**). “An Analytic Wavelet Transform With a Flexible Time–Frequency Covering,” IEEE Trans. Sig. Proc. **61**(5), 1131–1142, doi: `10.1109/tsp.2012.2232655`.

Bell, S. L., Allen, R., and Lutman, M. E. (**2002**). “An Investigation of the Use of Band–Limited Chirp Stimuli to Obtain the Auditory Brainstem Response,” Internat. J. Audiol. **41**(5), 271–278, doi: `10.3109/14992020209077186`.

Bernarding, C., Strauss, D. J., Latzel, M., and Corona-Strauss, F. I. (**2010**). “Non–Listening Effort Related Parameters in Auditory Discrimination Paradigms,” in *32th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 6682–6685, doi: `10.1109/iembs.2010.5626265`.

Bertrand, J., Bertrand, P., and Ovarlez, J. P. (**1990**). “Discrete Mellin Transform for Signal Analysis,” in *Int. Conf. Acoust. Speech Sig. Proc.*, Vol. 3, pp. 1603–1606, doi: `10.1109/icassp.1990.115730`.

Beylkin, G. (**1987**). “Discrete Radon Transform,” in *IEEE Trans. Acoust. Speech Sig. Proc.*, Vol. 35, pp. 162–172, doi: `10.1109/tassp.1987.1165108`.

Biacabe, B., Chevallier, J. M., Avan, P., and Bonfils, P. (**2001**). “Functional Anatomy of Auditory Brainstem Nuclei: Application to the Anatomical Basis of Brainstem Auditory Evoked Potentials,” Auris Nasus Larynx **28**(1), 85–94, doi: `10.1016/s0385-8146(00)00080-8`.

Bidelman, G. M., Pousson, M., Dugas, C., and Fehrenbach, A. (**2018**). “Test–Retest Reliability of Dual–Recorded Brainstem versus Cortical Auditory–Evoked Potentials to Speech,” J. Am. Acad. Audiol. **29**(2), 164–174, doi: `doi:10.3766/jaaa.16167`.

Bidelmann, G. M. (**2015**). “Towards an Optimal Paradigm for Simultaneously Recording Cortical and Brainstem Auditory Evoked Potentials,” J. Neurosci. Meth. **241**, 94–100, doi: `10.1016/j.jneumeth.2014.12.019`.

Binns, K. E., Grant, S., Withington, D. J., and Keating, M. J. (**1992**). “A Topographic Representation of Auditory Space in the External Nucleus of the Inferior Colliculus of the Guinea–Pig,” Brain Res. **589**(2), 231–242, doi: `10.1016/0006-8993(92)91282-J`.

Bizley, J. K. (**2017**). “Chapter 26 – Audition,” in *Conn's Translational Neuroscience*, edited by P. M. Conn (Academic Press, San Diego, USA), pp. 579–598, doi: `10.1016/b978-0-12-802381-5.00042-7`.

Bohórquez, J., and Özdamar, Ö. (**2006**). “Signal to Noise Ratio Analysis of Maximum Length Sequence Deconvolution of Overlapping Evoked Potentials,” J. Acoust. Soc. Am. **119**(5), 2881–2888, doi: `10.1121/1.2191609`.

Bohórquez, J., and Özdamar, Ö. (**2008**). “Generation of the 40–Hz Auditory Steady-State Response (ASSR) Explained Using Convolution,” Clin. Neurophysiol. **119**, 2598–2607, doi: `10.1016/j.clinph.2008.08.002`.

Brockhaus-Dumke, A., Mueller, R., Faigle, U., and Klosterkoetter, J. (**2008**). “Sensory Gating Revisited: Relation Between Brain Oscillations and Auditory Evoked Potentials in Schizophrenia,” Schizophr. Res. **99**(1–3), 238–249, doi: `10.1016/j.schres.2007.10.034`.

Buades, A., Coll, B., and Morel, J. (**2005**). “A Non–Local Algorithm for Image Denoising,” in *2005 IEEE Comp. Soc. Conf. Comp. Vis. Patt. Recog.*, pp. 60–65, doi: `10.1109/cvpr.2005.38`.

Budd, T. W., and Michie, P. (**1994**). “Facilitation of the N1 Peak of the Auditory ERP at Short Stimulus Intervals,” Neurorep. **5**, 2513–2516, doi: `10.1097/00001756-199412000-00027`.

Burkard, R. F., Eggermont, J. J., and Don, M. (**2006**). *Auditory Evoked Potentials: Basic Principles and Clinical Application*, 1 ed. (Lippincott Williams & Wilkins, Philadelphia, USA), isbn: `978-0781757560`.

Butzer, P. L., and Jansche, S. (**1997**). "A Direct Approach to the Mellin Transform," J. Fourier Anal. Applic. **3**(4), 325–376, doi: `10.1007/bf02649101`.

Campbell, R. A. A., and King, A. J. (**2004**). "Auditory Neuroscience: A Time for Coincidence?," Curr. Biol. **14**(20), R886–R888, doi: `10.1016/j.cub.2004.09.070`.

Cargnelutti, M., Cóser, P. L., and Biaggio, E. P. V. (**2017**). "LS CE-Chirp vs. Click in the Neuroaudiological Diagnosis by ABR," Brazil. J. Otorhinolaryngol. **83**(3), 313–317, doi: `doi.org/10.1016/j.bjorl.2016.04.018`.

Celesia, G. G., and Hickok, G. (**2015**). *The Human Auditory System: Fundamental Organization and Clinical Disorders*, 1 ed. (Elsevier, Amsterdam, Netherlands), isbn: `978-0444626301`.

Cohen, L. (**1993**). "The Scale Representation," IEEE Trans. Sig. Proc. **41**(12), 3275–3292, doi: `10.1109/78.258073`.

Corona-Strauss, F. I., Delb, W., Schick, B., and Strauss, D. J. (**2009**). "Phase Stability Analysis of Chirp Evoked Auditory Brainstem Responses by Gabor Frame Operators," IEEE Trans. Neural Sys. Rehabil. Eng. **17**(6), 530–536, doi: `10.1109/tnsre.2009.2032627`.

Corona-Strauss, F. I., Schick, B., Delb, W., and Strauss, D. J. (**2012**). "Notched–Noise Embedded Frequency Specific Chirps for Objective Audiometry Using Auditory Brainstem Responses.," Audiol. Res. **2**(1), 30–38, doi: `10.4081/audiores.2011.e7`.

Cree, M. J., and Bones, P. J. (**1993**). "Algorithms to Numerically Evaluate the Hankel Transform," Comput. Math. Applic. **26**(1), 1–12, doi: `10.1016/0898-1221(93)90081-6`.

Dau, T., Wegner, O., Mellert, V., and Kollmeier, B. (**2000**). “Auditory Brainstem Responses With Optimized Chirp Signals Compensating Basilar–Membrane Dispersion,” J. Acoust. Soc. Am. **107**(3), 1530–1540, doi: `10.1121/1.428438`.

Davis, H., Mast, T., Yoshie, N., and Zerlin, S. (**1966**). “The Slow Response of the Human Cortex to Auditory Stimuli: Recovery Process,” Electroencephalogr. Clin. Neurophysiol. **21**(2), 105–113, doi: `10.1016/0013-4694(66)90118-0`.

Dawson, G. D. (**1951**). “A Summation Technique for Detecting Small Signals in a Large Irregular Background,” J. Physiol. **115**(1), 2p–3p.

Dawson, G. D. (**1954**). “A Summation Technique for the Detection of Small Evoked Potentials,” Electroencephalogr. Clin. Neurophysiol. **6**, 65–84, doi: `10.1016/0013-4694(54)90007-3`.

De Sena, A., and Rocchesso, D. (**2007**). “A Fast Mellin and Scale Transform,” EURASIP J. Adv. Sig. Proc. **2007**(1), 1–9, doi: `10.1155/2007/89170`.

Delgado, R. E., and Özdamar, Ö. (**2004**). “Deconvolution of Evoked Responses Obtained at High Stimulus Rates,” J. Acoust. Soc. Am. **115**(3), 1242–1251, doi: `10.1121/1.1639327`.

Demiralp, T., and Ademoglu, A. (**2001**). “Decomposition of Event–Related Brain Potentials Into Multiple Functional Components Using Wavelet Transform,” Clin. Electroencephalogr. **32**(3), 122–138, doi: `10.1177/155005940103200307`.

Di, S., and Barth, D. S. (**1992**). “The functional anatomy of middle–latency auditory evoked potentials: Thalamocortical connections,” J. Neurophysiol. **68**(2), 425–431, doi: `10.1152/jn.1992.68.2.425`.

Donoho, D. L. (**1995**). “De–Noising by Soft–Thresholding,” IEEE Trans. Inf. Theo. **41**(3), 613–627, doi: `10.1109/18.382009`.

Eggermont, J. J. (**1993**). “Wiener and Volterra Analyses Applied to the Auditory System,” Hear. Res. **66**(2), 177–201, doi: `10.1016/0378-5955(93)90139-r`.

Eggermont, J. J. (**2010**). *The Auditory Cortex: The Final Frontier*, 97–127 (Springer, Boston, USA), doi: `10.1007/978-1-4419-5934-8_5`.

Ferree, T. C., and Hwa, R. C. (**2003**). "Power–Law Scaling in Human EEG: Relation to Fourier Power Spectrum," Neurocomp. **52–54**, 755–761, doi: `10.1016/s0925-2312(02)00760-9`.

Figueiredo, M. A. T., Dias, J. B., Oliveira, J. P., and Nowak, R. D. (**2006**). "On Total Variation Denoising: A New Majorization–Minimization Algorithm and an Experimental Comparison With Wavalet Denoising," in *2006 Int. Conf. Imag. Proc.*, pp. 2633–2636, doi: `10.1109/icip.2006.313050`.

Fisher, N. I. (**1993**). *Statistical Analysis of Circular Data*, 1 ed. (Cambridge University Press, Cambridge, UK), isbn: `978-0521568906`.

Fitzpatrick, D. C., Kuwada, S., and Batra, R. (**2002**). "Transformations in Processing Interaural Time Differences Between the Superior Olivary Complex and Inferior Colliculus: Beyond the Jeffress Model," Hear. Res. **168**(1), 79–89, doi: `10.1016/s0378-5955(02)00359-3`.

Fobel, O., and Dau, T. (**2004**). "Searching for the Optimal Stimulus Eliciting Auditory Brainstem Responses in Humans," J. Acoust. Soc. Am. **116**(4), 2213–2222, doi: `10.1121/1.1787523`.

Forte, A. E., Etard, O., and Reichenbach, T. (**2017**). "The Human Auditory Brainstem Response to Running Speech Reveals a Subcortical Mechanism for Selective Attention," ELife **6**, 1–12, doi: `10.7554/elife.27203`.

Francart, T., Wiebe, K., and Wesarg, T. (**2018**). "Interaural Time Difference Perception with a Cochlear Implant and a Normal Ear," J. Assoc. Res. Otolaryngol. **19**(6), 703–715, doi: `10.1007/s10162-018-00697-w`.

Fritz, J. B., Elhilali, M., David, S. V., and Shamma, S. A. (**2007**). "Auditory Attention – Focusing the Searchlight on Sound," Curr. Opinion Neurobiol. **17**(4), 437–455, doi: `10.1016/j.conb.2007.07.011`.

Gabor, D. (**1946**). “Theory of Communication,” J. Inst. Electr. Eng. **93**, 429–441.

Galambos, R., Makeig, S., and Talmachoff, P. J. (**1981**). “A 40–Hz Auditory Potential Recorded From the Human Scalp,” Proc. Nation. Academ. Sci. USA **78**(4), 2643–2647, doi: `10.1073/pnas.78.4.2643`.

Gelfand, S. A. (**2009**). *Essentials of Audiology*, 3 ed. (Thieme, New York, USA), isbn: `978-1604060447`.

Giroldini, W., Pederzoli, L., Bilucaglia, M., Melloni, S., and Tressoldi, P. (**2016**). “A New Method to Detect Event-Related Potentials Based on Pearson’s Correlation,” EURASIP J. Bioinform. Sys. Biol. **2016**(1), 1–13, doi: `10.1186/s13637-016-0043-z`.

Glasscock, M. E., Jackson, C. G., Josey, A. F., Dickins, J. R., and Wiet, R. J. (**1979**). “Brain Stem Evoked Response Audiometry in a Clinical Practice,” Laryngoscope **89**(7), 1021–1035.

Glendenning, K. K., Brunso-Bechtold, J. K., Thompson, G. C., and Masterton, R. B. (**1981**). “Ascending Auditory Afferents to the Nuclei of the Lateral Lemniscus,” J. Comp. **197**(4), 673–703, doi: `10.1002/cne.901970409`.

González–Trejo, E., Herrmann, D., Kohl, M. C., Corona-Strauss, F. I., and Strauss, D. J. (**2013**). “Paired Chirp Evoked Cortical Inhibition and its Behavioral Correlates in a Speech Intelligibility Task,” in *6th Int. IEEE Eng. Med. Biol. Soc. Conf. Neur. Eng.*, pp. 1350–1353, doi: `10.1109/ner.2013.6696192`.

González–Trejo, E., Kohl, M. C., Steinbach, A., Mögele, H., Pfleger, N., and Strauss, D. J. (**2015**). “Correlation Between Cortical Inhibition and Auditory Stream Segregation in a Driving Environment,” in *7th Int. IEEE Eng. Med. Biol. Soc. Conf. Neur. Eng.*, pp. 747–750, doi: `10.1109/ner.2015.7146731`.

Goodin, D. S. (**2012**). “Chapter 29 – Event–Related Potentials,” in *Aminoff’s Electrodiagnosis in Clinical Neurology*, edited by M. J. Aminoff, 6 ed. (W. B. Saunders, London, UK), pp. 633–650, doi: `10.1016/b978-1-4557-0308-1.00029-7`.

Goupillaud, P., Grossmann, A., and Morlet, J. (**1984**). "Cycle–Octave and Related Transforms in Seismic Signal Analysis," Geoexplor. **23**(1), 85–102, doi: `10.1016/0016-7142(84)90025-5`.

Grothe, B. (**2000**). "The Evolution of Temporal Processing in the Medial Superior Olive, an Auditory Brainstem Structure," Progr. Neurobiol. **61**(6), 581–610, doi: `10.1016/s0301-0082(99)00068-4`.

Haab, L., Lehser, C., Corona-Strauss, F. I., Bernarding, C., Seidler, H., Hannemann, R., and Strauss, D. J. (**2019**). "Implementation and Long–Term Evaluation of a Hearing Aid Supported Tinnitus Treatment Using Notched Environmental Sounds," IEEE J. Translat. Eng. Health Med. **7**, 1–9, doi: `10.1109/jtehm.2019.2897570`.

Hall, J. W. (**2007**). *New Handbook of Auditory Evoked Responses*, 1 ed. (Pearson, Boston, USA), isbn: `978-0205361045`.

Harte, J., Rønne, F., and Dau, T. (**2010**). "Modeling Human Auditory Evoked Brainstem Responses Based on Nonlinear Cochlear Processing," in *20th Int. Conf. Acoust.*, pp. 1–9.

Henkin, Y., Yaar-Soffer, Y., Givon, L., and Hildesheimer, M. (**2015**). "Hearing With Two Ears: Evidence for Cortical Binaural Interaction During Auditory Processing," J. Am. Acad. Audiol. **26**(4), 384–392, doi: `10.3766/jaaa.26.4.6`.

Holt, F. D. (**2015**). "Optimization and Modeling of Multicomponent Auditory Evoked Potentials Using Deconvolution and Stimulation Memory," Ph.D. Dissertation, University of Miami, url: `https://scholarlyrepository.miami.edu/oa_dissertations/1484/`.

Holt, F. D., and Özdamar, Ö. (**2014**). "Simultaneous Acquisition of High–Rate Early, Middle, and Late Auditory Evoked Potentials," in *36th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 1481–1484, doi: `10.1109/embc.2014.6943881`.

Holt, F. D., and Özdamar, Ö. (**2016**). "Effects of Rate (0.3–40/s) on Simultaneously Recorded Auditory Brainstem, Middle and Late Responses Using Deconvolution," Clin. Neurophysiol. **127**(2), 1589–1602, doi: `10.1016/j.clinph.2015.10.046`.

Hu, H., Kollmeier, B., and Dietz, M. (**2016**). "Suitability of the Binaural Interaction Component for Interaural Electrode Pairing of Bilateral Cochlear Implants," Adv. Exp. Med. Biol. **894**, 57–64, doi: `10.1007/978-3-319-25474-6_7`.

Huang, J.-H., Li, B., and Wang, T. (**2014**). "Optimization of CLAD Sequences Based on Solution–Space Contraction Strategy for Differential Evolution Algorithm," Chin. J. Electron. **42**(8), 1571–1576, doi: `10.3969/j.issn.0372-2112.2014.08.017`.

Hubel, D. H., Henson, C. O., Rupert, A., and Galambos, R. (**1959**). ""Attention" Units in the Auditory Cortex," Sci. **129**, 1279–1280, doi: `10.1126/science.129.3358.1279`.

International Electrotechnical Commission (**2007**). "Electroacoustics – Audiometric equipment – Part 3: Test signals of short duration." url: `https://webstore.iec.ch/publication/2773`, 2 ed. (Geneva, Swizerland), `IEC 60645-3:2007`.

Jääskeläinen, I. P., Ahveninen, J., Bonmassar, G., Dale, A. M., Ilmoniemi, R. J., Levänen, S., Lin, F.-H., May, P., Melcher, J., Stufflebeam, S., Tiitinen, H., and Belliveau, J. W. (**2004**). "Human Posterior Auditory Cortex Gates Novel Sounds to Consciousness," Proc. Nation. Academ. Sci. USA **101**(17), 6809–6814, doi: `10.1073/pnas.0303760101`.

Joris, P., and Yin, T. C. T. (**2007**). "A Matter of Time: Internal Delays in Binaural Processing," Trends Neurosci. **30**(2), 70–78, doi: `10.1016/j.tins.2006.12.004`.

Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., and Sejnowski, T. J. (**1999**). "Analyzing and Visualizing Single–Trial Event–Related Potentials," Adv. Neur. Inf. Proc. Sys. **11**, 118–124.

Kaas, J. H. (**2008**). "The Evolution of the Complex Sensory and Motor Systems of the Human Brain," Brain Res. Bull. **75**(2), 384–390, doi: 10.1016/j.brainresbull.2007.10.009.

Kaas, J. H., and Hackett, T. A. (**2000**). "Subdivisions of Auditory Cortex and Processing Streams in Primates," Proc. Nation. Acad. Sci. **97**(22), 11793–11799, doi: 10.1073/pnas.97.22.11793.

Kak, A. C., Slaney, M., and Wang, G. (**2002**). "Principles of Computerized Tomographic Imaging," Med. Phys. **29**(1), 107–107, doi: 10.1118/1.1455742.

Kern, K., Royter, V., Corona-Strauss, F. I., Mariam, M., and Strauss, D. J. (**2010**). "Habituation Analysis of Chirp vs. Tone Evoked Auditory Late Responses," in *32th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 6825–6828, doi: 10.1109/iembs.2010.5625954.

Kerr, C. C., Rennie, C. J., and Robinson, P. A. (**2008**). "Physiology–Based Modeling of Cortical Auditory Evoked Potentials," Biol. Cybernet. **98**(2), 171–184, doi: 10.1007/s00422-007-0201-1.

Klauke, I., Kohl, M. C., Hannemann, R., Kornagel, U., Strauss, D. J., and Corona-Strauss, F. I. (**2015**). "Impact of Monaural Frequency Compression on Binaural Fusion at the Brainstem Level," in *37th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 1646–1649, doi: 10.1109/embc.2015.7318691.

Klein, D. J., Depireux, D. A., Simon, J. Z., and Shamma, S. A. (**2000**). "Robust Spectrotemporal Reverse Correlation for the Auditory System: Optimizing Stimulus Design," J. Comp. Neurosci. **9**(1), 85–111, doi: 10.1023/a:1008990412183.

Klimesch, W., Schack, B., Schabus, M., Doppelmayr, M., Gruber, W., and Sauseng, P. (**2014**). "Phase–Locked Alpha and Theta Oscillations Generate the P1–N1 Complex and are Related to Memory Performance," Cogn. Brain Res. **19**(3), 302–316, doi: 10.1016/j.cogbrainres.2003.11.016.

Kohl, M. C., Schebsdat, E., Schneider, E. N., Niehl, A., Strauss, D. J., Özdamar, Ö., and Bohórquez, J. (**2019**a). "Fast Acquisition of Full–Range Auditory Event–Related Potentials Using an Interleaved Deconvolution Approach," J. Acoust. Soc. Am. **145**(540), doi: `10.1121/1.5087825`.

Kohl, M. C., Schebsdat, E., Schneider, E. N., and Strauss, D. J. (**2019**b). "Denoising of Single–Trial Event-Related Potentials by Shrinkage and Phase Regularization of Analytic Wavelet Filterbank Coefficients," in *9th Int. IEEE Eng. Med. Biol. Soc. Conf. Neur. Eng.*, pp. 251–254, doi: `10.1109/ner.2019.8717148`.

Kohl, M. C., and Strauss, D. J. (**2016**). "A Compact Representation for the Auditory Full–Range Response and its Fast Denoising Using an Image Filter Based on the Radon Transform," in *38th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 5877–5880, doi: `10.1109/embc.2016.7592065`.

Kral, A., and Majernik, V. (**1996**). "On Lateral Inhibition in the Auditory System," Gen. Physiol. Biophys. **15**(2), 109–127.

Kristensen, E., Guerin-Dugué, A., and Rivet, B. (**2017**). "Regularization and a General Linear Model for Event–Related Potential Estimation," Behav. Res. Meth. **49**(6), 2255–2274, doi: `10.3758/s13428-017-0856-z`.

Lehser, C., Hannemann, R., Corona-Strauss, F. I., Strauss, D. J., Haab, L., Seidler-Fallböhmer, B., and Seidler, H. (**2015**). "Dysfunctional Long Term Habituation to Exogeneous Tinnitus–Matched Sounds in Patients with High Tinnitus Distress," in *37th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 1967–1970, doi: `10.1109/embc.2015.7318770`.

Levine, R. A. (**1981**). "Binaural Interaction in Brainstem Potentials of Human Subjects," Annals of Neurology **9**(4), 384–393, doi: `10.1002/ana.410090412`.

Liégeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., and Chauvel, P. (**1994**). "Evoked Potentials Recorded From the Auditory Cortex in Man: Evaluation and Topography of the Middle Latency Components," Electroencephalogr. Clin. Neurophysiol. **92**(3), 204–214, doi: `10.1016/0168-5597(94)90064-7`.

Low, Y. F., and Strauss, D. J. (**2011**). "A Performance Study of the Wavelet-Phase Stability (WPS) in Auditory Selective Attention," Brain Res. Bull. **86**, 110–117, doi: `10.1016/j.brainresbull.2011.06.012`.

Mallat, S. (**1989**). "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," IEEE Trans. Patt. Anal. Mach. Int. **11**, 674–693, doi: `10.1109/34.192463`.

Malmivuo, J., and Plonsey, R. (**1995**). *Bioelectromagnetism : Principles and Applications of Bioelectric and Biomagnetic Fields*, 1 ed. (Oxford University Press, Oxford, UK), isbn: `978-0195058239`.

Mandyam, G., and Ahmed, N. (**1996**). "The Discrete Laguerre Transform: Derivation and Applications," IEEE Trans. Sig. Proc. **44**(12), 2925–2931, doi: `10.1109/78.553468`.

Mann, S., and Haykin, S. (**1991**). "The Chirplet Transform: A Generalization of Gabor's Logon Transform," in *Proc. Vision Interface '91*, pp. 205–212, issn: `0843-803x`.

Mariam, M., Delb, W., Corona-Strauss, F. I., Bloching, M., and Strauss, D. J. (**2009**). "Comparing the Habituation of the Late Auditory Potentials to Loud and Soft Sounds," Physiol. Meas. **30**(2), 141–153, doi: `10.1088/0967-3334/30/2/003`.

Mariam, M., Delb, W., Schick, B., and Strauss, D. J. (**2012**). "Feasibility of an Objective Electrophysiological Loudness Scaling: A kernel–based Novelty Detection Approach," Artific. Intell. Med. **55**(3), 185–195, doi: `10.1016/j.artmed.2012.03.004`.

McLaughlin, M., Verschooten, E., and Joris, P. (**2010**). "Oscillatory Dipoles As a Source of Phase Shifts in Field Potentials in the Mammalian Auditory Brainstem," J. Neurosci. **30**(40), 13472–13487, doi: `10.1523/jneurosci.0294-10.2010`.

McPherson, D. L., and Starr, A. (**1993**). "Binaural Interaction in Auditory Evoked Potentials: Brainstem, Middle– and Long–Latency Components," Hear. Res. **66**(1), 91–98, doi: `10.1016/0378-5955(93)90263-z`.

Meddis, R., Lopez-Poveda, E. A., Fay, R. R., and Popper, A. N. (**2010**). *Computational Models of the Auditory System*, 1 ed. (Springer, Boston, USA), isbn: `978-1441913708`.

Michelini, S., Arslan, E., Prosser, S., and Pedrielli, F. (**1982**). "Logarithmic Display of Auditory Evoked Potentials," J. Biomed. Eng. **4**(1), 62–64, doi: `10.1016/0141-5425(82)90028-0`.

Mohseni, H. R., Wilding, E. L., and Sanei, S. (**2007**). "Preprocessing of Event–Related Potential Signals via Kalman Filtering and Smoothing," in *15th Int. IEEE Conf. Digit. Sig. Proc.*, pp. 179–182, doi: `10.1109/icdsp.2007.4288548`.

Møller, A. R. (**2011**). *Hearing: Anatomy, Physiology and Disorders of the Auditory System*, 3 ed. (Plural Publishing, San Diego, USA), isbn: `978-1597564274`.

Moore, J. K. (**1987**). "The Human Auditory Brain Stem: A Comparative View," Hear. Res. **29**(1), 1–32, doi: `10.1016/0378-5955(87)90202-4`.

Mortezapouraghdam, Z., Haab, L., Corona-Strauss, F. I., Steidl, G., and Strauss, D. J. (**2015**). "Assessment of Long–Term Habituation Correlates in Event–Related Potentials Using a von Mises Model," IEEE Trans. Neural Syst. Rehabil. Eng. **23**, 363–373, doi: `10.1109/tnsre.2014.2361614`.

Müller, N., Schlee, W., Hartmann, T., Lorenz, I., and Weisz, N. (**2009**). "Top–Down Modulation of the Auditory Steady–State Response in a Task–Switch Paradigm," Front. Hum. Neurosci. **3**(1), 1–9, doi: `10.3389/neuro.09.001.2009`.

Murali, S., and Kulish, V. V. (**2007**). "Modeling of Evoked Potentials of Electroencephalograms: An Overview," Digit. Sig. Proc. **17**(3), 665–674, doi: `10.1016/j.dsp.2006.09.004`.

Mustaffa, I., Trenado, C., Schwerdtfeger, K., and Strauss, D. J. (**2010**). "Denoising of Single–Trial Matrix Representations Using 2D Nonlinear Diffusion Filtering," J. Neurosci. Meth. **185**(2), 284–292, doi: `10.1016/j.jneumeth.2009.09.017`.

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., and Winkler, I. (**2001**). "'Primitive Intelligence' in the Auditory Cortex," Trends Neurosci. **24**(5), 283–288, doi: `10.1016/s0166-2236(00)01790-2`.

Nelson, M. D., Hall, J. W., and Jacobson, G. P. (**1997**). "Factors Affecting the Recordability of Auditory Evoked Response Component Pb (P1)," J. Am. Acad. Audiol. **8**, 89–99.

O'Beirne, G. A., and Patuzzi, R. B. (**1999**). "Basic Properties of the Sound-Evoked Post–Auricular Muscle Response (PAMR)," Hear. Res. **138**(1–2), 115–132, doi: `10.1016/s0378-5955(99)00159-8`.

Oken, B. S., and Phillips, T. S. (**2009**). "Evoked Potentials: Clinical," in *Encyclopedia of Neuroscience*, edited by L. R. Squire (Academic Press, Oxford, UK), pp. 19–28, doi: `10.1016/b978-008045046-9.00587-8`.

Özdamar, Ö., and Bohórquez, J. (**2006**). "Signal–to–Noise Ratio and Frequency Analysis of Continuous Loop Averaging Deconvolution (CLAD) of Overlapping Evoked Potentials," J. Acoust. Soc. Am. **119**(1), 429–438, doi: `10.1121/1.2133682`.

Özdamar, Ö., Bohórquez, J., and Ray, S. S. (**2007**). "Pb(P1) Resonance at 40 Hz: Effects of High Stimulus Rate on Auditory Middle Latency Responses (MLRs) Explored using Deconvolution," Clin. Neurophysiol. **118**, 1261–1273, doi: `10.1016/j.clinph.2007.02.016`.

Özdamar, Ö., and Delgado, R. E. (**1996**). "Measurement of Signal and Noise Characteristics in Ongoing Auditory Brainstem Response Averaging," Ann. Biomed. Eng. **24**(6), 702–715, doi: `10.1007/bf02684183`.

Özdamar, Ö., Delgado, R. E., Yavuz, E., Thombre, K. V., and Acikgoz, N. (**2003**). "Deconvolution of Auditory Evoked Potentials Obtained at High Stimulus Rates," in *1st Int. IEEE Eng. Med. Biol. Soc. Conf. Neur. Eng.*, pp. 285–288, doi: `10.1109/cne.2003.1196816`.

Özdamar, Ö., and Kraus, N. (**1983**). "Auditory Middle–Latency Responses in Humans," Audiol. **22**(1), 34–49, doi: `10.3109/00206098309072768`.

Pantev, C., Elbert, T., Ross, B., Eulitz, C., and Terhardt, E. (**1996**). "Binaural Fusion and the Representation of Virtual Pitch in the Human Auditory Cortex," Hear. Res. **100**(1), 164–170, doi: `10.1016/0378-5955(96)00124-4`.

Pantev, C., Hoke, M., Lehnertz, K., Lütkenhöner, B., Anogianakis, G., and Wittkowski, W. (**1988**). "Tonotopic Organization of the Human Auditory Cortex Revealed by Transient Auditory Evoked Magnetic Fields," Electroencephalogr. Clin. Neurophysiol. **69**(2), 160–170, doi: `10.1016/0013-4694(88)90211-8`.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., and Hoke, M. (**1998**). "Increased Auditory Cortical Representation in musicians," Nat. **392**, 811–814, doi: `10.1038/33918`.

Peng, X., Chen, Y., Wang, T., Ding, L., and Tan, X. (**2017**a). "Noise Attenuation Estimation for Maximum Length Sequences in Deconvolution Process of Auditory Evoked Potentials," Comp. Mat. Meth. Med. **2017**, 1–9, doi: `10.1155/2017/3927486`.

Peng, X., Yuan, H., Chen, W., Wang, T., and Ding, L. (**2017**b). "New Metric for Optimizing Continuous Loop Averaging Deconvolution (CLAD) Sequences Under the 1/f Noise Model," PLOS ONE **12**(4), 1–24, doi: `10.1371/journal.pone.0175354`.

Picton, T. W. (**2010**). *Human Auditory Evoked Potentials*, 1 ed. (Plural Publishing, San Diego, USA), isbn: `978-1597563628`.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., and Taylor, M. J. (**2000**). "Guidelines for Using Human Event–Related Potentials to Study Cognition: Recording Standards and Publication Criteria," Psychophysiol. **37**(2), 127—152, doi: `10.1017/s0048577200000305`.

Picton, T. W., and Hillyard, S. A. (**1974**). “Human Auditory Evoked Potentials. II: Effects of Attention,” Electroencephalogr. Clin. Neurophysiol. **36**(2), 191–199, doi: `10.1016/0013-4694(74)90156-4`.

Picton, T. W., Hillyard, S. A., Krausz, H. I., and Galambos, R. (**1974**). “Human Auditory Evoked Potentials. I: Evaluation of Components,” Electroencephalogr. Clin. Neurophysiol. **36**(2), 179–190, doi: `10.1016/0013-4694(74)90155-2`.

Pólya, G. (**1920**). “Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem,” Mathematische Zeitschrift **8**(3), 171–181, doi: `10.1007/bf01206525`.

Ponton, C., Eggermont, J. J., Khosla, D., Kwong, B., and Don, M. (**2002**). “Maturation of Human Central Auditory System Activity: Separating Auditory Evoked Potentials by Dipole Source Modeling,” Clin. Neurophysiol. **113**(3), 407–420, doi: `10.1016/s1388-2457(01)00733-7`.

Presacco, A., Bohórquez, J., Yavuz, E., and Özdamar, Ö. (**2010**). “Auditory Steady–State Responses to 40–Hz Click Trains: Relationship to Middle Latency, Gamma Band and Beta Band Responses Studied With Deconvolution,” Clin. Neurophysiol. **121**, 1540–1550, doi: `10.1016/j.clinph.2010.03.020`.

Press, W. (**2006**). “Discrete Radon Transform Has an Exact, Fast Inverse and Generalizes to Operations Other Than Sums Along Lines,” Proc. Nation. Academ. Sci. USA **103**(51), 19249–19254, doi: `10.1073/pnas.0609228103`.

Price, K. V., Storn, R. N., and Lampinen, J. A. (**2005**). Nat. Comp. Ser. *Differential Evolution: A Practical Approach to Global Optimization*, 1 ed. (Springer, Heidelberg, Germany), isbn: `978-3540209508`.

Prosser, S., Arslan, E., and Michelini, S. (**1981**). “Habituation and Rate Effect in the Auditory Cortical Potentials Evoked by Trains of Stimuli,” Arch. Otorhinolaryngol. **233**(2), 179–187, doi: `10.1007/bf00453642`.

Quian Quiroga, R., and Garcia, H. (**2003**). "Single–Trial Event–Related Potentials With Wavelet Denoising," Clin. Neurophysiol. **114**(2), 376–390, doi: `10.1016/s1388-2457(02)00365-6`.

Quian Quiroga, R., Sakowitz, O. W., Basar, E., and Schürmann, M. (**2001**). "Wavelet Transform in the Analysis of the Frequency Composition of Evoked Potentials," Brain Res. Protoc. **8**(1), 16–24, doi: `10.1016/s1385-299x(01)00077-0`.

Radon, J. (**1917**). "Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten," Berichte Sächsische Akademie der Wissenschaften **69**, 262–267.

Ramachandran, V. (**2002**). *Encyclopedia of the Human Brain*, 1 ed. (Academic Press, Massachusetts, USA), isbn: `978-0122272103`.

Rauschecker, J. P., and Tian, B. (**2000**). "Mechanisms and Streams for Processing of "what" and "where" in Auditory Cortex," Proc. Nation. Acad. Sci. USA **97**(22), 11800–11806, doi: `10.1073/pnas.97.22.11800`.

Rees, A. (**2009**). *Encyclopedia of Neuroscience: Medial Geniculate Body*, 2009 ed. (Springer, Heidelberg, Germany), isbn: `978-3540237358`.

Rees, A., and Palmer, A. R. (**2010**). *The Oxford Handbook of Auditory Science. Volume 2: The Auditory Brain*, 1 ed. (Oxford University Press, Oxford, UK), isbn: `978-0199233281`.

Reichenbach, C. S., Braiman, C., Schiff, N. D., Hudspeth, A. J., and Reichenbach, T. (**2016**). "The Auditory–Brainstem Response to Continuous, Non–Repetitive Speech is Modulated by the Speech Envelope and Reflects Speech Processing," Front. Comput. Neurosci. **10**, 1–47, doi: `10.3389/fncom.2016.00047`.

Riedel, H., Granzow, M., and Kollmeier, B. (**2001**). "Single–Sweep–Based Methods to Improve the Quality of Auditory Brain Stem Responses Part II: Averaging Methods," Zeitschrift für Audiologie **40**(2), 62–85.

Riedel, H., and Kollmeier, B. (**2006**). "Interaural Delay–Dependent Changes in the Binaural Difference Potential of the Human Auditory Brain Stem Response," Hear. Res. **218**(1), 5–19, doi: `10.1016/j.heares.2006.03.018`.

Roederer, J. G. (**2009**). *The Physics and Psychophysics of Music: An Introduction*, 4 ed. (Springer, Heidelberg, Germany), isbn: `978-0387094700`.

Rousselet, G. (**2012**). "Does Filtering Preclude Us from Studying ERP Time–Courses?," Front. Psych. **3**(131), 1–9, doi: `10.3389/fpsyg.2012.00131`.

Rudin, L. I., Osher, S., and Fatemi, E. (**1992**). "Nonlinear Total Variation Based Noise Removal Algorithms," Physica D: Nonlin. Phenom. **60**(1), 259–268, doi: `10.1016/0167-2789(92)90242-f`.

Samar, V. J., Swartz, K. P., and Raghuveer, M. R. (**1995**). "Multiresolution Analysis of Event–Related Potentials by Wavelet Decomposition," Brain Cogn. **27**(3), 398–438, doi: `10.1006/brcg.1995.1028`.

Schebsdat, E., Kohl, M. C., Corona-Strauss, F. I., Seidler, H., and Strauss, D. J. (**2018**). "Free–Field Evoked Auditory Brainstem Responses in Cochlear Implant Users," Audiol. Res. **8**(2), 44–53, doi: `10.4081/audiores.2018.216`.

Schebsdat, E., Kohl, M. C., Geissler, G., Bohórquez, J., Strauss, D. J., and Chalupper, J. (**2019**). "Quantification of Binaural Interaction Facilitated By Interaural Latency Compensation in Bimodal Cochlear Implantees," IEEE Trans. Neural Sys. Rehabil. Eng. (under review).

Scherg, M., Vajsar, J., and Picton, T. W. (**1989**). "A Source Analysis of the Late Human Auditory Evoked Potentials," J. Cogn. Neurosci. **1**(4), 336–355, doi: `10.1162/jocn.1989.1.4.336`.

Scherg, M., and von Cramon, D. (**1986**). "Evoked Dipole Source Potentials of the Human Auditory Cortex," Electroencephalogr. Clin. Neurophysiol. **65**(5), 344–360, doi: `10.1016/0168-5597(86)90014-6`.

Schmidt, R. F., Lang, F., and Heckmann, M. (**2017**). *Physiologie des Menschen*, 31 ed. (Springer, Heidelberg, Germany), isbn: `978-3662541210`.

Schubert, J. K., González–Trejo, E., Retz, M., Roesler, M., Corona-Strauss, F. I., Steidl, G., Teuber, T., and Strauss, D. J. (**2014**). "Dysfunctional Cortical Inhibition in Adult ADHD: Neural Correlates in Auditory Event–Related Potentials," J. Neurosci. Meth. **235**, 181–188, doi: `10.1016/j.jneumeth.2014.06.025`.

Scott, B. H., Leccese, P. A., Saleem, K. S., Kibuchi, Y., Mullarkey, M. P., Fukushima, M., Mishkin, M., and Saunders, R. C. (**2017**). "Intrinsic Connections of the Core Auditory Cortical Regions and Rostral Supratemporal Plane in the Macaque Monkey," Cereb. Cortex **27**(1), 109–840, doi: `10.1093/cercor/bhv277`.

Sekiyama, K., Kanno, I., Miura, S., and Sugita, Y. (**2003**). "Auditory–Visual Speech Perception Examined by fMRI and PET," Neurosci. Res. **47**(3), 277–287, doi: `10.1016/s0168-0102(03)00214-1`.

Slugocki, C., Bosnyak, D., and Trainor, L. J. (**2017**). "Simultaneously–Evoked Auditory Potentials (SEAP): A New Method for Concurrent Measurement of Cortical and Subcortical Auditory–Evoked Activity," Hear. Res. **345**, 30–42, doi: `10.1016/j.heares.2016.12.014`.

Smith, E., Duede, S., Hanrahan, S., Davis, T., House, P., and Greger, B. (**2013**). "Seeing Is Believing: Neural Representations of Visual Stimuli in Human Auditory Cortex Correlate with Illusory Auditory Perceptions," PLOS ONE **8**(9), 1–7, doi: `10.1371/journal.pone.0073148`.

Stanishevskaya, I. (**2016**). "The Auditory Pathway of the Human Brain," url: `http://inessaskaya.com/portfolio/auditory-pathway/`, accessed: 04 Jan 2019.

Stein, B. E., Stanford, T. R., and Rowland, B. A. (**2009**). "The Neural Basis of Multisensory Integration in the Midbrain: Its Organization and Maturation," Hear. Res. **258**(1), 4–15, doi: `10.1016/j.heares.2009.03.012`.

Stockard, J. J., and Rossiter, V. S. (**1977**). “Clinical and Pathologic Correlates of Brain Stem Auditory Response Abnormalities,” Neurol. **27**(4), 316–316, doi: `10.1212/wnl.27.4.316`.

Storn, R. N. (**2007**). “Designing Nonstandard Filters with Differential Evolution,” in *Streamlin. Digit. Sig. Proc.*, pp. 25–32, doi: `10.1002/9780470170090.ch3`.

Storn, R. N., and Price, K. V. (**1997**). “Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces,” J. Glob. Optim. **11**, 341–359, doi: `10.1023/a:1008202821328`.

Strauss, D. J., Teuber, T., Steidl, G., and Corona-Strauss, F. I. (**2013**). “Exploiting the Self-Similarity in ERP Images by Nonlocal Means for Single–Trial Denoising,” IEEE Trans. Neural Sys. Rehabil. Eng. **21**, 576–583, doi: `10.1109/tnsre.2012.2220568`.

Suga, N., Gao, E., Zhang, Y., Ma, X., and Olsen, J. F. (**2000**). “The Corticofugal System for Hearing: Recent Progress,” Proc. Nation. Acad. Sci. USA **97**(22), 11807–11814, doi: `10.1073/pnas.97.22.11807`.

Suga, N., Ji, W., Ma, X., Tang, J., Xiao, Z., and Yan, J. (**2011**). *Corticofugal Modulation and Beyond for Auditory Signal Processing and Plasticity*, 313–352 (Springer, New York, USA), doi: `10.1007/978-1-4419-7070-1_11`.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., and Pernier, J. (**1996**). “Stimulus Specificity of Phase–Locked and Non–Phase–Locked 40–Hz Visual Responses in Human,” J. Neurosci. **16**(13), 4240–4249, doi: `10.1523/jneurosci.16-13-04240.1996`.

Taylor, J. R. (**1997**). *Introduction To Error Analysis: The Study of Uncertainties in Physical Measurements*, 2 ed. (University Science Books, New Jersey, USA), isbn: `978-0935702750`.

Ungan, P., Yagcioglu, S., and Goksoy, C. (**2001**). “Differences Between the N1 Waves of the Responses to Interaural Time and Intensity Disparities: Scalp Topography and Dipole Sources,” Clin. Neurophysiol. **112**(3), 485–498, doi: `10.1016/s1388-2457(00)00550-2`.

Upton, G., and Cook, I. (**2008**). *A Dictionary of Statistics*, 2 ed. (Oxford University Press, Oxford, UK), isbn: `978-0199541454`.

Verhulst, S., Bharadwaj, H. M., Mehraei, G., Shera, C. A., and Shinn-Cunningham, B. G. (**2015**). “Functional Modeling of the Human Auditory Brainstem Response to Broadband Stimulation,” J. Acoust. Soc. Am. **138**(3), 1637–1659, doi: `10.1121/1.4928305`.

Verkindt, C., Bertrand, O., Perrin, F., Echallier, J.-F., and Pernier, J. (**1995**). “Tonotopic Organization of the Human Auditory Cortex: N100 Topography and Multiple Dipole Model Analysis,” Electroencephalogr. Clin. Neurophysiol. **96**(2), 143–156, doi: `10.1016/0168-5597(94)00242-7`.

Villa, J., Rodríguez–Vera, R., Antonio Quiroga, J., de la Rosa, I., and González, E. (**2010**). “Anisotropic Phase–Map Denoising Using a Regularized Cost–Function With Complex–Valued Markov–Random–Fields,” Opt. Lasers Eng. **48**, 650–656, doi: `10.1016/j.optlaseng.2010.02.002`.

Wang, T., Özdamar, Ö., Bohórquez, J., Shen, Q., and Cheour, M. (**2006**). “Wiener Filter Deconvolution of Overlapping Evoked Potentials,” J. Neurosci. Meth. **158**, 260–270, doi: `10.1016/j.jneumeth.2006.05.023`.

Wang, Z., Maier, A., Leopold, D. A., Logothetis, N. K., and Liang, H. (**2007**). “Single–Trial Evoked Potential Estimation Using Wavelets,” Comp. Biol. Med. **37**(4), 463–473, doi: `10.1016/j.compbiomed.2006.08.011`.

Winer, J. A. (**2005**). “Decoding the Auditory Corticofugal Systems,” Hear. Res. **207**(1), 1–9, doi: `10.1016/j.heares.2005.06.007`.

Winer, J. A., and Lee, C. C. (**2007**). “The Distributed Auditory Cortex,” Hear. Res. **229**(1), 3–13, doi: `10.1016/j.heares.2007.01.017`.

Winer, J. A., Miller, L. M., Lee, C. C., and Schreiner, C. E. (**2005**). “Auditory Thalamocortical Transformation: Structure and Function,” Trends Neurosci. **28**(5), 255–263, doi: `10.1016/j.tins.2005.03.009`.

Woldorff, M. G. (**1993**). “Distortion of ERP Averages due to Overlap from Temporally Adjacent ERPs: Analysis and Correction,” Psychophysiol. **30**, 98–119, doi: `10.1111/j.1469-8986.1993.tb03209.x`.

Woldorff, M. G., and Hillyard, S. A. (**1991**). “Modulation of Early Auditory Processing During Selective Listening to Rapidly Presented Tones,” Electroencephalogr. Clin. Neurophysiol. **79**(3), 170–191, doi: `10.1016/0013-4694(91)90136-r`.

Xiao, Z., and Suga, N. (**2001**). “Modulation of Cochlear Hair Cells by the Auditory Cortex in the Mustached Bat,” Nat. Neurosci. **5**, 57–63, doi: `10.1038/nn786`.

Yeung, N., Bogacz, R., Holroyd, C. B., and Cohen, J. D. (**2004**). “Detection of Synchronized Oscillations in the Electroencephalogram: An Evaluation of Methods,” Psychophysiol. **41**(6), 822–832, doi: `10.1111/j.1469-8986.2004.00239.x`.

Zhang, F., Eliassen, J., Andersson, J., Scheifele, P., and Brown, D. (**2009**). “The Time Course of the Amplitude and Latency in the Auditory Late Response Evoked by Repeated Tone Bursts,” J. Am. Acad. Audiol. **20**(4), 239–250, doi: `10.3766/jaaa.20.4.4`.

Zirn, S., Angermeier, J., Arndt, S., Aschendorff, A., and Wesarg, T. (**2019**). “Reducing the Device Delay Mismatch Can Improve Sound Localization in Bimodal Cochlear Implant or Hearing–Aid Users,” Trends Hear. **23**, 1–13, doi: `10.1177/2331216519843876`.

Zwicke, P. E., and Kiss, I. (**1983**). “A New Implementation of the Mellin Transform and its Application to Radar Classification of Ships,” IEEE Trans. Patt. Anal. Mach. Int. **PAMI-5**(2), 191–199, doi: `10.1109/tpami.1983.4767371`.

List of Publications

Journal Papers (MEDLINE listed)

Kohl, M. C., Schebsdat, E., Schneider, E. N., Niehl, A., Strauss, D. J., Özdamar, Ö., and Bohórquez, J. (**2019**). ”Fast Acquisition of Full–Range Auditory Event–Related Potentials Using an Interleaved Deconvolution Approach,” J. Acoust. Soc. Am. **145**, 540–550, doi: `10.1121/1.5087825`.

Schebsdat, E.,[1] **Kohl, M. C.**,[1] Geissler, G., Bohórquez, J., Strauss, D. J., and Chalupper, J. (**2019**). ”Quantification of Binaural Interaction Facilitated By Interaural Latency Compensation in Bimodal Cochlear Implantees,” IEEE Trans. Neur. Sys. Rehabil. Eng. (under review).

Schebsdat, E., **Kohl, M. C.**, Corona–Strauss, F. I., Seidler, H., and Strauss, D. J. (**2018**). ”Free-Field Evoked Auditory Brainstem Responses in Cochlear Implant Users,” Audiol. Res. **8**(2), 44–53, doi: `10.4081/audiores.2018.216`.

Conference Proceedings (MEDLINE listed)

Kohl, M. C., and Strauss, D. J. (**2016**). ”A Compact Representation for the Auditory Full–Range Response and its Fast Denoising Using an Image Filter Based on the Radon Transform,” in *38th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 5877–5880, doi: `10.1109/embc.2016.7592065`.

Klauke, I., **Kohl, M. C.**, Hannemann, R., Kornagel, U., Strauss, D. J., and Corona–Strauss, F. I. (**2015**). ”Impact of Monaural Frequency Compression on Binaural Fusion at the Brainstem Level,” in *37th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 1646–1649, doi: `10.1109/embc.2015.7318691`.

---

[1] Both authors contributed equally to this work

Conference Proceedings (IEEE listed)

Kohl, M. C., Schebsdat, E., Schneider, E. N., and Strauss, D. J. (**2019**). ”Denoising of Single–Trial Event–Related Potentials by Shrinkage and Phase Regularization of Analytic Wavelet Filterbank Coefficients,” in *9th Int. IEEE Eng. Med. Biol. Soc. Conf. Neur. Eng.*, pp. 251–254, doi: `10.1109/ner.2019.8717148`.

González–Trejo, E., **Kohl, M. C.**, Steinbach, A., Mögele, H., Pfleger, N., and Strauss, D. J. (**2015**). ”Correlation Between Cortical Inhibition and Auditory Stream Segregation in a Driving Environment,” in *7th Int. IEEE Eng. Med. Biol. Soc. Conf. Neur. Eng.*, pp. 747–750, doi: `10.1109/ner.2015.7146731`.

González–Trejo, E., Herrmann, D., **Kohl, M. C.**, Corona–Strauss, F. I., and Strauss, D. J. (**2013**). ”Paired Chirp Evoked Cortical Inhibition and its Behavioral Correlates in a Speech Intelligibility Task,” in *6th Int. IEEE Eng. Med. Biol. Soc. Conf. Neur. Eng.*, pp. 1350–1353, doi: `10.1109/ner.2013.6696192`.

Acknowledgments

It is a well–known fact that a scientific work like the present dissertation doesn't write itself in a vacuum. Many people are owed my sincerest thanks at this point, each of them having contributed in their very own way to the success of this work, and it wouldn't have been possible without them.

First of all, I'd like to express my gratitude to my supervisor **Prof. Dr. rer. nat. Dr. rer. med. habil. Daniel J. Strauss**, who ultimately convinced me to graduate about a topic at the intersection of audiology, systems neuroscience and signal processing back when I wasn't yet aware of my inclination to this field. It is only after this endeavour that I can testify he was right all along, and I am very grateful that he provided me with an interesting *and* demanding work environment during my time at his chair.

Erik Schebsdat, M.Sc., by far my closest academic workmate – and fellow sufferer in countless hours of data analysis, out–of–the–box thinking and divisions by (well, *almost*) zero under the slightly unhealthy influence of repeated caffeination – deserves my thanks perhaps more than anyone else. If it were not for him, I'd probably have quit somewhere along the way. Also, I owe special thanks to **Elena N. Schneider, M.Sc.** and **Annika Niehl, M.Sc.**, who both assisted me multiple times in data acquisition when there was simply too much on my own schedule. Many thanks go to the whole team of the *Systems Neuroscience & Neurotechnology Unit*, namely **Patrick J. Schäfer, M.Sc.**, **Philipp Flotho, M.Sc.**, **Mayur Bhamborae, M.Sc.**, **Tobias Grün, M.Sc.**, **Andreas Schröer, M.Sc.**, **Dominic G. M. Laubner, M.Sc.** (who for reasons beyond my understanding stubbornly refused to believe that I would have the guts to sneak a full–blown biblical quotation somewhere in this document, hereby[1] to be proven wrong), **Dr. rer. med. Corinna Bernarding**, **Dr. rer. nat. Lars Haab** (special thanks to you for your guidance during the structuring of this manuscript), **Dr. rer. med. Farah I. Corona–Strauss** and of course all those not named here for their valuable suggestions and continuous support.

---

[1] *„Be helpful to all who may have doubts.”*, Jude 1:22, American Bible Society (1995)

I'd like to express my gratitude to **Jorge Bohórquez, Ph.D.**, **Özcan Özdamar, Ph.D.** and the team of the *Biomedical Engineering Department* at the *University of Miami* for having me as a visiting researcher at their chair, if only for a brief time. I yet have to experience their combined levels of friendly demeanour, professional expertise and the willingness to develop and share ideas and methods elsewhere, and it is largely due to our fruitful debates that I later was able to contribute methodological improvements to the field of deconvolution–based ERP acquisition myself.

My thanks also go to my now workmates **Gunnar Geissler, M.Sc.** and **Dr. Josef Chalupper** at *Advanced Bionics GmbH*, whom I already had the pleasure to work with and learn from during the research project *Bimodal Fusion*, kindly funded by the German Federal Ministry of Education and Research (BMBF). Being a partner in the same project, I owe special thanks to **Ing. Rainer Thie** of *Pilot Blankenfelde GmbH* for sharing his considerable empirical experience in favorable ERP acquisition parameters and, more importantly, his determination to actually *integrate* the algorithms developed in this work into a commercially available medical–grade hardware platform, making them available for clinical application in the very near future – to quote William James :

„A difference that makes no difference is no difference at all."

Last but not least, my sincerest gratitude goes to my dear friends and my beloved family – my partner **Johanna E. Schmitz**, my outstanding friend **Alexander Steinbach, M.Sc.** (above all for our many difficult yet illuminating conversations and the proofreading of this work) and my parents **Sylvia A. Kohl** and **Peter W. Kohl**. Thank you for bearing me during this arduous stage of my life.