



**KATHOLIEKE UNIVERSITEIT LEUVEN**  
FACULTEIT INGENIEURSWETENSCHAPPEN  
DEPARTEMENT ELEKTROTECHNIEK  
Kasteelpark Arenberg 10, 3001 Leuven (Heverlee)  
In samenwerking met:  
FACULTEIT GENEESKUNDE  
DEPARTEMENT NEUROWETENSCHAPPEN  
Herestraat 49, 3000 Leuven

**IMPROVING AUDITORY STEADY-STATE  
RESPONSE DETECTION USING MULTICHANNEL  
EEG SIGNAL PROCESSING**

Promoters:  
Prof. dr. J. Wouters  
Prof. dr. ir. M. Moonen

Proefschrift voorgedragen tot  
het behalen van het doctoraat  
in de ingenieurswetenschappen  
door  
**Bram VAN DUN**

December 2008





**KATHOLIEKE UNIVERSITEIT LEUVEN**  
FACULTEIT INGENIEURSWETENSCHAPPEN  
DEPARTEMENT ELEKTROTECHNIEK  
Kasteelpark Arenberg 10, 3001 Leuven (Heverlee)  
In samenwerking met:  
FACULTEIT GENEESKUNDE  
DEPARTEMENT NEUROWETENSCHAPPEN  
Herestraat 49, 3000 Leuven

**IMPROVING AUDITORY STEADY-STATE  
RESPONSE DETECTION USING MULTICHANNEL  
EEG SIGNAL PROCESSING**

Jury:

Prof. dr. ir. Y. Willems, voorzitter  
Prof. dr. J. Wouters, promotor  
Prof. dr. ir. M. Moonen, promotor  
Prof. dr. ir. S. Van Huffel  
Prof. dr. ir. H. Van hamme  
Prof. dr. ir. A.F.M. Snik – Radboud Univ. Nijmegen  
Prof. dr. T.W. Picton – Univ. of Toronto

Proefschrift voorgedragen tot  
het behalen van het doctoraat  
in de ingenieurswetenschappen  
door

**Bram VAN DUN**

U.D.C. 534.7

December 2008

© Katholieke Universiteit Leuven – Faculteit Toegepaste Wetenschappen  
Arenbergkasteel, B-3001 Heverlee (Belgium)

Alle rechten voorbehouden. Niets uit deze uitgave mag vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotocopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm or any other means without written permission from the publisher.

D/2008/7515/120

ISBN 978-94-6018-011-8

# Abstract

The ability to hear and process sounds is crucial. For adults, the inevitable ongoing aging process reduces the quality of the speech and sounds one perceives. If this effect is allowed to evolve too far, social isolation may occur. For infants, a disability in processing sounds results in an inappropriate development of speech, language, and cognitive abilities. To reduce the handicap of hearing loss in children, it is important to detect the hearing loss early and to provide effective rehabilitation. As a result, hearing of all newborns needs to be screened. If the outcome of the screening does not indicate normal hearing, more detailed hearing assessment is required. However, standard behavioral testing is not possible, so that assessment has to rely on objective physiological techniques that are not influenced by sleep or sedation. The last few decades, the use of auditory steady-state responses (ASSRs) has been investigated as an objective technique to assess hearing thresholds at different frequencies.

In this research project, we focus on reducing the required recording time of the ASSR technique and on improving its robustness against unwanted artifacts, generated by e.g. muscle activity, eye blinks, and electrode cable movement. This objective is achieved by processing multichannel electroencephalogram (EEG) recordings. First, we build a setup that allows us to apply custom made stimuli and to record multichannel EEG. Second, the effect of two multichannel processing techniques applied on these data is investigated. Both an independent component analysis (ICA) based technique and a multichannel Wiener filter (MWF) based approach show that a significant measurement time reduction is possible when compared with standard single channel recordings. Afterwards, the ICA- and MWF-based approaches are incorporated into a procedural multichannel framework that is constructed from elements of detection theory. It is shown that this detection theory based approach increases the number of detections significantly when compared with a noise-weighted single channel technique, in the case of artifact-rich EEG. Finally, the optimal electrode positions are determined for the recording of ASSRs originating mainly from the brainstem (and the auditory cortex). After processing with the multichannel EEG processing techniques presented in this work, these positions guarantee a close-to-optimal assessment of the subject's hearing thresholds.



# Korte Inhoud

De mogelijkheid om geluiden te horen en te verwerken is cruciaal voor zowel jong als oud. Voor kinderen betekent gehoorverlies een obstakel voor een normale spraak- en taalontwikkeling. Vooral voor hen is het belangrijk om dit gehoorverlies zo vroeg mogelijk op te sporen en er gepast op te reageren. Om deze reden zou het gehoor van alle pasgeborenen moeten worden gecontroleerd. Als het resultaat van deze ‘screening’ niet op een normaal gehoor wijst, is meer gedetailleerde gehoorschattting nodig. Het probleem hier is wel dat de standaard gebruikte gedragstesten niet kunnen gebruikt worden. Om deze reden moeten deze testen terugvallen op objectieve fysiologische technieken die niet beïnvloed worden door slaap of sedatie. De laatste decennia werd er gefocust op een techniek die gebruikt maakt van auditieve steady-state responsen (ASSR) om gehoordrempels te schatten op verschillende frequenties.

In dit onderzoeksproject proberen we de meettijd te verkorten van de ASSR-techniek en zijn robuustheid te vergroten tegen ongewenste artefacten zoals spieractiviteit. Dit doel wordt bereikt door meerkanaals verwerking van elektroencephalogram (EEG) metingen. Om te beginnen hebben we een opstelling gebouwd waarmee het mogelijk is om zelfgemaakte stimuli aan te bieden aan de proefpersoon en om meerkanaals EEG op te meten. Daarna wordt het effect van twee types meerkanaals signaalverwerking onderzocht die toegepast worden op deze meerkanaals data. Zowel het gebruik van onafhankelijke component analyse (ICA) als een meerkanaals Wiener filter (MWF) maakt het mogelijk om een significante meettijdreductie te bekomen ten opzichte van de standaard éénkanaals metingen. Nadien worden deze ICA- en MWF-gebaseerde benaderingen samengesmolten in een proceduraal meerkanaals raamwerk dat opgebouwd is met bouwstenen uit de detectietheorie. Er wordt aangetoond dat deze benadering het aantal detecties significant vergroot vergeleken met een ruisgewogen éénkanaals techniek wanneer EEG wordt gebruikt dat veel artefacten bevat. Om af te sluiten worden de optimale elektrodeposities bepaald voor het opmeten van ASSR die hoofdzakelijk gegenereerd worden in de hersenstam (en de auditieve cortex). Deze posities garanderen een bijna-optimale schatting van de gehoordrempels van de proefpersoon.





# Glossary

## Mathematical Notation

$a$	scalar $a$
$\mathbf{v}$	vector $\mathbf{v}$
$\mathbf{M}$	matrix $\mathbf{M}$
$\mathbf{M}^T$	transpose of matrix $\mathbf{M}$
$\mathbf{M}^*$	complex conjugate of matrix $\mathbf{M}$
$\mathbf{M}^H = (\mathbf{M}^*)^T$	Hermitian transpose of matrix $\mathbf{M}$
$\mathbf{M}^{-1}$	inverse of matrix $\mathbf{M}$
$\mathbf{X}(k : l, :)$	rows $k$ up to $l$ of matrix $\mathbf{X}$
$\mathbf{X}(:, k : l)$	columns $k$ up to $l$ of matrix $\mathbf{X}$
$\mathbf{x}(i)$	element $i$ of signal $\mathbf{x}$
$\mathbf{x}(t)$	time domain representation of signal $\mathbf{x}$
$X(f)$	frequency domain representation of signal $\mathbf{x}$
$\hat{a}, \hat{\mathbf{v}}, \hat{\mathbf{M}}$	estimate of a scalar $a$ , vector $\mathbf{v}$ , matrix $\mathbf{M}$
$\text{round}(x)$	round $x \in \mathbb{R}$ to the nearest integer
$a = b$	$a$ is equal to $b$
$a \sim b$	$a$ is similar to $b$
$a < b$	$a$ is smaller than $b$
$\mathbf{A} \equiv \mathbf{B}$	matrix $\mathbf{A}$ is equivalent to matrix $\mathbf{B}$
$\pm$	approximately
$ \cdot $	absolute value
$\ \cdot\ _2$	2-norm
$\otimes$	Kronecker product
$\mu_x$	mean of $x$
$\sigma_x$	standard deviation of $x$
$\text{GEVD}(\mathbf{M}, \mathbf{X})$	generalized eigenvalue decomposition of matrices $\mathbf{M}$ and $\mathbf{X}$
$\mathcal{E}\{\cdot\}$	expectation operator
$\text{vec}(\mathbf{M})$	stacks the columns of matrix $\mathbf{M}$ into one column vector $\mathbf{m}$
$\mathbf{v}_{\max}\{\mathbf{M}\}$	eigenvector associated with largest eigenvalue of matrix $\mathbf{M}$
$\arg\{\max_a f(x, a)\}$	value of $a$ for which the value of $f(x, a)$ attains its maximum

## Fixed Symbols

$\alpha$	unknown (ASSR source) amplitude
$\beta$	frequency modulation (FM) index
$\Delta\phi$	fixed phase difference
$\Lambda(\mathbf{z})$	likelihood ratio
$\phi$	fixed phase
$\varphi$	time variable phase
$\sigma^2$	noise power
$\omega = 2\pi f$	pulsation
$\Upsilon(\mathbf{z})$	sufficient statistic
$D$	duty cycle
$f$	frequency-domain variable
$f_c$	carrier frequency
$f_m$	amplitude modulation (AM) frequency
$f'_m$	frequency modulation (FM) frequency
$f_s$	sampling frequency
$H$	hypothesis
$m$	number of channels
$\mathcal{M}$	minimum set
$n$	number of data points
$p$	statistical $p$ -value
$P_r$	response power
$q$	number of independent components
$r$	correlation coefficient
$r_s$	Spearman correlation coefficient
$t$	time-domain variable
$U$	upper limit
$\mathbf{0}$	zero vector or zero matrix
$\mathbf{A}$	mixing matrix
$\mathbf{d}$	steering vector
$\mathbf{D}$	steering matrix
$\mathbf{I}$	identity matrix
$\mathbf{K}^{mn \times mn}$	$mn \times mn$ spatio-temporal noise covariance matrix
$\mathbf{K}_{\text{spat}}$	spatial noise covariance matrix
$\mathbf{K}_{\text{temp}}$	temporal noise covariance matrix
$\mathbf{n}$	noise vector
$\mathbf{N}$	noise matrix
$\mathbf{Q}$	orthogonal matrix
$\mathbf{R}$	upper triangle matrix
$\mathbf{s}$	signal vector
$\mathbf{S}$	signal matrix
$\mathbf{w}$	weighting vector
$\mathbf{w}_{\text{opt}}$	optimal weighting vector

<b>W</b>	separating matrix
<b>z</b>	observation vector
<b>Z<sup>n×m</sup></b>	$n \times m$ observation matrix

## Acronyms and Abbreviations

ABR	Auditory Brainstem Response
AEP	Auditory Evoked Potential
AM	Amplitude Modulation
ALR	Auditory Late Response
AMLR	Auditory Middle-Latency Response
AMFR	Amplitude Modulation Following Response
AUC	Area Under the ROC-Curve
ASSR	Auditory Steady-State Response
BOA	Behavioral Observation Audiometry
cf.	<b>confer</b> : compare with
CI	Cochlear Implant
CT2	Circular $T^2$ statistic
Cz	Vertex
dB	decibel
DC	Duty Cycle
eABR	electrical ABR
eASSR	electrical ASSR
eCochG	Electrocochleogram
EEG	Electroencephalogram
EFR	Envelope Following Response
e.g.	<b>exempli gratia</b> : for example
Fpz	Forehead
FFR	Frequency Following Response
FFT	Fast Fourier Transform
FM	Frequency Modulation
FN	False Negative
FP	False Positive
GEVD	Generalized EigenValue Decomposition
GUI	Graphical User Interface
HL	Hearing Level
HT2	Hotelling $T^2$ statistic
i.e.	<b>id est</b> : that is
IAFM	Independent Amplitude and Frequency Modulation
IC	Independent Component
ICA	Independent Component Analysis
ICR	Instrumental Conditional Reflex

ISO	International Organization for Standardization
JADE	Joint Approximate Diagonalization of Eigenmatrices
JCIH	Joint Committee on Infant Hearing
lMa	Left Mastoid
MASTER	Multiple Auditory STEady-state Responses
MINT	Multiple INTensity
MM	Mixed Modulation
MMSE	Minimum Mean Square Error
MSC	Magnitude Squared Coherence
MSE	Mean Squared Error
MVDR	Minimum Variance Distortionless Response
MWF	Multichannel Wiener Filter
OAE	Otoacoustic Emissions
Oz	Occiput
P300	Auditory P300 Response
PC	Phase Coherence
PCI	Peripheral Component Interconnect
PCMCIA	Personal Computer Memory Card International Association
peSPL	peak-equivalent Sound Pressure Level
PSD	Power Spectral Density
RECD	Real-Ear-to-Coupler Difference
RESPL	Real-Ear Sound Pressure Level
RLS	Recursive Least Squares adaptive filter
rMa	Right Mastoid
RMS	Root Mean Square
ROC	Receiver Operating Characteristic
SOMA	Setup ORL for Multichannel ASSR
SNR	Signal-to-Noise Ratio
SPL	Sound Pressure Level
SSEP	Steady-State Evoked Potential
TN	True Negative
TP	True Positive
VRA	Visual Reinforcement Audiometry
w.r.t.	with respect to

# Contents

<b>Abstract</b>	<b>i</b>
<b>Korte Inhoud</b>	<b>iii</b>
<b>Glossary</b>	<b>v</b>
<b>Contents</b>	<b>ix</b>
<b>Samenvatting</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	2
1.2 Hearing and hearing impairment . . . . .	3
1.2.1 The auditory system . . . . .	3
1.2.2 Hearing impairment . . . . .	6
1.3 Detection and intervention . . . . .	7
1.3.1 The need for early detection . . . . .	7
1.3.2 Hearing aids and cochlear implants . . . . .	7
1.4 Hearing threshold estimation . . . . .	8
1.4.1 Subjective methods . . . . .	9
1.4.2 Objective methods . . . . .	10

1.5	Auditory evoked potentials . . . . .	11
1.5.1	Electrocochleogram responses . . . . .	13
1.5.2	Auditory brainstem responses . . . . .	13
1.5.3	Auditory middle–latency responses . . . . .	16
1.5.4	Auditory late latency responses . . . . .	18
1.5.5	Auditory P300 responses . . . . .	18
1.6	Research objectives . . . . .	19
1.7	Thesis outline . . . . .	20
<b>2</b>	<b>Auditory Steady–State Responses</b>	<b>23</b>
2.1	Theoretical overview of the ASSR . . . . .	23
2.1.1	How it started . . . . .	24
2.1.2	Physiological model . . . . .	24
2.1.3	Stimulus parameters . . . . .	26
2.1.4	Recording EEG . . . . .	29
2.1.5	Multiple stimulus ASSR . . . . .	30
2.1.6	Comparison with the auditory brainstem response . . . . .	30
2.2	Techniques reducing recording time . . . . .	32
2.2.1	Stimuli eliciting higher responses . . . . .	33
2.2.2	Independent stimuli with multiple stimulus ASSR . . . . .	34
2.2.3	Intelligent placement of electrodes . . . . .	34
2.2.4	Averaging . . . . .	35
2.2.5	Adaptive filtering . . . . .	37
2.2.6	Statistics . . . . .	37
2.2.7	Smarter guesses of hearing thresholds . . . . .	39
2.3	Conclusions . . . . .	40

<b>3</b>	<b>A Flexible Research Platform for Multichannel ASSR Measurements</b>	<b>43</b>
3.1	Introduction . . . . .	44
3.2	Materials and methods . . . . .	45
3.2.1	Hardware setup . . . . .	45
3.2.2	Overview of the software . . . . .	48
3.2.3	Internal structure . . . . .	53
3.2.4	Evaluation . . . . .	56
3.3	Results . . . . .	57
3.3.1	Amplitude and noise levels . . . . .	57
3.3.2	Hearing threshold difference scores . . . . .	57
3.4	Discussion . . . . .	58
3.5	Conclusions . . . . .	60
<b>4</b>	<b>Improving ASSR Detection Using Independent Component Analysis</b>	<b>61</b>
4.1	Introduction . . . . .	62
4.2	Methods . . . . .	62
4.2.1	Independent component analysis . . . . .	62
4.2.2	Assumed model . . . . .	63
4.2.3	Experimental setup . . . . .	64
4.2.4	Stimuli . . . . .	66
4.2.5	Five ways to process the available EEG dataset . . . . .	66
4.2.6	Performance measures . . . . .	72
4.3	Results . . . . .	74
4.3.1	Multichannel ICA: 7 channels and 7 components . . . . .	75
4.3.2	Multichannel ICA: variable number of channels and ICs . . . . .	77
4.3.3	Multichannel ICA: effects of fixed separating matrix . . . . .	78

4.3.4	Single channel ICA . . . . .	81
4.3.5	Combining single channel and multichannel techniques . . . . .	82
4.4	Discussion . . . . .	82
4.4.1	Multichannel ICA: 7 channels and 7 components . . . . .	82
4.4.2	Multichannel ICA: variable number of channels and ICs . . . . .	83
4.4.3	Multichannel ICA: effects of fixed separating matrix . . . . .	84
4.4.4	Single channel ICA . . . . .	85
4.4.5	Multichannel ICA: the use of artifact rejection . . . . .	85
4.5	Conclusions . . . . .	86
<b>5</b>	<b>Improving ASSR Detection Using Multichannel Wiener Filtering</b>	<b>87</b>
5.1	Introduction . . . . .	87
5.2	Theoretical background . . . . .	88
5.2.1	ASSR signal model . . . . .	88
5.2.2	Simplifications of the model compared with real-life . . . . .	91
5.2.3	Maximizing the output-SNR of the ASSR . . . . .	92
5.2.4	Linking output-SNR maximization with the MWF . . . . .	93
5.3	Experimental setup . . . . .	94
5.4	Results . . . . .	95
5.5	Discussion . . . . .	96
5.6	Conclusion . . . . .	97
<b>6</b>	<b>A Procedural Framework for ASSR Detection</b>	<b>99</b>
6.1	Introduction . . . . .	100
6.2	Theoretical background . . . . .	100
6.2.1	Detection theory . . . . .	101
6.2.2	A detection theory framework for ASSR processing . . . . .	102



6.2.3	Exploiting ASSR stationarity and spatio-temporal EEG stationarity . . . . .	102
6.2.4	Better estimates of $\mathbf{K}$ using a duty cycle stimulus . . . . .	104
6.2.5	Alternative approaches for the calculation of $\mathbf{w}_{\text{opt}}$ . . . . .	106
6.2.6	Performance measures . . . . .	107
6.3	Simulations using multichannel EEG with artificial ASSRs . . . . .	108
6.3.1	Simulation design . . . . .	109
6.3.2	Processing schemes . . . . .	109
6.3.3	Comparison of the different schemes . . . . .	112
6.3.4	Varying the duty cycle $D$ . . . . .	114
6.3.5	Robustness against noise bursts . . . . .	115
6.4	Practical evaluation using real-life multichannel EEG . . . . .	118
6.4.1	Evaluation design . . . . .	119
6.4.2	Processing schemes for evaluation . . . . .	121
6.4.3	Results . . . . .	121
6.5	General discussion . . . . .	127
6.6	Conclusions . . . . .	128
<b>7</b>	<b>Optimal electrode selection for multichannel EEG based de- tection of auditory cortex and brainstem ASSRs</b>	<b>129</b>
7.1	Introduction . . . . .	130
7.2	Methods . . . . .	131
7.2.1	Brainstem stimulation . . . . .	133
7.2.2	Auditory cortex stimulation . . . . .	134
7.3	Results . . . . .	134
7.3.1	Brainstem stimulation . . . . .	135
7.3.2	Auditory cortex stimulation . . . . .	147
7.4	Discussion . . . . .	149

7.4.1	Brainstem stimulation . . . . .	150
7.4.2	Auditory cortex stimulation . . . . .	152
7.5	Conclusions . . . . .	152
<b>8</b>	<b>General Conclusions and Further Research</b>	<b>155</b>
8.1	General conclusions . . . . .	156
8.1.1	A multichannel platform for ASSR measurements . . . . .	157
8.1.2	Development of multichannel signal processing techniques improving ASSR detection . . . . .	157
8.1.3	Unifying multichannel techniques into a simplified frame- work for ASSR detection . . . . .	159
8.2	Suggestions for further research . . . . .	160
8.2.1	The multichannel measurement platform . . . . .	160
8.2.2	Multichannel EEG signal processing . . . . .	161
	<b>List of Publications</b>	<b>179</b>
	<b>Curriculum Vitae</b>	<b>183</b>

## Samenvatting

# Snellere detectie van auditory steady–state responsen door meerkanaals EEG–signaalverwerking

## Motivatie

De mogelijkheid om geluiden te horen en te verwerken is cruciaal voor zowel jong als oud. Wanneer volwassenen ouder worden, vermindert de kwaliteit van het geluid en in het bijzonder van de spraak die men waarneemt. Dit effect wordt omschreven als gehoorverlies. Als het gehoorverlies te ernstig wordt, kan sociale isolatie optreden. Voor kinderen betekent dit verlies een obstakel voor een normale spraak- en taalontwikkeling. Vooral voor hen is het belangrijk om dit gehoorverlies zo vroeg mogelijk op te sporen en er gepast op te reageren.

Deze oproep tot vroege gehoorprobleemdetectie heeft in verschillende delen van de wereld al navolging gekregen in de vorm van een systematische ‘screening’. In Vlaanderen wordt sinds 1998 ongeveer 97 % van de pasgeborenen op gehoorverlies getest door Kind & Gezin. Dit komt overeen met zowat 60.000 kinderen per jaar. Zodra er iets abnormaals wordt opgemerkt, wordt het kind doorverwezen naar gespecialiseerde centra voor een uitgebreide diagnose. In deze centra wordt een meer gedetailleerde analyse uitgevoerd van het vermoede gehoorverlies. Deze analyse wordt bekomen aan de hand van een aantal audiologische tests zoals *tympanometrie*, *otoakoestische emissies* en een methode

om (frequentiespecifieke) gehoordrempels te bepalen. Voor dit laatste geval zou normaalgezien gedragsaudiometrie gebruikt worden. Bij gedragsaudiometrie wordt aan de patiënt gevraagd of hij/zij een bepaalde stimulus al dan niet gehoord heeft. Deze techniek kan in dit geval echter niet toegepast worden gezien de jonge leeftijd van de doorgewezen patiënt. Om deze redenen is men hier vooral aangewezen op objectieve technieken die geen medewerking vereisen van de patiënt en die niet beïnvloed worden door slaap of sedatie.

De meest gebruikte objectieve techniek voor de schatting van gehoordrempels is de techniek die gebruik maakt van de klikgeëvokeerde *auditive hersenstamrespons* (ABR – ‘auditory brainstem response’). Het nadeel van deze techniek is dat gehoordrempels niet frequentiespecifiek kunnen worden bepaald. Deze informatie kan echter heel nuttig zijn om hoorapparaten af te stellen of om te beslissen of moet overgegaan worden tot cochleaire implantatie. Er bestaan varianten van de ABR die frequentiespecifieke informatie leveren, maar deze zijn zowel tijdrovend als veeleisend op het gebied van technische expertise.

Een mogelijke techniek die tegemoet komt aan dit probleem is de *auditive steady-state respons* techniek (ASSR – ‘auditory steady-state response’). Deze benadering maakt gebruik van continue tonen in plaats van transiënte kliks. Het is mogelijk om responsen op verschillende testfrequenties tegelijk te genereren en op te meten. Op deze manier kan een frequentiespecifieke gehoordrempelbepaling uitgevoerd worden in minder dan een uur, wat niet mogelijk is met een frequentiespecifieke ABR.

Natuurlijk zijn er niet enkel voordelen verbonden met de ASSR-techniek. De responsen die met deze benadering worden opgewekt zijn grootteordes kleiner dan de responsen opgewekt met de ABR-benadering. De techniek is gevoelig voor ruis, en artefacten kunnen de meting grondig verstoren of zelfs onmogelijk maken. Daarom is het denkbaar dat een ASSR-meting, en dus de hoordrempelbepaling, nog altijd te lang kan duren.

Het werk dat u nu voor zich hebt liggen, stelt verscheidene methodes voor om de meettijd te verkorten van deze ASSR-metingen. Het bespreekt de problemen die zich voordoen ten gevolge van een overvloed aan ruis en artefacten, en toont hoe deze in zekere mate kunnen worden verwijderd.

## Hoofdstuk 1: Inleiding

Het gehoor is één van de vijf zintuigen en maakt het mogelijk om geluid te detecteren. **Sectie 1.2** beschrijft het *auditief systeem* en de mogelijkheid dat het waarnemen van geluid en spraak gedeeltelijk of volledig wegvalt. Deze conditie wordt gedefinieerd als *gehoorverlies* en moet zo snel mogelijk gedetecteerd worden (**Sectie 1.3**). Langdurige gehoorproblemen kunnen namelijk

voor sociale isolatie zorgen bij volwassenen. Bij kinderen zorgt een verminderde gehoorperceptie voor een vertraagde spraak- en taalontwikkeling, leerstoornissen en voor emotionele en sociale problemen. Volgens Yoshinaga-Itano et al. (1998) hebben kinderen waarvan het gehoorverlies vóór de leeftijd van zes maanden wordt gedetecteerd betere taalcapaciteiten dan kinderen met een detectie ná de leeftijd van zes maanden. Aangezien ongeveer één op de duizend kinderen geboren worden met permanent bilateraal gehoorverlies, is een adequate ‘screening’ en interventie noodzakelijk. Daarom stelde de Amerikaanse Joint Committee on Infant Hearing (2000) een globale standaard voor die aangaf om alle kinderen te ‘screenen’ met objectieve fysiologische technieken vóór de leeftijd van één maand. Wanneer een gehoorprobleem wordt gedetecteerd, moet deze audiologisch en klinisch worden gediagnosticeerd vóór de leeftijd van drie maanden, en multidisciplinair geïntervenieerd vóór de leeftijd van zes maanden. Als het permanent gehoorverlies niet chirurgisch of met medicatie kan opgelost worden, kan gebruik gemaakt worden van een hoorapparaat of een cochleair implantaat. Dit lost het probleem niet op, maar verlicht het wel. Een *hoorapparaat* wordt ingezet bij matig tot zwaar gehoorverlies. Het versterkt en comprimeert het binnenkomend geluid. Bij zeer zwaar gehoorverlies of doofheid wordt de auditieve zenuw rechtstreeks elektrisch gestimuleerd met behulp van een *cochleair implantaat*. Het is belangrijk om de ernst van het gehoorverlies te kennen alvorens men overgaat tot het gebruik van een hoorapparaat of tot cochleaire implantatie. Bij de keuze voor een hoorapparaat is een correcte afstelling enkel mogelijk als zowel de ernst als de configuratie (de frequentiespecifieke informatie) van het gehoorverlies gekend is. Deze informatie wordt grafisch samengevat in een *audiogram*. **Sectie 1.4** beschrijft verschillende subjectieve en objectieve methodes om een audiogram te bepalen. Voor dit manuscript is een specifieke objectieve categorie van belang, de auditieve geëvokeerde potentialen.

*Auditieve geëvokeerde potentialen* (AEP – ‘auditory evoked potentials’) zijn responsen opgewekt door het auditief systeem, gaande van de oren, de auditieve zenuw tot de auditieve gebieden in de hersenen. Deze responsen zijn een reactie op een aangeboden auditieve of akoestische stimulus. Er zijn verschillende types, die worden beschreven in **Sectie 1.5**. Een AEP wordt geregistreerd door middel van een *elektroencefalogram* (EEG), waarbij elektroden op de scalp worden aangebracht en hersenpotentialen worden opgemeten. Een belangrijke soort van AEP is de *auditieve hersenstamrespons* (ABR). Deze wordt opgewekt in de hersenstam als reactie op een akoestische klik. Tegenwoordig wordt de ABR gebruikt voor de bepaling van gehoordrempels, waarbij gekeken wordt naar de aanwezigheid van een bepaalde golfvorm bij een gegeven klikintensiteit. Indien deze golfvorm (golf V) aanwezig is in het EEG, dan werd de klik waargenomen tot in de hersenstam. Door de intensiteit van de klik telkens te verlagen tot er geen golf V meer waargenomen wordt, kan een gehoordrempel bepaald worden. Een speciaal geval van de ABR is de *auditieve steady-state respons* (ASSR). Door de kliks van de ABR-techniek snel op elkaar te laten

volgen, vloeien de gegenereerde golfvormen in elkaar over, wat aanleiding geeft tot een regimetoestand, namelijk een *steady-state* respons. Een uitgebreide uiteenzetting van de ASSR wordt gegeven in Hoofdstuk 2.

## Hoofdstuk 2: De auditory *steady-state* respons

De eerste auditieve responsen werden ongeveer zeventig jaar geleden geregistreerd (Davis, 1939). In de veertig jaar nadien was het auditief onderzoek vooral toegespitst op akoestische kliks en de reactie van het auditief systeem hierop. Galambos et al. (1981) waren de eerste onderzoekers om significante resultaten te behalen met gemoduleerde tonen. De algemene verzamelnaam voor gemoduleerde tonen die een continue sinusoidale respons uitlokken, zijn *auditieve steady-state responsen* (ASSR).

**Sectie 2.1** introduceert de ASSR en bespreekt alle facetten. Een stimulus die een ASSR kan opwekken, bestaat meestal uit een draaggolf die gemoduleerd wordt met een modulator. Deze draaggolf exciteert de cochlea op een specifieke plaats tussen de apex (lage frequenties) en de basis (hoge frequenties). Grof genomen wordt enkel de modulator doorgegeven aan de auditieve zenuw. De responsen fluctueren mee aan de modulatiefrequentie van de modulator. Afhankelijk van de modulatiefrequentie liggen de bronnen van deze responsen grotendeels in de hersenstam (modulatiefrequenties vanaf  $\pm 75$  Hz) of in de auditieve cortex (lagere modulatiefrequenties). Dit heeft interessante praktische gevolgen. De ASSR met modulatiefrequenties vanaf 75 Hz zijn zeer geschikt voor een frequentiespecifieke gehoordrempelbepaling. Hier fungeert elke audiometrische testtoon (500, 1000, 2000 en 4000 Hz) als een draaggolf die gemoduleerd wordt aan een bepaalde modulatiefrequentie.

Een frequentiespecifieke gehoordrempelbepaling gebruik makend van de ASSR-techniek gaat als volgt. Er wordt vertrokken van een stimulusintensiteit die ver genoeg boven de verwachte gehoordrempel ligt. De responsen op deze (gemoduleerde) stimuli worden geregistreerd door middel van een EEG. Deze responsen zijn een soort van ‘label’ voor de akoestische stimuli. Als ze voorkomen in het EEG geeft dit aan dat de gemoduleerde stimulus de cochlea op een specifieke plaats heeft geëxciteerd, en dat de enveloppe (de modulator) van deze stimulus tot in de hersenstam is geraakt, waar ze de overeenkomstige generatoren heeft geactiveerd. Zodra deze respons wordt gedetecteerd in het EEG, kan de stimulusintensiteit worden verlaagd. De intensiteit waarop er nog net een respons aanwezig is, wordt gedefinieerd als de gehoordrempel. Het mooie aan deze techniek is dat er meerdere audiometrische frequenties tegelijkertijd kunnen getest worden, bijvoorbeeld vier aan het linkeroor en vier aan het rechteroor. Dit versnelt een frequentiespecifieke gehoordrempelbepaling met een factor twee tot drie vergeleken met het gebruik van een enkele stimulus (John et al., 2002a).

Alhoewel de huidige ASSR–techniek reeds een verbetering betekent ten opzichte van andere objectieve frequentiespecifieke gehoordrempelbepalingstechnieken, is de meettijd van de methode nog altijd behoorlijk lang. Meettijden kunnen in het optimale geval nog steeds oplopen tot één uur (Luts et al., 2006; Luts and Wouters, 2004). Wanneer de meetcondities niet optimaal zijn, zoals bijvoorbeeld bij een (jong) subject dat onrustig is, kan de duur van de meting sterk verlengen door de talrijk aanwezige storingen of artefacten in het EEG–signaal. **Sectie 2.2** geeft een overzicht van de reeds beschikbare methodes om de nodige meettijd te verkorten, wat belangrijk is om de ASSR–techniek klinisch toepasbaar te maken. De voorgestelde methodes zijn echter niet ontwikkeld met meerkanaals EEG in gedachten. Daarom focust dit werk zich op het ontwerp van signaalverwerkingstechnieken die zich richten op meerkanaals EEG–signalen. Op deze manier kan zowel de meettijd verkort worden, en de robuustheid tegen ongewenste storingen vergroot.

### Hoofdstuk 3: Een flexibel onderzoeksplatform voor meerkanaals ASSR–metingen

Alvorens kan overgegaan worden tot meerkanaals EEG–signaalverwerking, moet er natuurlijk eerst meerkanaals EEG beschikbaar zijn om te verwerken. Dit kan verzameld worden door onder andere zelf metingen te doen met behulp van commerciële apparatuur. Spijtig genoeg zijn er geen commerciële opstellingen beschikbaar die meer dan twee kanalen simultaan opmeten én tegelijkertijd geschikt zijn voor het registreren van ASSR, mede door de zware eisen die gesteld worden aan de kwaliteit van de EEG–versterker. Daarenboven laten dergelijke opstellingen meestal niet toe om zelfgemaakte stimuli aan te bieden of om een bepaalde stimulusintensiteit te overschrijden. Dit beschermt de patiënt namelijk tegen ongeoorloofd gebruik, maar beide opties zijn echter wel noodzakelijk voor klinisch onderzoek. Om deze redenen werd beslist om een eigen meerkanaals opstelling te bouwen die deze beperkingen niet heeft. Deze opstelling werd SOMA gedoopt, *Setup ORL voor Meerkanaals ASSR*. Deze in dit werk voorgestelde setup kan metingen uitvoeren tot acht EEG–kanalen tegelijk. Zelfgemaakte stimuli kunnen aangemaakt en binauraal aangeboden worden via ingelezen bestanden. De opstelling is mobiel, wat testen mogelijk maakt op andere locaties.

**Sectie 3.2** beschrijft de hardware en de software in meer detail. De hardware bestaat uit de relatief goedkope, maar kwalitatief uitstekende, meerkanaals *RME Hammerfall DSP Multiface II* geluidskaart. EEG–signalen worden versterkt met een achtkanaals lage–ruis versterker van *Jaeger–Toennies*. De software is geschreven in C++ en modulair uitbreidbaar. Naast de mogelijkheid om zelfgeconstrueerde stimuli aan te bieden, is de optie beschikbaar om tijdens een meting de intensiteiten van elke gemoduleerde draaggolf afzonderlijk aan

te passen. De opstelling SOMA wordt geëvalueerd in **Sectie 3.3**. De gehoordrempels van negen normaalhorenden worden geschat door middel van SOMA en een éénkanaals referentieopstelling van John and Picton (2000a). De resultaten geven aan dat beide systemen gelijkwaardig zijn in het schatten van gehoordrempels van normaalhorenden.

Momenteel wordt SOMA gebruikt in vier projecten binnen ExpORL, waarvan de twee belangrijkste even aangehaald worden. Het eerste project wordt uitgevoerd in een klinische omgeving, meer specifiek de operatiezaal van het UZ Leuven, waar buisjes in de oren worden geplaatst bij kinderen onder zes jaar. Zowel ASSR- als ABR-gehoordrempelschattingen worden hier bepaald en met elkaar vergeleken. Uiteindelijk is het de bedoeling te evalueren of de ASSR-techniek in deze omgeving iets kan bijdragen. Het tweede onderzoek focust op ASSR van gemoduleerde spraakachtige stimuli en de eventuele mogelijkheid tot vroegtijdige detectie van sommige vormen van dyslexie (Alaerts et al., 2007a).

## Hoofdstuk 4: Het verbeteren van ASSR-detectie met onafhankelijke component analyse

In het verdere verloop van dit werk wordt ingegaan op de toepassing van meerkanaals EEG-sigitaalverwerking op ASSR-data. Een eerste techniek die wordt besproken, is gebaseerd op *onafhankelijke component analyse* (ICA – ‘independent component analysis’). ICA maakt het mogelijk om de onbekende factoren van multivariate statistische data te bepalen. Hierbij wordt naar componenten gezocht die zowel statistisch onafhankelijk zijn als niet-Gaussisch (Comon, 1994; Hyvärinen et al., 2001). Indien volledige statistische onafhankelijkheid tussen de gevonden bronnen niet kan gegarandeerd worden, zoekt het ICA-algoritme naar bronnen die statistisch gezien zo verschillend mogelijk zijn. De rationale achter de toepassing van ICA op ASSR-data is gebaseerd op de veronderstelling dat sinusoidale (platykurtische) ASSR statistisch immens verschillen van het bijna-Gaussische EEG. Mochten er ASSR aanwezig zijn in het EEG, dan moet ICA in staat zijn deze te scheiden van de andere (ongewenste) hersengolven in het EEG. **Sectie 4.2** beschrijft dit model en deze assumpties. Verder wordt een experimentele studie beschreven met acht normaalhorenden om de tijdswinst te bepalen door ICA toe te passen op zowel éénkanaals als meerkanaals ASSR-data. In dit hoofdstuk wordt voorlopig enkel gefocust op ASSR-data die bekomen werd op intensiteiten aanzienlijk boven de gehoordrempel. Voor het effect van meerkanaals technieken op gehoordrempelbepaling wordt verwezen naar Hoofdstuk 7.

**Sectie 4.3** toont dat tussen -1 en 23 % snelheidswinst kan geboekt worden per subject voor intensiteiten boven de gehoordrempel wanneer een éénkanaals procedure wordt toegepast op een standaard normaalgewogen referentiekanaal. In-



dien de beschikbare data wordt uitgebreid naar meerdere kanalen, dan constateert men het volgende. Er is een limiet op het maximaal aantal kanalen dat kan toegevoegd worden tot de performantie niet meer stijgt per extra toegevoegd kanaal. De limiet voor deze dataset komt overeen met vijf kanalen en de actieve elektroden worden best geplaatst op de achterkant van het hoofd. Deze resultaten worden ook bevestigd in Hoofdstukken 5 en 7. Wanneer ICA wordt toegepast in deze laatste elektrodenconfiguratie, bedraagt de tijdswinst per subject tussen -2 en 63 % vergeleken met het standaard normaalgewogen referentiekanaal bij intensiteiten boven de gehoordrempel. Deze resultaten geven aan dat de mogelijke tijdswinst sterk kan variëren tussen subjecten onderling, iets wat typerend is voor ASSR-metingen in het algemeen. Dit betekent dat EEG-kanalen die optimaal zijn voor het ene subject niet automatisch goede ASSR-metingen opleveren voor een ander subject (Hoofdstuk 7). Door de resultaten hierboven te combineren, kan nog een extra performantiestijging bekomen worden. Wanneer de tijdswinst vergeleken wordt tussen metingen aan hoge intensiteiten (met dus grote ASSR in het EEG) en metingen aan lagere intensiteiten (met kleine ASSR), dan kan geconcludeerd worden dat de tijdswinst in het laatste geval kleiner is. Dit wordt veroorzaakt door het slecht presteren van de ICA-techniek bij lage signaal-ruis-verhoudingen (Hyvärinen et al., 2001).

## Hoofdstuk 5: Het verbeteren van ASSR-detectie met meerkanaals Wiener filteren

In Hoofdstuk 4 werd meerkanaals EEG verwerkt door middel van ICA. Een nadeel van ICA is dat deze techniek geen gebruik maakt van op voorhand gekende informatie, behalve de *veronderstelling* dat de onbekende bronnen onafhankelijk zouden zijn. Een meerkanaals techniek die wel gebruik maakt van extra informatie is *meerkanaals Wiener filteren* (MWF), met voorafgaande QR-factorisatie. Deze techniek neemt de gekende modulatiefrequentie mee in rekening waarmee de ASSR-stimulus is gemoduleerd (**Sectie 5.2**). Hiervoor wordt een fysisch ASSR-model beschreven dat het geregistreerde signaal aan de scalpelektroden modelleert als een combinatie van een instantane en verzwakte versie van een (vereenvoudigde) ASSR-bron in de hersenen, en ongewenste ruis afkomstig van het EEG en andere processen. Er wordt aangetoond dat voor de ASSR-toepassing beschreven in dit werk het meerkanaals Wiener filter identiek is aan de oplossing van een uitgangs-SNR maximalisatie.

In **Secties 5.3** en **5.4** wordt het experiment van Hoofdstuk 4 herhaald met de MWF-gebaseerde techniek in plaats van de ICA-gebaseerde methode. De resultaten zijn gelijkaardig. Deze overeenkomst wordt bevestigd in Parra and Sajda (2003) en Hoofdstuk 6. De grootste verschillen tussen de twee technieken is dat in het geval van ICA de data voorverwerkt moet worden door uitmid-

deling om de initiële signaal–ruis–verhouding hoog genoeg te krijgen. Dit is niet nodig bij de MWF–gebaseerde methode. Daarnaast vereist de MWF–gebaseerde techniek geen kunst– en vliegwerk in de vorm van extra artificiële ‘kanalen’ om de variabiliteit van de resultaten onder controle te houden.

## Hoofdstuk 6: Een proceduraal raamwerk voor ASSR–detectie

**Sectie 6.2** unificeert de technieken van Hoofdstuk 4 en 5 in een proceduraal raamwerk voor ASSR–detectie dat zijn inspiratie vindt in elementen van de detectietheorie. Aan de hand van dit raamwerk kan een meerkanaals signaalverwerkingsstrategie voor EEG worden ontwikkeld die geëvalueerd wordt in Secties 6.3 en 6.4.

**Sectie 6.2.1** introduceert het concept van de *voldoende statistiek* vanuit de detectietheorie. Door dit domein te combineren met de wereld van de ASSR, leidt **Sectie 6.2.2** een *voldoende statistiek* af die geschikt is voor ASSR–detectie. Deze statistiek maakt het mogelijk om voordeel te halen uit de spatiotemporele karakteristieken van de EEG–ruis (**Sectie 6.2.3**), waarvan eventueel een betere schatting kan bekomen worden met een aangepaste stimulus (**Sectie 6.2.4**). De hier voorgestelde *voldoende statistiek* gebaseerde techniek kan gelinkt worden aan zowel de ICA–gebaseerde afleiding uit Hoofdstuk 4 als de MWF–gebaseerde techniek van Hoofdstuk 5 (**Sectie 6.2.5**).

**Sectie 6.3** gebruikt simulaties met korte stukken EEG en artificiële ASSR om de meerkanaals technieken van Hoofdstukken 4 en 5 te vergelijken met de nieuwe, op het raamwerk gebaseerde, methode. De vergelijking met éénkanaals methodes wordt beschreven in Sectie 6.4. Voor de artificiële simulaties wordt zevenkanaals EEG zonder stimuli gebruikt van tien verschillende subjecten voor een totale duur van ongeveer twee uur. Een ASSR wordt toegevoegd met variabele amplitude en fase. De in dit hoofdstuk voorgestelde methode is identiek aan de MWF–gebaseerde techniek van Hoofdstuk 5 en gelijkaardig aan de ICA–gebaseerde methode van Hoofdstuk 4. Spatiaal gecorreleerde en spatiaal ongecorreleerde reeksen van ruisstoten worden toegevoegd aan de artificiële ASSR+EEG combinatie om de robuustheid van de verscheidene methodes tegen ruis te beoordelen. Bij spatiaal gecorreleerde ruis presteren alle meerkanaals technieken gelijkaardig. Bij spatiaal ongecorreleerde ruis daarentegen presteert de voorgestelde methode in dit hoofdstuk beter.

De simulaties in Sectie 6.3 geven nog geen correct beeld van de prestaties ten opzichte van de bestaande éénkanaals technieken. **Sectie 6.4** vergelijkt daarom de meerkanaals technieken met een goed presterende éénkanaals techniek, de ruisgewogen uitmiddeling van een éénkanaals EEG–signaal. Hiervoor wordt

achtkanaals EEG gebruikt van tien normaalhorenden, opgenomen op twee verschillende manieren. Eerst werd EEG opgenomen waarbij gevraagd werd om zo rustig mogelijk te blijven liggen. Dit type EEG bevat weinig storingen (of artefacten). Nadien werd EEG opgenomen waarbij de proefpersoon een gecontroleerde reeks hoofdbewegingen moest uitvoeren. Dit geeft aanleiding tot EEG met een grote hoeveelheid artefacten. Beide types EEG worden gebruikt in de analyses. Meerkanaals en éénkanaals signaalverwerkingstechnieken zijn gelijkaardig in performantie wanneer toegepast op EEG met weinig artefacten. Echter, wanneer de in dit werk voorgestelde meerkanaalstechnieken worden toegepast op EEG met veel artefacten, stijgt het aantal responsdetecties significant vergeleken met de ruisgewogen éénkanaals referentietechniek van John et al. (2001a).

## Hoofdstuk 7: Optimale elektrodenkeuze voor meerkanaals EEG gebaseerde detectie van hersenstam en auditieve cortex ASSR

Na de validatie van het meerkanaals raamwerk in Hoofdstuk 6, wordt in dit hoofdstuk onderzocht waar de elektroden best geplaatst worden op de schedel. Deze vraagstelling wordt beantwoord voor ASSR gegenereerd in de hersenstam (met modulatiefrequenties tussen 80 en 110 Hz) en voor ASSR gegenereerd in de auditieve cortex (met een modulatiefrequentie van 10 Hz). Hierbij wordt de beste elektrodenpositie bepaald indien slechts één EEG-kanaal beschikbaar is, en de beste elektrodenkeuze gezocht wanneer meerdere kanalen worden gecombineerd door middel van meerkanaals signaalverwerking. Bij de uitwerking van deze vragen wordt rekening gehouden met het voornemen om zo weinig mogelijk elektroden te moeten gebruiken, en om de gehoordrempelschatting in het geval van de hersenstam-ASSR zo goed mogelijk te benaderen voor zoveel mogelijk subjecten. In het geval van auditieve cortex-ASSR wordt naar een zo groot mogelijke signaal-ruis-verhouding gestreefd voor zoveel mogelijk subjecten.

Voor hersenstam-ASSR wordt aangetoond dat in het geval van een éénkanaals systeem best geopteerd wordt voor de achterkant van het hoofd om elektroden te plaatsen, met voorkeur voor de vertex-occiput positie (**Sectie 7.3.1**). Dit geeft aanleiding tot de beste schatting van de gemiddelde gehoordrempels uitgemiddeld over alle proefpersonen. Wanneer meerdere elektroden kunnen geplaatst worden, in combinatie met meerkanaals signaalverwerking, worden deze best geplaatst op de Oz, P3 en de rechter mastoïd (samen met de vertex als referentie). Deze combinatie garandeert voor alle proefpersonen in deze studie een bijna-optimale schatting van de gehoordrempels. Deze combinatie is ook significant robuuster tegen artefacten wanneer vergeleken wordt met een ruisgewogen éénkanaals vertex-occiput referentie.

**Sectie 7.3.2** toont dat voor ASSR gegenereerd in de auditieve cortex de volgende combinatie een bijna-optimale signaal-ruis-verhouding garandeert voor 80 % van de proefpersonen: de contralaterale mastoid, F4, F3 en het voorhoofd (samen met de vertex als referentie). Dit zijn vooral elektroden aan de voorzijde van het hoofd. Wanneer slechts één kanaal beschikbaar is, wordt best geopteerd voor elektroden op de achterkant van het hoofd, met voorkeur voor de beide mastoïden. Dit lijkt een contradictie die wordt weerlegd in de volgende paragraaf.

**Sectie 7.4** haalt aan dat deze resultaten kunnen verklaard worden door de achterliggende werking van de gebruikte meerkanaals techniek. Deze techniek maximaliseert de signaal-ruis-verhouding van de ASSR waar naar gezocht wordt door EEG-kanalen te combineren en gemeenschappelijke ruisinformatie uit te buiten. Praktisch gezien wordt het EEG-kanaal gekozen met de grootste signaal-ruis-verhouding voor de gezochte ASSR. Door extra kanalen toe te voegen, wordt getracht de EEG-ruis te verlagen. In het geval van de auditieve cortex-ASSR is dit snel in te zien. Hier wordt het kanaal gekozen met de grootste signaal-ruis-verhouding (de contralaterale mastoïd). Extra kanalen worden toegevoegd met weinig respons, maar met een hoge ruiscorrelatie. In het geval van de metingen met hersenstam-ASSR wordt meestal de vertex-occiput als eerste kanaal genomen. Deze wordt echter niet gecombineerd met kanalen vooraan het hoofd, hoewel dit logischer zou geweest zijn. Vooraan het hoofd zijn de ASSR inderdaad ook klein, maar de ruiscorrelatie tussen elektroden vooraan en achteraan het hoofd is echter te laag om van nut te zijn voor de meerkanaals signaalverwerking.

## Hoofdstuk 8: Besluit en suggesties voor verder onderzoek

Dit werk onderneemt een zoektocht naar technieken die de meettijd van de ASSR-techniek verkorten en de robuustheid van deze techniek tegen onvermijdelijke artefacten vergroten. Dit onderzoeksproject kan gesplitst worden in drie onderdelen. Eerst wordt een meerkanaals meetplatform beschreven om de ASSR te registreren, samen met het EEG, voor klinische en onderzoeksdoeleinden. Aan de hand van dit meetplatform worden studies uitgevoerd die meerkanaals signaalverwerkingstechnieken evalueren die voorgesteld worden in het tweede deel van dit manuscript. Dit tweede deel focust vooral op meettijdreductie. De meerkanaals technieken voorgesteld in het tweede deel worden ondergebracht in een vereenvoudigd proceduraal raamwerk van het derde deel. Dit raamwerk is opgebouwd vertrekkend vanuit de detectietheorie en het biedt de mogelijkheid om een meerkanaals signaalverwerkingsstrategie te construeren. Dit laatste deel is meer georiënteerd op optimale elektrodenplaatsing, artefactrobuustheid en een verhoging van het aantal ASSR-detecties. Deze ver-

beteringen impliceren natuurlijk eveneens een meettijdreductie.

De globale winst die mogelijk is met de bevindingen van dit onderzoeksproject wordt in deze paragraaf samengevat. Deze winst kan volledig worden toegeschreven aan het gebruik van meerdere EEG-kanalen, in combinatie met geschikte meerkanaals signaalverwerkingstechnieken. Voor de duidelijkheid wordt het globale resultaat hier beschreven voor EEG met weinig artefacten (bij optimale meetcondities), en voor EEG met veel artefacten (te vermijden, maar meer realistische meetcondities). In de praktijk zal een EEG-meting meestal een combinatie zijn van deze twee EEG-vormen. Wanneer de meetcondities optimaal zijn, is er een significante meettijdreductie (tot 60 %) mogelijk bij intensiteiten boven de gehoordrempel. Wanneer de intensiteiten rond gehoordrempelniveau zweven, is er in principe geen gemiddelde meettijdvermindering meer mogelijk, getuige de niet-significante verschillen tussen de gehoordrempels bepaald met een éénkanaals EEG en met een meerkanaals EEG bij gelijke meetduur. Het gebruik van meerdere EEG-kanalen tegelijk geeft hier wel het voordeel dat voor elk individueel subject een bijna-optimale registratie kan gegarandeerd worden van de gezochte respons. Wanneer het EEG gecontamineerd is met artefacten, is de meettijdreductie bij intensiteiten boven de gehoordrempel waarschijnlijk gelijkaardig of mogelijk nog groter dan bij EEG met weinig artefacten, gezien de significante vermeerdering van het aantal responsdetecties door de meerkanaals techniek. Op gehoordrempelniveau zal deze reductie nog steeds aanwezig zijn, aangezien artefacten er in het éénkanaals geval voor zorgen dat EEG-data systematisch moet weggegooid worden, wat onvermijdelijk langere metingen oplevert. In het meerkanaals geval kan de EEG-data echter gewoon behouden worden.

**Sectie 8.2.1** beschrijft een aantal mogelijke verbeteringen en uitbreidingen van het meerkanaals meetplatform SOMA. De meerkanaals technieken beschreven in dit werk zouden uiteindelijk in reële tijd kunnen worden berekend tijdens een ASSR-meting. Op deze manier kan het werkelijke praktische voordeel van deze technieken worden bepaald.

Momenteel worden de impedanties van de elektroden enkel bij het begin en het einde van een testsessie gemeten. Deze zouden continu geobserveerd kunnen worden. Dit is reeds mogelijk in commerciële produkten, en zou de kwaliteit van de ASSR-metingen aardig kunnen verhogen. Hetzelfde argument kan aangehaald worden voor het continu waarnemen van de geluidsintensiteit aan het trommelvlies door middel van een microfoon. Aangezien deze intensiteit varieert door het volume van het oorkanaal en de positie van de luidspreker, is het voordelig om te weten wat nu de exacte intensiteit is waarmee de stimuli worden aangeboden.

Het meetplatform SOMA kan uitgebreid worden voor eABR- en eASSR-metingen. Op deze manier kunnen cochleaire implantaten (CI) eventueel objectief ingesteld worden. De combinatie van ASSR en CI is wel uitzonderlijk uitdagend

te noemen doordat de zwakke ASSR moeten gerecupereerd worden uit door het CI gegenereerde elektrische artefacten die verscheidene grootteordes groter zijn.

**Sectie 8.2.2** geeft aan dat het voorgestelde meerkanaals raamwerk significant beter presteert met EEG dat veel artefacten bevat, vergeleken met éénkanaals technieken. Toch zijn er een aantal aspecten voor verbetering vatbaar. Het model dat het raamwerk gebruikt is een vereenvoudigde versie van de werkelijkheid. Er wordt uitgegaan van slechts één ASSR-bron in de hersenen, terwijl er minstens twee aanwezig zijn bij normale volwassenen (Herdman et al., 2002). De amplitude van de ASSR varieert onder invloed van de attentie van het subject, terwijl het model in dit werk uitgaat van een constante ASSR-amplitude. Daarnaast kan de temporele ruiscovariantiematrix misschien nog preciezer worden geschat, terwijl deze nu enkel een gescaalde diagonaalmatrix is.

Dit werk bespreekt de optimale elektrodenkeuze voor ASSR met modulatiefrequenties tussen 80 en 110 Hz, en met een modulatiefrequentie van 10 Hz. Het zou interessant zijn om ook andere modulatiefrequenties te exploreren, gaande van een paar Hertz tot frequenties boven 110 Hz. De resultaten kunnen meer inzicht verschaffen in de positie en grootte van de ASSR-bronnen. Nieuwe signaalverwerkingstechnieken zouden kunnen ontwikkeld worden die niet het verschil tussen ASSR- en gedragsdrempels minimaliseren, maar de variantie op deze verschillen. Dit is gebaseerd op een andere interpretatie van het begrip ‘gehoordrempelschatting’, waarbij gekeken wordt naar de toegelaten afwijking.

Drie toepassingen in het bijzonder zouden aardig wat voordeel kunnen halen uit een doordachte elektrodenplaatsing in combinatie met meerkanaals signaalverwerking. In navolging van van der Reijden et al. (2005) zouden interessante meerkanaals studies kunnen uitgevoerd worden op jonge kinderen. Door elektroden intelligent te plaatsen zouden storingen die vaak optreden bij metingen bij jonge kinderen grotendeels verwijderd kunnen worden. Hetzelfde voordeel kan behaald worden bij metingen met beengeleiders. Deze apparaten zorgen voor veel stimulusartefacten. Optimale plaatsing van de elektroden zouden deze artefacten kunnen reduceren. Zoals reeds aangehaald is de combinatie van ASSR en cochleaire implantaten een bijzondere uitdaging door de sterke aanwezigheid van elektrische CI-artefacten. Een geslaagde toepassing van de meerkanaals technieken van dit werk zou een grote stap kunnen betekenen op de weg naar CI-implantatie op (zeer) jonge leeftijd. Momenteel kunnen CI nog bijna niet gefit worden op jonge leeftijd door de hoge moeilijkheidsgraad van dit proces. Om deze reden wordt nu nog geopteerd voor een hoorapparaat. Daarenboven is het uiterst belangrijk om een betrouwbare schatting te verkrijgen van de gehoordrempels van de jonge patiënt. Deze schatting stuurt de beslissing om tot cochleaire implantatie over te gaan. Eénmaal geïmplanteerd kan namelijk niet meer op deze beslissing worden teruggekomen.

# Chapter 1

## Introduction

In the first section of this introductory chapter a motivation is given for the techniques that will be developed in the forthcoming chapters of the thesis.

In Section 1.2 a short overview of the auditory system is presented. If this system does not function optimally, one refers to hearing impairment.

Hearing impairment needs to be detected as soon as possible. This is the case for adults but especially for infants. Section 1.3 describes the need for early detection and introduces some devices, hearing aids and cochlear implants, to alleviate the permanent hearing problem when other approaches fail.

These devices need to be set up ('fitted') in a way that they compensate optimally for the hearing problem of the hearing impaired. This can be accomplished based on a reliable estimation of the hearing thresholds. Section 1.4 describes several subjective and objective techniques to obtain such a hearing threshold estimation.

An important class of objective hearing threshold estimation techniques for the topic of this thesis are auditory evoked potentials. Section 1.5 presents a detailed overview of a response evoked in the auditory system by a click, a short-duration stimulus or a stimulus with a periodic nature.

Section 1.6 will cover the research objectives of this thesis.

An outline and an overview of the different chapters of the thesis will be presented in Section 1.7.

## 1.1 Motivation

The ability to hear and process sounds is crucial. For adults, the inevitable ongoing aging process reduces the quality of the speech and sounds one perceives. If this effect is allowed to evolve too far, social isolation may occur. For infants, a disability in processing sounds results in an inappropriate development of speech, language, and cognitive abilities. To reduce the handicap of hearing loss in children, it is important to detect the hearing loss early and to provide effective rehabilitation.

In the case of infants, the process of hearing problem detection ('screening') has taken systematic forms in several parts of the world. In Flanders (Belgium) for example, hearing of about 97 % of all newborns is screened by the Flemish public agency Kind & Gezin, which accounts for about 60,000 infants a year. If the outcome of the screening does not indicate normal hearing, the infant is referred for further diagnosis. The hearing status of the referred infant has to be determined more quantitatively, but standard behavioral techniques are not yet sufficiently applicable at young age. Hearing assessment has to rely on objective physiologic techniques that are not influenced by sleep or sedation.

In Belgium, the most commonly used technique for hearing threshold estimation is the click-evoked *auditory brainstem response* (ABR). This technique is limited however by the fact it only gives an indication of the degree of the hearing loss. If frequency specific information is required, e.g. for the fitting of an hearing aid or for the decision whether a cochlear implant is appropriate, the standard ABR technique does not suffice. There are some variants of the ABR that return more frequency specific information, but these are time consuming and technically demanding.

In response to the shortcomings of the ABR technique, the *auditory steady-state response* (ASSR) technique was developed. This technique uses continuous rather than transient stimuli and enables the recording of frequency dependent responses to several carrier frequencies simultaneously. As a result, it is possible to get a frequency specific hearing threshold estimation in less than an hour. This is much faster than what is obtainable with a frequency specific ABR. Moreover, an additional benefit is the objective detection that facilitates interpretation.

Unfortunately, the ASSR technique also has its shortcomings. The electrophysiological responses one likes to record are much smaller than those of the ABR technique. The technique is very susceptible to noise and artifacts that could disrupt the measurement. Therefore it still could take a very long time to record responses and to give a reliable estimation of the patient's hearing thresholds. Time that generally is not available in standard clinical settings.



This thesis will present several techniques that shorten the duration of auditory steady-state response measurements. It addresses the problem of noise and artifacts appearing in the measurements and shows how to discard these.

## 1.2 Hearing and hearing impairment

*Hearing* is one of the traditional five senses, and refers to the ability to detect sound. In humans and other vertebrates, hearing is performed primarily by the auditory system: sound is detected by the ear and transduced into nerve impulses that are perceived by the brain. This section describes the most important aspects of the auditory system and its possibility of malfunctioning, also known as hearing impairment.

### 1.2.1 The auditory system

The *auditory system* can be divided into four parts: the outer ear, the middle ear, the inner ear and the central auditory nervous system. The outer ear consists of the pinna and the ear canal. Sound waves arrive at the pinna and travel to the tympanic membrane through the ear canal. The middle ear functions as a converter of sound waves into mechanical vibrations, using the tympanic membrane, the malleus, the incus and the stapes. These mechanical vibrations are transmitted to the fluid in the inner ear, by ways of the oval window. The inner ear contains the vestibular system (the sensors for balance) and the cochlea (the sensors for hearing). The central auditory system includes all of the complex interconnections in the auditory system beyond the cochlea, including the auditory cortex (Yost, 2000).

Figure 1.1 shows the *cochlea*, which roughly can be considered as an auditory filter bank (Moore, 2003). The frequency specificity, or tonotopy, of this filter bank is a result of the design of the basilar membrane, which divides the cochlea along its length. The basilar membrane is narrower and stiffer near the base, which makes it more susceptible to high frequencies. Lower frequencies tend to excite the basilar membrane more at the apex side. Hair cells are aligned over this membrane and the outer hair cells are in contact with the tectorial membrane (Figure 1.2). When the basilar membrane moves up and down, the hair cells are bent, which leads to the generation of action potentials in the neurons of the auditory nerve.

The *auditory nerve fibers* form a highly organized system of connections which connects the haircells in the inner ear with the brainstem and the auditory cortex higher on. These fibers respond better to specific frequencies than to others, which induces an extra aspect of frequency specificity. The nerve fibers also show phase-locking (Rose et al., 1967). Neural firings tend to occur at a

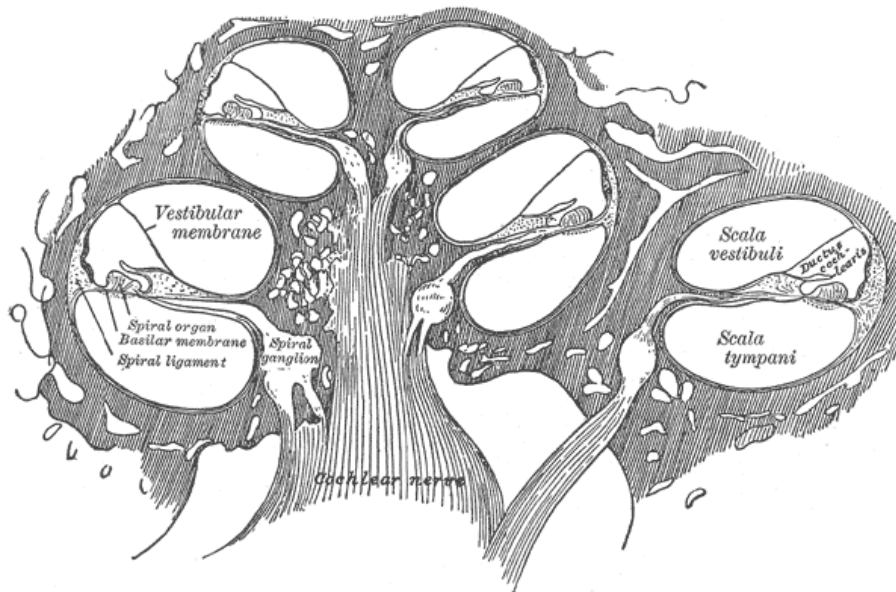


Figure 1.1: Diagrammatic longitudinal section of the cochlea. From Gray (1918).

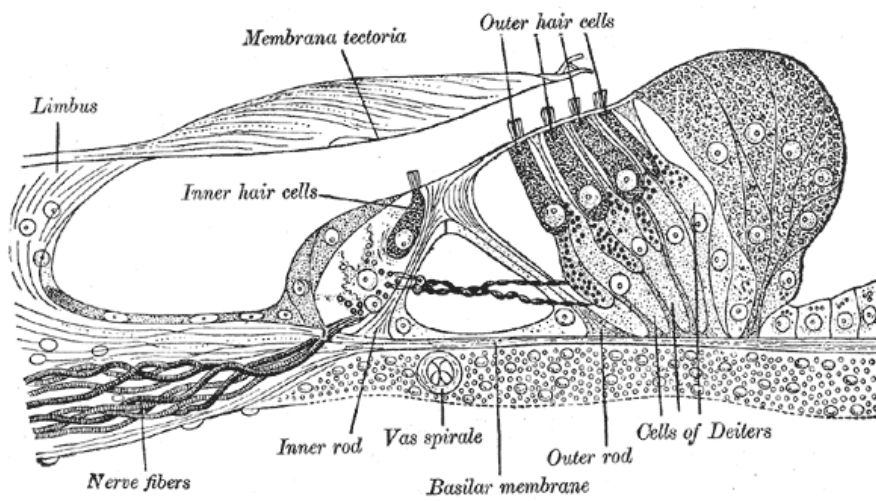


Figure 1.2: Haircells, the tectorial and the basilar membrane. From Gray (1918).

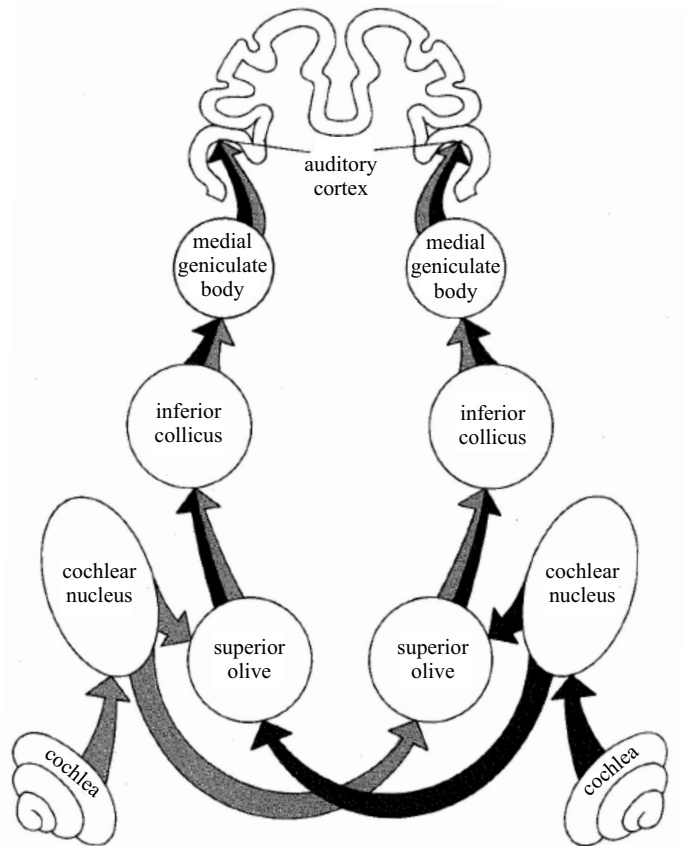


Figure 1.3: Highly schematic diagram of the ascending (afferent) pathways of the central auditory system from both cochleas to the auditory cortex. From Yost (2000).

particular phase of the stimulating waveform, so that there is temporal regularity in the firing pattern of a neuron in response to a periodic stimulus.

Figure 1.3 illustrates in schematic form the principal connections of the ascending or afferent auditory system. These are auditory paths going from the cochlea toward the cortex. The term ‘ascending’ implies that also descending pathways from the auditory cortex to the cochlea exist. These pathways are not discussed here, but descending fibers appear to have an inhibitory action on electrophysiological responses of the cochlea. After the neural pulses leave the cochlea, they travel to the *cochlear nucleus*. In the cochlear nucleus, the cochlear partition is completely tonotopically duplicated from the base to the apex. Its function is not clear yet, but it is assumed that the cochlear nucleus

refines the code for sound provided by the auditory periphery, like e.g. processing complex spectral information or echo suppression. From the cochlear nucleus, pathways lead to the *olivary complexes*, both contralaterally (on the opposite side of the brain compared to the side of stimulus application) as ipsilaterally (on the same side). This implies most bilateral representation (information from both ears) occurs at this point and above, a feature needed for e.g. localization of acoustic events in space. From the superior olive, neural impulses are transmitted to the *inferior colliculus*, the *medial geniculate body* and the *auditory cortex*. These regions are assumed to combine information from different processes that occur lower at the brainstem, allowing to arrange information for complex auditory pattern recognition (Yost, 2000).

### 1.2.2 Hearing impairment

*Hearing impairment* is generally described using three parameters, namely the type, the degree and the configuration of the hearing loss.

Three types of hearing loss can be identified: conductive, sensorineural and mixed. It refers to which part of the hearing system is damaged. With *conductive hearing loss* the outer or middle ear are obstructed. This causes a reduced sound level reaching the inner ear. One speaks of *sensorineural hearing loss* when the inner ear (i.e., cochlear hearing loss) or the auditory nerve pathway (i.e., retrocochlear hearing loss) is damaged. Sounds are not only attenuated, but also distorted. As broken hair cells or neural damage cannot be repaired, surgery or medication will not remedy this type of hearing impairment. The combination of conductive and sensorineural hearing loss is referred to as *mixed hearing loss*.

Depending on the amount of outer/middle ear obstruction and/or hair cell damage, the hearing loss varies from *mild* (26–45 dBHL), *moderate* (46–55 dBHL), *moderately severe* (56–70 dBHL), *severe* (71–90 dBHL), and *profound hearing loss* or deafness (more than 90 dBHL). The *hearing level* (HL) suffix is a relative scale with its zero defined by the standard audiograms of a group of normal-hearing young adults (ISO 389, 1998).

The configuration of the hearing loss defines the degree of hearing loss at each frequency. Possible configurations are *high-frequency*, *low-frequency*, *flat* or *cookie-bite hearing loss*. This last configuration represents a loss at middle frequencies. An extreme case is the *notch hearing loss*, caused by extreme stimuli at a small frequency band (e.g. industrial noise). This research project focuses on optimizing a method that determines the configuration of the hearing loss objectively.

## 1.3 Detection and intervention

In the unfortunate case of hearing impairment, the degree and the configuration of the hearing loss needs to be determined to allow intervention in an appropriate way. For children, an early detection is necessary. The longer a hearing problem is not remedied, the larger the delay will become in language development, together with other undesired developmental consequences. For adults, hearing loss can result in social isolation and should also be detected as soon as possible. This section treats the need for detection of hearing problems and two possible ways of compensation, i.e. hearing aids and cochlear implants.

### 1.3.1 The need for early detection

The impact of hearing impairment is immense, both for (aging) adults as for children (Morgan-Jones, 2001). For children however, hearing loss is a silent, hidden handicap. If undetected and untreated, it can lead to delayed speech and language development, learning problems, social and emotional problems (Northern and Downs, 2001). Yoshinaga-Itano et al. (1998) show that children whose hearing losses were identified by six months of age demonstrated significantly better receptive and expressive language skills than did children whose hearing losses were identified after the age of six months. As one in a thousand newborns are affected by permanent bilateral hearing loss (Mason and Herrmann, 1998; Mehl and Thomson, 2002), an adequate screening and reaction are definitely necessary to avoid problems in later life.

In a response to the demand for a global standard recommendation concerning early hearing problem detection with infants, the American Joint Committee on Infant Hearing (2000) states that all neonates should be screened with objective physiologic measures before the age of one month. Moreover, in case of failed hearing screening, an appropriate audiological and medical diagnosis should be made before the age of three months. All infants with confirmed permanent hearing loss should receive multidisciplinary intervention by the age of six months. Currently many regions in the world have implemented a highly covering screening program. Since 1998, the Dutch speaking part of Belgium (Flanders) is internationally leading in systematically screening, diagnosing and intervening hearing problems with neonates (Van Kerschaver and Stappaerts, 2004).

### 1.3.2 Hearing aids and cochlear implants

Permanent hearing impairment can be repaired through surgery or medication in some cases. If this approach fails, alternative solutions only can alleviate the hearing problem, not solve it. Currently these alternatives are the use of a hearing aid or a cochlear implant. A hearing aid is used in cases of mild to severe

hearing loss by compensation through sound amplification and compression. A cochlear implant actually bypasses the hair cells by stimulating the auditory nerve directly in case of profound hearing loss or deafness.

The *hearing aid* records the sound signals in the acoustic environment through one or more microphones (Dillon, 2001). These sound signals generally are a mixture of a speech source, generated by the speaker the hearing aid user is listening to, and one or more interfering sources (e.g., traffic noise, other speakers, ...). The recorded signals are (digitally) processed e.g. by amplifying the input signal's frequency bands at which the user has elevated hearing thresholds, or by noise reduction or compression. The resulting output signal is applied to the ear canal using a loudspeaker.

The *cochlear implant* may restore the perception of sound for patients who (almost) have no hearing left, in case the auditory nerve is still intact (Clark, 2003). The auditory nerve is stimulated directly by means of an intra-cochlear electrode array, which is implanted by a surgeon. The acoustic sound signals are recorded through one or more microphones, present in a behind-the-ear hearing aid. The sound signals are digitized, processed in a digital sound processor and decomposed in frequency bands. The filtered signal components are transformed into electric current pulses, that are transmitted wirelessly to the electrodes in the cochlea via a coil implanted under the mastoid of the patient. These electrodes stimulate the auditory nerve electrically at different places in the cochlea.

The decision whether a hearing aid or a cochlear implant should be applied can only be made based on an estimation of the degree of the hearing loss. A hearing aid can be fitted correctly only when sufficiently accurate information is available about both the hearing loss degree and configuration. The next section describes several ways to obtain a subjective or objective assessment of these parameters that can be graphically represented by an audiogram.

## 1.4 Hearing threshold estimation

An *audiogram* is a graphical representation of how well a certain person can perceive different sound frequencies. In an audiogram, a hearing threshold level is defined for each audiometric frequency. A *hearing threshold* is theoretically the level at which a (pure tone) stimulus is just sufficient to produce a sensation or an effect, in a noiseless environment. In practice however, determined hearing thresholds can differ greatly from the actual thresholds, due to age of the subject, the used method and the acoustic environment where the test is conducted. In this section different methods of hearing threshold estimation are described. These methods can be divided into two main classes: subjective and objective methods.

### 1.4.1 Subjective methods

Modern technology has greatly increased the number of options available to test the hearing, as shown later on in Section 1.4.2. However, regardless of how sophisticated testing techniques become, there will always be need for the behavioral hearing evaluation, since many of the newer procedures require expensive equipment or lengthy time commitment. For children two types of subjective methods can be defined: techniques without reinforcement (*behavioral observation audiometry* – BOA) and procedures based on reinforcement of the subject's responses (*visual reinforcement audiometry* – VRA) (Northern and Downs, 2001). The behavioral thresholds of older children and adults are determined using the modified Hughson–Westlake pure tone audiometry.

#### Behavioral observation audiometry

*Behavioral observation audiometry* (BOA) is typically limited to infants between six and twelve months of age. Noisemakers and sound field signals as acoustic stimuli are used to evoke an active response from an infant passively involved in the task at hand. Benefits of this method exist in the efficiency in time required and the lack of need for specialized equipment. Critics argue that the technique is useful for initial hearing screening, but not for the establishment of specific hearing threshold data, as interest in the stimuli fades away rapidly. A common type of BOA is the Ewing–test, generally administered to infants with the age of eight months.

#### Visual reinforcement audiometry

To avoid the rapid loss of interest in the applied stimuli, *visual reinforcement audiometry* (VRA) introduces a form of reward by flashing lights immediately following the response of the child looking forward the light. As the test becomes more interesting, toddlers up to two years old can be tested this way. For older children (up to the age of four), the *instrumental conditioned reflex* (ICR) integrates audiometry with game situations.

#### Pure tone audiometry

When the observed subject is able to respond unambiguously to the question whether a certain stimulus has been heard, one can use the modified Hughson–Westlake pure tone technique as a means to determine the subject's audiogram. The method is applied for each separate audiometric frequency.

### 1.4.2 Objective methods

Subjective methods have one big drawback. When a subject is not able or does not want to cooperate actively, the possibility to get a reliable response subjectively is marginal. In the former case a subject is not physically or mentally capable to respond, e.g. children younger than six months or the mentally disabled. In the latter case, subjects who simulate or aggravate hearing loss can modify their responses. To counteract this drawback, objective methods can be used. These methods do not expect active cooperation of the tested subject. In this section three objective ways to estimate hearing thresholds are described. *Tympanometry* and *oto-acoustic emissions* (OAEs) measure objective parameters of the middle and inner ear. *Auditory evoked potentials* (AEPs) are responses on the auditory path evoked by an acoustic stimulus. AEPs are discussed in detail in Section 1.5.

#### Tympanometry

*Tympanometry* is an objective technique for measuring the compliance or mobility of the tympanic membrane and middle ear as a function of mechanically varied air pressures in the external auditory canal. The compliance of the tympanic membrane at specific air pressures is plotted on a graph known as a tympanogram. Tympanic membrane mobility is of particular interest, since almost any pathology located on or medial to the eardrum will influence its movement. The compliance of the tympanic membrane is at its maximum when air pressures on both sides of the eardrum are equal. If these air pressures are unequal, the eardrum's mobility is lower and causes heightened hearing thresholds. Unequal pressures are explained by fluid in the middle ear or problems with the eustachian tube (Northern and Downs, 2001). The 1000 Hz tympanogram is especially relevant for the neonatal population (Alaerts et al., 2007b).

#### Otoacoustic emissions

*Otoacoustic emissions* (OAEs) are low-level, mostly inaudible sounds produced by the inner ear and arise by a number of different cellular mechanisms. Kemp (1978) shows that OAEs could be detected in the human external ear canal following stimulation with clicks. OAEs provide evidence of a normal functioning cochlea and are reduced by or non-detectable with a hearing loss of more than 30 dBHL (Kemp et al., 1986). Nowadays, OAEs are widely used as a screening test, as an alternative for the auditory brainstem response (ABR), which is described in Section 1.5.2 (Berg et al., 2005; Capua et al., 2007; Norton et al., 2000).



### Auditory evoked potentials

An *auditory evoked potential* (AEP) is “activity (a response) within the auditory system (the ear, the auditory nerve, or auditory regions of the brain) that is produced or stimulated (evoked) by sounds (auditory or acoustic stimuli)” (Hall, 2007). Basically an auditory stimulus is applied at the ear and the (faint) response, together with other brain processes, is recorded through an *electroencephalogram* (EEG) using electrodes on the scalp. By repeating this procedure and using signal averaging, combined with the use of appropriate filtering, activity different from the response can be reduced and the response can be made visible.

A more extensive review of this type of objective audiometry is important for the general structure of this manuscript and the comprehension of its motivation. Therefore the following section will focus on auditory evoked potentials in detail.

## 1.5 Auditory evoked potentials

Five types of AEPs can be distinguished based on their latency, the delay between stimulus onset and response. The whole path of the auditory nerve can be covered this way, ranging from responses in the cochlea itself to loci in the auditory cortex far beyond the brainstem. An overview is given in Figure 1.4. The *electrocochleogram* (eCochG) is the first AEP ever recorded (1930). The response arises from the cochlea and the eighth (auditory) cranial nerve, close to the cochlea. This results in a small latency, approximately two to three milliseconds after stimulation. The *auditory brainstem response* (ABR), currently mainly used for hearing screening and hearing threshold estimation, originates from the eighth nerve and the brainstem and has a latency between 1.5 and 5.5 milliseconds. Age, gender and stimulus intensity affect these figures considerably. The *auditory middle-latency response* (AMLR), 12 to about 60 milliseconds, the *auditory late response* (ALR), about 60 to 250 milliseconds and the *auditory P300 response* (P300), around 300 milliseconds, are all electrophysiological processes that give insight into higher-level auditory functioning. These responses originate from the thalamus (early AMLR), auditory cortex (AMLR and ALR) and beyond (P300), according to Hall (2007).

A special class of auditory evoked responses are *auditory steady-state responses* (ASSRs). An auditory steady-state response is an *envelope following response* (EFR), a physical response that falls into step with the periodic envelope of a stimulus. The AEPs described in the previous paragraph are generally evoked by a stimulus click or a short duration stimulus. By applying this stimulus at a repetitive rate, responses overlap and a periodic response to the repetition frequency is established. To make a distinction between responses evoked directly

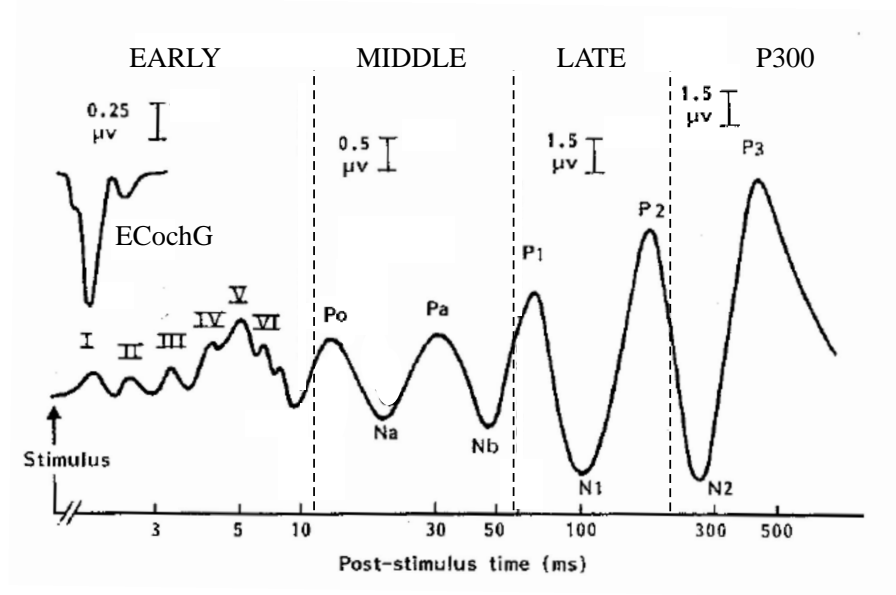


Figure 1.4: Early, middle and late latency components of the auditory evoked potential, together with the electrocochleogram and P300 component. Adapted from McCormick (1993).

by a stimulus with a specific carrier frequency, or by the envelope (modulator) of a certain carrier, we use the following nomenclature in this manuscript. Responses evoked directly by a stimulus with a specific carrier frequency are called *frequency following responses* (FFRs). If the response is evoked by the modulator of a certain stimulus, we refer to envelope following responses, and more specifically to auditory steady-state responses in the case the modulator is constant in frequency.

Auditory steady-state responses have been discovered in the early eighties of the previous century (Galambos et al., 1981). Since then, extensive research has been conducted which resulted in a clinical applicable technique that can test more than one frequency at both ears simultaneously (Lins and Picton, 1995). Nowadays, the ASSR technique is used as a diagnostic technique in highly specialized clinical environments. The technique is not replacing, but contributing to the standard audiological test bench including tympanometry, OAE and ABR. A detailed overview of the ASSR is given in Chapter 2. The following sections describe the five different AEP types. Where applicable, the auditory steady-state response evoked by a repetitive application of the stimulus evoking these AEPs is also presented.

### 1.5.1 Electrocochleogram responses

The *electrocochleogram* (ECochG) response occurs within two to three milliseconds after an abrupt stimulus. It arises from the cochlea and the eighth (auditory) cranial nerve. The ECochG response consists of three components, the *cochlear microphonic* (CM), the *summating potential* (SP), and the most important component, the *action potential* (AP, the negative peak in Figure 1.4). This potential is a combined action potential of many nerve fibers and equivalent with wave I of the auditory brainstem response. A major difference between the ECochG response and all subsequent responses is that the site of the recording electrode(s) is not on the scalp, but as close to the cochlea as possible. Reason for this is that the ECochG is a near-field AEP. Medical applications of the ECochG are the contribution to the diagnosis of Meniere's disease and intraoperative monitoring of cochlear and eighth nerve activity during surgery that puts the auditory system at risk. It is not a test of hearing (Hall, 2007).

### 1.5.2 Auditory brainstem responses

Since the classic paper of Jewett and Williston (1971) published in *Brain*, research focusing on the *auditory brainstem response* (ABR) has boomed. Nowadays the main application of the ABR are hearing threshold estimation, newborn hearing screening and neurodiagnosis for problems with the auditory nerve up to the brainstem.

When a sufficient number (1000 or more) of acoustic click stimuli is applied to the ear, an ABR trace can be recorded with electrodes at the scalp by averaging the responses to these clicks. Figure 1.5 shows several of these ABR traces after stimulation by a number of clicks at different intensities. As soon a click is presented to the ear, responses to this click are evoked in several nuclei along the auditory nerve. Based on their latencies up to ten milliseconds, the positive peaks have been chosen to be denoted with Roman numbers (I, II, III, IV and V). Each peak is generated in a different part of the auditory nerve or brainstem. Their latencies increase with decreasing intensities. For audiological purposes, peak/wave V is the most prominent. It is the biggest response present and its latency gives one information about the auditory nerve up to the inferior colliculus (Moller, 1994).

The stimulus is applied to the ear using headphones, insert phones or bone vibrators. The click stimulus, generally in the order of 100 microseconds long, is applied several tens of times a second. As the physical response location is dependent on the stimulated ear, the electrode placement varies accordingly. The active electrode is placed on the forehead, the reference electrode on the ipsilateral mastoid and the ground electrode on the contralateral mastoid. If one tests the other ear, only the electrode leads of the reference and ground electrodes need to be changed. This way, no extra electrodes have to be positioned

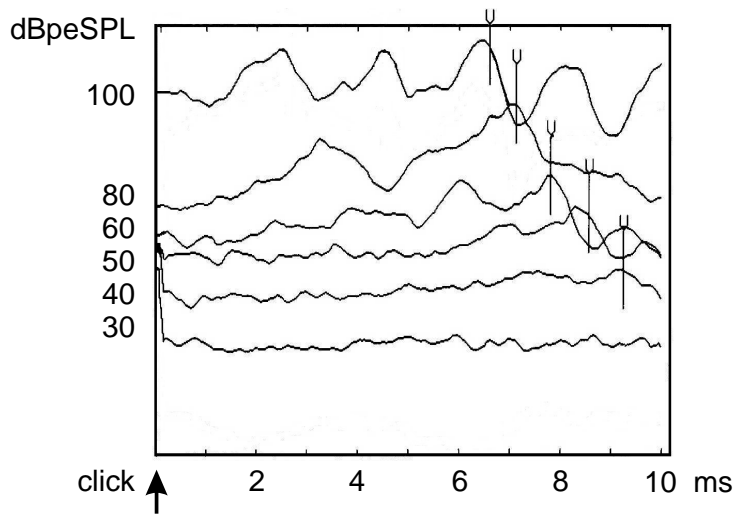


Figure 1.5: Auditory brainstem response (ABR) traces as a reaction to a click stimulus, here depicted at different intensities. The click stimulus intensities are measured in decibel peak-equivalent Sound Pressure Level (dBpeSPL), which lies roughly 35–40 dB above the decibel normal-Hearing Level (dBnHL) unit. dBnHL is a biological calibration, taking the thresholds of normal-hearing young adults obtained with the ABR-stimuli as a zero-reference (Stapells et al., 1982).

on the skull. The recorded EEG is amplified by a factor of 100,000, bandpass filtered between 30 and 3000 Hz and the first 15 milliseconds of the response are withheld. By repeating the stimulus up to 4000 times, a noise-reduced average of the response can be acquired (Hall, 2007).

By modifying the intensity of the click stimulus, hearing thresholds can be determined (Figure 1.5). As long as the starting intensity is high enough, wave V is clearly visible. By lowering the stimulus intensity until wave V is not distinguishable anymore, the hearing threshold is determined by the last intensity a wave V was still present. This method offers a relatively fast way to determine hearing thresholds. Both ears can be reliably tested in 45 minutes or less (Gorga et al., 2006; Picton et al., 1994; Stapells and Oates, 1997). On the other hand, the ABR technique suffers from three serious drawbacks. A first drawback is the lack of frequency specificity of the technique, which is necessary for an accurate fitting of hearing aids and for general diagnostic purposes. The click stimulus covers the whole frequency spectrum with the frequencies containing the highest energy lying around 2000 to 4000 Hz (Gorga et al., 2006). This fact avoids an accurate guess of the threshold at a specific frequency and, worse, can underestimate hearing thresholds at lower frequencies around 500

Hz (Gorga et al., 2006; Picton et al., 1994; Stapells and Oates, 1997). Secondly, the maximum presentation level of a click stimulus is limited, which makes it difficult to differentiate between severe and profound hearing losses. Thirdly, the response detection is based on visual inspection of the waveforms. For adequate response detection, decision-making of an experienced clinician is required.

Over the years different modifications of the standard ABR method have evolved to cope with the lack of frequency specificity. The use of *tone bursts* instead of clicks concentrates the energy more in the frequency region of interest (Beattie and Torre, 1997; Gorga et al., 2006; Johnson and Brown, 2005; Rance et al., 2006). These methods give a better estimation of the hearing thresholds at specific frequencies. However, measurement time increases as a hearing threshold estimation procedure has to be conducted for each separate frequency. Moreover, the spread of energy around each of these transient, sudden onset signals may become problematic. The spread of excitation to lower- and higher-frequency regions will be greater when stimulus intensity is increased. Secondly, the use of a *chirp* that evokes higher responses, is described in Dau et al. (2000); Elberling et al. (2007); Fobel and Dau (2004); Junius and Dau (2005); Wegner and Dau (2002). This modified stimulus compensates for the travelling time of lower frequencies exciting the apex of the cochlea. All points on the basilar membrane reach maximum excitation simultaneously for this stimulus. A disadvantage of chirps is the fact that the stimulus is designed based on theoretical models and therefore not necessarily optimal for any given individual subject. The method is promising but needs further research, especially with extra studies on children.

A third alternative could be the use of *noise masking*. The noise, which is presented simultaneously with the click or tone burst, restricts the regions of the basilar membrane that are capable of contributing to the ABR, by selectively masking certain regions that are outside the region to be stimulated (Beattie and Kennedy, 1992; Oates and Purdy, 2001). This procedure is subject to a lot of discussion however. Three variants exist. First, clicks can be masked using highpass noise. This way, all auditory fibers with a characteristic frequency above the cutoff frequency do not contribute to the response. It works well for a hearing sensitivity assessment around 500 Hz (Purdy et al., 1989), but at higher frequencies estimates are not reliable as responses at lower frequencies also contribute. This drawback can be avoided by the use of notched-noise masking (Beattie and Kennedy, 1992; Stapells and Oates, 1997). The stimuli are presented simultaneously with notch-filtered broadband noise. As a result, the basilar membrane is only stimulated at the frequencies without noise. Unfortunately, by limiting the amount of stimulus effectively exciting the basilar membrane, responses are small and a long measurement time is required. A final technique makes use of derived responses (Oates and Stapells, 1997). The response in highpass noise at one cutoff frequency is subtracted from the re-

sponse obtained in the presence of highpass noise with a higher cutoff frequency. The result is a derived response, representing the narrowband contributions to the response from portions of the basilar membrane located between the two cutoff frequencies. In general, noise masking results in several unknown technical and physiologic variables (Hall, 2007). Masking parts of the cochlea using noise does not guarantee the stimulus not evoking a response at these masked regions. The effect of masking noise in the region of stimulation is unknown, together with the interactions between masking noise and stimulus. As a result, large scale application is prohibited by the long measurement times, the procedure complexity and a required interpretation by trained clinicians.

By repeatedly applying the stimulus above, or by modulating noise or a pure frequency tone, an auditory steady-state response pulsating at the modulation frequency is generated. Depending on the modulation frequency, this response is generated in the auditory cortex and/or the brainstem. For modulation frequencies above about 75 Hz, most of the ASSR is generated in the brainstem. Above 95 Hz, the source lies entirely in the brainstem. This is depicted in Figure 1.6 of Purcell et al. (2004).

It is thus assumed that ASSRs generated at 75 Hz or above are mainly originating from the brainstem. As the brainstem generator's true delay between stimulus and response is estimated to be 7.3 ms (Figure 1.6), an ASSR at 75 Hz or above may mainly be considered as a (continuous) form of the ABR. The physically measured delay of the response between 75 and 90 Hz lies between 11 and 19 milliseconds however (Picton et al., 2001; Purcell et al., 2004). This can be explained by the fact that the true delay of part of the generators lying in the auditory cortex is 29 milliseconds (Figure 1.6). As a result, the combined ASSR (with frequencies of 75 Hz and higher) originating from generators in the auditory cortex and/or brainstem has a delay somewhere in between 7.3 and 29 ms. The higher the modulation frequency, the smaller the physically measured delay. These values are confirmed by Alaerts et al. (2008) with mean latencies in adults between 24.3 ms (82 to 86 Hz) and 18.0 ms (106 to 110 Hz). They also indicate that latencies in infants are on average 9.5 ms longer than in adults, and that they decrease with age (on average 2.0 ms for infants between 3–8 weeks of age compared with infants  $\leq 0$  weeks of age).

### 1.5.3 Auditory middle-latency responses

The *auditory middle-latency response* (AMLR) is a response observed between about 12 and 60 milliseconds, following the auditory brainstem response. The major sequential peaks in Figure 1.4 are labelled N (negative voltage waves) and P (positive voltage waves), and alphabetically in lowercase (Na, Pa and Nb). P0 is not a true component of the AMLR. Generators are suggested to be in the auditory thalamus and primary auditory cortex. Clinical applications are based on detection possibilities of central nervous system dysfunctions above

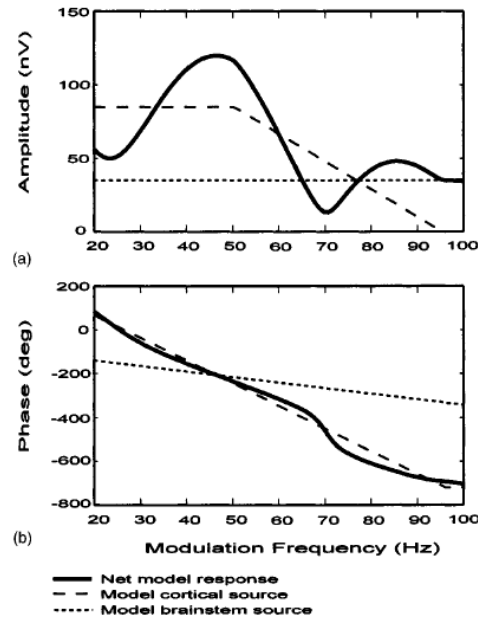


Figure 1.6: Model of ASSR amplitude and phase w.r.t. modulation frequency resulting from the sum of two sources. Hypothetical cortical and brainstem sources are shown with long dashed and short dashed lines, respectively. Amplitude is shown in panel A, and phase in panel B. The cortical source has a true delay of 29 ms, and constant amplitude from 20 to 50 Hz. The amplitude decreases linearly to zero from 50 to 95 Hz. The brainstem source has a true delay of 7.3 ms and a constant amplitude. The continuous thick line indicates the sum of the two sources as might be recorded at scalp electrodes. From Purcell et al. (2004).

level of the brainstem (Hall, 2007).

Auditory steady-state responses generated around 40 Hz are mainly originating in the auditory cortex. Delays are similar to those of the AMLR as the true delay lies between 24.3 and 29.3 ms (Purcell et al., 2004). Findings of Bohórquez and Ozdamar (2008) indicate that the generation of the 40 Hz ASSRs can be explained successfully by the superposition of the ABR and MLR waves generated at that stimulation rate. Na – Pa and Nb – Pb components of the MLR contribute about equally (45% each), while the wave V of the ABR contributes a lesser amount (10%).

### 1.5.4 Auditory late latency responses

*Auditory late latency responses* (ALRs) are recorded in a time period from about 60 to 250 milliseconds after acoustic stimulation. In comparison to earlier responses, the amplitude of the ALR is large. The response is composed of waves P1, N1, P2 and N2. Exact generators are not known, but presumably arise from the auditory cortex. A maximum response is typically obtained for moderate versus high-intensity stimuli. It is susceptible to the state of arousal of the subject, which means that sleep affects the response amplitude significantly. ALRs are mainly used as an electrophysiological assessment of higher-level auditory central nervous system functioning (Hall, 2007).

Assuming the presence of a pattern that the true delay in milliseconds of an ASSR is the reciprocal of its frequency in Hertz, auditory steady-state responses evoked by a stimulus with a modulation of 10 Hz should have a true delay of around 100 milliseconds. This delay value places 10 Hz ASSRs in the region of the ALR. This value is confirmed by Stapells et al. (1987) with reported latencies between 75 and 153 ms. A more recent study of Herdman et al. (2002) shows that the 12 Hz response is a combination of brainstem responses (with a latency varying between 21 and 58 ms depending on the direction of projection) and the auditory cortex (delays ranging between 154 and 168 ms). Here the cortical components of the response should be related to the P1, N1 and P2 components of the ALR. Whatever being the real delay value, the main conclusion is that the 10 Hz response originates mainly somewhere from the auditory cortex, but sources in the brainstem may still be present. The 10 Hz response is a research topic in Chapter 7.

### 1.5.5 Auditory P300 responses

The *auditory P300 response* is the largest and latest AEP resulting from a click stimulus. Its major peak is labelled P3 (as the third positive wave) or P300 (as it appears about 300 milliseconds after stimulus). Its generator appears to be in the medial temporal lobe, beyond the auditory cortex. The P300 response is optimally evoked by unpredictable, infrequent acoustic stimuli presented randomly with a probability of 15 to 20 %, in an ‘oddball’ test paradigm. This type of test paradigm offers acoustic stimuli in a frequent condition, but occasionally a different, oddball, stimulus is presented instead. The amplitude of the P300 is subject to the oddball probability and the amount of attention the subject is directing to the task. This way, electrophysiological assessment of higher-level auditory processing can be achieved (Hall, 2007).



## 1.6 Research objectives

It is possible to determine a frequency specific hearing threshold estimation much faster using the auditory steady-state response technique than with the (frequency specific) ABR technique. The possibility of objective detection facilitates the interpretation by using a detection algorithm instead of an experienced clinician. Unfortunately, the technique is very susceptible to noise and artifacts that could disrupt the measurement. Therefore it still could take a very long time to record responses and to give a reliable estimation of the patient's hearing thresholds. For these reasons, our interest shifted to the research of auditory steady-state responses to exploit all the benefits the ASSR technique offers and to reduce the impact of its flaws.

The intensity of use of the auditory steady-state response technique has increased significantly over the last decade. More and more clinical environments add ASSR-based hearing threshold assessment to their objective test batteries, next to e.g. tympanometry, oto-acoustic emissions and the auditory brainstem response. The technique still has its skeptics however. This skepticism is not without any basis. Measuring auditory steady-state responses reliably requires specialized equipment, trained personnel and lots of patience. Especially for the target population of the ASSR technique, which are newborns and infants, measurement conditions need to be optimal. This optimality mainly encompasses a subject that is relaxed or sleeping. Awake subjects generate more muscle artifacts, eye blinks, electrode cable movement. Unfortunately, this condition is not fulfilled most of the time, and it gets more difficult with increasing age of the young subject.

To address these problems, the ASSR technique basically needs to improve on two points, while still guaranteeing precise and accurate frequency specific audiogram estimations. First, the time that is needed to record the necessary EEG data should be reduced as much as possible. The longer a measurement lasts, the more difficult a subject can be kept quiet. This total duration of a test is also very determining for widespread application in the clinical field. Second, the robustness of ASSR recordings against artifacts should be increased. If the majority of the data needs to be discarded because of too much noise in the recordings, measurements are extended unnecessarily. In this research project, an approach is presented that may provide an answer to both points.

More specifically, the following objectives were aimed for:

- To build and evaluate a robust experimental test platform for recording auditory steady-state responses for research and clinical applications. This platform should provide a means to record more information simultaneously compared to current single channel setups. This immediately implies multichannel recordings with more than three electrodes.

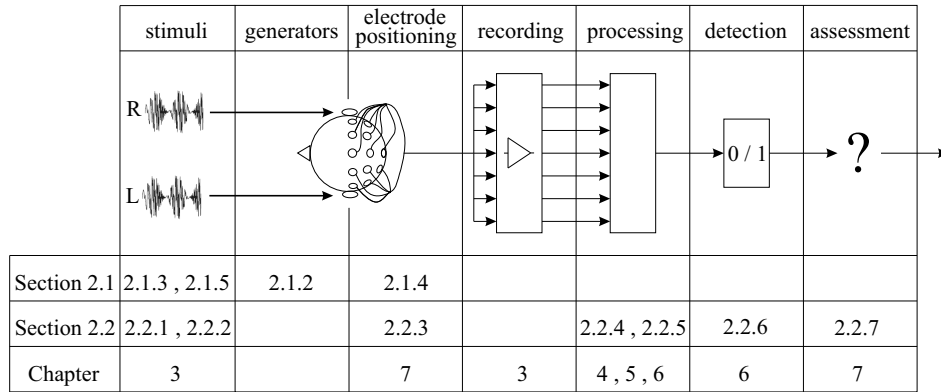


Figure 1.7: Overview of the thesis and the position of the chapters and sections in the general picture.

- To develop and evaluate a multichannel processing scheme that improves auditory steady-state response detection and thus reduces measurement time.
- To develop and evaluate a multichannel processing scheme that improves the robustness against artifacts of the ASSR technique and thus implicitly reduces measurement time accordingly.
- To propose optimal electrode positions on the scalp for best assessment of ASSR thresholds, subject to as few electrodes necessary as possible.

## 1.7 Thesis outline

Figure 1.7 shows a graphical overview of the structure of the thesis, positioning the different chapters in the general picture.

In **Chapter 2**, an overview is given of the auditory steady-state response. Its history, physiological model, stimulus and recording parameters are discussed. Different existing ASSR recording time reduction techniques available are presented.

In **Chapter 3**, a research platform is developed for multichannel ASSR recordings. The setup allows multichannel measurements and the use of own custom made stimuli which is a requirement for the studies described in the rest of the thesis. The mobile setup is based on an inexpensive multichannel RME soundcard and software is written in C++. Both hardware and software of the setup are described. An evaluation study with nine normal-hearing adults will

show that no significant performance differences exist between a reference and the proposed platform for standard recordings. The material presented in this chapter has been published in Van Dun et al. (2008c).

**Chapter 4** presents the first multichannel processing technique that is investigated in this thesis, based on independent component analysis (ICA). This multichannel technique basically allows one to find the underlying factors from multivariate statistical data by looking for components that are both statistically independent, and non-Gaussian. As ASSRs are statistically different from the surrounding EEG noise, it is expected that ASSR detection is improved after processing. ICA is applied to seven-channel data containing ASSRs from eight normal-hearing adults. First, ICA-processed multichannel data containing ASSRs is compared with the most common single channel ASSR technique. Second, the optimal number of input channels, the optimal electrode positions and the optimal number of independent components are reported. Third, the performance of an ICA-based procedure applied to single channel data is considered. Finally, a combination of previous techniques is presented, together with the quantitative recording time reduction that is possible. The material presented in this chapter has been published in Van Dun et al. (2006, 2007a).

**Chapter 5** describes the results of the application of multichannel Wiener filtering (MWF) – with QR factorization – to multichannel data containing ASSRs. This approach adds extra information to the multichannel signal processing equations. The sinusoidal nature of the ASSR makes it possible to search for a specific frequency, equal to the modulation frequency used in the stimulus. The MWF technique is applied to seven-channel data containing ASSRs of eight normal-hearing adults. Results are compared with the most common single channel ASSR technique and with multichannel data processed using ICA. The possible measurement time reduction using this technique is presented. The material presented in this chapter has been published in Van Dun et al. (2007b).

In **Chapter 6**, a procedural framework for ASSR detection is worked out that allows the development of a multichannel processing strategy, starting from a detection theory approach. It is shown that a sufficient statistic can be calculated that best captures the amount of useful signal in the recorded data. The framework can be linked with multichannel Wiener filtering and independent component analysis. First, simulations are conducted using seven-channel EEG and artificial ASSRs to evaluate the performance of the different multichannel techniques in this research project. Afterwards, eight-channel data containing ASSRs from ten normal-hearing adults are used to evaluate the framework practically and to compare its performance with existing single channel processing approaches. For EEG containing numerous artifacts, multichannel techniques can clearly offer significant improvement over single channel techniques. The material presented in this chapter has been published in Van Dun et al. (2008b).

**Chapter 7** probes further on the framework presented in Chapter 6. Based on eight-channel EEG measurements of ten normal-hearing adults, the EEG derivation with the best estimation of the behavioral hearing thresholds for ASSRs with modulation frequencies between 80 and 110 Hz (originating mainly from the brainstem) is determined. For ASSRs originating mainly from the auditory cortex (with a modulation frequency of 10 Hz), the EEG derivation with the highest SNR is determined similarly. The EEG channel combination for brainstem ASSRs with the best estimation of the behavioral hearing thresholds for as many subjects as possible is presented, taking into account that the number of required electrodes should be as low as possible. For auditory cortex ASSRs, the EEG channel combination with the highest SNR for a maximum number of subjects is given, with a similar restriction on the number of used electrodes. Finally, the robustness of the proposed multichannel processing technique against artifacts is evaluated when using this optimal channel combination for brainstem responses. The material presented in this chapter has been published in Van Dun et al. (2008a).

**Chapter 8** comprises a general discussion of the results of this research. We end with an overview of future research directions.

## Chapter 2

# Auditory Steady–State Responses

This chapter covers the basic theory of auditory steady–state responses and contains an overview of currently available recording techniques. Section 2.1 covers the theoretical aspects of the ASSR. It includes its history, assumed model and the path from stimulus to response, recording and processing. Section 2.2 focuses on seven different techniques currently available that were proposed to reduce recording time. None of these techniques are developed specifically for multichannel ASSR recordings however. Signal processing on multichannel ASSR recordings is the main topic of this thesis. Chapter conclusions are provided in Section 2.3.

### 2.1 Theoretical overview of the ASSR

This section covers the theoretical aspects of the auditory steady–state response. A much more detailed overview can be found in the review paper of Picton et al. (2003). The PhD text of Luts (2005) also provides an interesting overview, but focuses more on the clinical relevance of the ASSR technique for newborns.

Section 2.1.1 introduces the auditory steady–state response by describing its history. In Section 2.1.2, its physiological model and its sources in the brain are given. ASSRs are evoked by stimuli, generally modulated carriers. By changing the stimulus parameters, different responses can be elicited (Section 2.1.3) which are recorded using an electroencephalogram (Section 2.1.4). The ASSR technique that currently is used the most is the multiple stimulus ASSR tech-

nique from Section 2.1.5. Finally, Section 2.1.6 describes the advantages and the drawbacks of the ASSR technique compared with the auditory brainstem response technique.

### 2.1.1 How it started . . .

Responses to acoustic stimuli in general have been recorded for about 70 years (Davis, 1939). In these early days, rather straightforward averaging procedures were developed to lower the background EEG and to improve the ratio between the observed signal and noise (Geisler, 1960). The cited work describes a direct link between acoustic clicks and scalp recorded potentials, with the use of an ‘average–response computer’, as these responses are generally too small to be detected in the electroencephalogram (EEG) directly. When high–frequency sinusoids are modulated with a lower modulation frequency and applied to the ear, the auditory neuron fibers fire ‘phase locked’ (in synchrony) to the envelope (Moller, 1974). The first significant auditory potential study with modulated tones was conducted by Galambos et al. (1981) on modulation frequencies of 40 Hz. From that moment on numerous studies have been conducted. Definitions that are being used are the *amplitude modulation following response* (AMFR by Kuwada et al. (1986)), the *steady–state evoked potential* (SSEP by Cohen et al. (1991)) and the *envelope following response* (EFR by Dolphin and Mountain (1992)). The currently most popular term is the *auditory steady–state response* (ASSR) and will be used throughout this manuscript. A steady–state response is a response ‘whose constituent discrete frequency components remain constant in amplitude and phase over an infinitely long time period’ (Regan, 1989). Whereas transient responses (e.g. auditory brainstem responses) are evoked by stimuli that do not have a periodic nature, steady–state responses are evoked by repetitive stimuli that are presented continuously or at a stimulus rate that is fast enough to cause an overlap of the consecutive responses.

### 2.1.2 Physiological model

A physiological model for auditory steady–state responses is described in Lins and Picton (1995). The clear presentation of the model in this key paper provides a good way to understand the underlying properties of the ASSR and its generation (Figure 2.1).

When a modulated sinusoid with a certain carrier frequency arrives at its corresponding frequency region of the *basilar membrane*, the hairs of the inner hair cells attached to this specific part of the basilar membrane are bent. The bending of these hair cells modify their membrane potential. The transfer function of this potential is asymmetric and nonlinear: larger when bending away from the basal body (depolarization) than when bending towards it (polarization). For large deflections, saturation occurs. The *inner hair cells* activate the

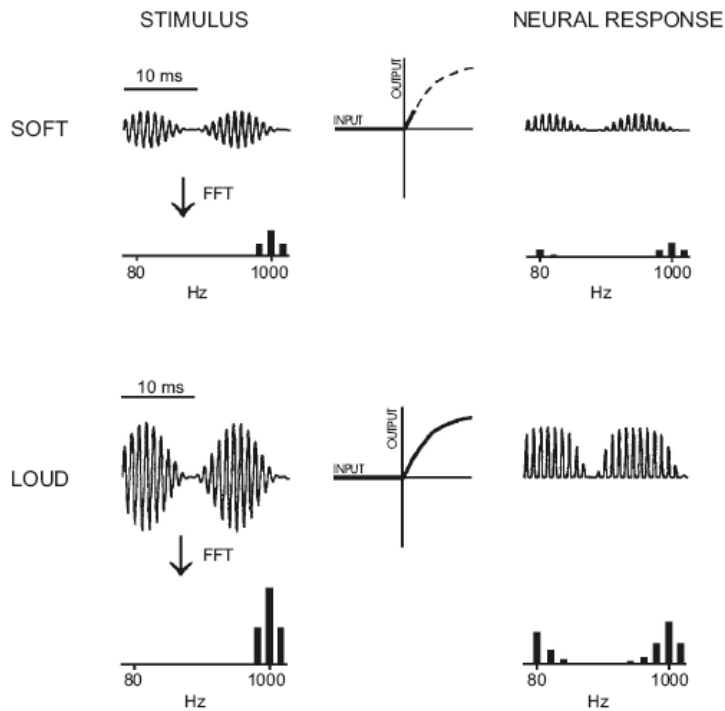


Figure 2.1: A simple model for compressive rectification. Time and frequency domain of an amplitude modulated signal and its rectification are provided for two types of intensity (soft and loud). The rectification causes the output of the cochlea to have a spectral component at the modulation frequency. Compression occurs when depolarization saturates because of loud stimuli. From Lins et al. (1995).

ganglion cells that compose the auditory nerve. Here a complete rectification occurs as only depolarization of the inner hair cells activates the action potentials in the ganglion cell. Larger deflections cause faster firing of these cells. This way, the resulting *transfer function* is nonlinear and saturating towards high deflections. The nonlinear transfer function of the hair cells immediately explains the transfer of the envelope of the modulated signal to the auditory nerve, as a rectified AM signal also carries energy on its modulation frequency. The carrier frequency itself is also transmitted to the *auditory nerve*, but the phase-locked firing capability of the auditory nerve cells deteriorates fast above 1000 Hz (Batra et al., 1986).

The exact sources of the auditory steady-state response and their location are still subject of discussion. It is assumed that roughly two separate sources compose the ASSR recorded at the scalp. The relative contribution of each source depends on the modulation frequency of the stimulus. Generally ASSR generated by stimuli with a modulation frequency below  $\pm 60$  Hz originate mainly from the *auditory cortex* while ASSRs above  $\pm 80$  Hz have their main source in the *brainstem*. Purcell et al. (2004) nicely illustrate this in their Figure 1.6 together with a possible explanation of the observations. The cortical source, with a declining amplitude for increasing frequencies, constructively interferes (around 40 and 90 Hz) and destructively interferes (around 20 and 70 Hz) with the brainstem source. These findings are backed by other studies. At high modulations, ASSRs are not affected by state of arousal (Aoyagi et al., 1993; Cohen et al., 1991). However at lower modulation frequencies, amplitude decreases during sleep or sedation (Kuwada et al., 1986). This implies a cortical generator. Several dipole source analyses show indeed that 88 Hz responses are generated for the most part in the brainstem and 39–40 Hz responses in the cortex (Herdman et al., 2002; Purcell et al., 2004; Reyes et al., 2005; Ross and Pantev, 2004). The low-frequency 10 Hz response originates mainly somewhere from the auditory cortex, but sources in the brainstem may still be present (Herdman et al., 2002).

The usefulness of this knowledge becomes apparent when auditory functions have to be evaluated up to a certain point of the auditory pathway. Responses to stimuli with higher modulations only suggest normality up to the brainstem, while those at lower modulation frequencies imply normal function at higher parts of the auditory nerve (Herdman et al., 2002).

### 2.1.3 Stimulus parameters

Auditory steady-state responses are evoked by any signal that is modulated by a sinusoid. For matter of convenience we refer to such a modulated signal as an *ASSR stimulus*. Examples of signals that can be modulated, and that are currently subject to research, are other sinusoids, (speech weighted) noise and babble noise (Alaerts et al., 2007a). The following theoretical derivation is valid for a sinusoid (the *carrier*) that is modulated by another sinusoid (the *modulator*). When other signals like e.g. noise are modulated, the following formulae still can be used. These signals can be decomposed in their composing sinusoids and thus can be seen as different carriers modulated by the same modulator, eliciting a composed ASSR with that modulation frequency.

A standard single-carrier ASSR stimulus  $x_{\text{ASSR}}$  consists of a sinusoidal carrier with carrier frequency  $f_c = \frac{\omega_c}{2\pi}$  that is amplitude modulated with a modulation rate  $\omega_m = 2\pi f_m$  and/or frequency modulated with a modulation rate  $\omega'_m = 2\pi f'_m$ . It can be represented by



$$x_{\text{ASSR}} = \underbrace{[1 + m \sin(\omega_m t + \phi)]}_{\text{AM component}} \cdot \underbrace{\sin(\omega_c t + \varphi)}_{\text{modulated carrier}} \quad (2.1)$$

with

$$\varphi = \underbrace{\beta \sin(\omega'_m t + \phi')}_{\text{FM component}} \quad (2.2)$$

and

$$\beta = \frac{\Delta\omega}{\omega'_m} \quad (2.3)$$

A maximum AM depth  $m$  ( $0 \leq m \leq 1$ ) implies a full amplitude modulation with a modulation rate  $\omega_m$ . For frequency modulation, the modulation index  $\beta$  is the maximal change in frequency  $\Delta\omega$  from the carrier modulation rate  $\omega_c$ , relative to the modulation rate  $\omega'_m$ . In general the AM frequency  $f_m = \frac{\omega_m}{2\pi}$  and FM frequency  $f'_m = \frac{\omega'_m}{2\pi}$  are taken identical. When in this case  $\Delta\phi = \phi' - \phi$  is taken equal to  $-\frac{\pi}{2}$ , the maxima of AM and FM coincide and the dynamic range of the modulated carrier  $\omega_c$  is maximal.

In this section, the effect of these parameters is discussed.

### Carrier frequency $f_c$

The *carrier frequency*  $f_c$  of an ASSR stimulus defines the location where the cochlea is excited. The major part of all ASSR studies covered in the references of this thesis investigates the audiological frequencies between 500 and 4000 Hz because of the major relevance for communication, speech understanding and speech recognition. Some recent studies also try to cover the frequencies beyond those used in classical audiometry, more specifically at 250 and 12000 Hz. The results here attest to the ability to record ASSRs to a wide range of carrier frequencies but also suggest that accuracy of threshold estimation suffers toward the audiometric extremes (Petitot et al., 2005; Tlumak et al., 2007).

### Modulation frequency $f_m$

The *modulation frequency*  $f_m$  modulates the carrier frequency  $f_c$  and acts as a ‘label’ by which the effect of the carrier frequency up to the brainstem (or cortical regions) is recognized, as described in Section 2.1.2. The studied range of modulation frequencies for ASSR purposes varies between 2 and 600 Hz

(Alaerts et al., 2007a; Dimitrijevic et al., 2004; Kuwada et al., 1986; Picton et al., 1987; Purcell et al., 2004; Rees et al., 1986). It is shown in the earlier studies that the magnitude of the ASSR decreases with increasing modulation frequency, except for the presence of two local increases at 40 and 90 Hz (Figure 1.6). The EEG noise also drops with higher modulation frequencies, which indicates that these 40 and 90 Hz regions are interesting as the SNR increases (Picton et al., 2003).

The first modulation frequency zone to be investigated thoroughly was around 40 Hz (Galambos et al., 1981). In adults, the responses in this modulation frequency region were the largest and could be detected easier than those in the 90 Hz region. However, sleep and sedation affects the magnitude of the response (Cohen et al., 1991; Plourde and Picton, 1990). Moreover, responses at 40 Hz are difficult to measure in neonates, as parts of the 40 Hz response are generated more cortically compared to 90 Hz responses. Cortical responses need a longer maturation and are generally not fully present at the earliest stages of human development in the first few months after birth (Maurizi et al., 1990; Stapells et al., 1988).

For this reason, focus was shifted more to modulation frequencies of 90 Hz. Responses are smaller, but less influenced by sleep (Cohen et al., 1991; Plourde and Picton, 1990). Reliable measurements in children, especially used for audiometry, are possible (Aoyagi et al., 1993, 1994; Cohen et al., 1991; Rance et al., 1995; Rickards et al., 1994). However, one has to keep in mind that ASSR at 90 Hz still are not mature up to the age of six weeks and thus are different from those observed in older subjects (Alaerts et al., 2008; Cone-Wesson et al., 2002; Rance and Tomlin, 2006; Savio et al., 2001).

### Modulation type (AM / FM) and depth ( $m$ ) or index ( $\beta$ )

The simplest ASSR stimuli are sinusoidal tones that are only amplitude modulated. This is the case for  $\varphi$  being constant in (2.1). The larger the AM depth  $m$  in (2.1), the greater the response that the stimulus elicits.

The use of only FM stimuli is possible too (Maiste and Picton, 1989; Picton et al., 1987). When  $m$  is taken equal to zero in (2.1), only the modulation index  $\beta$  in (2.3) determines the amount of modulation. The larger  $\beta$ , the larger the responses the stimulus generates. However, when taken too large, the frequency specificity of the stimulus decreases as the width of the stimulus in the frequency spectrum transcends one critical band of the cochlea.

The *mixed modulation* (MM) technique involves both the use of amplitude and frequency modulation at the same modulation frequency  $f_m = f'_m$ . Each modulation type excites independent responses which are added together. When  $\Delta\phi = \phi' - \phi$  is taken equal to  $-\frac{\pi}{2}$ , the largest responses are obtained (John and Picton, 2000a). Currently the combination of both AM (maximum of  $m = 100$

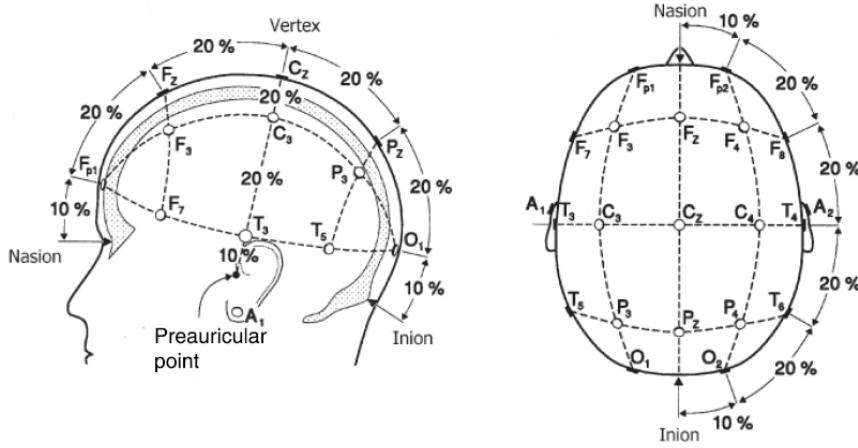


Figure 2.2: The international 10–20 EEG system for electrode placement. From Malmivuo and Plonsey (1995).

%) and FM ( $\Delta\omega = 20\text{--}25\%$  of  $\frac{\omega_c}{2}$ ) stimuli is mostly used as this parameter setting induces the largest responses without losing frequency specificity (Cohen et al., 1991; John et al., 2001b). When  $f_m \neq f'_m$ , one speaks of *independent amplitude and frequency modulation* (IAFM) according to Dimitrijevic et al. (2001).

### 2.1.4 Recording EEG

Potential differences are generated between electrodes positioned on the scalp surface when areas in the brain are active. The EEG records these potential differences as a function of time using electrodes that are placed in standard positions. Figure 2.2 presents the 10–20 international system electrode placement (Malmivuo and Plonsey, 1995). The electrode positions are constructed by dividing the line between the nasion and inion into 10 % or 20 % intervals. The line between the preauricular points is divided similarly. In this work we apply the 10–20 international system of electrode placement, extended with two positions behind the ears, the mastoids (lMa and rMa).

For standard single channel ASSR recordings, the *active electrode* (+) is mostly placed at the vertex (Cz) or at the high forehead (Fpz). The *reference electrode* (-) is placed at the neck or occiput (Oz). The place of the *ground electrode* is more variable: one of the mastoids, the right clavicle or Pz (John et al., 2001a; Lins and Picton, 1995; Luts and Wouters, 2005; van der Reijden et al., 2004).

### 2.1.5 Multiple stimulus ASSR

Lins and Picton (1995) were the first to publish a key paper that describes the possibility to evoke more than one ASSR at a time. Figure 2.3 shows different carriers that can be applied simultaneously to one or both ears when modulating each carrier with a different modulation frequency. The responses that appear in the EEG can be considered as labels. These labels identify the fact that the modulation frequency corresponding to a certain carrier has generated a response on the auditory path. This knowledge opened the way for relatively fast objective audiometry (Lins et al., 1996).

The major benefit of this multiple stimulus technique is the measurement time reduction that is possible. Although one would expect a time reduction by a factor of eight in the case of eight simultaneously tested carriers, this number more comes down to a factor of two or three, which is still a significant improvement (John et al., 2002a). The reason for this phenomenon lies in the loss of ASSR amplitude because of the interaction of the combined stimuli in the auditory nerve (Lins and Picton, 1995) or overlap on the basilar membrane (Picton et al., 2003). These effects are amplified when using stimuli with intensities above 75 dB SPL (John et al., 1998; Lins and Picton, 1995; Lins et al., 1996). When modulation frequencies are used that are less than 1.3 Hz apart or carrier frequencies closer together than one octave, the same effect applies, even at lower intensities (John et al., 1998).

Modulation frequencies around 40 Hz give much greater interactions and this indicates that the dichotic multiple stimulus technique is not useful for this frequency range (John et al., 1998). More recent data however, with a multiple stimulus applied at only one ear, shows an advantage compared to a single stimulus application at low and moderate intensities (Fontaine and Stapells, 2007). When applying at even lower modulation rates, e.g. 14 Hz, the amount of interaction between stimuli even decreases again (Armstrong and Stapells, 2007). Although the earlier studies may lead to the premature conclusion that the interaction between adjacent frequencies is greater at higher levels of the auditory system (Picton et al., 2003), this tendency for lower modulation frequencies should not be generalized without additional research.

### 2.1.6 Comparison with the auditory brainstem response

Auditory steady-state responses are often compared to the auditory brainstem response. Reason for this is the similarity of both techniques in assessing hearing thresholds. Four advantages for the ASSR compared to the ABR can be found.

- ASSRs are frequency specific. A tested frequency only needs to be modulated with a predefined modulation frequency that acts as a label to be

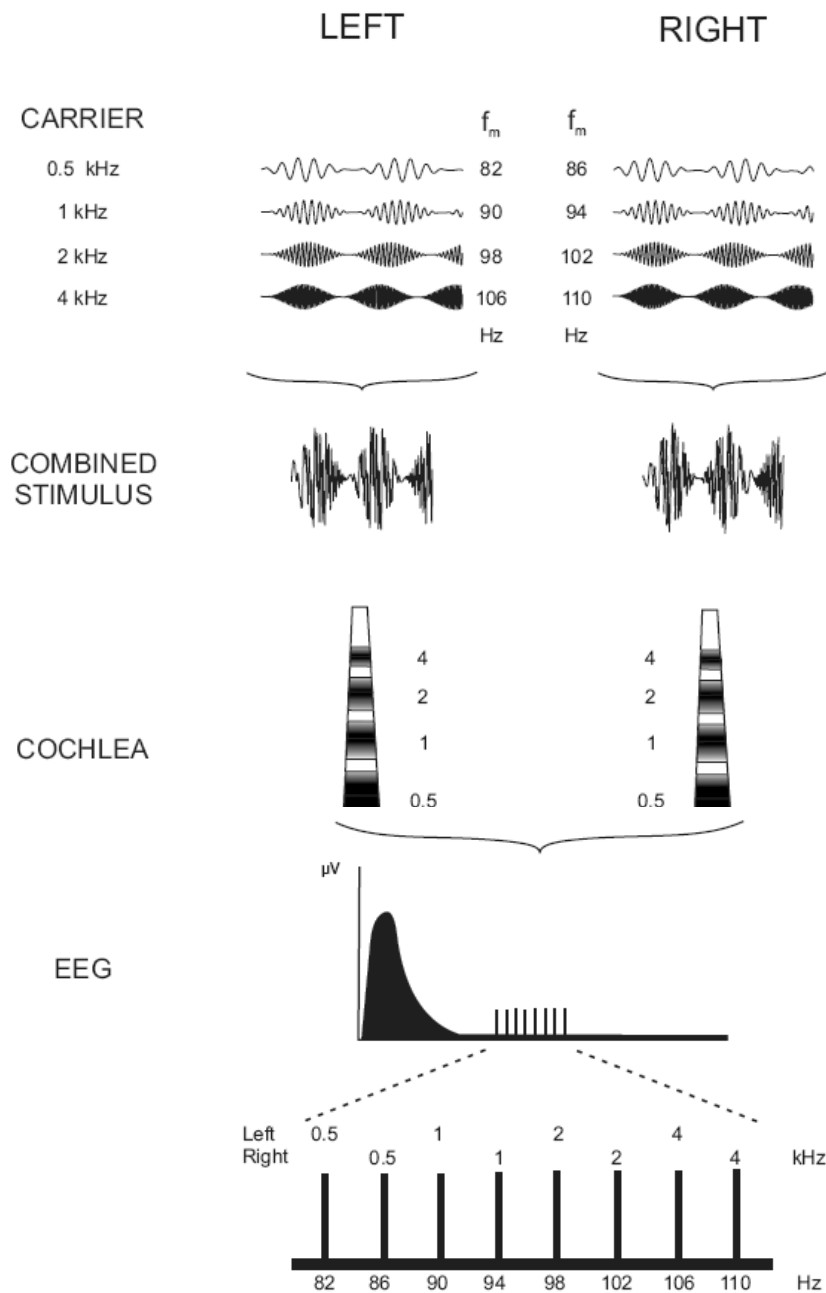


Figure 2.3: Multiple stimulus ASSR. Both ears are stimulated simultaneously by a combined stimulus that consists of four carriers per ear. Each carrier is modulated with a different modulation frequency  $f_m$ . Each separate carrier excites the cochlea at its corresponding carrier frequency. The modulation is passed through to the auditory nerve. The response appears in the EEG, originating mainly in the brainstem for the current example. Each separate modulation frequency acts as a label for its corresponding carrier. From Luts (2005).

retrieved in the EEG. The carrier frequency excites the corresponding part of the cochlea to be tested. The resulting response thus only gives information about the frequency of interest. As the standard ABR technique concerns a click-evoked response, the entire cochlea is excited and no frequency specific information can be retrieved from the response. In a special case of ABR however, tone bursts can be used as stimuli, as described in Section 1.5.2. This technique improves the frequency specificity, but as a trade-off measurement time significantly increases. With the ASSR technique, it is technically possible to test a specific audiometric frequency without the need of filtering part of the signal (as with ABR). In practice however, highpass and lowpass filters are still applied to the signal to reduce amplifier overloads due to artifacts.

- Different carrier stimuli can be applied simultaneously, as long as different modulation frequencies are chosen to distinguish (Lins and Picton, 1995). This induces a possible time reduction of a factor two to three (John et al., 2002a).
- The maximum test intensity of the ASSR stimulus can be higher than the maximum intensity of the ABR stimulus. ASSRs involve continuous stimuli, that generate less peak energies than its ABR counterparts, using clicks. Hearing thresholds can be determined at higher intensities, allowing differentiation between deafness and severe hearing loss.
- ASSRs emerge in the frequency spectrum as solitary peaks compared to surrounding noise in neighboring frequency bins. By comparing the response power with the mean noise power in those adjacent bins, an objective statistical method can be used. The ABR technique still relies on the subjective evaluation of experienced clinicians.

This is not the complete picture however, as the ABR also has its benefits compared to the ASSR technique.

- ASSRs are faint responses (order of nanoVolts) compared to ABR (order of hundreds of nanoVolts). Therefore, ASSRs are much more difficult to detect what results in longer measurement times and technically more specialized measurements.

## **2.2 Techniques reducing recording time**

The largest drawback of the ASSR technique is the extensive recording time needed for reliable hearing threshold estimation. Generally measurements last between 45 minutes and one hour (Luts et al., 2006; Luts and Wouters, 2004).

It would be of general interest if this duration could be lowered significantly. Literature shows several efforts to accomplish this goal. A general overview is provided in Figure 1.7 and situates some techniques in the general picture and the structure of the thesis. This section covers these techniques. In Sections 2.2.1 and 2.2.2, new stimuli are discussed that evoke larger responses or make use of the available recording time more intelligently. Next, these responses should be recorded optimally using an appropriate electrode placement (Section 2.2.3). Sections 2.2.4 and 2.2.5 review a couple of basic single channel signal processing techniques to enhance the recorded signal. Section 2.2.6 describes the statistics that can be used for detection. Section 2.2.7 uses the acquired information to assess the hearing thresholds one is looking for.

### 2.2.1 Stimuli eliciting higher responses

A common stimulus used for ASSR generation is the 100 % AM stimulus, with  $m = 1$  and  $\beta = 0$  in (2.1), (2.2) and (2.3). Other stimuli have been proposed that evoke higher responses than this basic stimulus.

A more efficient stimulus that already has been mentioned in Section 2.1.3, is the mixed modulation (MM) stimulus that combines both AM and FM. Generally  $m$  is taken equal to 1 and  $\Delta\omega = 0.2\frac{\omega_c}{2}$ . The modulation may also follow an exponential sine function (John et al., 2002a). For this purpose, the AM component of (2.1) can be extended to

$$\text{AM component} = 1 + 2m \left( \left( \frac{1 + \sin(\omega_m t + \phi)}{2} \right)^N - 0.5 \right) \quad (2.4)$$

When the exponent  $N$  is taken equal to 2, together with  $m = 1$  and  $\beta = 0$  in (2.2), the use of the so called *exponential modulation* also increases response amplitudes. Both methods, mixed and exponential modulation, boost ASSR amplitudes between 15 and 30 % individually. This is the case for adults (John et al., 2001b, 2002a) as well as for infants (John et al., 2004). When  $\beta$  and  $N$  are further increased, the frequency specificity of the stimulus lowers and neighboring frequency bands in the cochlea are stimulated. A recurring theme in most of the studies that search for better stimuli is the increase of side bands neighboring the investigated carrier frequency. Both FM as exponential modulation increase the number of side bands. Other stimuli that seriously exploit this technique are described in Elberling et al. (2007), Riquelme et al. (2006), and Stürzebecher et al. (2001). The latter study claims to increase ASSR amplitudes by 65 %.

### 2.2.2 Independent stimuli with multiple stimulus ASSR

If the intensities of the stimuli can be manipulated individually within a maximum mutual range of 20 dB, a shorter measurement session is possible. This technique is referred to as the *multiple intensity* (MINT) technique (John et al., 2002b). Larger differences in intensity cause interactions between stimuli. With this manipulation, the measurement protocol has the possibility to stop collecting data for a certain carrier frequency when its response is already significant. This can occur before the predefined amount of data for the current intensity is collected or before all other responses to their corresponding frequencies have become significant. The intensity of the specific carrier frequency is immediately reduced as soon as a response is detected (Mühler et al., 2005). An extensive study covering the amount of time reduction that is practically possible using this method has not been published yet at this moment. A research project in our lab environment is currently studying this topic using the SOMA program described in Chapter 3, which has the capability of changing individual intensities at runtime.

### 2.2.3 Intelligent placement of electrodes

The Cz–Oz or Cz–neck derivations are commonly the derivations mostly used for adults to record the responses evoked by the type of stimuli described in the previous section (John et al., 2001a; Lins and Picton, 1995; Luts and Wouters, 2005). The choice for these derivations lies in the general experience that the mean ASSR thresholds are lower with this EEG derivation compared to other derivations. This experience is confirmed for the Cz–Oz derivation by van der Reijden et al. (2004), who use a multichannel setup to search for the EEG derivations offering the highest signal-to-noise (SNR) values and thus the lowest ASSR thresholds. van der Reijden et al. (2004) also indicate that this Cz–Oz derivation does not always guarantee the highest SNR and that sometimes other derivations (like e.g. Cz–mastoid, Cz–Pz and Cz–neck) could offer better SNRs. This is backed by the observation that ASSR thresholds (and their SNRs) are largely variable across subjects (John et al., 2001a; Lins and Picton, 1995; Luts and Wouters, 2005). According to van der Reijden et al. (2005), the highest SNR for infants younger than six months are mainly recorded from the mastoids ipsilateral to the ear of stimulation referenced to Cz. Thus, the derivation with the highest SNR is subject dependent and cannot be predicted beforehand. Recording EEG using multiple derivations simultaneously, combined with an appropriate algorithm to process these recordings, should be able to select the ‘best’ channel (or combination of channels) for as many subjects as possible. Appropriate multichannel processing schemes for this matter are described in Chapters 5 and 6. A practical evaluation of these schemes is described in Chapter 6. An efficient electrode placement is proposed in Chapter 7.



### 2.2.4 Averaging

After recording EEG data, the data need to be processed as already described shortly in the previous section. ASSR processing, and evoked potential processing in general, mostly involves a basic form of averaging after splitting up the available EEG data stream into smaller blocks (e.g. *sweeps*) to increase the SNR of the observed signal. The oldest technique is mean averaging, while over the years several new techniques have emerged that try to take into account the level of noise in each data block.

*Mean averaging* assumes that recorded single sweeps contain a stationary signal superimposed on randomly occurring noise. Therefore, the synchronous summation of responses improves the signal component while reducing noise. The averaging process is a consistent estimator of a signal in noise because it has no bias and the variance decreases by increasing the number of sweeps averaged. This approach has some limitations however. Mean averaging is based on the principle that the signal in a response is constant and phase locked to the stimulus, whereas the noise is stationary and random with no phase locking to the stimulus. Unfortunately, noise is nonstationary as well. Depending on the severity of deviations from these normal assumptions, mean averaging may produce suboptimal extraction of the signal from noise. In mean averaging, each sweep is treated as having the same amount of noise and contributing equally to the final average. During recording however, owing to the patient's changes in position or relaxation state, the level of noise in some sweeps may be higher than in others. Mean averaging disregards this fact. This problem is especially apparent when averages are disturbed by data blocks with extremely high noise levels (*artifacts*).

A rough, but effective, solution for artifacts is the use of *artifact rejection* (Pantev and Khvoles, 1984; Picton et al., 1983). Small data blocks exceeding a certain amplitude level are not used for analysis. This reject level can be determined beforehand and kept fixed throughout the measurements. Alternatively, the reject level can be calculated afterwards as a certain percentile of the noise distribution of the dataset. Real-time analysis is not possible in this case however, but the reject level is more tailored to the subject (John et al., 2001a).

*Sample-weighted averaging* copes with the abrupt processing of artifact rejection (Hoke et al., 1984; Lütkenhöner et al., 1985). Small data blocks, *epochs*, are weighted with the variance of the epoch in the time domain. The whole sample (response + noise) is taken into account. The idea behind this method is that noise in an epoch produces a bigger variance compared to a response in that epoch. Epochs are taken instead of whole sweeps to better characterize the individual variances of the epochs and to follow sudden changes in the signal. In the end the epochs are combined to a sweep and the sweeps are summed together. Each epoch is divided by the sum of the weights of the epoch the

final epoch is composed with. This can be expressed as follows (adapted from Lütkenhöner et al. (1985) by John et al. (2001a)):

$$\mathbf{a}(t) = \sum_{j=1}^N w_j \mathbf{x}_j(t) \quad (2.5)$$

where  $\mathbf{a}(t)$  is the weighted average waveform across the time points of the epoch  $\mathbf{x}_j(t)$ ,  $N$  is the number of epochs being summed together, and  $w_j$  is the weighting factor for the  $j$ -th epoch,

$$w_j = \hat{\xi}_j^{-2} \left( \sum_{k=1}^N \hat{\xi}_k^{-2} \right)^{-1} \quad (2.6)$$

where  $\hat{\xi}_k^2$  is an estimate of the variance of the epoch  $\mathbf{x}_k$ ,

$$\hat{\xi}_k^2 = \frac{\sum_{i=1}^M \mathbf{x}_k^2(i)}{M} \quad (2.7)$$

where  $M$  is the number of data points in the epoch  $\mathbf{x}_k$ .

A variant of sample weighted averaging is *noise-weighted averaging*. Whereas the former takes both response and noise power into account, the latter only estimates the noise power in the frequency region of interest. Both techniques are similar (Dobie and Wilson, 1994), although some reports show a benefit for sample-weighted averaging (John et al., 2001a). It is indicated anyway that weighted averaging is preferred above normal averaging or artifact rejection.

Theoretically, an approximation of the optimal weights  $\hat{\mathbf{w}}_{opt}$  for averaging can be deduced (Davila and Mobin, 1992),

$$\hat{\mathbf{w}}_{opt} = \frac{\bar{\mathbf{X}}\mathbf{X}^T}{\|\bar{\mathbf{X}}\mathbf{X}^T\|} \quad (2.8)$$

with  $\mathbf{X}^T$  being an array of  $M$  single sweep waveforms  $\mathbf{x}_i(t) = \mathbf{s}_i(t) + \mathbf{n}_i(t)$  and  $\bar{\mathbf{X}} = \mu_\sigma \mathbf{s}(t) + \mu_n$  with

$$\mu_\sigma = \frac{1}{M} \sum_{i=1}^M \sigma_i^2 \quad (2.9)$$

$$\mu_n = \frac{1}{M} \sum_{i=1}^M \mathbf{n}_i(t) \quad (2.10)$$

$\sigma_i^2$  are the variances of the signals  $\mathbf{s}_i$ . This result is based on a SNR maximization by means of the generalized eigenvalue decomposition (GEVD) of the signal and the noise correlation matrices. Its approach is similar to the application of the multichannel Wiener filter in Chapter 5.

### 2.2.5 Adaptive filtering

Another way of processing data is the use of *adaptive filtering*. Tang and Norcia (1995) show through simulations that a significant improvement is possible when applying an adaptive recursive least-squares (RLS) filter to EEG data. An input EEG signal containing an artificial signal is used as the primary input. Two reference signals (sinusoid and cosinusoid) are weighted so as to minimize the squared error between the primary input and the predicted response, which is the difference between the primary input and the weighted reference signals. The amplitude and phase of the evoked response are estimated by coherent averaging of the weights. They both are compared to a threshold to decide for detection.

### 2.2.6 Statistics

The processing in the previous subsections will return an output that may or may not contain a response. The decision whether a response is present in the processed recorded data is made using statistical techniques. Myriads of studies have been conducted to determine which statistical evaluation method achieves the highest detection performance for steady-state evoked potentials. The three most important statistical methods are the F-test, magnitude squared coherence (MSC) and phase coherence (PC). Other methods exist like circular  $T^2$  (CT2) and Hotelling  $T^2$  (HT2), but these are merely deviations from one of the three approaches. An overview of the important three is given below.

- *F-test*. The F-test is a statistical test in which the test statistic has an F-distribution if the null hypothesis is true. It is a measure of the ratio of the variation explained by the model (systematic variation) and the variation explained by unsystematic factors (unsystematic variation). More specifically applied to ASSR, it is assumed that the activity in the adjacent frequency bins around the response is random and uncorrelated, with an equal variance and a mean of zero in the real and imaginary dimensions. The surrounding frequency bins are only assumed to be uncorrelated asymptotically in the case of colored noise (Brillinger, 2001; Van Trees, 2001). The power of the response relative to the mean power of the adjacent frequency bins is denoted as the F-ratio. If this F-ratio transcends a predefined threshold of a  $F(2, 2N)$  distribution, a response is considered to be present (2 and  $2N$  being the degrees of freedom with  $N$  the number of noise bins) (Lins et al., 1996).

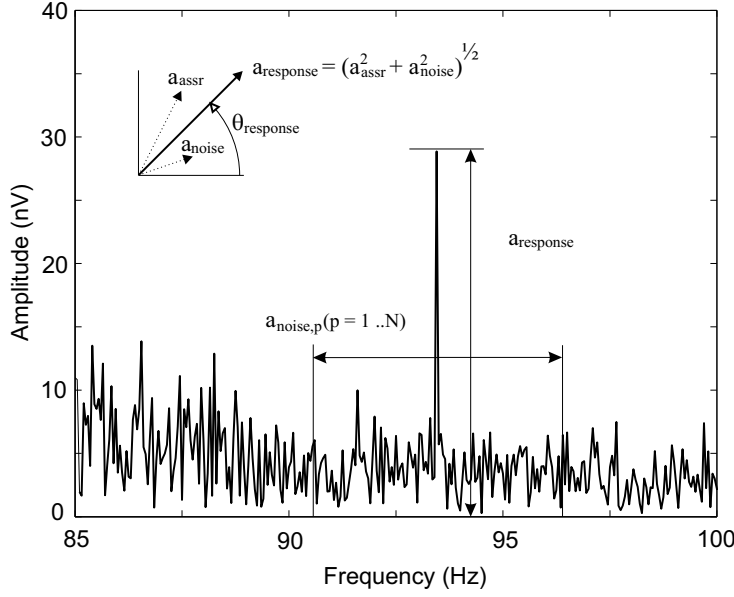


Figure 2.4: Frequency domain representation of an observed ASSR in the EEG.

In most ASSR studies, an FFT analysis is carried out prior to response detection (John and Picton, 2000a; Lins and Picton, 1995; Luts et al., 2006; Valdes et al., 1997). An example of an FFT from EEG data containing an ASSR is displayed in Figure 2.4.  $a_{\text{response}}$  is the amplitude of the observed response, which is a vectorial summation of both the amplitude of the real (unknown) ASSR response  $a_{\text{assr}}$  and the amplitude of the noise in the same frequency bin  $a_{\text{noise}}$ . By calculating the mean noise power  $\sigma^2$  in the adjacent noise bins  $p$  with  $p = 1 \dots N$ , the noise power in the frequency bin that also contains the ASSR is estimated.

The detection is based on the ratio between the response power  $P_r$  at the modulation frequency and the mean noise power  $\sigma^2$  in  $N$  neighboring frequency bins at each side,

$$F_{\text{ratio}} = \frac{P_r}{\sigma^2} = \frac{a_{\text{response}}^2}{\frac{1}{N} \sum_{p=1}^N a_{\text{noise},p}^2} \quad (2.11)$$

with  $a_{\text{response}}$  the amplitude of the modulation frequency bin and  $a_{\text{noise},p}$  the noise amplitude in the  $p$ -th adjacent frequency bin (John and Picton, 2000a).

- *Magnitude squared coherence* (MSC). The MSC technique is a measure of the degree to which system output (such as an AEP) is determined by a specified input, as a function of frequency. For a linear system, MSC

measures the proportion of response power attributable to a given stimulus. For non-linear systems, e.g. the topic of this thesis, the technique still can be used since the main interest is not in measuring coherence per se, but in determining whether a response is present. An estimated coherence value above the appropriate critical value does this, whether the response is linear or not. MSC is calculated as  $\gamma_{xy}^2(f)$  and is compared to a predefined threshold to determine response detection (Dobie and Wilson, 1989),

$$\gamma_{xy}^2(f) = \frac{|G_{xy}(f)|^2}{G_{xx}(f)G_{yy}(f)} \quad (2.12)$$

with  $G_{xx}(f)$ ,  $G_{yy}(f)$  and  $G_{xy}(f)$  the (cross) power spectral density (PSD) of input  $\mathbf{x}(t)$  and output  $\mathbf{y}(t)$ ,

$$G_{xx}(f) = |X(f)|^2 \quad (2.13)$$

$$G_{xy}(f) = X^*(f)Y(f) \quad (2.14)$$

- *Phase Coherence* (PC). Opposite to the F-test and the MSC, which use both phase and magnitude information of the signal, this test only calculates the phase coherence  $R_\theta$  by taking into account phase information  $\theta_{\text{response}}$  of the response (Figure 2.4). Higher values indicate a lower probability that the phase is changing randomly over  $q$  data blocks (Picton et al., 2001; Stapells et al., 1987).

$$R_\theta = \frac{1}{q} \sqrt{\left(\sum_{i=1}^q \cos \theta_{\text{response},i}\right)^2 + \left(\sum_{i=1}^q \sin \theta_{\text{response},i}\right)^2} \quad (2.15)$$

It can be concluded that all these different methods have essentially the same performance (Dobie and Wilson, 1994, 1996; Valdes et al., 1997). If a study claims to show a significant difference however, the improvement is marginal.

### 2.2.7 Smarter guesses of hearing thresholds

This last category does not exactly reduce measurement time, but suggests to use the time available in a smarter way, allocating relatively more interest on intensities near hearing thresholds.

The common way to determine hearing thresholds is the use of *threshold bracketing*. Discrete intensities are applied for a predefined period or until a response has been found for that specific intensity. The examination starts at an intensity high enough to evoke a response. The intensity is decreased until a response

can no longer be recognized. The threshold is defined as the lowest intensity where the response can be detected.

A second method records responses at one or more intensities above threshold. When the response amplitude is plotted as a function of stimulus intensity, this relation may be extended towards zero response amplitude through *extrapolation*.

When applying sweeps with continuously varying intensities, one refers to the *intensity sweep technique* (Picton et al., 2007). The response is recorded as the intensity is slowly but continuously swept from below to above hearing threshold levels. By decreasing the slope of the ‘stimulus intensity over time’ function at hearing threshold intensities, relatively more time is attributed to this critical region.

The main problem of the extrapolation and intensity sweep techniques is that the relationship between response amplitude and intensity is non-linear. Lins et al. (1995) show that the amplitude-intensity functions in normal subjects have different slopes for different intensity ranges. This effect is even greater in subjects with sensorineural hearing loss. It might be possible to prevent these mistakes by limiting the extrapolation to near-threshold data and eliminating data where the slope changes significantly (Picton et al., 2003).

To finalize it should be stressed that any deviation from a test protocol one is familiar testing with needs to be investigated thoroughly before being used clinically. It should be checked whether the false detection rate (the error rate) does not increase significantly when altering the test protocol. According to Luts et al. (2008), even the change from measurements with a fixed length to measurements with variable lengths can give rise to an explosion of the error rate. Measurements with variable lengths stop the measurement after a certain number of significant data blocks (‘sweeps’) have been recorded. This way, recordings seem to speed up. However, in most variations of this paradigm the detection threshold should be made more stringent accordingly which will reduce or cancel out the benefit of the possibility to stop earlier in a measurement.

## 2.3 Conclusions

This chapter describes both the ASSR technique, its applications and ways to increase efficiency in determining a subject’s hearing thresholds.

Sections 2.1.1 and 2.1.2 review the history of the auditory steady-state response and its physiological model. ASSRs are generated along the auditory path as a response to auditory stimuli applied to the ear. These stimuli are basically

modulated sinusoids, with the modulation frequencies emerging in the EEG, indicating that the modulated carrier has been perceived up to the response generator. These generators lie in the auditory cortex and/or the brainstem, depending on the modulation frequency of the stimuli (Section 2.1.3). Section 2.1.4 shows how the responses they generate are recorded using electrodes on the scalp in the form of an EEG. By varying the stimulus parameters larger responses can be evoked. A combined stimulus, containing more than one modulated carrier at a time, can reduce measurement time significantly (Section 2.1.5). This way more than one carrier frequency can be tested simultaneously. Section 2.1.6 concludes with a comparison between the ASSR and the auditory brainstem response.

Section 2.2 covers different ways of increasing test efficiency. According to Section 2.2.1, the stimulus itself can be optimized. By modifying its parameters such that more energy is concentrated within the frequency band of interest, larger responses can be evoked. When the separate components of a combined stimulus are modified independently, intensities can be lowered per stimulus component to better match the audiometric configuration of the test subject (Section 2.2.2). Section 2.2.3 discusses the smart placement of electrodes on positions like the Cz–Oz configuration for adults or the mastoid–Cz combination for infants. This will increase the chance of picking up larger responses compared to other electrode positions. After recording the responses, some basic techniques for signal improvement and detection have been used like signal averaging and statistics. In Section 2.2.4, it can be concluded that the best choice is noise-weighted averaging combined with any detection statistic (Section 2.2.6). Also, adaptive filtering can improve detection of sinusoids in EEG noise (Section 2.2.5). Section 2.2.7 shows that hearing threshold estimation may be more efficient when using amplitude extrapolation or sweeps of continuously changing intensity. The positive effects of all these methods can be summarized into a general measurement time reduction. They were not developed with multichannel recordings in mind however. This thesis will focus on multichannel signal processing strategies for auditory steady-state response measurements.





## Chapter 3

# A Flexible Research Platform for Multichannel ASSR Measurements

The current possibilities of commercially available auditory steady-state response recording devices are mostly limited to avoid unintentional misuse and to guarantee patient safety as such. Some setups e.g. do not allow the application of high intensities or the use of custom generated stimuli via files. Moreover, most devices generally only allow data collection using maximum two EEG channels. The freedom to modify and extend the accompanying software and hardware is very restricted or nonexistent. As a result, these devices are not well suited for research and several clinically diagnostic purposes. In this chapter, a research platform for multichannel ASSR measurements is presented, referred to as SOMA (*Setup ORL for Multichannel ASSR*). The setup allows multichannel measurements and the use of custom made stimuli which is a requirement for the studies described in the rest of this thesis. The mobile setup is based on an inexpensive multichannel RME soundcard and software is written in C++.

Section 3.1 describes the general background and the motivation for building the setup. In Section 3.2, both hardware and software of the setup are described. Section 3.3 presents the results of an evaluation study with nine normal-hearing subjects that will show that no significant performance differ-

---

The material presented in this chapter has been published in ‘Van Dun, B., Verstraeten, S., Alaerts, J., Luts, H., Moonen, M., and Wouters, J. (2008), “A flexible research platform for multi-channel auditory steady-state response measurements,” *J. Neurosci. Meth.*, 169, 239–248’.

ences exist between a reference and the proposed platform. These observations are discussed in Section 3.4. Section 3.5 ends this chapter with some conclusions.

### 3.1 Introduction

Currently available devices use different implementations and features but basically they are all based on one of the following two techniques. Firstly, the binaural multistimulus MASTER-technique (Multiple Auditory STEady-state Responses) developed by and based on the research of John and Picton (2000a) at the Rotman Research Institute, University of Toronto. The MASTER-technique was firstly implemented in the MASTER-system of Natus Medical Inc. (formerly Bio-logic Systems Corp.). Secondly, the monaural single stimulus Audera-technique, based on research at the Department of Otolaryngology, The University of Melbourne (Rickards, Rance and Cone-Wesson, et al). First implementation of the Audera-technique occurred in equipment manufactured by ERA Systems, Ltd., and Grason-Statler Inc. (GSI) afterwards. The research platform described in this chapter is based on the binaural multistimulus MASTER-technique; however, it does away with some of the limitations of commercial devices. These limitations sometimes prevent the use of specific functionalities that are necessary for advanced clinical diagnostic and research studies (like testing more than one modulated stimulus simultaneously at intensities above 80 dB SPL, the use of self-created stimuli or multichannel EEG recording). The research platform for multichannel ASSR measurements described here, is referred to as SOMA (Setup ORL for Multichannel ASSR). SOMA uses a standard high-quality multichannel RME soundcard and is written in C++.

The possibility to use custom made stimuli is included. The need for stimuli other than the standard multiple stimulus ASSR is demonstrated by the following studies. Stürzebecher et al. (2006) compensate the travelling wave delay on the basilar membrane by a combined stimulus. Alaerts et al. (2007a) use modulated speech-type stimuli and Picton et al. (2007) work with amplitude modulated tones that vary their intensity over time. SOMA offers the feature to independently modify intensities of multiple carrier frequency stimuli during measurement. Each time an independent stimulus is altered, its corresponding buffer is cleared and the EEG collection for that specific stimulus is started over again. This feature was described first by John et al. (2002b) in Section 2.2.2 and makes a significant measurement time reduction possible (Mühler et al., 2005). The use of multichannel measurements offers extra benefits as fixed single channel electrode positions do not guarantee maximum response recording. van der Reijden et al. (2004, 2005) describe that one should consider different electrode positions for adults and infants. Additional features can be added

with little effort to SOMA. This makes the platform flexible and modularly extensible. Multichannel signal processing, as described in Chapters 4 to 7, can eventually be carried out during measurement. Automatic decision protocols, suggested by John et al. (2002b), can be implemented to make a fully automatic hearing threshold determination possible. The setup is kept portable as the external multichannel RME soundcard is able to cover the needed dynamic range for stimulus presentation, without the need for extra attenuators or audiometers.

Currently, the SOMA setup is being used in four applications. Firstly, SOMA is in use in a clinical environment, more specifically in an operation room during ear tube surgery with sedated young children (age < six years old). Both ASSR and auditory brainstem response (ABR) measurements are carried out. Based on the results of both techniques the subject's hearing thresholds can be estimated objectively. SOMA is also being used in ongoing research. In a second application, SOMA is used for research focusing on the presence of ASSRs evoked by modulated speech-type stimuli (Alaerts et al., 2007a). Thirdly, the possibility to alter stimuli independently during measurement from Section 2.2.2 is investigated. Fourthly, SOMA is used for research purposes on multichannel evoked potentials, as described in Chapters 4 to 7.

## 3.2 Materials and methods

This section describes the hardware and software used for the SOMA setup (Sections 3.2.1 and 3.2.2), its internal structure (Section 3.2.3) and the method of evaluation (Section 3.2.4).

### 3.2.1 Hardware setup

The core of the hardware setup is the RME Hammerfall DSP Multiface II soundcard ([www.rme-audio.com](http://www.rme-audio.com)). The choice for this device instead of a standard data acquisition card was driven by the needs to have an inexpensive, multichannel, portable, high-end, and plug-and-play device. To illustrate, EEG measurements already can be conducted without much effort using the Adobe Audition program. Moreover, the presence yet of several devices in the author's research environment, the available experience and the reusability of C++ software code added to the decision factors.

The RME device is equipped with eight input and eight output channels that can be used simultaneously for playback and recording. It is able to sample up to 96 kHz per channel. For the SOMA setup, the lowest possible sampling rate of 32 kHz was chosen, as both the acoustic stimuli and the EEG data can be represented using even lower sampling rates. The solid lines in

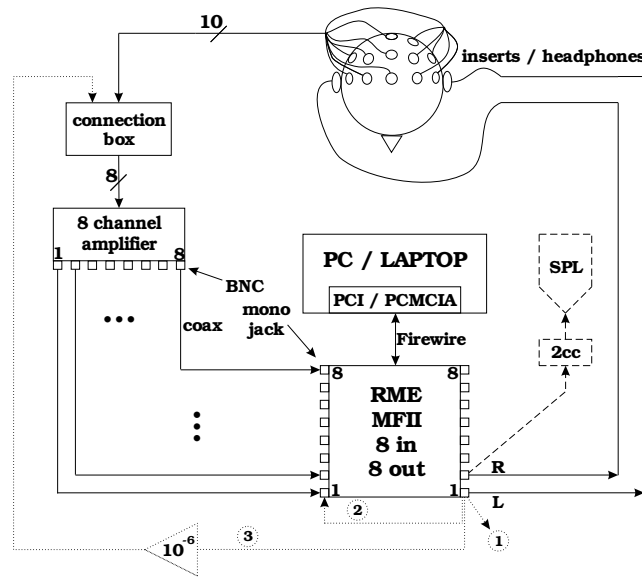


Figure 3.1: Overview of the hardware in solid lines. The SOMA software generates the 64 bit stimuli online or reads in a custom made 64 bit `.sbin`-file and sends these to the multichannel RME soundcard after conversion to 24 bit precision. The first two output channels operate the acoustic transducers. Maximum ten electrodes placed on the skull collect the subject's EEG. The ten EEG potentials are converted such that amplification is possible by a eight-channel differential amplifier. The amplified EEG channels are read in through the RME soundcard inputs and sent back to the SOMA software for further analysis. For a single channel setup only three electrodes, no connection box and a single channel differential amplifier are needed. The dashed lines show the acoustic calibration of the right one of two output channels using a 2cc-coupler, an artificial ear, and a sound level meter (for insert phones in this case). The dotted lines indicate EEG calibration in three steps: (1) voltage measurement of full scale calibration signal out of left output channel, (2) calibration of selected input channel, (3) calibration of connection box / amplifier loop using an attenuated signal.

Figure 3.1 depict the RME soundcard in combination with other hardware. The SOMA software generates the 64 bit stimuli online or reads in a 64 bit `.sbin`-file constructed earlier in e.g. Matlab. The `.sbin`-format was chosen instead of 24 bit `.wav`-files as conversion errors between SOMA software and Matlab were experienced. Depending on the use of a PC or a laptop, the stimuli are reduced to 24 bit precision and are sent to the external RME multichannel soundcard through a PCI or PCMCIA cardbus and a Firewire link. The first two of the eight output channels operate the acoustic transducers,

left and right. In the multichannel case, the subject's EEG is collected using maximum ten electrodes (eight active, one reference and one ground) placed on the skull. A connection box converts the ten EEG potentials to eight channels to be amplified by an eight-channel low-noise differential amplifier. As each differential amplifier channel requires an active, a reference and a ground input, the connection box just distributes the potentials on the reference and ground electrodes to the eight channels. The eight active electrode potentials are guided directly to the active input of each amplifier channel. In the case of a single channel measurement, only three electrodes are sufficient and no connection box is needed. As an example Stanford Research Systems, Grass or Nicolet single channel low-noise differential amplifiers have been used in different ASSR-studies (Luts et al., 2006; Picton et al., 2007; Small and Stapells, 2006). For multichannel recordings, e.g. Jaeger-Toennies (Chapters 4 to 7) or Nihon Kohden (van der Reijden et al., 2005) amplifiers are available. The amplified EEG signals are read in by the RME soundcard inputs as a 24 bit signal. This signal is sent back to the SOMA program via the Firewire link and PCI / PCMCIA cardbus for further real-time processing. The raw EEG is stored in a 64 bit `.sbin`-format and can be processed offline in e.g. Matlab.

The dynamic range (112 dB(A)) of the output channels, used for stimulus presentation, seems rather low to provide an accurate stimulus to external transducers at all desired intensities (10 – 100 dB SPL). It is considered sufficient for the current application however. The maximum intensity per AM/FM modulated carrier is set to 100 dB SPL (Etymotic Research ER-3A insert phones calibrated acoustically using a Brüel & Kjær Sound Level Meter 2260 in combination with a 2-cc coupler DB0138 and an artificial ear 4152 (see Sections 3.2.2 and 3.2.4). Each carrier can be attenuated to 0 dB SPL in the software (see Section 3.2.3). In practice, an attenuation of 100 dB can pose a threat to accurate signal presentation due to the proximity of the noise floor. Figure 3.2 shows the power spectral density of a 100 % amplitude modulated (89.84 Hz) 1000 Hz carrier at an intensity of 10 dB SPL (calibrated), which signifies a 90 dB attenuation from maximum intensity. The effective dynamic range for this signal, defined by the magnitude of the side bands compared to the magnitude of the highest peak of the noise floor in the one-third octave band around the 1 kHz carrier, is more than 25 dB. It can be argued that the dynamic range of 25 dB is still too low, as appropriate speech (42 dB) or music (75 dB) presentation needs a much higher dynamic range (Everest, 2001). The application described here makes use of modulated carriers, exciting particular places on the basilar membrane of the cochlea. The noise in the same one-third octave band is likely masked by the stimulus (Moore, 2003). Moreover, the noise within the one-third octave band of the modulated carrier lies below the hearing threshold of most subjects. The noise in the other one-third octave bands does not influence the perception of the signal by the ear. To summarize, no extra attenuators are necessary to ensure an accurate stimulus presentation between 10 and 100 dB SPL. This improves the mobility of the setup. A stimulus intensity of 0

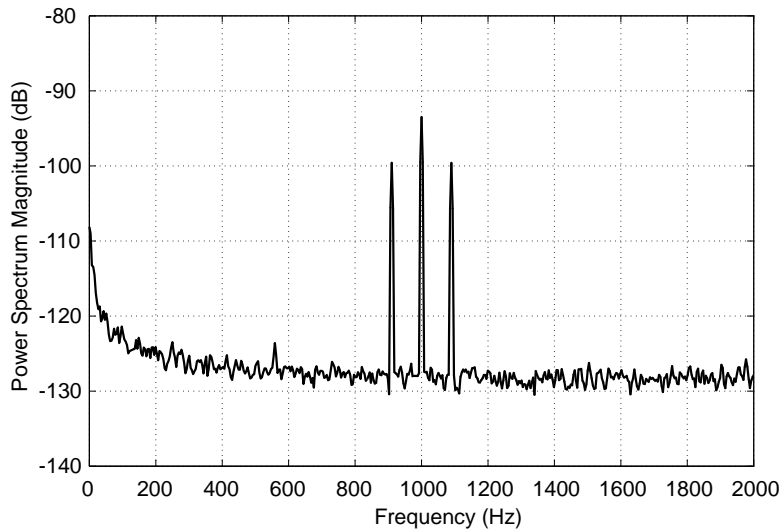


Figure 3.2: Power spectral density of a 100 % amplitude modulated (89.84 Hz) 1000 Hz carrier at an intensity of 10 dB SPL (calibrated) using 128 ms blocks (sampling rate 32 kHz). An unmodulated 1000 Hz sinusoid at 100 dB SPL has its peak at 0 dB.

dB SPL is possible. However one cannot claim for sure that the dynamic range of 15 dB is sufficient for stimuli at this intensity. For audiometric purposes, and especially hearing threshold determination, the possibility for testing at 0 dB SPL is not highly required. In a similar way, 110 dB SPL testing intensities are not really needed either and is a maximum intensity of 100 dB SPL considered as sufficient.

### 3.2.2 Overview of the software

Basically the process can be divided into four steps: calibration, experiment configuration, measurement and real-time analysis. All configuration and calibration files are stored in `.m`-files from the Matlab environment. Custom made stimuli and recorded EEG files are stored in a 64 bit `.sbin`-format. These files can be used for setting up experiments without the SOMA program and automatic offline data analysis afterwards.

#### Calibration

Before starting any measurements, it is necessary to calibrate all data that is sent out (the acoustic stimuli) and read in (the amplified EEG). The dashed lines in Figure 3.1 indicate the acoustic calibration setup. For insert earphones,

a sound level meter is used in combination with a 2-cc coupler and an artificial ear. For headphones, the meter is connected to an artificial ear only. Two types of stimuli can be calibrated and used: stimuli generated by the software and custom made stimuli. The generated stimuli have a fixed range of parameter types, but are suitable to perform standard ASSR measurements. These stimuli are a combination of up to four separate components per ear. For each component, carrier frequency, modulation frequency, AM depth, FM depth and phase difference between the AM maximum and the FM maximum can be chosen. After parameter setting, each separate component is calibrated acoustically at 70 dB SPL. Internally, this calibration factor is stored as the amount of attenuation (dB) needed for a stimulus at full scale intensity to reach 70 dB SPL. All parameter and calibration data are stored in a `.m`-file. During stimulus presentation while running an experiment, a 64 bit combined stimulus is created online based on the stimulus parameters of the different components in this file. Each component is multiplied individually with its calibration factor and the intensity factor given up by the user in the experiment. Finally the individual components are summed to a combined stimulus per ear. The 64 bit stimulus is then reduced to 24 bit precision and sent to the external RME soundcard. Custom made stimuli need to be provided in the 64 bit `.sbin` format. The calibration process is similar to the technique above. The stimulus is played back through the insert or head phones and its intensity is calibrated acoustically at 70 dB SPL. The calibration value is stored for each custom made stimulus together with other stimuli and their calibration values in a single `.m`-file.

The dotted lines in Figure 3.1 show the three steps of the calibration process for EEG potentials.

1. The left output channel was arbitrarily chosen to produce the calibration signal. As an EEG signal is generally expressed in Volt, the calibration signal also should be expressed in this unit. At full scale intensity (0 dB attenuation), the RME outputs produce a 4.86 V zero-to-peak signal when a sinusoid with a calibration frequency of 89.84 Hz is applied. This calibration frequency lies in the center of the frequency range of interest for brainstem evoked auditory steady-state responses. For stimuli with other modulation frequencies, different calibration frequencies should be chosen.
2. A calibration factor is a way to compensate for a difference between the signal one expects and the signal one measures after propagation through an uncalibrated system. The output channel in step (1) generates a stimulus with a known magnitude. This step checks how a full scale signal is perceived by one of the input channels by sending a signal directly from the output channel to the tested input channel.

3. If the reference magnitude (e.g. 4.86 V zero-to-peak) of a full scale stimulus is known at the tested input channel, the magnitude of all other signals arriving at the input channel can be derived. The final step attenuates a full scale signal from the output channel by a factor of  $10^{-6}$ . The attenuated signal propagates through the connection box and the amplifier to the tested input channel. If the amplifier uses an amplification factor of 50,000, the expected signal magnitude at the input channel should be one-twentieth of the full scale magnitude. All deviations from this value can be compensated using the calibration factor.

Both the full scale magnitude of the calibration signal (Volt) and the calibration factor the input signal has to be multiplied with to reach its correct value are stored in a `.m` Matlab file. The duration of the calibration signal should be taken long enough (e.g. one minute) to average out noise influences.

Both calibration files (stimulus and EEG) keep track of a history with calibration values at a certain time. This way abnormal changes in calibration parameters can be detected more easily. Different hardware profiles can be stored in the same file. A hardware profile is a predefined combination of hardware, e.g. amplifier, connection box, RME device, insert phones. This way one can easily switch between different hardware configurations when necessary without having to recalibrate by selecting a default hardware profile.

### Experiment configuration

Figure 3.3 shows the experiment configuration screen. Two different stimulus types can be chosen. Firstly, SOMA can generate stimuli online. All parameter and calibration data stored in a `.m`-file from Section 3.2.2 are loaded. Only initial signal component intensities still have to be selected before starting the experiment. Secondly, one can choose to use custom made stimuli from any 64 bit `.sbin` source file on disk (left and/or right). If the corresponding calibration factor is not present in the calibration file constructed in Section 3.2.2, a warning will be given and application of the custom stimulus will not be possible. A separate stimulus calibration file, identical to the one used for stimulus generation, has to be constructed beforehand to indicate which frequencies have to be examined in the online analysis. Here, other parameters in the configuration file are discarded. All intensities are in dB SPL when properly calibrated. The appropriate EEG calibration can be selected. Other parameters that can be set are the number of recorded EEG channels (up to eight, the maximum number of RME soundcard input channels), the artifact rejection level (in nanoVolt), the possibility to change carrier intensities online during measurement, and the number of sweeps to be recorded. Patient information can be entered.



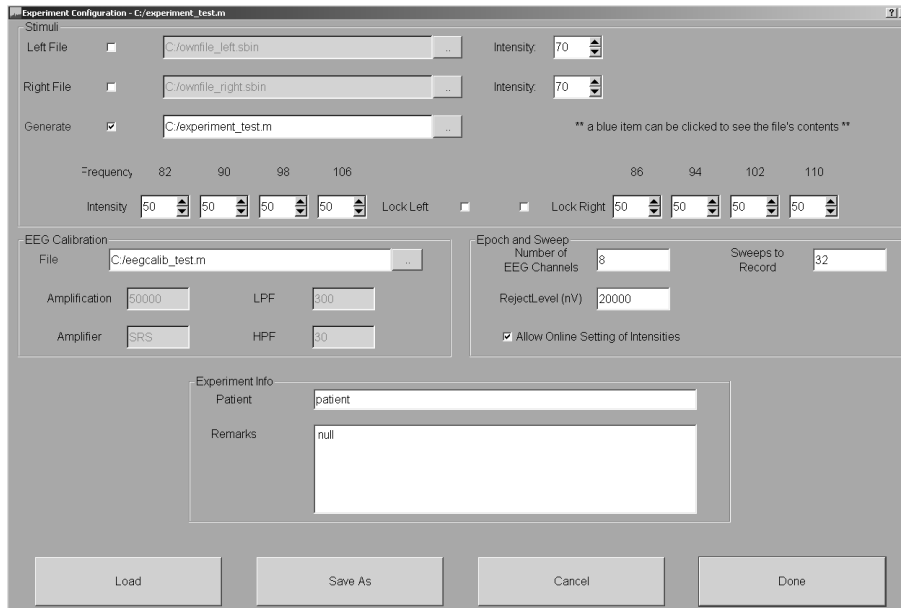


Figure 3.3: The experiment configuration screen allows the user to select the used stimuli, their intensity, calibration and other relevant parameters, as well as to enter subject information.

### Measurement and real-time analysis

During the experiment, visual information is returned by the SOMA software in four windows.

Figure 3.4 displays the most recently recorded EEG epoch. An epoch is a data block unit of 1.024 seconds, as defined in John and Picton (2000a). Epochs that do not meet the artifact rejection criterion are rejected. Additional information about the amplifier is given. The monitored EEG channel can be selected. An option is included to view all EEG channels at once.

Figure 3.5 shows the frequency spectrum of the average of all recorded sweeps. A sweep consists of a predefined number of accepted epochs, e.g. 16, which defines a sweep at 16.384 seconds. This window provides a visual representation, with the modulation frequencies of interest, next to the surrounding (noise) frequencies. The displayed EEG channel can be selected. A RMS noise value is calculated in a certain frequency range to make noise level monitoring possible.

Figure 3.6 shows a frame with statistical information. Analysis measures are displayed for each applied modulated carrier. The modulated carriers that are

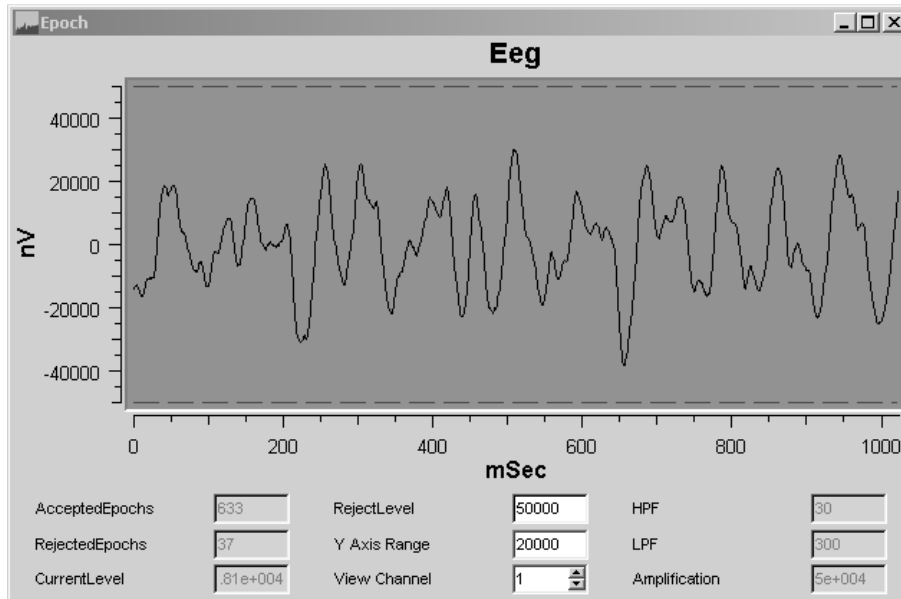


Figure 3.4: The most recently collected EEG epoch. An epoch not meeting the artifact rejection criterion is rejected for further analysis. The monitored EEG channel can be selected. Additional information about the amplifier is given.

examined are the same carriers that are used in the calibration file for stimulus generation from Section 3.2.2. These analyses include phase and amplitude of the modulation frequency, the RMS amplitude of the surrounding noise bins, and their corresponding SNR value. The  $p$ -values on the eighth row are the most important figures that determine the statistical presence of a response.  $p$ -values below a certain criterium (generally  $p = 0.05$  for a standard single channel measurement) are considered to indicate that an auditory steady-state response at that specific modulation frequency is present. An interesting feature makes it possible to alter intensities of modulated carriers independently while running the experiment. For this reason, a separate EEG buffer for each carrier is implemented. This way the intensity of a single modulated carrier can be altered without influencing other carriers. The upper row of Figure 3.6 shows eight spin boxes, each one corresponding to the intensity of its modulated carrier. Once the spin box of a modulated carrier is modified (e.g. when significance has been reached and the user changes the spin box in response), the carrier's intensity changes the moment the next sweep starts. The buffer of the specific modified carrier is cleared and a new EEG data stream starts filling the buffer. EEG data of previous intensities is not used for the calculation of the spectrum or statistical information. The EEG buffers of the other carriers are kept unchanged. The bottom row of Figure 3.6 indicates the number

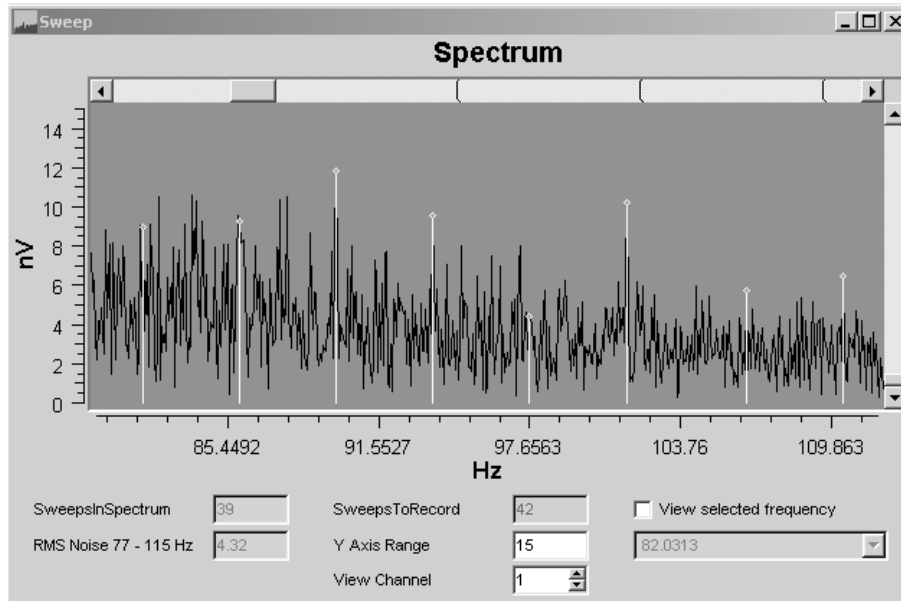


Figure 3.5: The spectrum of the average of all measured sweeps. Highlighted frequencies correspond to the modulation frequencies of interest. These frequencies can be observed in higher detail. The displayed EEG channel can be chosen. The RMS noise level between a certain predefined frequency range makes it possible to monitor the amount of noise in the current experiment.

of sweeps that have been averaged for a specific intensity. The moment the intensity of the modulated carrier changes, its value is reset to zero.

The fourth window functions as the main screen (not displayed here). Besides the basic navigational and recording functions it is possible to replay experiments conducted earlier, to reset or redo a current experiment and to show a  $p$ -value tracker over time.

### 3.2.3 Internal structure

The SOMA program is written in C++, with focus on object-oriented design, modularity and testability. Three frameworks are used: the commercial *Qt* library ([www.trolltech.com](http://www.trolltech.com)) for the main graphical interface, the *Qwt* library ([qwt.sourceforge.net](http://qwt.sourceforge.net)) for the plots, and the internally developed *streamapp* library which provides a general event handling interface, multi-threading and hardware (soundcard, file) abstraction.

Left					Right			
60	60	60	60	Intensity( dB SPL )	60	60	60	60
500	1000	2000	4000	Carrier( Hz )	500	1000	2000	4000
82	90	98	106	Modulation( Hz )	86	94	102	110
105	95.3	191	145	Phase( o )	321	109	58.6	39.2
8.96	11.8	4.38	5.71	Amplitude( nV )	9.24	9.54	10.2	6.46
5.76	4.56	3.65	2.89	Noise( nV )	5.36	4.05	3.35	2.74
3.86	3.29	1.6	5.91	SNR( dB )	4.7	7.44	9.7	7.47
<b>0.0902</b>	<b>0.00142</b>	<b>0.238</b>	<b>0.0216</b>	P-Value	<b>0.0543</b>	<b>0.00443</b>	<b>0.000124</b>	<b>0.00426</b>
39	39	39	39	Sweeps in Average	39	39	39	39

Figure 3.6: Statistical information of the modulation frequencies of interest. If a  $p$ -value is smaller than 0.05 in a standard single channel measurement, a response is decided to be significant. The intensities of the modulated carrier frequencies can be modified while running an experiment. The bottom row indicates the number of sweeps that have been collected thus far for the corresponding modulated carrier frequencies. It resets to zero the moment the intensity of the individual carrier changes, independently from the intensity of other carriers.

The general design strategy is kept as simple as possible (Figure 3.7). All SOMA classes (called modules further on) derive from the *SomaModule* interface. This interface defines methods for event-based communication; an event consists of an event ID and event data. When a module wants to communicate, it sends an event to the router. The router looks up the receivers for the given event ID and delivers the event to them. This modular approach allows for easy extensibility. A new module only has to be registered with the router for the events it wants to receive. It can send any event to the router without having to know the modules that will receive the event.

The signal flow can be described as follows. The graphical interface (GUI modules), soundcard and threads are initialized on application startup using settings from the main configuration file. When an experiment has been configured, all modules are initialized. During the experiment, the *soundcard* module plays stimuli through the soundcard, while recording the EEG signal. The recorded signal is downsampled and placed in a buffer. When the buffer is full, the data is sent to the epoch module. This module uses a processor to amplify the collected epoch (sent to the epoch GUI for display) and to decide whether the epoch is accepted or not. An accepted epoch is placed in an epoch buffer. When this epoch buffer is full, the data is sent to the sweep module, which calculates the spectrum and statistical parameters. Results are sent to sweep and statistical GUIs for display. Additional modules save all data and user in-

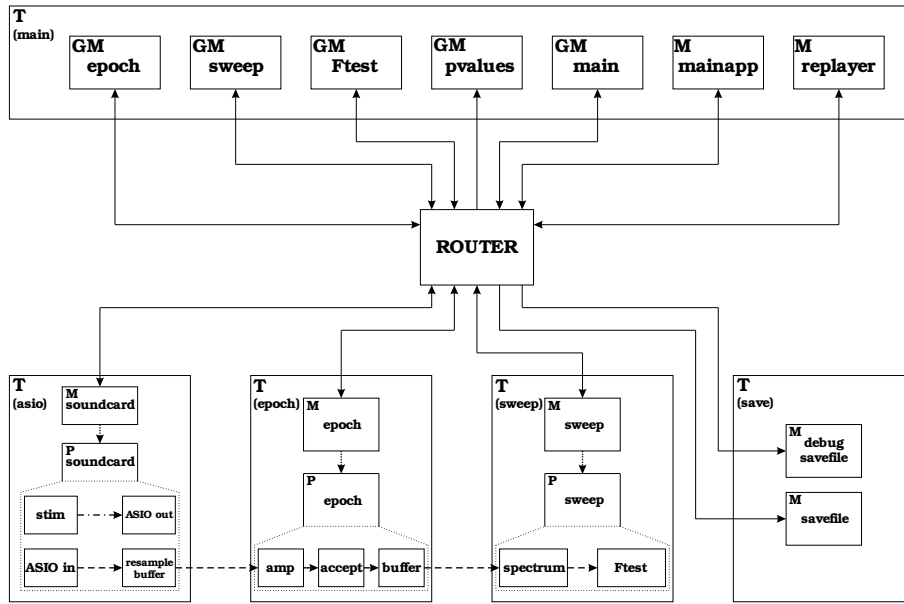


Figure 3.7: Overview of the internal software structure. All modules are depicted together with the threads executing their code. The actual (EEG) signal flow is displayed using dashed lines; the dash-dotted line correspond to audio signals. Full lines show event communication between modules. A dotted line indicates that a module 'uses' its created processor. Letter codes are used for the building blocks: T (thread), M (module), GM (GUI module), P (processor or input/output stream). The epoch GUI corresponds to Figure 3.4, the sweep GUI to Figure 3.5, the Ftest GUI to Figure 3.6 and the main GUI to the not-displayed main window. The pvalues GUI shows the  $p$ -value as a function of the number of processed sweeps, when activated.

teractions (e.g. changed intensities and epoch rejection levels), allowing offline analysis. SOMA can load a recorded experiment and review it one epoch at a time. This makes it possible to fine-tune results and signifies a great aid in in-depth analysis.

In combination with the buffers, the organization in five threads assures that the audio thread runs in parallel with the processing load from epoch, sweep and save threads, or the screen updates in the main thread, without interference. This setup ensures that no collected samples are lost during the switching of the buffers, calculation of the statistics or updating of the display.

### 3.2.4 Evaluation

A single channel study has been set up to evaluate SOMA. Hearing thresholds estimated with SOMA were compared with those estimated using the MASTER software research version (John and Picton, 2000a), installed on a PC in combination with a National Instruments NIDAQ PCI-card and a clinical audiometer Madsen Orbiter 922. This setup may be considered as a gold standard. Based on this study the practical use and benefits of the SOMA program in combination with the RME soundcard are evaluated.

Nine normal hearing subjects (age range 20–29 years) participated in the study. Behavioral thresholds were determined at 0.5, 1, 2 and 4 kHz with a 5 up–10 down Hughson–Westlake method using a clinical audiometer Madsen Orbiter 922 and identical ASSR stimuli as in the actual test. The MASTER–software provided the calibrated stimuli as an input for the audiometer. The behavioral thresholds were less than or equal to 20 dBHL. All experiments were carried out in a double–walled soundproof room with Faraday–cage. Subjects were asked to lie down on a bed with eyes closed and to relax or sleep.

Two combined stimuli with four 100% amplitude modulated (AM) & 20% frequency modulated (FM) carrier frequencies each, were applied to each ear. The carrier frequencies were the same for both ears, namely 0.5, 1, 2 and 4 kHz. The modulation frequencies were taken close to respectively 82, 90, 98 and 106 Hz for the left ear, and 86, 94, 102 and 110 Hz for the right ear. The stimuli were applied to Etymotic Research ER–3A insert phones for acoustic subject stimulation. The eight separate signals were calibrated at 70 dB SPL, using a Brüel & Kjær Sound Level Meter 2260 in combination with a 2-cc coupler DB0138 and an artificial ear 4152. Stimuli were applied at an intensity of 50 dB SPL and lowered by 10 dB after a trial of approximately 10 minutes per intensity (32 sweeps, each sweep corresponding to 16.384 seconds). If no response was found at 50 dB SPL, one additional trial was carried out at 60 dB SPL. MASTER and SOMA measurements were conducted in the same session and were randomized to avoid influence of test order.

3M Red Dot electrodes were placed on the Cz–Pz–Oz positions (active, ground, reference electrode) according to the international 10–20 system from Figure 2.2 (Malmivuo and Plonsey, 1995). They were placed on the subject’s scalp after the skin was abraded with Nuprep abrasive skin prepping gel. A conductive paste was used to keep the electrodes in place and to avoid that inter–electrode impedances exceeded 5 k $\Omega$  at 30 Hz. The electrodes were connected to a low–noise Stanford Research Systems SR560 amplifier. The single EEG channel was amplified ( $\times$  50,000) and bandpass filtered between 30 and 300 Hz (6 dB/octave). The sampling rate was set equal to 1000 Hz for SOMA and 1250 Hz for MASTER. Artifact rejection was set initially at 20  $\mu$ V, so that approximately 5% of the recorded data was discarded. The acquired signals were divided into data blocks of 16.384 seconds (one sweep) and normally averaged.

An FFT analysis was carried out and a response was considered present if the F-ratio statistic of (2.11) showed a significant difference (F-test with 2 and  $2N$  degrees of freedom,  $p < 0.05$ ) between the response power and the mean noise power of  $N = 120$  neighboring frequency bins, approximately 3.7 Hz (60 bins) at each side (John and Picton, 2000a).

### 3.3 Results

#### 3.3.1 Amplitude and noise levels

Figure 3.8 shows the response and noise amplitudes per carrier frequency averaged over all nine subjects for a stimulus intensity of 40 dB SPL for measurements of 32 sweeps. The noise amplitude per carrier frequency is estimated around its corresponding modulation frequency using the RMS value of  $N = 120$  neighboring frequency bins, approximately 3.7 Hz (60 bins) at each side. Response and noise values of the same carrier frequency (left and right) are taken together for analysis, e.g. 1 kHz in the figure shows both response and noise values of modulation frequencies 90 and 94 Hz. Lower stimulus intensities show a similar behavior.

To estimate the general noise level during measurement, complementary to the calculation of the noise levels around each modulation frequency conducted above, RMS noise values over a greater frequency range (70 – 110 Hz) and over all completely available intensities (10 – 40 dB SPL) were calculated for 32 sweep measurements. Their means and standard deviations are 5.06 (1.68) nV and 4.99 (1.82) nV for the MASTER and SOMA setup respectively, showing no significant difference using a paired samples t-test,  $p > 0.05$ . Signal-to-noise ratios do not show significant differences using a paired samples t-test for each separate frequency between the two methods,  $p > 0.05$ .

#### 3.3.2 Hearing threshold difference scores

Table 3.1 presents the mean differences between the measured ASSR thresholds and the corresponding behavioral thresholds. A paired samples t-test does not show any significant differences for each separate frequency between the two methods,  $p > 0.05$ .

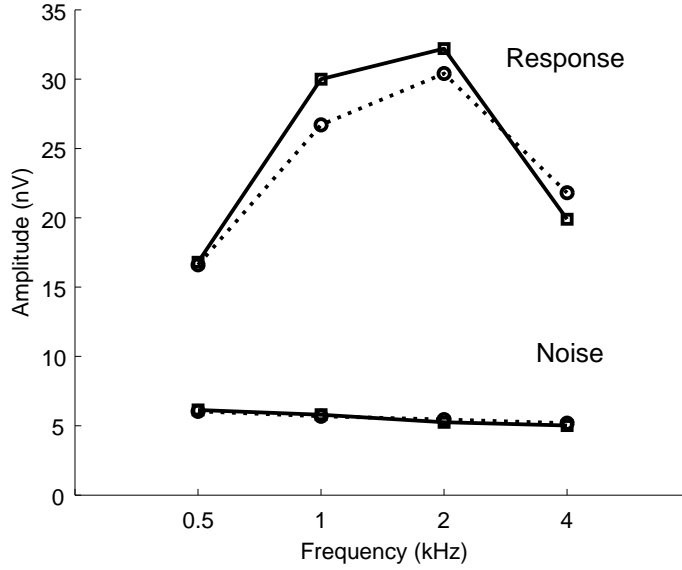


Figure 3.8: Mean ASSR and noise amplitudes (nV) for 40 dB SPL data of nine subjects for measurements of 32 sweeps. MASTER: solid line & square marker. SOMA: dotted line & circle marker. The noise amplitude per frequency is estimated around its corresponding modulation frequency using the RMS value of  $N = 120$  neighboring frequency bins, approximately 3.7 Hz (60 bins) at each side. Response and noise values of the same carrier frequency (left and right) are taken together for analysis, e.g. 1 kHz in the figure shows both response and noise values of modulation frequencies 90 and 94 Hz. Error bars are omitted for clarity. Standard deviations (nV) are 11.9, 16.5, 16.1, 8.7 and 8.9, 11.9, 14.3, 9.1 for MASTER and SOMA response amplitudes respectively. For noise amplitudes the standard deviations (nV) are 2.5, 2.6, 2.6, 2.6 and 2.5, 2.6, 2.7, 2.9 for MASTER and SOMA respectively.

### 3.4 Discussion

This chapter focusses on the development of a setup for multichannel auditory steady-state response measurements, called SOMA – Setup ORL for Multichannel ASSR. Both the hardware and software implementation are described, together with an evaluation study on nine normal-hearing subjects.

The urge for the development of an alternative research setup, next to the currently available commercial devices, was created by the need for a mobile, multichannel, inexpensive, flexible and modular extensible research platform, capable of processing custom made stimuli and changing carrier frequency intensities on the fly. Most commercial devices do not provide more than two-



channel EEG measurements. The possibility to read in own stimuli or to alter stimulus and recording parameters is very limited or non-existent. Adding extra functionalities without referring to the manufacturer is not possible. If research is the main purpose, commercial devices often do not fulfill the researcher's expectations although these devices are highly suitable in their respect for use in clinical environments.

SOMA uses a standard high-quality multichannel RME soundcard. This soundcard based choice is motivated by the fact that the RME soundcard is relatively inexpensive compared to standard multichannel data acquisition cards, by the use of several RME devices in the author's research environment, the available experience built up through this presence and the availability of C++ software code for the device. A portable SOMA prototype setup using a RME soundcard was built up in a very short time in combination with multichannel recording software, like e.g. Adobe Audition. The prototype showed that stimulus presentation over a wide dynamic range needed for ASSR measurements was possible using the RME soundcard and that it was capable of recording multichannel amplified EEG potentials. The sufficient dynamic range makes it possible to avoid the use of attenuators or audiometers which keeps the setup portable.

The most important advantages of the setup are the use of multiple EEG channels, the use of custom stimuli and the flexibility for research studies. Two-channel measurements are available in the most recent ASSR devices of e.g. Natus, Interacoustics, Grason-Statler Inc. or Intelligent Hearing Systems. The use of more than two channels however, is not available commercially. Some devices are more flexible than others, but in general the possibilities to change experiment stimuli or parameters is limited. Minor advantages are the portability, the possibility to change independent intensities on the fly (also implemented by Interacoustics) and the modular extensibility, which makes the user independent from a commercial supplier. The biggest limitation of the SOMA setup is the intensity range of 10 to 100 dB SPL (90 dB dynamic range). This seems rather low compared to other devices, but it is considered sufficient for hearing threshold determination for both normal-hearing as severely hearing impaired subjects.

Table 3.1 can be compared to studies with other multiple frequency ASSR thresholds in normal hearing subjects (Dimitrijevic et al., 2002; Herdman and Stapells, 2001; Luts and Wouters, 2005; Perez-Abalo et al., 2001), that use the same MASTER-system as in this study, except for Perez-Abalo et al. (2001) who use the Audix system from Neuronix S.A. Standard deviations of the difference scores are comparable to the current data. The difference scores are lower than those achieved by MASTER and SOMA in the present study, except for the data in Luts and Wouters (2005), which are similar. Reasons for the lower difference scores in the other three studies can be found in elevated hearing thresholds or longer test durations (Luts and Wouters, 2005).

Table 3.1: Mean difference scores and standard deviations of the ASSR threshold (measured at 10 dB precision) and the corresponding behavioral threshold (measured at 5 dB precision) for 18 normal-hearing ears. Intensities of 50, 40, 30, 20 and 10 dB SPL have been applied in experiments of 32 sweeps each (approximately 10 minutes). In some cases, a 60 dB SPL intensity has been used to confirm a threshold.

Carrier (Hz)	500	1000	2000	4000	All
MASTER (dB)	22 ± 11	15 ± 8	14 ± 7	17 ± 8	17 ± 9
SOMA (dB)	24 ± 9	19 ± 9	15 ± 8	16 ± 9	18 ± 9

### 3.5 Conclusions

Section 3.1 describes the general background and the motivation for building the setup. Commercially available setups do not provide the specifications required for the studies in this thesis. Therefore it is decided to build a custom setup. In Section 3.2, both hardware and software of the setup are described. SOMA makes use of an inexpensive multichannel soundcard that is controlled by software written in C++. Section 3.3 presents the results of an evaluation study with nine normal-hearing subjects that shows that no significant performance differences exist between the proposed platform and a reference platform (the MASTER platform from John and Picton (2000a)). These observations are discussed in Section 3.4.

It can be concluded that the SOMA program, in combination with the RME multichannel soundcard, can be used to assess ASSR hearing thresholds reliably. SOMA presents a flexible and modularly extensible high-end mobile ASSR test platform, that allows multichannel measurements, the use of own stimuli and independent intensity changes. It is not restricted by the limitations of commercial software and is thus suited for research and several clinically diagnostic purposes. Although the validation study was conducted using only single channel measurements, an extension to multiple channels will not compromise these conclusions.

In the following chapters the SOMA setup is used to apply generated and custom made stimuli and to record multichannel EEG. This multichannel EEG is processed using different multichannel signal processing techniques.

## Chapter 4

# Improving ASSR Detection Using Independent Component Analysis

Section 2.1 showed that auditory steady-state responses are a reliable assessment technique for hearing threshold estimation. Unfortunately, ASSR measurements can last a long time. To reduce recording time to about 45 to 60 minutes, Section 2.2 offered several solutions. Although being effective in most of the cases, these methods were not developed with multichannel EEG measurements in mind. The following chapters will address the processing of multichannel EEG data. This way, measurement times can be reduced significantly and robustness against artifacts can be increased.

The first multichannel processing technique discussed in this thesis is *independent component analysis* (ICA). After a short introduction in Section 4.1, Section 4.2 introduces the theory of ICA, its assumed ASSR model, the experimental setup and the used evaluation method. In Section 4.3, the theory is applied to real-life EEG data. Firstly, the available seven-channel data of eight normal-hearing subjects are processed based on ICA (using the JADE algorithm) and results are compared to those from the most common single channel ASSR technique (John and Picton, 2000a). It will be shown that ICA significantly improves detection for measurements between 30 and 60 dB SPL. Secondly, the optimal number of input channels, the optimal electrode positions

---

The material presented in this chapter has been published in ‘Van Dun, B., Wouters, J., and Moonen, M. (2007), “Improving auditory steady-state response detection using independent component analysis on multi-channel EEG data,” *IEEE Trans. Biomed. Eng.*, 54(7), 1220–1230’.

and the optimal number of independent components are reported. Thirdly, by fixing the separating matrix  $\mathbf{W}$ , calculated from data with a high SNR, a performance improvement may be expected. This assumption is evaluated. Fourthly, the performance of an ICA-based procedure applied to single channel data is considered. Finally, a combination of previous techniques is presented. Section 4.4 discusses these results. Section 4.5 ends this chapter with some conclusions.

## 4.1 Introduction

Over the last decade, independent component analysis (Hyvärinen et al., 2001) has appeared as a powerful signal analysis tool for a variety of industrial, medical and even financial applications (Back and Weigend, 1997; Bounkong et al., 2003; De Lathauwer et al., 2000; Fang and De-Shuang, 2005). ICA allows finding the underlying factors from multivariate statistical data by looking for components that are both statistically independent, and non-Gaussian. In this chapter, the possible use of ICA in ASSR detection is investigated.

## 4.2 Methods

This section describes the methods used in this study. In Section 4.2.1, the theory of independent component analysis is presented. When ICA is applied to the recorded multichannel EEG containing ASSRs, a model needs to be assumed (Section 4.2.2). Sections 4.2.3 and 4.2.4 describe the used setup and stimuli for the evaluation study. The single channel reference method and the way the multichannel EEG is processed using ICA is described in Section 4.2.5. Finally, the used performance evaluation procedure is described in Section 4.2.6.

### 4.2.1 Independent component analysis

Independent component analysis (ICA) is a blind source separation technique that is used to find a latent structure underneath a set of observations (Comon, 1994; Hyvärinen et al., 2001). This underlying structure comes in the form of unknown sources or independent components (ICs). The ICA general model is

$$\mathbf{X} = \mathbf{f}(\theta, \mathbf{S}) \quad (4.1)$$

with  $\mathbf{X}^T = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_m]$  a matrix with  $m$  observations  $\mathbf{x}_i$  and  $\mathbf{f}$  an unknown function with parameters  $\theta$  that operates on statistically independent underlying variables  $\mathbf{S}^T = [\mathbf{s}_1 \ \mathbf{s}_2 \ \dots \ \mathbf{s}_q]$  with  $q \leq m$ . If  $\mathbf{f}$  is a linear function, a special

case of the above equation is obtained, namely

$$\mathbf{X} = \mathbf{A}\mathbf{S} \quad (4.2)$$

with  $\mathbf{A}$  an  $m \times q$  mixing matrix. The pseudoinverse of  $\mathbf{A}$  is defined as the separating matrix  $\mathbf{W}$ ,

$$\mathbf{W} = \mathbf{A}^+ \quad (4.3)$$

This formula states that each of the observations  $\mathbf{x}_i$  is a linear combination of a set of  $q$  underlying ICs  $\mathbf{s}_j$ :

$$\mathbf{x}_i = a_{i1}\mathbf{s}_1 + \dots + a_{iq}\mathbf{s}_q \quad \text{for } i = 1 \dots m \quad (4.4)$$

ICA-algorithms estimate both  $\mathbf{S}$  and  $\mathbf{A}$ .

The most important assumption of ICA is that the components, linearly combining into observations, are mutually independent of each other. The fundamental problem is how to assess the independence of the components. The more common approach assumes the distributions of the ICs to be as far from normal Gaussian as possible. This idea is fed by the inverse of the central limit theorem, which states that the distribution of a sum of independent variables shifts to a normal (Gaussian) distribution when the number of variables goes to infinity (Trotter, 1959). To make this approach practically usable, different approximate measures of non-Gaussianity have been developed. By maximizing such a measure, a matrix  $\mathbf{S}$  can be constructed numerically. The ICA-algorithm that has been used in the rest of this study is the *joint approximate diagonalization of eigenmatrices* (JADE) algorithm (Cardoso and Soloumiac, 1993). This algorithm has advantages over e.g. FastICA (Hyvärinen, 1999) that practically suffers much more from local optima, leading to the calculation of different ICs and separating matrices  $\mathbf{W}$  when the algorithm is run several times over the same dataset. However, alternative algorithms could have returned a higher performance than JADE, like the recent MILCA algorithm (Stögbauer et al., 2004). It is possible to tailor the ICA-algorithm to specific needs of the problem using the Bayes theorem (Knuth, 1999). An extension of ICA to underdetermined mixtures is also a possible approach (Comon, 2004; Deville et al., 2004). At the time of this research, JADE was considered a proven technique that was tested on several applications, while the new methods above did not have those benefits that much.

### 4.2.2 Assumed model

In Section 4.2.3, a seven-channel setup will be described that records linear combinations of an unknown number  $Q$  of latent sources. When eight ASSRs are present one can assume eight latent sources to be ASSRs, as these sources are assumed to be independent from each other. This assumption is not correct anatomically when intensity, number of responses, carrier- and modulation frequency are varied (Lins and Picton, 1995). However, when all these parameters

are kept fixed, the assumption holds as each modulation frequency excites a different part of the brainstem (Picton et al., 2003). This assumption is also supported by our own experience that ICA-application on real EEG data shows that not all ASSRs are projected on one single independent component.

It is assumed in this model that each ASSR is generated by only one source. If not, the number of ASSR sources will be larger than eight. This however does not impact on the model and conclusions. Condensing this in a generative model, one obtains  $\mathbf{X} = \mathbf{B}\mathbf{S} + \mathbf{N}$ :

$$\begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_7 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1Q} \\ \vdots & \vdots & \ddots & \vdots \\ b_{71} & b_{72} & \dots & b_{7Q} \end{bmatrix} \begin{bmatrix} \mathbf{s}_{\text{ASSR}_1} \\ \mathbf{s}_{\text{ASSR}_2} \\ \vdots \\ \mathbf{s}_{\text{ASSR}_7} \\ \mathbf{s}_{\text{ASSR}_8} \\ \mathbf{s}_{\text{noise}_1} \\ \vdots \\ \mathbf{s}_{\text{noise}_{Q-8}} \end{bmatrix} + \begin{bmatrix} \mathbf{n}_1 \\ \vdots \\ \mathbf{n}_7 \end{bmatrix} \quad (4.5)$$

where  $\mathbf{s}_{\text{ASSR}_i}$  ( $i = 1, \dots, 8$ ) are ASSR sources,  $\mathbf{s}_{\text{noise}_j}$  ( $j = 1, \dots, Q - 8$ ) are muscle artifacts, eye blinks, brain processes, ... and  $\mathbf{n}_k$  ( $k = 1, \dots, 7$ ) are external noise sources like amplifier noise and e.g. line noise picked up by the electrode cables.

When observing (4.5), each row  $k$  in  $\mathbf{B}$  gives information about the SNR of a certain ASSR in the corresponding observation  $\mathbf{x}_k$ . Therefore, it is possible to look for the  $\mathbf{x}_k$  with the highest SNR for each ASSR source. After application of ICA ( $\hat{\mathbf{S}} = \mathbf{W}\mathbf{X}$ ),  $\mathbf{B}$  is replaced by  $\hat{\mathbf{B}} = \mathbf{W}\mathbf{B}$ . The simulations from Section 4.3 are expected to show that  $\hat{\mathbf{B}}$  returns a better SNR for an ASSR in certain components from  $\hat{\mathbf{S}}$  than  $\mathbf{B}$  for an ASSR in the observation matrix  $\mathbf{X}$ . It is likely that the ICA-technique can assess components  $\hat{\mathbf{S}}$  this way that will be more useful for detection than the original observations  $\mathbf{X}$ , based on the assumption that ASSR sources have a platykurtic distribution, while the EEG noise sources have a more mesokurtic one (close to Gaussian).

### 4.2.3 Experimental setup

The ASSR measurements were conducted in a sound-proof Faraday cage. The recording electrode placement can be found in Table 4.1, in accordance with the international 10-20 system of Figure 2.2 (Malmivuo and Plonsey, 1995). All seven active electrodes were referenced to the common electrode, which was placed on the forehead. The Kendall electrodes were placed on the subject's scalp after the skin was abraded with Nuprep abrasive skin prepping

Table 4.1: Recording electrode positions for a seven-channel setup. All channels are referenced to the common electrode on the forehead. The configuration for the single channel reference method is denoted using bold typeface.

Channel	Position
<b>1</b>	<b>Oz</b>
2	P4
3	P3
<b>4</b>	<b>Cz</b>
5	F4
6	F3
7	Pz
common <b>ground</b>	forehead <b>left mastoid</b>

gel. A conductive paste was used to keep the electrodes in place and to avoid that inter-electrode impedances exceeded  $5\text{ k}\Omega$  at 30 Hz. The electrodes were connected to a low-noise Jaeger-Toennies multichannel amplifier. Each EEG channel was amplified ( $\times 10,000$ ), bandpass filtered between 70 and 120 Hz (6 dB/octave) and finally software highpass filtered at 75 Hz (60 dB/octave). The amplified EEG signals were read using an RME Hammerfall DSP Multiface multichannel sound card and recorded using Adobe Audition 1.0 at a sampling rate of 32 kHz and downsampled to 250 Hz. Downsampling does not influence the performance of the ICA-algorithm, but greatly improves its efficiency. An alternative for downsampling could exist in the use of these extra samples to create additional artificial channels. All offline processing was performed using Matlab R14. The sound card was also used to generate the stimuli (see below). In the end, the stimulation and recording equipment used here is an older version of, but very similar to, the setup described in Chapter 3. An artifact rejection protocol was used, where all epochs (data blocks of 256 samples) greater than  $20\ \mu\text{V}$  in absolute value were rejected.

Eight normal-hearing volunteers (age range 22–33 years) participated in the study. Their behavioral hearing thresholds were less than 20 dB SPL for octave audiometric frequencies between 500 and 4000 Hz. Subjects were asked to lie down on a bed with eyes closed and to relax or sleep. Four trials with a length of 48 sweeps each (approximately 13 minutes) were conducted: at 60, 50, 40 and 30 dB SPL respectively. These intensities were chosen as the main goal of the study was to decrease the measurement duration, and not hearing threshold assessment in general. At the end of the session, behavioral thresholds were determined at 0.5, 1, 2 and 4 kHz with a 5 up–10 down method using modulated sinusoids.

#### 4.2.4 Stimuli

Two stimuli with four 100% amplitude modulated (AM) carrier frequencies each, were applied to each ear. The carrier frequencies were the same for both ears, namely 0.5, 1, 2 and 4 kHz. The modulation frequencies were taken close to respectively 82, 90, 98 and 106 Hz for the left ear, and 86, 94, 102 and 110 Hz for the right ear. To obtain an integer number of modulation frequency cycles in one data block of 256 samples (1.024 seconds), the previous values had to be corrected slightly (John and Picton, 2000a). This stimulus configuration was used with four out of eight subjects. The other four subjects received a reduced stimulus set, with only 0.5 & 4 kHz applied to the left ear and 1 & 2 kHz to the right ear. This difference in stimulus sets does not influence the following results.

Stimuli were created using Matlab R14 and played using Adobe Audition 1.0 at a sampling rate of 32 kHz. An RME Hammerfall DSP Multiface multichannel sound card sent the stimuli to Etymotic Research ER-3A insert phones for subject stimulation. The eight separate signals were calibrated at 70 dB SPL, using a Brüel & Kjær Sound Level Meter 2260 in combination with a 2-cc coupler DB138.

#### 4.2.5 Five ways to process the available EEG dataset

This subsection describes five different ways to process the available multichannel EEG. These methods are all evaluated in this chapter. The first method serves as the (standard) reference method all other methods are compared with. The third method is a combination of the first and the second method. The fifth method is a combination of methods 1, 2 and 4. It will be shown that these combinations of methods will increase performance compared with the individual methods.

##### Method 1: Single channel reference method

All ICA-processed EEG data in this chapter is compared with a gold standard, a single channel reference method that is most commonly used in clinical practice and described in literature (for an extensive review of single channel recording techniques, we refer to Picton et al. (2003)). For adults this single channel reference technique is the placement of the active and reference electrodes at Cz (inion) on top of the head and Oz (occiput) at the back of the head just above the base of the skull. The ground electrode position is not relevant.

According to Table 4.1, channel 1 (Oz) and channel 4 (Cz) are both referenced to the common electrode at the forehead. When channel 1 (Oz-forehead) is subtracted from channel 4 (Cz-forehead), a new EEG channel is obtained



between electrodes Cz and Oz. This new channel is defined as the reference channel for single channel measurements. The EEG data in channels 1 and 4 were artifact rejected ( $20 \mu V$ ) according to Section 4.2.3 before taking the difference of both channels.

Enough data was collected such that the data stream could be divided into  $N = 48$  sweeps. One sweep is exactly 16.384 seconds long (or  $M = 4096$  samples), based on parameters also used in John and Picton (2000a). Initial recording lengths were such that the division in 48 sweeps was always possible after artifact rejection. This implies that measurements with subjects generating a lot of artifacts are longer compared with measurements with subjects that do not produce that many artifacts. Data is averaged over  $N$  sweeps with equal weights (normal averaging) according to the steps below.

- The single channel EEG recordings are divided in 48 data blocks (sweeps) with a length of 4096 samples (16.384 seconds).
- Each sweep  $s_n$  ( $n = 1 \dots 48$ ) is averaged together with all preceding sweeps  $s_1$  to  $s_{n-1}$ . This creates 48 averaged sweeps  $\bar{s}_n$  ( $n = 1 \dots 48$ ) with a length of 4096 samples. Averaged sweep  $\bar{s}_1$  is identical to sweep  $s_1$ . Averaged sweep  $\bar{s}_{48}$  is an average of all sweeps  $s_1$  to  $s_{48}$ . This averaging step is necessary to increase the SNR of the ASSR to an acceptable level.
- For each processed (averaged) sweep  $\bar{s}_n$ , the F-ratio (SNR) of each modulation frequency is determined using (2.11). When 8 modulation frequencies are used, 8 F-ratios are calculated.

All ICA-processed results in this chapter (and the results obtained after MWF-processing in Chapter 5) are compared with this reference method.

### Method 2: Multichannel ICA

The sketchy algorithm below processes  $m$ -channel EEG data with a length of  $NM$  samples.  $N$  is the number of sweeps (data blocks) the data will be divided into. In this chapter the data is divided into  $N = 48$  sweeps. One sweep is 16.384 seconds long (or  $M = 4096$  samples), based on parameters also used in John and Picton (2000a). This  $m \times NM$  EEG data matrix is, after artifact rejection and an averaging step, processed using the ICA-technique from Section 4.2.1 to obtain  $q$  independent components (ICs). For clarity, the description will explicitly show the process using  $N = 48$  sweeps and  $M = 4096$  samples. The number of channels  $m$  and the number of ICs  $q$  are kept variable.

- The  $m$ -channel EEG recordings are divided in 48  $m$ -channel data blocks (sweeps) with a length of 4096 samples (16.384 seconds).

- Each  $m$ -channel sweep  $s_n$  ( $n = 1 \dots 48$ ) is averaged together with all preceding sweeps  $s_1$  to  $s_{n-1}$ . This creates 48 averaged  $m$ -channel sweeps  $\bar{s}_n$  ( $n = 1 \dots 48$ ) with a length of 4096 samples. Averaged sweep  $\bar{s}_1$  is identical to sweep  $s_1$ . Averaged sweep  $\bar{s}_{48}$  is an average of all sweeps  $s_1$  to  $s_{48}$ . This averaging step is necessary to increase the SNR of the ASSR to an acceptable level. Without averaging, the ICA-technique will fail, in contrast with the MWF-method from Chapter 5 that actually does not need prior averaging.
- The JADE algorithm takes one averaged  $m$ -channel sweep  $\bar{s}_n$  as an input.
- $q$  ICs are calculated based on the  $m$ -channel averaged sweep  $\bar{s}_n$  (with a length of 4096 samples). Each IC also has a length of 4096 samples.
- For each IC in a processed (averaged) sweep, the F-ratio (SNR) of each modulation frequency is determined using (2.11). When 8 modulation frequencies are used, 8 F-ratios are calculated. Consequently, when  $q$  ICs are available,  $8q$  F-ratios are calculated for each processed (averaged) sweep  $\bar{s}_n$ .
- For each modulation frequency within the same processed (averaged) sweep  $\bar{s}_n$ , the largest F-ratio is taken out of  $q$  calculated F-ratios.

This procedure is not used directly in this chapter, but is improved first as described by method 3.

### Method 3: Multichannel ICA combined with single channel reference method (method 1 + 2)

The first simulations using method 2 show a major variation over the different subjects, with several subjects performing worse than the single channel reference (method 1). The general improvement is marginal. To avoid this effect, methods 1 and 2 are combined. In particular, the best F-ratio out of  $q + 1$  F-ratios for each modulation frequency is taken:  $q$  F-ratios from the  $q$  ICs calculated using method 2 with  $m$  EEG-channels as an input and one extra F-ratio from the original single channel reference using method 1.

One should be aware that this combination of both the single- and multichannel approach does not necessarily ensure a better performance, compared to the single channel reference using method 1. As it is important to keep the specificity of the combined processing (method 3) equal to the specificity of the reference (method 1), the single channel data should truly be viewed as an extra channel, which raises the detection threshold accordingly due to the multiple testing aspect (Section 4.3.1).

The results obtained by applying this procedure on multichannel EEG data is described in Sections 4.3.1, 4.3.2 and 4.3.3.

#### Method 4: Single channel ICA combined with single channel reference (method 1)

Some research already has been conducted on the single channel case of independent component analysis (Davies and James, 2007; James and Lowe, 2003; Warner and Proudler, 2003). Single channel ICA contradicts the intuition that ICA is only suitable for processing of multichannel measurements. Recently, it has been suggested by Davies and James (2007) that single channel ICA (SCICA) can identify and separate sources successfully only when these sources have disjoint spectral support. However, if one-channel data is divided in different channels (say, two), it may still be worth the effort to check whether the algorithm may be able to extract two components with the following characteristics. The first component may be the ASSR of interest, which is presumed to be present in both (divided) channels, in the form of a sinusoid at a certain frequency. The second component may be the background EEG noise. One may argue that the non-stationary EEG noise varies too much over the two created channels. However, the chance that the statistical properties of the EEG noise from the first channel are totally different from those of the EEG noise of the second channel, is considered low.

A limitation of the ICA-technique is that the number of  $q$  ICs cannot exceed the number of  $m$  observations or channels. Otherwise the ICs are not identifiable because  $\mathbf{A}$  is not invertible. Therefore, no more than one IC can be estimated from single channel data, and then – quite trivially – this IC would be equal to the original data. To avoid this problem, the available data has to be split up to create  $m - 1$  extra artificial channels, so that  $m$  channels are available to calculate  $q$  ICs. Other approaches may also be used to create extra channels. First, when the available single channel EEG data is oversampled, extra channels can be created by downsampling the data multiple times, with a shift by an integer number of samples for each new downsampling step. Second, an interesting procedure is described in Davies and James (2007). Extra channels are created by shifting the available single EEG channel a couple of times by one sample each (thus creating a Toeplitz matrix with ‘EEG channels’). After application of a multichannel ICA-algorithm, the ICs are clustered with a clustering algorithm. Based on the acquired clusters, reconstruction filters are calculated that can be applied on the original single channel EEG data.

A schematic overview is shown in Figure 4.1 of the used single channel ICA-technique applied to a one-channel EEG data stream with a length of  $NM$  samples. This EEG data stream is the same (reference) channel (Cz-Oz) as used in method 1.  $N$  is the number of sweeps (data blocks) the data will be divided into, after artefact rejection ( $20 \mu V$ ). In this chapter the data is divided into  $N = 48$  sweeps. One sweep is 16.384 seconds long (or  $M = 4096$  samples), based on parameters also used in John and Picton (2000a). For the sake of simplicity, the description will explicitly show the process using  $N = 48$  sweeps and  $M = 4096$  samples per sweep. The number of channels  $m$  and

ICs  $q$  is taken equal to 2. This has been observed to be the optimal number of channels and ICs for the single channel case (Section 4.3.4). Other values for  $m$  and  $q$  degrade performance significantly.

The single channel procedure is as follows:

- *Step 1)* An  $48 \times 4096$  matrix is constructed from the data stream originating from the single channel recording system described in Section 4.2.3. No averaging has been carried out yet. This matrix is divided in  $m = 2$  parts ('channels' or 'observations'), by interleaving the odd and even sweeps, with  $\frac{N}{m} = 24$  sweeps per part, each sweep 4096 samples long. Part 1 contains sweeps  $s_1, s_3, \dots, s_{47}$  and Part 2 contains all even-numbered sweeps.
- *Step 2)* Sweeps  $s_1, s_3, \dots, s_{2k-1}$  from Part 1 are averaged and stored in sweep  $\bar{s}_k$  of an averaged matrix ( $k = 1, \dots, 24$ ). Sweeps  $s_2, s_4, \dots, s_{2k}$  from Part 2 are averaged and stored in sweep  $\bar{s}_{24+k}$  of that averaged matrix ( $k = 1, \dots, 24$ ). The resulting averaged matrix has the same structure as the matrix from Step 1.
- *Step 3)* Independent component analysis is performed 24 times, each time using a group of  $m = 2$  sweeps  $\bar{s}_k$  and  $\bar{s}_{24+k}$  from the matrix from Step 2 with  $k = 1, \dots, 24$ . Each element from this group of sweeps shares the property that the element is constructed by averaging  $k$  sweeps from the matrix from Step 1. For each  $k$ , the JADE algorithm takes in  $m = 2$  observations and returns  $q = 2$  ICs, which are linear combinations of sweeps  $\bar{s}_k$  and  $\bar{s}_{24+k}$ .
- *Step 4)* For each  $k$ , every IC from Step 3 is Fourier transformed and the F-ratio (SNR) is calculated of the modulation frequency of interest using (2.11). Because ICA does not order its components in any way (rows in  $\mathbf{A}$  can be permuted randomly), it is unknown which component has the highest F-ratio of a certain modulation frequency. For each modulation frequency, all F-ratios (SNR) from all  $q = 2$  ICs are calculated and the largest one is chosen.
- *Step 5)* After the largest F-ratio is chosen, a dimension problem arises. If only one F-ratio is calculated out of  $m = 2$  sweeps from an  $N = 48$ -sweep data stream, only  $\frac{N}{m} = 24$  F-ratios are available. To compensate for this, each IC has to be copied 2 times to make comparison possible with the single channel reference from method 1, which still contains 48 sweeps after averaging. As such, the resulting sweep from an ICA-operation on 2 sweeps replaces those 2 sweeps by the resulting sweep and its copy.

The first simulations using single channel ICA only show no performance increase. However, when the single channel ICA-approach is combined with

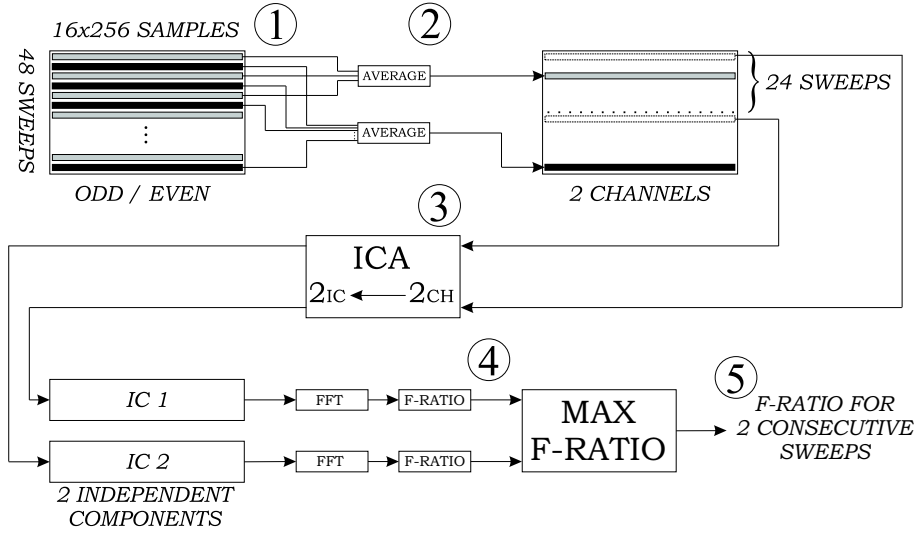


Figure 4.1: Data flow for ICA applied on single channel data. Artificial input channels are created by splitting the available data in  $m = 2$  different parts. These parts are used to calculate  $q = 2$  independent components.

method 1, there is some improvement (we refer to Section 4.3.4 for the results). In particular, the best F-ratio out of 3 F-ratios for each modulation frequency is taken: 2 F-ratios from the 2 ICs calculated using the method described in this subsection with the single channel reference (Cz-Oz) as an input and one extra F-ratio from the original single channel reference (Cz-Oz) using method 1.

The results obtained after application of this procedure are described in Section 4.3.4.

#### Method 5: Multichannel ICA combined with single channel reference method and single channel ICA (method 1 + 2 + 4)

The procedures from method 3 (which itself is a combination of methods 1 and 2, results described in Sections 4.3.1, 4.3.2 and 4.3.3) and method 4 (results in Section 4.3.4) are combined. The best F-ratio out of  $q + 3$  F-ratios for each modulation frequency is taken:  $q$  F-ratios from the  $q$  ICs calculated using method 2, one F-ratio from the original single channel reference using method 1, and two F-ratios obtained with method 4 (here, 2 ICs are extracted from single channel EEG data).

The results obtained by applying this procedure on multichannel EEG data is described in Section 4.3.5.

### 4.2.6 Performance measures

For evaluation of the proposed techniques, two methods have been chosen. The first one uses *receiver operating characteristic* (ROC) curves. It has the possibility to indicate statistically significant differences between techniques. However, ROC-curves do not give insight in the absolute benefit from one method compared to another one in terms of measurement time. Therefore, a second measure represents the amount of time that is needed to obtain a significant response. The drawback here is that the results do not have a statistical meaning and only give an indication for what is possible for an individual subject.

#### ROC-curves

In order to evaluate the above described techniques, receiver operating characteristic (ROC) curves were calculated from 8 subjects (Green and Swets, 1966; Hanley and McNeil, 1982). The curves were constructed as follows for a certain number of averaged sweeps:

1. Select 50 decision criteria  $p_i$ , represented by p-values that vary between 0.9 and  $10^{-15}$ .
2. For each decision criterium  $p_i$ , calculate the *sensitivity*  $\frac{TP}{TP+FN}$  and *specificity*  $\frac{TN}{FP+TN}$  over all measurements (4 intensities per subject), using 16 modulation frequencies of which 8 were used as control frequencies, as in these frequencies it was assured that only noise was present. A true-positive (TP) is a correct assessment from the algorithm that a response is present in reality. A false-negative (FN) is an uncorrect assessment that no response is present, while in reality a response is there. True-negatives (TN) and false-positives (FP) are the dual cases. All TP, TN, FP and FN are summed together over the same decision criterium  $p_i$  and used in the above equations.
3. The ROC-curve is built up plotting the ‘sensitivity’ as a function of ‘1-specificity’. Each  $p_i$  defines a new point of the curve.

After plotting all 50 points of the ROC-curve, the area below the curve is calculated. The area under the curve was used as a measure of detection accuracy. This procedure is carried out for each evaluated method. These calculations were carried out each time an additional sweep was collected and averaged with previous sweeps, so that the performance could also be analyzed on a time based scale.

To statistically compare different ROC-areas, a Z-test was carried out. The z-value can be calculated using (Hanley and McNeil, 1983):

$$z = \frac{A_1 - A_2}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2r\sigma_1\sigma_2}} \quad (4.6)$$

with  $A_1$  the ROC-area calculated for method 1 (the single channel reference) and  $A_2$  the ROC-area calculated for all other methods (single- or multichannel ICA). Here,  $\sigma_i$  is the standard deviation of  $A_i$  and  $r$  is the correlation coefficient between the data obtained from method 1 and the other methods. The  $\sigma_i$  are calculated using

$$\sigma_i = \sqrt{\frac{X_1 + X_2 + X_3}{n_A n_N}} \quad (4.7)$$

with

$$\begin{aligned} X_1 &= A_i(1 - A_i) \\ X_2 &= (n_A - 1)(Q_1 - A_i^2) \\ X_3 &= (n_N - 1)(Q_2 - A_i^2) \\ n_N &= TP + FN \\ n_A &= TN + FP \\ Q_1 &= \frac{A_i}{2 - A_i} \\ Q_2 &= \frac{2A_i^2}{1 + A_i} \end{aligned}$$

where  $n_A$  are all ‘abnormal’ (negative in reality) cases and  $n_N$  all ‘normal’ (positive in reality) cases. To find  $r$ , a value must be looked up in Table I of (Hanley and McNeil, 1983) using  $r_N$ ,  $r_A$ ,  $A_1$  and  $A_2$ . Here,  $r_N$  is the Pearson Product-moment correlation between all processed normal data ( $n_N$ ) from method 1 and the other methods. In the same way  $r_A$  is calculated with all abnormal data ( $n_A$ ). As an alternative,  $r$  can be taken as the mean of  $r_A$  and  $r_N$  which produces a maximum error on  $r$  of 10 %.

### Effective measurement time reduction

ROC-curves provide a theoretical means to assess different methods. To evaluate the practical applicability of a method, a measure for the effective benefit can be obtained by counting the number of processed sweeps until a response is first-detected. The difference between these numbers obtained with the two methods is an indication for a practical improvement or decline. In this study, we define a first-detection to be valid when the response is significantly present for 3 consecutive sweeps for methods 3 and 4, and 6 consecutive sweeps for method 5. In Chapter 5, this evaluation method is also used for an MWF-based approach. Here, 8 consecutive sweeps are required for detection. This quantity is based on the condition that there is no improvement allowed for noise frequencies when comparing methods 3, 4, 5 and the MWF-based approach from Chapter 5, with the single channel reference (method 1).

Due to the statistical multiple testing effect, the detection threshold (a  $p$ -value which is calculated by applying an  $F$ -test as described in Section 2.2.6 on the finally selected (maximum)  $F$ -ratio) needs to be made more stringent when adding more ICs for detection. The  $p$ -values for detection are the following:  $p = 0.050$  (method 1),  $p = 0.0065$  (method 3),  $p = 0.020$  (method 4),  $p = 0.0053$  (method 5), and  $p = 0.00025$  (MWF-based approach of Chapter 5). One needs to keep in mind this evaluation method is patient dependent and relies much on the used detection criterion.

### 4.3 Results

This section presents the results when applying ICA to the recorded single- and multichannel EEG. Section 4.3.1 will show the effect of extracting seven independent components out of seven EEG channels. Section 4.3.2 describes what happens when the number of EEG channels and the number of independent components are varied. Section 4.3.3 introduces the fixed separation matrix. These sections describe the results obtained using implementations of method 3, with method 1 used for comparison purposes. Section 4.3.4 shows ICA applied to only one EEG channel, an application of method 4. Finally, multichannel and single channel ICA are combined using method 5 in Section 4.3.5.

Figures 4.2, 4.4, 4.5, 4.6 and 4.7 show the area under the ROC-curve as a function of the number of averaged sweeps for different methods, compared with the single channel reference (method 1). The reference method is the standard single channel MASTER setup (John and Picton, 2000a) with artifact rejection at  $20\mu V$  and with the amplified difference between Cz and Oz as EEG signal. This setup is described in more detail in the first part of Section 4.2.5. A paired  $Z$ -test was carried out (Section 4.2.6) to compare the different techniques statistically using ROC-curves. The dotted lines in Figures 4.2, 4.6 and 4.7 denote two standard deviations of the ROC-areas. When they do not overlap, a significant difference is present (94.7 % significance interval, two standard deviations). Important for clinical application is the fact that, with the current sampling rate of 250 Hz, a real-time calculation during measurement is possible. Every 16.384 seconds, a new sweep is read. On a high-end PC (Pentium 4), the calculation time of the seven-channel ICA does not exceed five seconds. As a result, each new sweep can be downsampled, processed online and visualized before the next one is collected.



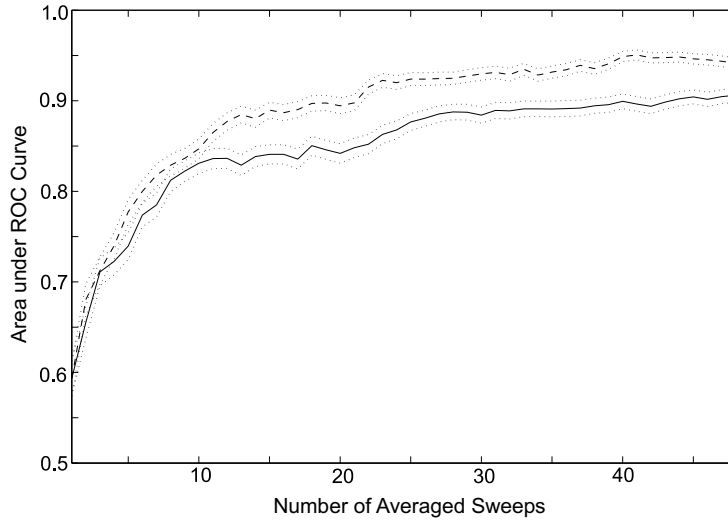


Figure 4.2: ‘area under ROC–curve’ versus ‘number of averaged sweeps’: Method 3 with  $q = 7$  independent components and  $m = 7$  channels (dashed); Method 1, single channel reference (solid). The dotted lines denote two standard deviations.

### 4.3.1 Multichannel ICA (method 3): seven channels and seven extracted components

#### ROC

Figure 4.2 shows the results of method 3 with  $q = 7$  independent components and  $m = 7$  channels. A significant difference (two standard deviations) between the single channel reference (method 1) and multichannel ICA–configuration (method 3) is present from sweep 11 on. The best performance of the single channel reference (method 1) is obtained after 48 sweeps of data collection. In contrast, a significantly better performance than the single channel reference (method 1) will ever achieve, is reached after 23 sweeps by the multichannel ICA–technique (method 3). This can be interpreted as a measurement time reduction of 52 %, in terms of ROC–area.

It is important to know that this improvement is only caused by the use of the ICA–technique, and not by the benefit of just using multiple channels instead of one. If the ICA–technique is omitted, for example by substituting the separating matrix  $\mathbf{W}$  by the identity matrix  $\mathbf{I}$ , the ROC–areas of methods 1 and 3 in Figure 4.2 will coincide. This condition represents similar performance, which of course needs to be avoided.

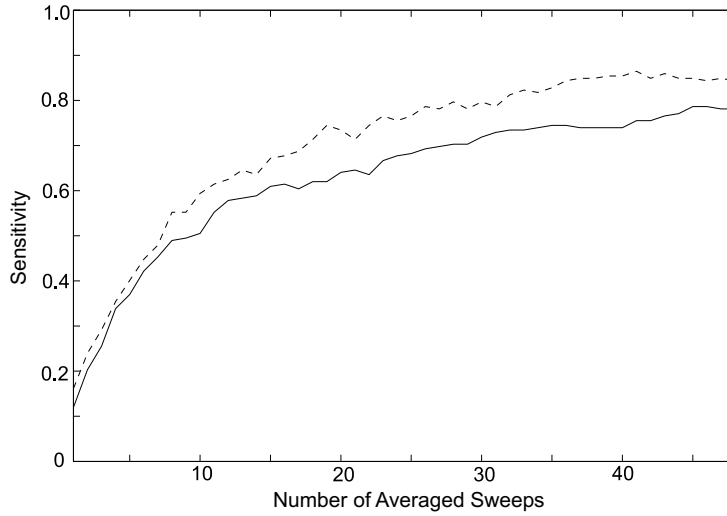


Figure 4.3: ‘sensitivity’ versus ‘number of averaged sweeps’: Method 3 with  $q = 7$  independent components and  $m = 7$  channels (dashed,  $p = 0.0065$ ); Method 1, single channel reference (solid,  $p = 0.050$ ). The specificity is taken equal to 95.0 %.

### Specificity and sensitivity

Because ROC-areas are a global evaluation using an integration over a range of decision criteria (p-values), a biased representation is possible. Therefore, the sensitivity and specificity is checked for a fixed p-value most used in literature. Using the single channel reference method, a p-value of 5 % corresponds to a specificity, or true-negative rate, of 95 %. This means that from 100 situations without a response, five will still be interpreted as if a response is present, due to noise influences (5% false-positive rate). If more channels are evaluated simultaneously, which is the case for the ICA-technique, the multiple testing aspect raises the number of false-positives if the p-value is kept constant. To avoid this, the decision criterion for response detection is made more stringent to achieve the same sensitivity. In the case of method 3, the effect of selecting the best out of 7 ICs and 1 reference channel forces the p-value to be lowered by a factor of 7.7 (from  $p = 0.050$  to  $p = 0.0065$ ). The reason for not applying a full Bonferroni correction (Bland and Altman, 1995) of a factor of eight can be explained by the fact that there is still some dependence left between the different ICs and the additional channel. This dependence causes the full correction by a factor of eight to be too extreme (Stürzebecher et al., 2005). Figure 4.3 shows the sensitivity for both methods at a specificity of 95 %, which provides a more realistic view of detection performance. The detection criterion is lowered to  $p = 0.0065$  for method 3, which still outperforms the reference method.

Table 4.2: Time reduction (in %) per subject for methods 3, 4, 5, and the MWF-approach of Chapter 5. Figures relative to the single channel reference (method 1). Conditions of detection are described in Section 4.2.6.

Subject	1	2	3	4	5	6	7	8	mean
method 3	8	-1	30	-2	13	8	63	19	17
method 4	-1	16	10	8	23	23	4	2	11
method 5	13	13	30	24	18	6	58	15	22
MWF (Chap 5)	9	4	12	5	2	-9	62	22	13

Table 4.3: Time reduction (in %) per intensity for methods 3, 4, 5, and the MWF-approach of Chapter 5. Figures relative to the single channel reference (method 1). Conditions of detection are described in Section 4.2.6.

Intensity (dB SPL)	30	40	50	60
method 3	25	8	22	32
method 4	7	4	19	29
method 5	16	15	41	45
MWF (Chap 5)	9	14	14	37

### Effective measurement time reduction

A comparison of method 3 with method 1, per subject, intensity and carrier frequency, is shown in Tables 4.2, 4.3 and 4.4. On average a mean detection time decrease of 17 % is obtained. The use of method 3 yields a major decrease of the detection time for one subject (63 %). For two subjects no improvement is obtained, taking into account that the standard deviation on the measurement time difference for noise frequencies is almost 5 % (with a mean of 0 %). This shows that these numbers vary considerably between subjects. They can be used to indicate some underlying trends. The data concerning intensity and carrier frequency dependencies show a possible correlation between detection time reduction and signal-to-noise ratio (SNR). Higher intensities and frequencies with physically larger responses (like 1 and 2 kHz carriers) are more prone to faster detection using ICA.

### 4.3.2 Multichannel ICA (method 3): variable number of channels and extracted components

Figure 4.4 shows the performance of a seven-channel system decomposed into  $q$  ICs using method 3. It is observed that performance is maximized when the number of ICs is taken as large as possible ( $q = 7$ ) and decreases when the

Table 4.4: Time reduction (in %) per carrier frequency for methods 3, 4, 5, and the MWF-approach of Chapter 5. Figures relative to the single channel reference (method 1). Conditions of detection are described in Section 4.2.6.

Carrier (Hz)	500	1000	2000	4000
method 3	15	25	31	12
method 4	8	9	3	27
method 5	10	33	31	25
MWF (Chap 5)	1	25	23	15

number of ICs goes down. The number of responses present in the data set does not influence this observation.

Figure 4.5 shows the obtained results as a function of the number of channels used in method 3. The order of Table 4.1 is respected to construct the figure, except for the two-channel case (usage of channel 1 and channel 4, corresponding to channel 4 minus channel 1 of the single channel reference method). An  $m$ -channel ICA thus uses the first  $m$  channels of Table 4.1. It can be observed that a saturation effect appears after 5 channels. There is no improvement when using a multichannel recording with more than 5 channels (or 7 electrodes). After permuting through all possible combinations of 5 channels in Table 4.1, the most optimal combination is obtained if 4 electrodes are located on the back of the head (Oz, P4, P3, Cz). According to the current dataset, the fifth electrode should be placed at F4. When other combinations are chosen, performance significantly degrades. Focussing on individual subjects in terms of effective measurement time reduction, only 2 out of 8 show a slightly (non-significantly) better performance when the contralateral electrode F3 is taken instead of F4. The same result is noticed for 1 out of 4 frequencies and for 1 out of 4 intensities. This does not imply an interaction between electrode position and intensity or frequency is present.

### 4.3.3 Multichannel ICA (method 3): effects of fixed separating matrix

This section illustrates the effects of keeping separating matrix  $\mathbf{W}$  fixed per subject or for all subjects. ICA is performed once and  $\mathbf{W}$  is calculated at highest intensity (60 dB SPL) and after a measurement of 48 sweeps. Afterwards no ICA is applied and the same  $\mathbf{W}$  is used for all other calculations of the ICs.

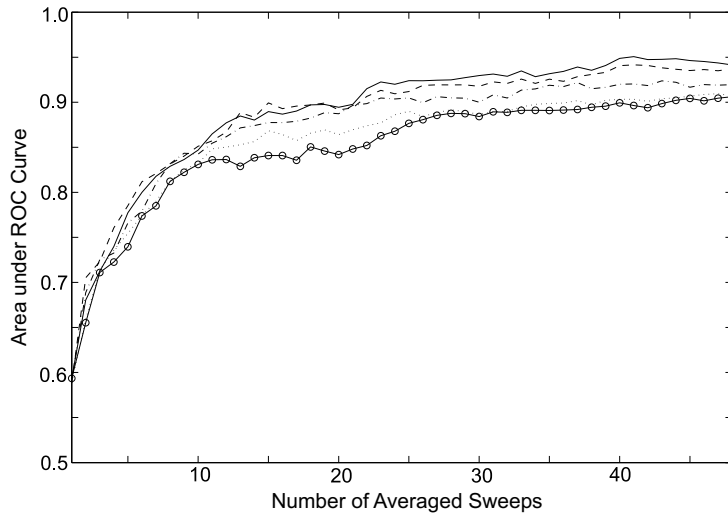


Figure 4.4: ‘area under ROC-curve’ versus ‘number of averaged sweeps’. Method 3 with  $m = 7$  channels and  $q$  independent components:  $q = 7$  (solid),  $q = 6$  (dashed),  $q = 4$  (dashdot),  $q = 2$  (dotted); Method 1, single channel reference (solid-circle).

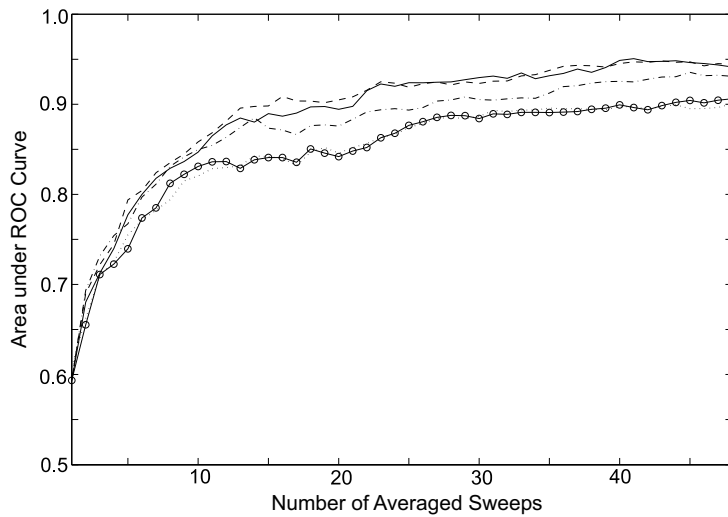


Figure 4.5: ‘area under ROC-curve’ versus ‘number of averaged sweeps’. Method 3 with  $m = q$  channels and  $m = q$  independent components:  $m = 7$  (solid),  $m = 5$  (dashed),  $m = 4$  (dashdot),  $m = 2$  (dotted); Method 1, single channel reference (solid-circle).

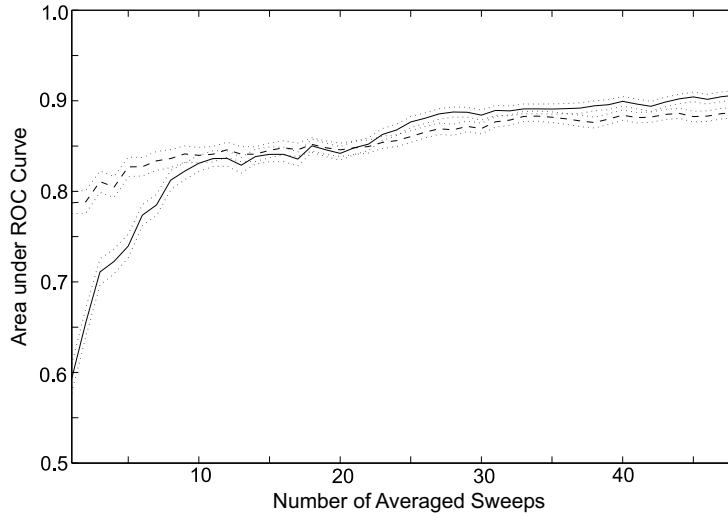


Figure 4.6: ‘area under ROC–curve’ versus ‘number of averaged sweeps’: Method 4 (dashed); Method 1, single channel reference (solid). The dotted lines denote two standard deviations.

#### Fixed separating matrix for all subjects

One fixed  $\mathbf{W}$  matrix (random subject, 60 dB SPL, 48 sweeps) is used to compute the ICs of all subjects. In general, no improvement is noticed compared to the single channel reference (method 1). However, for the dataset used here, some  $\mathbf{W}$  matrices (e.g.  $\mathbf{W}$  calculated from subject 3), return a significantly better result close to the performance of method 3 from Section 4.3.1. However, it is not possible to predict a priori an optimal  $\mathbf{W}$  matrix (or a subject that provides this matrix) that maximizes the performance of multichannel ICA.

#### Fixed separating matrix for one subject

The effect of keeping  $\mathbf{W}$  fixed for each subject separately is significant for some cases. However, performance is significantly worse than the results from Section 4.3.1.

#### Fixed IC for one subject

By additionally fixing the component with the highest SNR for a certain modulation frequency and for the same subject (together with a fixed  $\mathbf{W}$ ), the multiple testing aspect is avoided. The detection criterion for data processed with method 3 rose from  $p = 0.0065$  to  $p = 0.03$  (fixed IC with largest F–ratio together with the F–ratio of the single channel reference (method 1)) for

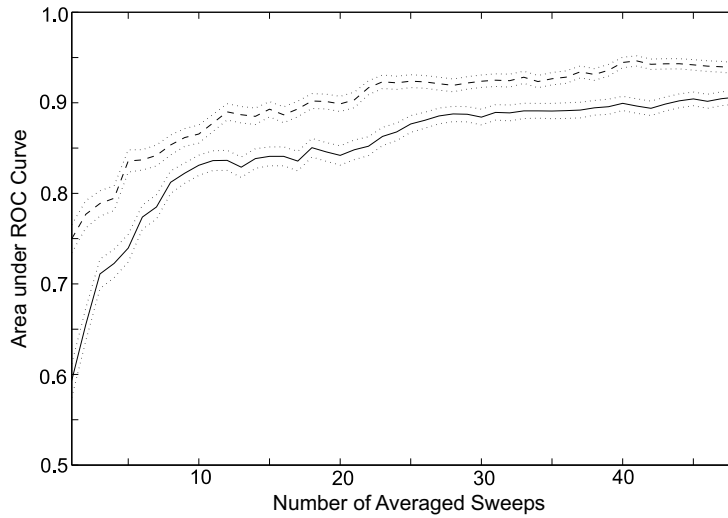


Figure 4.7: ‘area under ROC–curve’ versus ‘number of averaged sweeps’: Method 5 with  $q = 7$  independent components and  $m = 7$  channels (dashed); Method 1, single channel reference (solid). The dotted lines denote two standard deviations.

a specificity of 95.0 %. However, no significant gain compared to the single channel reference (method 1) is observed.

#### 4.3.4 Single channel ICA combined with method 1

Following only the single channel ICA–technique, no improvement compared to the single channel reference (method 1) is observed. However, when the F–ratios obtained with the single channel ICA–technique are combined with the F–ratio obtained from the single channel reference (method 1), a significant benefit is visible for a lower number of averaged sweeps. Further investigation shows that the use of ICA is not even necessary here. If the separating matrix  $\mathbf{W}$  is replaced by an identity matrix  $\mathbf{I}$ , results are almost identical. One can conclude that for a single channel setup it is sufficient to divide the available channel in two parts, add the single channel reference (method 1) and then take the best result. The ICA–contribution can be removed here.

Figure 4.6 displays the results for method 4. The best results are achieved with division into  $m = 2$  parts. A severe degradation of performance is observed when the signal is divided in more than two parts. When one of two parts originates from a different simultaneous channel, no improvement is noticed anymore. The first 9 sweeps show a significant improvement. A drawback of this method is the significantly worse performance for a higher number of

averaged sweeps. For a more practical approach, Tables 4.2, 4.3 and 4.4 show a comparison of method 4 with method 1 per subject, intensity and carrier frequency. A measurement time reduction of almost 11 % is possible.

### 4.3.5 Combining single channel and multichannel techniques

#### ROC

Comparing Figure 4.2 and Figure 4.6, the idea of combining both methods comes forward. Figure 4.7 shows the ROC-curve of method 5 that combines methods 1, 2 and 4. The benefits of both methods merge into a scheme that provides significant improvement over the whole range of averaged sweeps.

#### Effective measurement time reduction

When comparing method 5 with method 3, the detection criterion is reduced from  $p = 0.0065$  to  $p = 0.0053$ . This corrected decision criterion is used to construct the result for method 5 in Tables 4.2, 4.3 and 4.4. It confirms a further reduction of measurement time. For this dataset, the responses for all individual subjects can be detected up to 58 % and on average about 22 % faster. No negative values are present which however does not imply that negative values are not possible anymore. All conclusions from Section 4.3.2, stating that the optimal number of channels is equal to five, still hold.

## 4.4 Discussion

This section discusses the results from Section 4.3. The structure of the first four subsections (Sections 4.4.1 to 4.4.4) is identical to the structure of the previous section. Section 4.4.5 discusses the effect of artifact rejection on independent component analysis.

### 4.4.1 Multichannel ICA (method 3): seven channels and seven extracted components

A major improvement in detection speed is possible when recorded multichannel EEG data is preprocessed using the ICA-technique. ASSRs are sinusoidal which is characterized by a platykurtic distribution, while EEG noise sources can be considered to have mesokurtic, or Gaussian, distributions. This is in line with the central limit theorem explaining that the combined sum of all noise sources will take on a Gaussian distribution if the number of noise sources goes to infinity (Trotter, 1959). As ICA is trying to separate the sources that



have optimally different distributions (statistical independence), it is probable that the algorithm is indeed able to separate the sources because of these characteristics. This separation ability is influenced by the amount of noise that is present in the used measurements. ICA performs better when less noise is present (Hyvärinen et al., 2001). The lower the SNR, the smaller the measurement time reduction. This is reflected in Tables 4.3 and 4.4, where especially measurements at higher intensities and with carrier frequencies that evoke larger responses (1 and 2 kHz) tend to return a larger measurement time reduction. More noise induces a worse estimation of the separating matrix  $\mathbf{W}$ . This is especially obvious when not using the averaging step which initially pre-processes the data solely to increase the SNR. This indicates that ICA would only be useful when applied to data with high SNR, like data obtained at relatively high intensities above hearing threshold.

Two observations can be deduced from Table 4.2: the possibility to perform worse than the single channel reference (method 1) for some subjects and the large variance on the individual subject measurement time reduction. Firstly, it is clearly possible that the SNR of a response can trigger a detection criterion for the single channel case, but not the criterion for the multichannel case, which is much more stringent. This behavior can be considered as a sort of ‘noise floor’ which could even result in a measurement time reduction for frequencies with no response present. As such, response measurement time increases which lie within the noise floor variance ( $\pm 5\%$ ) can be neglected. Secondly, the large variance (data ranging from -2 to 63% and already mentioned in Section 4.2.5) is caused by the sometimes failing ability of the single channel method to detect responses, even at high intensities. In practice, this can happen from time to time. It is generally assumed that the Cz–Oz electrode combination is the best one for detecting responses, as confirmed by the numerous studies using this electrode combination described in the review paper of Picton et al. (2003). However, this assumption is not always correct and should be investigated further, as shown in van der Reijden et al. (2004, 2005) and Chapter 7.

#### 4.4.2 Multichannel ICA (method 3): variable number of channels and extracted components

Connecting many electrodes to the subject’s head, is not very practical. To reduce preparation time, it is considered an advantage if a method requires fewer measurement electrodes. For this dataset, the ICA–technique applied to EEG containing ASSRs reaches a saturation level at 5 channels (7 electrodes). This could indicate that it is sufficient to condense all estimated sources  $\hat{\mathbf{S}}$  from Section 4.2.2 in only five major components  $\hat{\mathbf{s}}_i$ , each component being a combination of several ASSRs and noise sources. If more channels are used, no performance improvement is observed anymore. This effect can be explained by the fact that a sixth channel still adds extra information to the complete

system, but this additional gain is annihilated by the extra increase of false-positives (false detections) caused by the addition of an extra channel. The best electrode positions are located on the back of the head. This is expected because of the higher SNR that is obtained at these positions (van der Reijden et al., 2005). For this dataset, the additional fifth channel should be located in between the four channels on the back of the head and the forehead, more specifically on position F4. The choice for the symmetrical electrode position F3 degenerates the performance.

#### 4.4.3 Multichannel ICA (method 3): effects of fixed separating matrix

##### Variability of the separating matrix

The reconstruction of the assumed underlying generators  $\hat{\mathbf{S}}$  from Section 4.2.2 is performed by multiplying the computed separating matrix  $\mathbf{W}$  with the signal matrix  $\mathbf{X}$ . As a result, the obtained ICs are an optimal linear combination of the recorded signals in a non-Gaussian and independent sense. The elements of matrix  $\mathbf{W}$  are determined by several factors, like the position and quality of the electrodes, the location and orientation of the underlying sources inside, and the physical properties of the subject's skull. During one subject session with different intensities, these parameters are sufficiently stationary. This is reflected in reoccurring ICs with similar structure (combinations of different responses) for the same intensity and even for different intensities, as long as one and the same subject is considered. Almost the same holds for the separating matrices  $\mathbf{W}$ . When comparing two  $\mathbf{W}$  matrices that originate from different datasets, but close to each other on a time scale, the matrix coefficients differ only slightly with sometimes permuted rows. The similarities of these matrices are caused by the high correlation between an averaged sweep and its successor. This behavior is not observed when comparing  $\mathbf{W}$  matrices from the same subject but at different intensities.

##### Fixed separating matrix for all subjects

Keeping  $\mathbf{W}$  fixed for all subjects yields good performance as long as the correct  $\mathbf{W}$  is chosen. However, the major problem is the choice of  $\mathbf{W}$ . For some separating matrices, a performance close to that of the standard use of ICA is reached. This could indicate that an optimal separating matrix can be calculated a priori for a multichannel setup. However, this matrix can only be used for a certain electrode configuration and data set, if it already can be found in the first place. If so, the calculation for each new averaged sweep of the  $\mathbf{W}$  coefficients will likely perform better.

### Fixed separating matrix for one subject

The ROC-curves show that fixing  $\mathbf{W}$  calculated from data with the largest assumed SNR from each subject, can reduce detection time significantly. However, this improvement is still significantly smaller than the improvement obtained when ICA is applied every time to each collected, averaged, sweep. The former results show that the derived ICs represent corresponding real physical entities that change between subjects and intensities, but stay rather constant for the same intensity.

### Fixed IC for one subject

Fixing ICs per modulation frequency should be avoided. Although it seems interesting without the multiple testing aspect (and therefore a less stringent detection criterion needs to be used), it is not advisable to link a component to a modulation frequency immediately. ICA does not order its ICs. Therefore, a fixed IC cannot be chosen beforehand. One can try to solve this problem by introducing an ordering system, like sorting ICs according to their energy. This will not always work, especially at lower SNRs. This observation can be explained by (4.5). Depending on the variable composition of  $\mathbf{S}$ , fixed ICs or fixed  $\mathbf{W}$  will give variable results, without the existence of an optimal solution.

#### 4.4.4 Single channel ICA combined with method 1

In general, ICA does not work on single channel data. This is expected because ICA is essentially a multichannel technique. When multiple channels are recorded simultaneously, ICA uses the relatively high correlation between the different channels to separate mesokurtic background noise and leptokurtic responses. When only one channel is available, the different parts apparently do almost not have any correlation at all because EEG signals are non-stationary. Our assumption in Section 4.2.5 does not seem to be valid. This is confirmed by Davies and James (2007) who show that independent processes must be bandlimited with disjoint spectral support. This is not the case with ASSRs that are superimposed on the EEG and thus overlap. The better performance at lower SNRs when splitting the single channel data cannot be explained yet.

#### 4.4.5 Multichannel ICA: the use of artifact rejection

In this chapter, artifact rejection is used to define the reference approach as close as possible to the clinically standard single channel technique. However, no noise-weighted averaging is used already. This is done in Chapters 6 and 7. For all methods from Section 4.2.5, epochs that contain samples with a value higher than  $20 \mu V$  in absolute value, are rejected. In theory, ICA can separate external sources (muscles, eye blinks, ...) easily from the actual EEG that is

being monitored. In practice, the benefits of ICA have already been confirmed when noisy subject data are considered, as long as enough channels are used (Jung et al., 2000). It seems that the number of channels is quite low for this purpose in the used setup of this manuscript. It seems better to use artifact rejection instead of separating artifacts into separate components. However, the single channel reference method is generally incapable of extracting any useful information from data contaminated with much noise and artifacts, which is indeed experienced e.g. by clinical people who have to test restless babies. When multichannel recordings are used, we observe that even the noisiest input data to some extent still can be processed to usable ASSRs, even with a relatively small number of available channels. The robustness against artifacts is investigated in Chapters 6 and 7, using other techniques.

## 4.5 Conclusions

In this chapter it is shown that ICA is a valuable tool to separate the additive background noise from the EEG-waveform of interest, namely the ASSR. After a short introduction in Section 4.1, Section 4.2 describes the theoretical aspects of independent component analysis, the used evaluation setup and stimuli. A specific model is assumed and performance measures are described.

Section 4.3 shows that ICA applied on single and multichannel recordings yields a significantly better performance than the clinically used single channel reference technique for data obtained at intensities above hearing threshold. For single channel measurements a time reduction up to 23 % for a single subject has been acquired. For multichannel EEG measurements there is a significant measurement time reduction possible of 52 % for 48-sweep measurements compared to the single channel reference technique. For individual subjects, an improvement up to 63 % in measurement time has been recorded. When both single- and multichannel techniques are combined, performance can be improved even more. This ultimate combination generally guarantees a significant improvement for all measurement durations. However, multichannel ICA is not always capable of reducing measurement time for each individual subject, illustrating variability between subjects of ICA applied to ASSR and of ASSR in general. Five-channel ICA yields the same performance as the seven-channel ICA. Electrodes are best placed on the back of the head. This is important for the clinical applicability of the described technique. When the separating matrix  $\mathbf{W}$  is kept fixed for each subject separately, a significant improvement is observed for some cases. Keeping  $\mathbf{W}$  fixed for all subjects improves detection speed significantly, as long as an optimal  $\mathbf{W}$  is found a priori. This, however, is not always guaranteed. Section 4.4 discusses the results found in Section 4.3.

## Chapter 5

# Improving ASSR Detection Using Multichannel Wiener Filtering

This chapter is organized as follows. After a short introduction in Section 5.1, the theoretical background of this chapter is described in Section 5.2. In Section 5.3, the used experimental setup and analysis methods are described. Section 5.4 is dedicated to the obtained results. In Section 5.5, a comparison is made with independent component analysis applied to the same data set in the previous chapter. Finally, conclusions are given in Section 5.6.

### 5.1 Introduction

In the previous chapter, it was shown that the use of independent component analysis (ICA) could reduce measurement time significantly compared to single channel methods. One drawback of the ICA-technique is that ICA does not use any prior information except for the assumption of independence of the underlying non-Gaussian sources. Another technique is proposed in this chapter that incorporates a priori knowledge: multichannel Wiener filtering with QR factorization. The sinusoidal nature of the ASSR makes it possible to search for a specific frequency, equal to the modulation frequency used in the stimulus.

---

The material presented in this chapter has been published in ‘Van Dun, B., Wouters, J., and Moonen, M. (2007), “Multi-channel Wiener filtering based auditory steady-state response detection,” *Proc. of the 32nd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Honolulu, HI, USA, vol. 2, 929–932’.

## 5.2 Theoretical background

In this section, a general ASSR signal model is presented in Section 5.2.1, together with the simplifications it assumes (Section 5.2.2). Section 5.2.3 introduces the concept of SNR-maximization for ASSRs. This concept can be linked with the multichannel Wiener filter, as explained in Section 5.2.4.

### 5.2.1 ASSR signal model

A simplified ASSR signal model can be given as

$$\mathbf{Z} = \alpha \mathbf{S} \mathbf{D}^T + \mathbf{N} \quad (5.1)$$

with  $\mathbf{Z}^{n \times m}$  an observation matrix,  $\alpha$  a scalar representing the ASSR source amplitude proportional to the applied stimulus intensity,  $\mathbf{S}^{n \times 2}$  a desired signal matrix with columns representing the ASSR (sinusoid & cosinusoid, oscillating at the known modulation frequency),  $\mathbf{D}^{m \times 2}$  a steering matrix,  $\mathbf{N}^{n \times m}$  an additive noise matrix,  $n$  the number of data points and  $m$  the number of measured EEG channels.

In (5.1), given only  $\mathbf{Z}$  and  $\mathbf{S}$ ,  $\alpha \mathbf{D}$  and  $\mathbf{N}$  can be estimated using a QR-factorization:

$$\left[ \begin{array}{cc|c} \overbrace{\mathbf{S}}^{\mathbf{S}} & & \\ \hline \mathbf{s}_1^{n \times 1} & \mathbf{s}_2^{n \times 1} & \mathbf{Z}^{n \times m} \end{array} \right] = \mathbf{Q}^{n \times (m+2)} \mathbf{R}^{(m+2) \times (m+2)} \quad (5.2)$$

with

$$\mathbf{Q} = \left[ \begin{array}{cc|c} \mathbf{s}_1 & \mathbf{s}_2 & \mathbf{Q}_*^{n \times m} \end{array} \right], \quad \mathbf{R} = \left[ \begin{array}{ccc} 1 & 0 & \hat{\mathbf{d}}_1^T \\ 0 & 1 & \hat{\mathbf{d}}_2^T \\ \mathbf{0} & \mathbf{0} & \mathbf{R}_*^{m \times m} \end{array} \right] \quad (5.3)$$

$\mathbf{Z}$  can then be written as

$$\mathbf{Z} = \left[ \begin{array}{cc} \overbrace{\mathbf{S}}^{\mathbf{S}} \\ \hline \mathbf{s}_1 & \mathbf{s}_2 \end{array} \right] \overbrace{\left[ \begin{array}{c} \hat{\mathbf{d}}_1^T \\ \hat{\mathbf{d}}_2^T \end{array} \right]}^{\hat{\mathbf{D}}^T} + \overbrace{\mathbf{Q}_* \mathbf{R}_*}_{\hat{\mathbf{N}}} \quad (5.4)$$

When  $\mathbf{S}$  contains an integer number of periods of the ASSR,  $\mathbf{S}$  is orthogonal. This is guaranteed in e.g. the experimental Section 5.3 by correcting the used modulation frequencies such that an integer number of ASSR periods is present in each column of the measurement matrix  $\mathbf{Z}$  in (5.1). From this orthogonality of  $\mathbf{S}$ , it follows that

$$\hat{\mathbf{D}}^T = (\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T \mathbf{Z} = \mathbf{S}^T \mathbf{Z} \quad (5.5)$$

which corresponds to a least-squares estimation of  $\alpha \mathbf{D}^T$ ,

$$\hat{\mathbf{D}} = \arg \left\{ \min_{\mathbf{D}} \|\mathbf{Z} - \mathbf{S}\hat{\mathbf{D}}^T\|_2^2 \right\} \quad (5.6)$$

The ASSR generator has an unknown source amplitude  $\alpha$  which depends on the stimulus level. After propagation through the skull, the distribution of the recorded ASSR on the electrodes, present in the observation matrix  $\mathbf{Z}$ , can be described by the steering matrix  $\mathbf{D}$ . The steering matrix  $\mathbf{D}$ , in contrast with the additive EEG noise, is usually assumed to be stationary as it is merely a representation of the source position, its directivity pattern, the electrode positions and the propagation attenuation from source to electrode. Physically no measurable delay occurs between the ASSR source and the electrodes (order of nanoseconds). Therefore the ASSR delay difference and hence the ASSR phase difference between electrodes is considered zero. On the other hand, the delay between ASSR stimulus and response is measurable (order of milliseconds) but is unknown to the observer due to several physical parameters of the subject. The ASSR phase  $\varphi$  at the electrodes is thus unknown but equal in all channels. The assumptions above may be simplifications of reality. Any consequences hereof are discussed in Section 5.2.2.

Under these assumptions the (exact) steering matrix  $\mathbf{D}$  is a rank-1 matrix

$$\underbrace{\mathbf{D}^T}_{2 \times m} = \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix} \underbrace{\mathbf{d}^T}_{1 \times m} \quad (5.7)$$

so that (5.1) can be rewritten as

$$\mathbf{Z} = \alpha \mathbf{S} \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix} \mathbf{d}^T + \mathbf{N} \quad (5.8)$$

and taking

$$\mathbf{s} = \mathbf{S} \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix} \quad (5.9)$$

such that

$$\mathbf{Z} = \alpha \mathbf{s} \mathbf{d}^T + \mathbf{N} \quad (5.10)$$

where now  $\mathbf{s}$  and  $\mathbf{d}$  are vectors, and  $\varphi$  corresponds to the ASSR phase, which needs to be estimated.

Based on (5.10), a joint least-squares minimization of  $\varphi$  and  $\mathbf{d}$  can be given as

$$\hat{\varphi} = \arg \left\{ \min_{\hat{\varphi}, \hat{\mathbf{d}}} \|\mathbf{Z} - \hat{\mathbf{s}} \hat{\mathbf{d}}^T\|_2^2 \right\} \quad (5.11)$$

with  $\hat{\mathbf{d}}$  corresponding to a least-squares estimation of  $\alpha \mathbf{d}^T$ ,

$$\hat{\mathbf{d}}^T = (\mathbf{s}^T \mathbf{s})^{-1} \mathbf{s}^T \mathbf{Z} = \mathbf{s}^T \mathbf{Z} \quad (5.12)$$

For a given (optimal)  $\hat{\varphi}$ , the corresponding optimal  $\hat{\mathbf{d}}$  is the solution of a least-squares problem,

$$\hat{\mathbf{d}}^T = \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix}^T \mathbf{S}^T \mathbf{Z} \quad (5.13)$$

Substituting (5.9) and (5.13) in (5.11), gives

$$\hat{\varphi} = \arg \left\{ \min_{\hat{\varphi}} \left\| \mathbf{Z} - \mathbf{S} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix}^T \mathbf{S}^T \mathbf{Z} \right\|_2^2 \right\} \quad (5.14)$$

$$= \arg \left\{ \min_{\hat{\varphi}} \left\| \underbrace{(\mathbf{I} - \mathbf{S} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix}^T \mathbf{S}^T)}_{\text{projection matrix}} \mathbf{Z} \right\|_2^2 \right\} \quad (5.15)$$

The factor between brackets is a projection matrix  $\mathbf{P}$  ( $\mathbf{P}^T = \mathbf{P}$  and  $\mathbf{P}^T \mathbf{P} = \mathbf{P}$ ), such that

$$\hat{\varphi} = \arg \left\{ \min_{\hat{\varphi}} \left\| \mathbf{Z}^T (\mathbf{I} - \mathbf{S} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix}^T \mathbf{S}^T) \mathbf{Z} \right\|_2 \right\} \quad (5.16)$$

$$= \arg \left\{ \min_{\hat{\varphi}} \left\| \mathbf{Z}^T \mathbf{Z} - \hat{\mathbf{D}} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix}^T \hat{\mathbf{D}}^T \right\|_2 \right\} \quad (5.17)$$

$$= \arg \left\{ \max_{\hat{\varphi}} \left\| \hat{\mathbf{D}} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix}^T \hat{\mathbf{D}}^T \right\|_2 \right\} \quad (5.18)$$

$$= \arg \left\{ \max_{\hat{\varphi}} \left\| \hat{\mathbf{D}} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix} \right\|_2^2 \right\} \quad (5.19)$$

or alternatively, including a prewhitening transformation

$$\hat{\varphi} = \arg \left\{ \max_{\hat{\varphi}} \left\| (\hat{\mathbf{N}}^T \hat{\mathbf{N}})^{-\frac{1}{2}} \hat{\mathbf{D}} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix} \right\|_2^2 \right\} \quad (5.20)$$

Based on  $\hat{\varphi}$ , one can then compute

$$\hat{\mathbf{d}} = \hat{\mathbf{D}} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix} \quad (5.21)$$

A spatio-temporal noise covariance matrix  $\mathbf{K}^{mn \times mn}$  can be defined as

$$\mathbf{K} = \mathcal{E} \{ \mathbf{n} \mathbf{n}^T \} \quad (5.22)$$

with  $\mathbf{n} = \text{vec}(\mathbf{N})$ . Here  $\mathcal{E}\{\cdot\}$  is the expected value operator and the  $\text{vec}(\cdot)$  operator stacks the columns of a matrix  $\mathbf{X}$  into one column vector  $\mathbf{x} = \text{vec}(\mathbf{X})$ .

If the spatial and temporal correlations are separable, as will be observed here, the spatio-temporal noise covariance matrix  $\mathbf{K}$  can be written as (Johnson and Dudgeon, 1993)

$$\mathbf{K} = \mathbf{K}_{\text{spat}} \otimes \mathbf{K}_{\text{temp}} \quad (5.23)$$



where  $\otimes$  represents the Kronecker product, with a spatial noise covariance matrix

$$\mathbf{K}_{\text{spat}} = \mathcal{E}\{\mathbf{N}^T\mathbf{N}\} \quad (5.24)$$

and a temporal noise covariance matrix

$$\mathbf{K}_{\text{temp}} = \mathcal{E}\{\mathbf{N}\mathbf{N}^T\} \quad (5.25)$$

Based on (5.4) the noise covariance matrices can be estimated as

$$\hat{\mathbf{K}}_{\text{spat}} = \hat{\mathbf{N}}^T\hat{\mathbf{N}}, \quad \hat{\mathbf{K}}_{\text{temp}} = \hat{\mathbf{N}}\hat{\mathbf{N}}^T \quad (5.26)$$

### 5.2.2 Simplifications of the model compared with real-life

The model presented in this chapter is a simplification of reality. It assumes only one intracranial source. Healthy adults, however, have at least two intracranial sources (Herdman et al., 2002; Purcell et al., 2004). This has some implications. Formula (5.8) should then be reformulated as

$$\mathbf{Z} = \sum_{i=1}^Q \alpha_i(t) \mathbf{S} \begin{bmatrix} \cos \varphi_i(t) \\ \sin \varphi_i(t) \end{bmatrix} \mathbf{d}_i^T + \mathbf{N} \quad (5.27)$$

with the observation matrix  $\mathbf{Z}$  containing  $Q$  sources from an ASSR with a specific modulation frequency. It is assumed in (5.8) that an ASSR source is constant in amplitude  $\alpha$  and phase  $\varphi$  over time. This cannot be guaranteed however. For example, ASSR amplitudes mainly originating from the auditory cortex vary significantly with the state of arousal (Galambos et al., 1981). With more than one ASSR source, the assumption that the resulting phase  $\varphi$  is equal at all electrodes is not valid either, unless these sources are exactly in phase.

The concerns above are assumed to have only a small effect. According to Purcell et al. (2004), the general two sources from an ASSR with a specific modulation frequency are a main source in the brainstem and a main source in the auditory cortex. The modulation frequencies used in this study are chosen in an extreme region: the region with the brainstem source being dominant (80–110 Hz). Therefore, the approximation by the model is assumed to have reasonable validity.

The biggest concern is the estimation of the ASSR amplitude  $\alpha_i(t)$ , being incorporated in the least-square estimation of  $\alpha_i(t)\mathbf{D}_i$ , as  $\alpha_i(t)$  is estimated on the entire observation matrix  $\mathbf{Z}$  using (5.1). This way, fluctuating amplitudes cannot be captured. However, using less data compromises the assessment of  $\alpha_i(t)$ , as the ASSR amplitude is already small compared to the noise.

### 5.2.3 Maximizing the output–SNR of the ASSR

A possible way to improve the detection possibility of an ASSR, is to increase its signal-to-noise ratio (SNR). Assuming equal phase  $\varphi$  over all electrodes, the weight vector  $\mathbf{w}$  is searched for that jointly maximizes the energy of the projection of the observation matrix  $\mathbf{Z}$  on the plane formed by the desired signal vector  $\mathbf{S}$  (the signal plane), together with a minimization of the projection of  $\mathbf{Z}$  on the plane orthogonal to  $\mathbf{S}$  (the noise plane),

$$\max_{\mathbf{w}} \frac{\|\mathbf{S}\mathbf{S}^T(\mathbf{Z}\mathbf{w})\|_2^2}{\|(\mathbf{I} - \mathbf{S}\mathbf{S}^T)(\mathbf{Z}\mathbf{w})\|_2^2} \quad (5.28)$$

Note that we have assumed that  $\mathbf{S}^T\mathbf{S} = \mathbf{I}$  such that

$$\max_{\mathbf{w}} \frac{\|\mathbf{S}^T(\mathbf{Z}\mathbf{w})\|_2^2}{\|(\mathbf{I} - \mathbf{S}\mathbf{S}^T)(\mathbf{Z}\mathbf{w})\|_2^2} = \max_{\mathbf{w}} \frac{\|(\mathbf{S}^T\mathbf{Z})\mathbf{w}\|_2^2}{\|(\mathbf{Z} - \mathbf{S}\mathbf{S}^T\mathbf{Z})\mathbf{w}\|_2^2} \quad (5.29)$$

By knowing that  $\hat{\mathbf{N}} = \mathbf{Z} - \mathbf{S}\mathbf{S}^T\mathbf{Z}$ , and by using (5.4) and (5.5), one can write

$$\max_{\mathbf{w}} \frac{\|\hat{\mathbf{D}}^T\mathbf{w}\|_2^2}{\|\mathbf{R}_*\mathbf{w}\|_2^2} = \max_{\mathbf{w}} \frac{\mathbf{w}^T\hat{\mathbf{D}}\hat{\mathbf{D}}^T\mathbf{w}}{\mathbf{w}^T\mathbf{R}_*\mathbf{R}_*\mathbf{w}} \quad (5.30)$$

$$= \max_{\bar{\mathbf{w}}} \frac{\bar{\mathbf{w}}^T\mathbf{R}_*^{-T}\hat{\mathbf{D}}\hat{\mathbf{D}}^T\mathbf{R}_*^{-1}\bar{\mathbf{w}}}{\bar{\mathbf{w}}^T\bar{\mathbf{w}}} \quad (5.31)$$

A solution for  $\bar{\mathbf{w}} = \mathbf{R}_*\mathbf{w}$  can be found as the eigenvector corresponding to the largest eigenvalue of  $\mathbf{R}_*^{-T}\hat{\mathbf{D}}\hat{\mathbf{D}}^T\mathbf{R}_*^{-1}$  in (5.31). The multichannel observation matrix  $\mathbf{Z}$  can now be transformed to a single channel measurement vector  $\mathbf{z} = \mathbf{Z}\mathbf{w}$  with

$$\mathbf{w} = \mathbf{R}_*^{-1}\mathbf{v}_{\max}\left\{\mathbf{R}_*^{-T}\hat{\mathbf{D}}\hat{\mathbf{D}}^T\mathbf{R}_*^{-1}\right\} \quad (5.32)$$

with  $\mathbf{v}_{\max}\{\cdot\}$  the eigenvector associated with the largest eigenvalue.

The weight vector  $\mathbf{w}$  thus recombines the rows of observation matrix  $\mathbf{Z}$  into one channel that maximizes the SNR for the signal component  $\mathbf{S}$  with the specified modulation frequency  $f$ .

This method can be extended to more than one modulation frequency. By introducing extra  $\mathbf{s}_i$  with other frequencies  $f_i$  in (5.4), the output channel can be optimized in SNR-sense for all defined  $\mathbf{s}_i$ . For example, when  $I - 1$  extra modulation frequencies  $f_i$  are added,  $\mathbf{Q}$  becomes an  $n \times (m + 2I)$  matrix and  $\mathbf{R}$  an  $(m + 2I) \times (m + 2I)$  matrix,

$$\mathbf{Q} = \left[ \begin{array}{ccc|c} \mathbf{s}_1 & \dots & \mathbf{s}_{2I} & \mathbf{Q}_*^{n \times m} \end{array} \right], \quad \mathbf{R} = \left[ \begin{array}{ccc|c} \mathbf{I}^{2I \times 2I} & & & \begin{array}{c} \hat{\mathbf{d}}_1^T \\ \vdots \\ \hat{\mathbf{d}}_{2I}^T \end{array} \\ \hline \mathbf{0} & \dots & \mathbf{0} & \mathbf{R}_*^{m \times m} \end{array} \right] \quad (5.33)$$

$\mathbf{R}_*$  does not contain information on the modulation frequencies  $f_i$ . The steering matrix  $\mathbf{D}$  is an  $(m \times 2I)$  matrix. To calculate the optimal weight vector  $\mathbf{w}_i$  for a desired signal matrix  $\mathbf{S}_i$  containing a (co)sinusoid with modulation frequency  $f_i$ , (5.32) is recomputed with  $\mathbf{R}_*$  from (5.33) and with selected columns  $\mathbf{D}(:, 2i - 1 : 2i)$ .

#### 5.2.4 Linking output–SNR maximization with the multi-channel Wiener filter

If the phase  $\varphi$  is estimated, steering matrix  $\mathbf{D}$  can be reduced to a steering vector  $\mathbf{d}$  according to (5.21), and the desired signal matrix  $\mathbf{S}$  to a desired signal vector  $\mathbf{s}$  based on (5.9),

$$\hat{\mathbf{d}} = \hat{\mathbf{D}} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix}, \quad \hat{\mathbf{s}} = \mathbf{S} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix} \quad (5.34)$$

Using (5.4) and (5.34), a Wiener filter  $\mathbf{w}_{\text{MSE}}$  can be constructed,

$$\mathbf{w}_{\text{MSE}} = \mathbf{R}_{zz}^{-1} \mathbf{r}_{zs} \quad (5.35)$$

with  $\mathbf{R}_{zz} = \mathcal{E}\{\mathbf{Z}^T \mathbf{Z}\}$  and  $\mathbf{r}_{zs} = \mathcal{E}\{\mathbf{Z}^T \mathbf{s}\} = \mathcal{E}\{\mathbf{d}\}$ . Because the signal component  $\mathbf{S}$  and the noise component  $\mathbf{N}$  from (5.4) are considered to be uncorrelated,  $\mathbf{R}_{zz} = \mathbf{R}_{dd} + \mathbf{K}_{\text{spat}}$  with  $\mathbf{R}_{dd} = \mathcal{E}\{\mathbf{d}\mathbf{d}^T\}$  and  $\mathbf{K}_{\text{spat}} = \mathcal{E}\{\mathbf{N}^T \mathbf{N}\}$  according to (5.24).

$$\hat{\mathbf{w}}_{\text{MSE}} = (\hat{\mathbf{R}}_{dd} + \hat{\mathbf{K}}_{\text{spat}})^{-1} \hat{\mathbf{d}} \quad (5.36)$$

The Wiener filter solution  $\mathbf{w}_{\text{MSE}}$  can be rewritten as the product of a real-valued scalar and the solution of an output–SNR maximization problem  $\mathbf{w}_{\text{SNR}}$  (Monzingo and Miller, 1980; Simmer et al., 2001). By applying the Sherman–Morrison–Woodbury formula,

$$(\mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T + \mathbf{A}^{-1})^{-1} \equiv \mathbf{A} - \mathbf{A}\mathbf{B}(\mathbf{C} + \mathbf{B}^T\mathbf{A}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{A} \quad (5.37)$$

which is also known as the matrix inversion lemma, with  $\mathbf{A} = \mathbf{K}_{\text{spat}}^{-1}$ ,  $\mathbf{B} = \hat{\mathbf{d}}$  and  $\mathbf{C} = 1$ , the following identity can be written,

$$\hat{\mathbf{w}}_{\text{MSE}} = \left( \hat{\mathbf{K}}_{\text{spat}}^{-1} - \frac{\hat{\mathbf{K}}_{\text{spat}}^{-1} \hat{\mathbf{d}} \hat{\mathbf{d}}^T \hat{\mathbf{K}}_{\text{spat}}^{-1}}{1 + \hat{\mathbf{d}}^T \hat{\mathbf{K}}_{\text{spat}}^{-1} \hat{\mathbf{d}}} \right) \hat{\mathbf{d}} \quad (5.38)$$

$$= \left( 1 - \frac{\hat{\mathbf{d}}^T \hat{\mathbf{K}}_{\text{spat}}^{-1} \hat{\mathbf{d}}}{1 + \hat{\mathbf{d}}^T \hat{\mathbf{K}}_{\text{spat}}^{-1} \hat{\mathbf{d}}} \right) \hat{\mathbf{K}}_{\text{spat}}^{-1} \hat{\mathbf{d}} \quad (5.39)$$

$$= \underbrace{\frac{1}{1 + \hat{\mathbf{d}}^T \hat{\mathbf{K}}_{\text{spat}}^{-1} \hat{\mathbf{d}}}}_{\text{real-valued scalar}} \hat{\mathbf{w}}_{\text{SNR}} \quad (5.40)$$

with

$$\hat{\mathbf{w}}_{\text{SNR}} = \hat{\mathbf{K}}_{\text{spat}}^{-1} \hat{\mathbf{d}} \quad (5.41)$$

As a result, the solution of the SNR-maximization expressed by (5.32) is also a multichannel Wiener filter. This is only valid when the steering matrix  $\mathbf{D}$  is a rank-1 matrix, as expressed by (5.7). Otherwise the factor in (5.40) is no scalar anymore, but becomes a subband or frequency domain post-filter (Simmer et al., 2001). When this happens, the SNRs of the filtered results by applying  $\mathbf{w}_{\text{SNR}}$  and  $\mathbf{w}_{\text{MSE}}$  to the observation matrix  $\mathbf{Z}$  are not equal anymore.

### 5.3 Experimental setup

The setup and stimuli used in this study are identical to those described in Sections 4.2.3 and 4.2.4. The applied procedure is similar to method 2 from Section 4.2.5. No need for the addition of extra components or channels as in method 3 from Section 4.2.5 is necessary. This immediately shows the benefit of the MWF-technique compared with the ICA-technique from Chapter 4.

The sketchy algorithm below processes  $m$ -channel EEG data with a length of  $NM$  samples.  $N$  is the number of sweeps (data blocks) the data will be divided into. In this chapter the data is divided into  $N = 48$  sweeps. One sweep is 16.384 seconds long (or  $M = 4096$  samples), based on parameters also used in John and Picton (2000a). This EEG data matrix is, after artifact rejection and an averaging step, processed using the MWF-technique (equivalent to a SNR maximization) from Section 5.2.3 to obtain one combined EEG channel that maximizes the SNR of the modulation frequency one is looking for. For clarity, the description will explicitly show the process using  $N = 48$  sweeps and  $M = 4096$  samples. The number of channels  $m$  is kept variable.

- The  $m$ -channel EEG recordings are divided in 48  $m$ -channel data blocks (sweeps) with a length of 4096 samples (16.384 seconds).
- Each  $m$ -channel sweep  $s_n$  ( $n = 1 \dots 48$ ) is averaged together with all preceding sweeps  $s_1$  to  $s_{n-1}$ . This creates 48 averaged  $m$ -channel sweeps  $\bar{s}_n$  ( $n = 1 \dots 48$ ) with a length of 4096 samples. Averaged sweep  $\bar{s}_1$  is identical to sweep  $s_1$ . Averaged sweep  $\bar{s}_{48}$  is an average of all sweeps  $s_1$  to  $s_{48}$ .
- The MWF-technique represented by (5.32) takes one averaged  $m$ -channel sweep  $\bar{s}_n$  as an input.
- One resulting combined sweep is calculated based on the  $m$ -channel averaged sweep  $\bar{s}_n$  (with a length of 4096 samples).

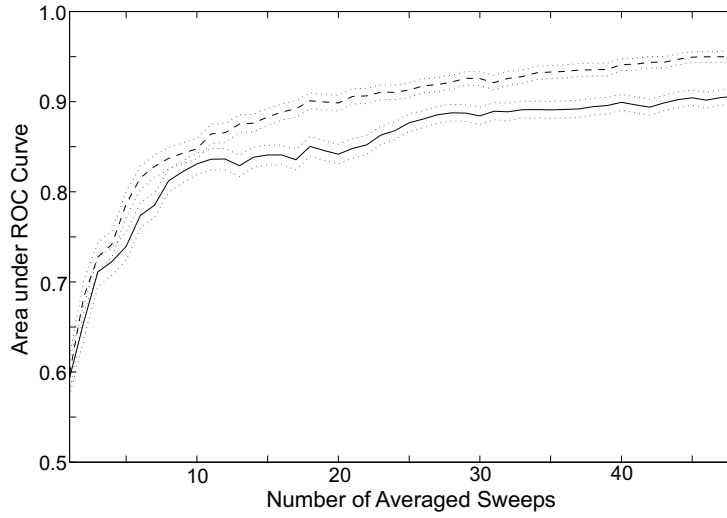


Figure 5.1: ‘area under ROC–curve’ versus ‘number of averaged sweeps’: MWF–procedure applied to a seven–channel data set (dashed); Method 1, single channel reference (solid). The dotted lines denote two standard deviations.

- For each processed (averaged) sweep, the F–ratio (SNR) of each modulation frequency is determined using (2.11). When 8 modulation frequencies are used, 8 F–ratios are calculated for each processed (averaged) sweep  $\bar{s}_n$ .

In order to evaluate the single channel and multichannel technique, *receiver operating characteristic* (ROC) curves and the effective measurement time reduction are calculated from 8 subjects as described in Section 4.2.6.

## 5.4 Results

Figure 5.1 shows the results of the MWF–procedure on a seven–channel EEG data set. For a sufficiently large number of sweeps, a significant performance increase is obtained compared to the single channel reference (method 1 from Chapter 4). A problem with this configuration is its poor performance for a small number of sweeps (left–hand side of Figure 5.1). The calculation of the eigenvectors is found to be sensitive to noise.

When the results per subject are considered in Table 4.2, an average measurement duration reduction of 13 % is observed, varying between a 9 % measurement duration increase and a possible 62 % decrease. Higher intensities and

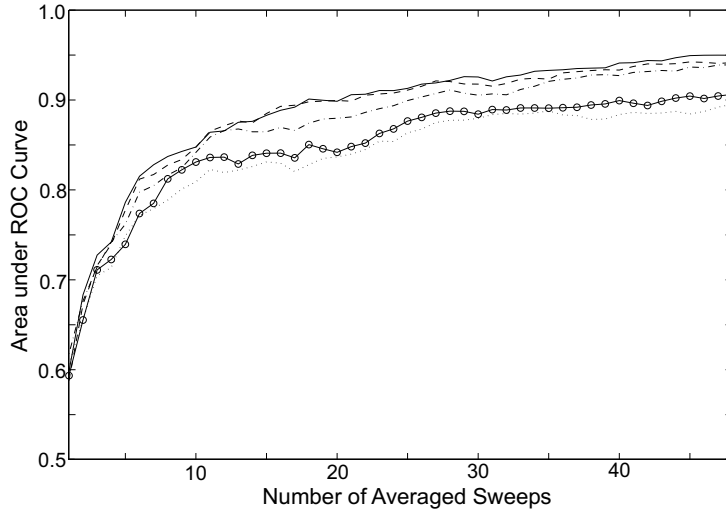


Figure 5.2: ‘area under ROC–curve’ versus ‘number of averaged sweeps’. MWF–procedure applied to an  $m$ –channel data sets:  $m = 7$  (solid),  $m = 5$  (dashed),  $m = 4$  (dashdot),  $m = 2$  (dotted); Method 1, single channel reference (solid–circle).

frequencies with physically larger responses (e.g. with 1 and 2 kHz carriers) are more prone to faster detection (Tables 4.3 and 4.4). The one case with the measurement duration increase lies closely to the noise floor of 6.6 % ( $2\sigma$  around mean measurement duration reduction of 0 % for noise frequencies).

While measurement duration can be reduced significantly with the multichannel approach, connecting many electrodes to the subject’s head is not very practical. To reduce preparation time, it is considered an advantage if a method requires fewer measurement electrodes. Figure 5.2 shows that performance is only marginally compromised by using five instead of seven channels for this data set. Adding extra (i.e. more than five) simultaneous EEG channels does not further reduce measurement duration. It is expected that a more efficient electrode placement may increase performance beyond five channels. This assumption is investigated in Chapter 7.

## 5.5 Discussion

Chapter 4 showed already that the application of independent component analysis (ICA) on ASSR measurements provides performance benefits over single channel methods. In contrast with the MWF–procedure, that exploits the

known modulation frequency, ICA does not use any prior information except for the assumption of independence of the underlying non-Gaussian sources. Nevertheless, it cannot be proven with this data set that the MWF-procedure performs significantly better than ICA. However, when both techniques are compared, the huge difference between the ICA- and the MWF-based method is that the former needs additional components to keep the variability of the results under control. This is not necessary for the MWF-technique, which can be applied directly to the EEG-data. Additionally, the ICA-technique needs prior averaging. It will be shown in Chapter 6 that the MWF-procedure does not need averaging of the signal before processing. Another benefit of the MWF-technique is that a considerably smaller number of code lines is needed and thus less operations per processed sweep. This is appealing for real-time processing.

As both techniques are essentially source separation techniques, they are rather similar. In literature, the link between (generalized) eigenvalue decomposition and blind source separation is already known (Parra and Sajda, 2003). This is also confirmed in Chapter 6. The poor performance for lower SNR values can be observed from the results of both methods and is confirmed by ICA-theory (Hyvärinen et al., 2001). This effect is worse however for ICA-based processing. The saturation effect for five channels is also present in the same, ICA-processed, data set described in Chapter 4. Similar behavior is observed with the MWF-procedure. These similarities are further investigated and confirmed in the theoretical analysis and simulations of Chapter 6.

## 5.6 Conclusion

After a short introduction in Section 5.1, Section 5.2 presents an ASSR signal model that is used to derive the multichannel Wiener filter technique. In Section 5.3, the used experimental setup and analysis methods are described. Section 5.4 shows that a significant improvement in detection speed with a mean of 13 % is possible when recorded multichannel EEG data at intensities above hearing threshold are preprocessed using a multichannel Wiener filtering (MWF) based technique with a priori knowledge through QR factorization. However, the MWF-technique is performing poorly when applied to low SNR data. For this data set, a saturation level is reached at five channels (seven electrodes). Results are similar to those from the same, ICA-processed, data set described in the previous chapter. In Section 5.5, a comparison is made with independent component analysis applied to the same data set in the previous chapter. The next chapter will explain these similarities by fitting them in a theoretical framework and will confirm them using simulations.





## Chapter 6

# A Procedural Framework for ASSR Detection

In Chapters 4 and 5, two multichannel techniques were introduced that processed the recorded multichannel EEG with the main goal to reduce recording time. Results from both techniques were found to be similar. This chapter derives a general procedural multichannel framework for auditory steady-state response detection, originating from the realms of detection theory. Based on this framework, a multichannel processing strategy can be developed starting from a detection theory approach.

Section 6.1 covers a short introduction. It will be shown in Section 6.2 that a *sufficient statistic* can be calculated that best captures the amount of ASSR in the recorded data. This sufficient statistic can exploit spatio-temporal information present in the EEG measurements. The resulting approach can be linked with the multichannel Wiener filter (MWF) approach and independent component analysis (ICA) approach from Chapters 4 and 5. This detection theory based technique is investigated in detail in two ways.

First, simulations are conducted using multichannel EEG combined with artificial ASSRs. These simulations in Section 6.3 are an intermediary step towards the practical evaluation of the approach in Section 6.4. The effects of varying response amplitudes, of ASSRs with a duty-cycle, and of spatially (un)correlated noise bursts appearing in the EEG channels are presented.

---

The material presented in this chapter has been accepted for publication in *IEEE Trans. Biomed. Eng.* as Van Dun, B., Rombouts, G., Wouters, J., and Moonen, M., “A procedural framework for ASSR detection”.

Second, the practical evaluation of the newly developed sufficient statistic based approach is conducted in Section 6.4 using multichannel data from ten normal-hearing adults. Two types of EEG are used: EEG with few artifacts and EEG with a significant number of (controlled) artifacts. The performance is evaluated and compared with an efficient single channel processing technique, which is noise-weighted averaging from John et al. (2001a), and with multichannel approaches as MWF and ICA.

Section 6.5 discusses the results and Section 6.6 finishes this chapter with some conclusions.

## 6.1 Introduction

Responses to acoustic stimuli in general are being recorded for about 80 years (Davis, 1939). In these early days, rather straightforward averaging procedures were developed to lower the background EEG and to improve the ratio between the observed signal and noise (Geisler, 1960). Nowadays, more advanced detection techniques for certain classes of auditory responses have been considered, like those based on e.g. wavelets and Bayesian networks for auditory brainstem responses (Bradley and Wilson, 2004; Zhang et al., 2006). When it comes to auditory steady-state response detection however, only a relatively small number of studies is available. Improvements have been achieved with techniques like adaptive regularized least-squares (RLS) filtering (Tang and Norcia, 1995), noise-weighted averaging (John et al., 2001a), and independent component analysis / multichannel Wiener filtering on multichannel EEG data (Chapters 4 and 5). However, a *general* procedural approach for the detection of auditory steady-state responses in EEG noise has never been presented.

## 6.2 Theoretical background

This section presents a short overview of detection theory (Section 6.2.1) and its application to ASSR detection (Section 6.2.2), including the exploitation of the spatio-temporal structure of the multichannel EEG (Section 6.2.3). A new type of ASSR, using a duty cycle stimulus, is introduced in Section 6.2.4. Links with other multichannel techniques are presented in Section 6.2.5. Finally, Section 6.2.6 describes the performance measurement of the different processing schemes in the simulations of Section 6.3.

### 6.2.1 Detection theory: detecting signals with unknown amplitude in noise

Detection theory is a means to quantify the ability to discern between signal and noise (Green and Swets, 1966). As the auditory steady-state response is buried in noise, originating from sources both inside and outside the skull, an approach that finds its origins in the realm of detection theory seems to be a valid one.

Assume that a target signal  $\tilde{\mathbf{s}}$  has an unknown amplitude  $\alpha$ . The hypotheses  $H_i$  can be stated as

$$H_0 : \mathbf{z} = \mathbf{n} \quad (6.1)$$

$$H_1 : \mathbf{z} = \alpha\tilde{\mathbf{s}} + \mathbf{n}, \quad \alpha = ? \quad (6.2)$$

with  $\mathbf{z}$  a vector containing a number of observations. As the signal's waveform  $\tilde{\mathbf{s}}$  is known exactly, the noise  $\mathbf{n}$  is the only stochastic component,

$$p_{\mathbf{z}|H_i}(\mathbf{z}|H_i) = p_{\mathbf{n}}(\mathbf{z} - \tilde{\mathbf{s}}_i) \quad (6.3)$$

where  $\tilde{\mathbf{s}}_0 = \mathbf{0}$  and  $\tilde{\mathbf{s}}_1 = \alpha\tilde{\mathbf{s}}$ .

This way, only the distribution of the noise is of importance. Assume the noise is colored and Gaussian with a distribution

$$p_{\mathbf{n}}(\mathbf{n}) = \frac{1}{\sqrt{\det(2\pi\mathbf{K})}} \exp\left(-\frac{1}{2}\mathbf{n}^T\mathbf{K}^{-1}\mathbf{n}\right) \quad (6.4)$$

The likelihood ratio  $\Lambda(\mathbf{z})$  can be defined as

$$\Lambda(\mathbf{z}) = \frac{p_{\mathbf{n}}(\mathbf{z} - \alpha\tilde{\mathbf{s}})}{p_{\mathbf{n}}(\mathbf{z})} \quad (6.5)$$

$$= \frac{\exp\left(-\frac{1}{2}(\mathbf{z} - \alpha\tilde{\mathbf{s}})^T\mathbf{K}^{-1}(\mathbf{z} - \alpha\tilde{\mathbf{s}})\right)}{\exp\left(-\frac{1}{2}\mathbf{z}^T\mathbf{K}^{-1}\mathbf{z}\right)} \quad (6.6)$$

This likelihood ratio is compared to a certain threshold  $\eta$  for detection.

$$\Lambda(\mathbf{z}) \underset{H_0}{\overset{H_1}{\geq}} \eta \quad (6.7)$$

If the left hand side is larger than  $\eta$ ,  $H_1$  is chosen, and  $H_0$  otherwise.

The logarithm of both sides is taken and after simplification, assuming that  $\alpha > 0$ , one can write (Johnson and Dudgeon, 1993)

$$\mathbf{z}^T\mathbf{K}^{-1}\tilde{\mathbf{s}} \underset{H_0}{\overset{H_1}{\geq}} \frac{1}{\alpha} \ln \eta + \alpha\tilde{\mathbf{s}}^T\mathbf{K}^{-1}\tilde{\mathbf{s}} \quad (6.8)$$

The left hand side of (6.8) is defined as the *sufficient statistic*  $\Upsilon(\mathbf{z})$ ,

$$\Upsilon(\mathbf{z}) = \mathbf{z}^T\mathbf{K}^{-1}\tilde{\mathbf{s}} \quad (6.9)$$

The right hand side of (6.8) is the threshold term. Despite its explicit dependence on a variety of factors, it is sufficient to determine the threshold term by specifying a false-alarm probability only, as a *uniformly most powerful test* exists. This special case arises when (as in the current case) both the sufficient statistic and one of the hypotheses (here  $H_0$ ) does not depend on the unknown parameter  $\alpha$ . Otherwise, the detection problem cannot be solved without inserting some value for  $\alpha$  (Johnson and Dudgeon, 1993).

### 6.2.2 A detection theory framework for ASSR processing

The ASSR signal model of (5.10), can be rewritten as

$$\mathbf{z} = \alpha \overbrace{\mathbf{d} \otimes \mathbf{s}}^{\tilde{\mathbf{s}}} + \mathbf{n} \quad (6.10)$$

with  $\mathbf{z} = \text{vec}(\mathbf{Z})$  and  $\mathbf{n} = \text{vec}(\mathbf{N})$  respectively.

The sufficient statistic formula (6.9) can be applied here (by identifying (6.10) with (6.2)),

$$\Upsilon_A(\mathbf{z}) = \mathbf{z}^T \mathbf{K}^{-1} \tilde{\mathbf{s}} \quad (6.11)$$

Replacing  $\mathbf{K}$ , by using (5.23), and  $\tilde{\mathbf{s}}$  by using (6.10), one can write

$$\Upsilon_A(\mathbf{z}) = \mathbf{z}^T (\mathbf{K}_{\text{spat}} \otimes \mathbf{K}_{\text{temp}})^{-1} (\mathbf{d} \otimes \mathbf{s}) \quad (6.12)$$

$$= \text{vec}(\mathbf{Z})^T [(\mathbf{K}_{\text{spat}}^{-1} \mathbf{d}) \otimes (\mathbf{K}_{\text{temp}}^{-1} \mathbf{s})] \quad (6.13)$$

$$= \text{vec}[(\mathbf{K}_{\text{temp}}^{-1} \mathbf{s})^T \mathbf{Z} (\mathbf{K}_{\text{spat}}^{-1} \mathbf{d})] \quad (6.14)$$

$$= \mathbf{s}^T \mathbf{K}_{\text{temp}}^{-T} \mathbf{Z} \mathbf{K}_{\text{spat}}^{-1} \mathbf{d} \quad (6.15)$$

Substituting the estimates based on (5.4), (5.20) and (5.26) then leads to a sufficient statistic  $\hat{\Upsilon}_A(\mathbf{z})$  suitable for ASSR detection,

$$\hat{\Upsilon}_A(\mathbf{z}) = \begin{bmatrix} \cos \hat{\varphi} & \sin \hat{\varphi} \end{bmatrix} \mathbf{S}^T \hat{\mathbf{K}}_{\text{temp}}^{-T} \mathbf{Z} \hat{\mathbf{K}}_{\text{spat}}^{-1} \hat{\mathbf{D}} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix} \quad (6.16)$$

### 6.2.3 Exploiting ASSR stationarity and spatio-temporal EEG stationarity

In the following paragraphs, both  $\mathbf{K}_{\text{spat}}$  and  $\mathbf{K}_{\text{temp}}$  are discussed. Different assumptions on the stationarity of the EEG noise lead to different detection procedures. In the first two paragraphs, two approaches to determine  $\mathbf{K}_{\text{spat}}$  are considered. In the third and last paragraph, the structure of  $\mathbf{K}_{\text{temp}}$  is studied.

**$\mathbf{K}_{\text{spat}}$  constant throughout the experiment**

The steering vector  $\mathbf{d}$  may be assumed constant during the same experiment, as described in Section 5.2.1. When the EEG noise is stationary over all channels throughout the experiment,  $\mathbf{K}_{\text{spat}}$  is constant too. The factor  $\mathbf{K}_{\text{spat}}^{-1} \mathbf{d}$  in (6.15) then represents an optimal weight vector  $\mathbf{w}_{\text{opt}}$  that combines the  $m$  EEG channels (columns of observation matrix  $\mathbf{Z}$ ) into one channel, such that the signal-to-noise ratio (SNR) of the combined signal  $\mathbf{Z}\mathbf{w}_{\text{opt}}$  is maximized,

$$\mathbf{w}_{\text{opt}} = \mathbf{K}_{\text{spat}}^{-1} \mathbf{d} \quad (6.17)$$

In (6.16), an estimate of  $\mathbf{w}_{\text{opt}}$  is used, namely

$$\hat{\mathbf{w}}_{\text{opt}} = \hat{\mathbf{K}}_{\text{spat}}^{-1} \hat{\mathbf{D}} \begin{bmatrix} \cos \hat{\varphi} \\ \sin \hat{\varphi} \end{bmatrix} \quad (6.18)$$

When  $\hat{\varphi}$  is computed with (5.20), this  $\hat{\mathbf{w}}_{\text{opt}}$  corresponds to the multichannel Wiener filter (MWF) solution in (5.32), where  $\mathbf{K}_{\text{spat}}$  is equal to  $\mathbf{R}_*^T \mathbf{R}_*$ .

This optimal weight vector can be expressed alternatively as

$$\mathbf{w}_{\text{opt}} = \mathbf{v}_{\max} \left\{ \text{GEVD}(\mathbf{D}\mathbf{D}^T, \mathbf{K}_{\text{spat}}) \right\} \quad (6.19)$$

with  $\text{GEVD}(\cdot)$  the generalized eigenvalue decomposition and  $\mathbf{v}_{\max}\{\cdot\}$  the eigenvector associated with the largest eigenvalue (see also Section 6.2.5).

 **$\mathbf{K}_{\text{spat}}$  varying during an experiment**

When it is assumed that noise sources emerge and disappear uncorrelated over time on the different recorded channels during an experiment,  $\mathbf{K}_{\text{spat}}$  cannot be considered constant anymore and so (6.16) needs to be modified. If the noise in  $\mathbf{Z}$  is stationary only in blocks of  $T$  samples (rows) and if the noise is assumed uncorrelated between such blocks, then  $\Upsilon_A(\mathbf{z})$  can be calculated as the sum of the sufficient statistics  $\Upsilon_{A,i}(\mathbf{z}_i)$ , with each block  $\mathbf{Z}_i$  processed separately,

$$\Upsilon_A(\mathbf{z}) = \sum_{i=1}^{\frac{n}{T}} \mathbf{s}_i^T \mathbf{K}_{\text{temp},i}^{-T} \mathbf{Z}_i \mathbf{K}_{\text{spat},i}^{-1} \mathbf{d} \quad (6.20)$$

$$= \sum_{i=1}^{\frac{n}{T}} \Upsilon_{A,i}(\mathbf{z}_i) \quad (6.21)$$

For each  $\Upsilon_{A,i}(\mathbf{z}_i)$  an approximation as in (6.16) can then be substituted. An indication for the length of  $T$  could be 64 or 128 milliseconds as will be calculated in Section 6.3.3, which may be a good guess of the stationarity of EEG in the frequency region of brainstem ASSRs.

When implemented practically, the steering vector  $\mathbf{d}$  is estimated on the entire observation matrix  $\mathbf{Z}$  as in (5.4), (5.20) and (5.21) to exploit the ASSR stationarity, encompassing ASSR source position, directivity pattern, electrode positions and propagation attenuation. Each stationary block  $\mathbf{Z}_i$  generates a  $\mathbf{K}_{\text{spat},i}$ , as it is assumed that the spatial covariance matrix of each block  $\mathbf{Z}_i$  is different. An optimal weight vector  $\mathbf{w}_{\text{opt},i}$  is calculated for each stationary block  $\mathbf{Z}_i$  separately, and (6.21) sums all  $\hat{\mathbf{Y}}_{A,i}(\mathbf{z})$  into one sufficient statistic  $\hat{\mathbf{Y}}_A(\mathbf{z})$ .

### Structure of $\mathbf{K}_{\text{temp}}$

The temporal noise covariance matrix  $\mathbf{K}_{\text{temp}}$  expresses the correlation of noise samples within the same channel. The simplest case is when  $\mathbf{K}_{\text{temp}}$  is a diagonal matrix  $\sigma^2 \mathbf{I}$ , i.e. when it is assumed that the noise is white with power  $\sigma^2$ , uncorrelated and stationary within the data block. In a more general case  $\mathbf{K}_{\text{temp}}$  is not a diagonal matrix, but assumed to have a Toeplitz structure, such that the correlation between noise samples, and thus the ‘color’ of the noise spectrum, can be incorporated. If an experiment is not entirely stationary, it can be divided in stationary blocks. The practical calculation of  $\mathbf{K}_{\text{temp}}$  is described in Sections 6.3.3 and 6.4.2.

## 6.2.4 Better estimates of $\mathbf{K}$ using a duty cycle stimulus

In the previous section,  $\mathbf{K}$  is calculated using (5.22) with the estimates of  $\mathbf{N}$  based on a QR-factorization in (5.2) to remove the influence of the ASSR present in the observation matrix  $\mathbf{Z}$  in (5.1). However, a better assessment of  $\mathbf{N}$  is available in parts of the data without any stimulus present. To obtain this type of observation matrix, with an ASSR not present at all times, the desired signal matrix  $\mathbf{S}$  in (5.1), representing the ASSR, is modified.  $\check{\mathbf{S}}$  is defined as

$$\check{\mathbf{S}}^{n \times 2} = \begin{bmatrix} \bar{\mathbf{S}}^{\tau \times 2} \\ \mathbf{0}^{(T-\tau) \times 2} \\ \bar{\mathbf{S}}^{\tau \times 2} \\ \vdots \end{bmatrix} \quad (6.22)$$

with  $\tau$  the duration that the desired signal is non-zero and  $\bar{\mathbf{S}}$  the desired signal matrix  $\mathbf{S}$  with reduced length  $\tau$ . Hence,  $T$  is the length of one block containing exactly one duty cycle period  $[\bar{\mathbf{S}}^T \quad \mathbf{0}^T]^T$  of the duty cycle (DC-)ASSR  $\check{\mathbf{S}}$ . The duty cycle  $D = \frac{\tau}{T}$  is the proportion of time during which the desired signal is active (not to be confused with the signal covariance matrix  $\mathbf{D}$  from (5.1)). One has to take care  $\check{\mathbf{S}}$  stays orthogonal, even at small values of  $\tau$ , by checking whether the active part of the desired signal  $\bar{\mathbf{S}}$  contains an integer number of periods. In the simulations of Section 6.3.4,  $T$  will be taken equal to 128 milliseconds and  $\tau$  varies between 0 and 1 with steps of  $\frac{1}{12}$ , or consequently

steps of one signal period (in the simulations, the modulation frequency  $f_m$  is taken equal to 93.75 Hz, or 12 periods in 128 milliseconds).

The observation vector  $\mathbf{Z}$  in (5.1) can be rewritten as

$$\mathbf{Z} = \alpha \check{\mathbf{S}} \mathbf{D}^T + \mathbf{N} \quad (6.23)$$

$$= \alpha \begin{bmatrix} \check{\mathbf{S}} \\ \mathbf{0} \\ \check{\mathbf{S}} \\ \vdots \end{bmatrix} \mathbf{D}^T + \begin{bmatrix} \bar{\mathbf{N}}_k \\ \bar{\mathbf{N}}_{k-1} \\ \bar{\mathbf{N}}_{k-2} \\ \vdots \end{bmatrix} \quad (6.24)$$

with  $\bar{\mathbf{N}}_{k-j}$  an  $m$ -column additive noise matrix with the number of rows proportional to the duty cycle  $D$  of  $\check{\mathbf{S}}$ . Subscript  $k-j$  (with  $j = 0, \dots, \frac{2n}{T} - 1$ ) is the index of one of two parts of a block  $\mathbf{Z}_i$  (out of  $\frac{n}{T}$  blocks).

While it is not obvious to produce a covariance matrix estimate for this problem, an estimate for a covariance matrix for a similar problem is easy to obtain. The observation vector  $\mathbf{Z}$  is split into two parts:  $\mathbf{Z}_s$  containing the components in which the desired signal vector is present, and  $\mathbf{Z}_n$  in which only noise components are present.

$$\mathbf{Z}_s = \alpha \begin{bmatrix} \check{\mathbf{S}} \\ \check{\mathbf{S}} \\ \check{\mathbf{S}} \\ \vdots \end{bmatrix} \mathbf{D}^T + \begin{bmatrix} \bar{\mathbf{N}}_k \\ \bar{\mathbf{N}}_{k-2} \\ \bar{\mathbf{N}}_{k-4} \\ \vdots \end{bmatrix} \quad (6.25)$$

$$\mathbf{Z}_n = \alpha \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \end{bmatrix} \mathbf{D}^T + \begin{bmatrix} \bar{\mathbf{N}}_{k-1} \\ \bar{\mathbf{N}}_{k-3} \\ \bar{\mathbf{N}}_{k-5} \\ \vdots \end{bmatrix} \quad (6.26)$$

Under the similar assumption as in the previous section that each  $T \times m$  data block  $\mathbf{Z}_i$  is uncorrelated with its neighboring blocks, the covariance matrix  $\mathbf{K}_n$  of  $\mathbf{Z}_n$  can be calculated as

$$\mathbf{K}_n = \begin{bmatrix} \bar{\mathbf{n}}_{k-1} \bar{\mathbf{n}}_{k-1}^T & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{n}}_{k-3} \bar{\mathbf{n}}_{k-3}^T & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \bar{\mathbf{n}}_{k-5} \bar{\mathbf{n}}_{k-5}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ddots \end{bmatrix} \quad (6.27)$$

with  $\bar{\mathbf{n}} = \text{vec}(\bar{\mathbf{N}})$ .

Due to the stationarity assumption for a  $T \times m$  data block  $\mathbf{Z}_i$ , also assumed in the previous section,  $\mathbf{K}_n$  of  $\mathbf{Z}_n$  is equal to the noise covariance matrix  $\mathbf{K}_s$  of  $\mathbf{Z}_s$ . The noise characteristics can be determined in the parts where no desired signal

is present, and because of a block assumed to be stationary for  $T$  samples, they remain valid during the adjacent ‘response+noise’ and ‘noise only’ parts of one block  $\mathbf{Z}_i$ . Now, (6.20) can be reformulated as

$$\Upsilon_A(\mathbf{z}) = \sum_{i=1}^{\frac{n}{T}} \bar{\mathbf{s}}_i^T \mathbf{K}_{n,\text{temp},i}^{-T} \mathbf{Z}_{s,i} \mathbf{K}_{n,\text{spat},i}^{-1} \mathbf{d} \quad (6.28)$$

with  $\mathbf{K}_{n,\text{temp},i}$  and  $\mathbf{K}_{n,\text{spat},i}$  the estimates of the temporal and spatial noise covariance matrices for each block  $\mathbf{Z}_i$  (with  $\mathbf{K}_{n,\text{temp},i} \otimes \mathbf{K}_{n,\text{spat},i} = \mathbf{K}_{n,i}$ ).

For each ‘signal+noise’ subblock (with length  $\tau$ ) of block  $\mathbf{Z}_i$ , being stationary and uncorrelated with other blocks,  $\mathbf{K}_{n,\text{spat}}$  can be calculated based on the noise-only data with length  $T - \tau$  in that block  $\mathbf{Z}_i$ . This estimate of  $\mathbf{K}$  should be more precise than the estimates of  $\mathbf{K}$  in (5.22).

The practical calculation of  $\mathbf{D}$ ,  $\mathbf{K}_{\text{spat}}$  and  $\mathbf{K}_{\text{temp}}$  for the DC-ASSR approach is described in Section 6.3.4.

### 6.2.5 Alternative approaches for the calculation of $\mathbf{w}_{\text{opt}}$

The weight vector  $\mathbf{w}_{\text{opt}}$ , as defined in (6.17), combines the  $m$ -channel EEG signal into one channel. It is interesting to note that there exist alternative approaches to calculate a weight vector  $\mathbf{w}'_{\text{opt}}$ , while still reflecting a high degree of similarity with (6.17). For example, an optimal weight vector can also be expressed by

$$\mathbf{w}'_{\text{opt}} = \mathbf{W}\mathbf{U} \quad (6.29)$$

with  $\mathbf{U}$  a ‘selection matrix’ selecting the weight column  $\mathbf{w}'_{\text{opt}}$ . Here,  $\mathbf{w}'_{\text{opt}}$  is defined as the column of  $\mathbf{W}$  that produces a combined signal (independent component) with the largest SNR, computed using (2.11). Matrix  $\mathbf{W}$  represents the separating matrix from (4.3) for an independent component analysis (ICA) approach (Hyvärinen et al., 2001) and can be calculated using (Parra and Sajda, 2003),

$$\mathbf{W} = \text{GEVD}(\mathbf{Q}, \mathbf{K}_{\text{spat}}) \quad (6.30)$$

with  $\mathbf{Q}^{m \times m}$  a ‘cross statistics’ matrix containing the sum of the fourth order cumulants of observation matrix  $\mathbf{Z}$ ,

$$\mathbf{Q} = \mathcal{E}\{\mathbf{Z}^H \mathbf{Z} \mathbf{Z}^H \mathbf{Z}\} - \mathbf{Z}^T \mathbf{Z} \sum \text{diag}(\mathbf{Z}^T \mathbf{Z}) - \mathcal{E}\{\mathbf{Z}^T \mathbf{Z}\} \mathcal{E}\{\mathbf{Z}^H \mathbf{Z}^*\} - \mathbf{Z}^T \mathbf{Z} \mathbf{Z}^T \mathbf{Z} \quad (6.31)$$

Here, (6.31) is valid if the sources are assumed to be non-Gaussian and independent. For other assumptions about the sources, the expression for  $\mathbf{Q}$  takes on different forms (Parra and Sajda, 2003).

The observations in (6.19) and (6.30) show that both approaches can be expressed by a GEVD. They seem to link the theoretical aspects of independent



component analysis (Hyvärinen et al., 2001) and multichannel Wiener filtering (Simmer et al., 2001), and are a possible explanation for the similar results when applying ICA and MWF on e.g. auditory steady-state responses (Chapters 4 and 5).

### 6.2.6 Performance measures

A possible metric for the evaluation of detection performance is the area under the *receiver operating characteristic* (ROC) curve, described in detail in Section 4.2.6. An ROC-curve is a graphical plot of the proportion of correct detections (sensitivity) versus the proportion of false alarms (1 - specificity) for a binary classifier system as its discrimination threshold is varied. The *area under the ROC-curve* (AUC) has an important statistical property: the AUC of a classifier is equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance (Fawcett, 2004). A first advantage of the AUC is that it can be calculated, together with its standard deviations, by means of a small amount of data, as e.g. applied in Chapters 4 and 5. A second advantage is the fact that the AUC does not assume any predefined form of the ROC-curve or, equivalently, the signal and noise distributions.

A major drawback however is the need for experiments with the whole range of discrimination thresholds to construct the ROC-curve. Especially for the problem of ASSR detection a well-founded objection can be raised that only a false alarm rate of 5% and its corresponding specificity are relevant (Picton et al., 2003). The use of other points in the ROC-space only skews the interpretation of the results. As an alternative, the AUC can be linked with an equivalent evaluation method, based on *d-prime*, which is a distance measure between the noise and the signal distribution (Green and Swets, 1966; Macmillan and Creelman, 1991). Generally *d-prime* is taken as the distance between the means of normal signal and noise distributions, normalized by (a combination of) their standard deviations, but definitions may vary. In the case of non-normal distributions however, the use of means and standard deviations becomes insufficient. For the evaluation of the simulations in Section 6.3, a custom tailored *d-prime* is used. The 95 percentile of the noise distribution acts as the decision threshold to determine the *P* percentile of the signal distribution (Figure 6.1). All noise values above the 95 noise percentile are falsely categorized as a signal. All signal values below this threshold are not detected, although definitely present. The lower the value *P*, the higher the detection performance of the observed method to detect a signal in noise, while keeping the chance level for false detections equal to 5%.

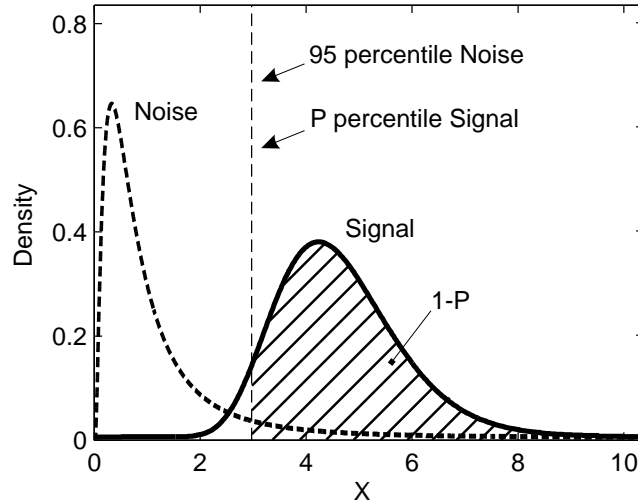


Figure 6.1: Performance evaluation using a custom tailored  $d$ -prime.  $X$  is an observation value with a certain probability density. The 95 percentile of the noise distribution acts as the decision threshold to determine the  $P$  percentile of the signal distribution. All noise values above the 95 noise percentile are falsely categorized as a signal. All signal values below this criterium are not detected, although definitely present. The lower the value  $P$ , the higher the detection performance of the observed method to detect a signal in noise, while keeping the chance level for false detections equal to 5%.

### 6.3 Simulations using multichannel EEG with artificial ASSRs

This section describes a *simulation* setup in Section 6.3.1 that evaluates the framework and the sufficient statistic based ASSR detection approach developed in Section 6.2. Using varying assumptions on its parameters from Section 6.3.2, its performance is compared in Section 6.3.3 with the independent component analysis based approach from Chapter 4 and the multichannel Wiener filtering based approach from Chapter 5. The possible benefits of duty cycle ASSR are discussed in Section 6.3.4. Section 6.3.5 describes the effect of the presence of spatially (un)correlated noise bursts. As a first, incomplete, comparison with a single channel setup, the EEG channel with the largest ASSR SNR is taken without artifact rejection. To complete single channel comparison, Section 6.4 proposes an *evaluation* setup based on real-life multichannel measurements on subjects, evaluating the framework's performance compared with similar multichannel techniques and a noise-weighted common EEG channel.

### 6.3.1 Simulation design

A seven-channel EEG from ten normal-hearing adults was recorded without any acoustic stimulus. Subjects were asked to lie down on a bed with eyes closed and to relax or sleep to minimize the number of artifacts. The electrodes were placed as depicted in Table 4.1, in accordance with the international 10–20 system from Figure 2.2 (Malmivuo and Plonsey, 1995). Recordings were made using a sampling rate of  $f_s = 1000$  Hz and a low-noise Jaeger–Toennies multichannel amplifier with an amplification of 10,000 or 50,000, with bandpass filtering between 70 and 120 Hz (6 dB/octave).

The available EEG ( $\sim 13$  minutes per subject, about 2 hours of EEG data in total) is divided into 950 seven-channel experiments of 8192 samples (8.192 seconds) per channel. A single reference frequency of 93.75 Hz (96 periods in each 1.024 seconds) with varying source amplitude  $\alpha$ , a fixed steering vector  $\mathbf{d}$  throughout all experiments, and a fixed ASSR phase  $\varphi$  within one experiment, is used to construct all artificial ‘EEG + ASSR’ data in this study.

The amplitudes of the ASSR  $\alpha\mathbf{d}$  (at the scalp electrodes) are scaled versions of the following steering vector  $\mathbf{d} = [-0.238 \ 0.327 \ 0.175 \ -0.187 \ 0.726 \ -0.588 \ 1.000]^T$  for all experiments. The values were chosen randomly in a uniform  $[-1, 1]$  range, and scaled to fix the largest channel amplitude equal to 1. The choice of the values in  $\mathbf{d}$  does not represent realistic distributions of the auditory steady-state response on the scalp. However, results are similar when other instances of  $\mathbf{d}$  are taken, meaning that the performance order is similar. The ASSR phases  $\varphi$  were constant over the seven channels. For each experiment however,  $\varphi$  was varied randomly in a uniform  $[0, 2\pi]$  range (for the rationale, we refer to Section 6.2.3).

### 6.3.2 Processing schemes

Ten different processing schemes are evaluated in this chapter. Five of them are based on the sufficient statistic  $\Upsilon_A(\mathbf{z})$  of (6.20), with varying restrictions on  $\mathbf{K}_{\text{spat}}$  and  $\mathbf{K}_{\text{temp}}$ . In this simulation section, some schemes are not evaluated yet (schemes 1, 2, 8 and 9). They will be covered in Section 6.4 but are already included in the list here to concentrate all schemes into one place.

#### Single channel, without artifact rejection – Scheme 0

The channel with the largest SNR (channel 7 according to Section 6.3.1) is used as a reference to define the best single channel performance possible without artifact rejection. Detection is managed using the FFT analysis from (2.11). This scheme is only evaluated in Section 6.3.

### Single channel, with artifact rejection – Scheme 1

This is the standard approach to process data containing ASSRs (John and Picton, 2000a; Lins and Picton, 1995; Luts et al., 2006). The provided EEG data stream is divided into 32 sweeps (data blocks of 16.384 seconds) and averaged. Detection is managed using the FFT analysis from (2.11). This scheme is only evaluated in Section 6.4 as artifact rejection cannot be incorporated in the simulations of Section 6.3 due to the short data segments.

### Single channel, noise-weighted averaging – Scheme 2

Each unfiltered epoch (data block of 1.024 seconds) is transformed to the frequency domain using an FFT. The average power  $P_i$  between 77 and 115 Hz is computed after removing the power at the 8 frequencies at which responses occurred. The time domain epoch is then weighted with the average power  $P_i$  and concatenated with the preceding epochs to form sweeps. Each epoch of the final summed sweep is then divided by the sum of the weights  $P = \sum P_i^{-1}$  of the epochs that has been combined to form that particular epoch (adapted from John et al. (2001a)). A more detailed description of the basics is available in Section 2.2.4. Detection is managed using the FFT analysis from (2.11). This scheme is only evaluated in Section 6.4 as noise-weighted averaging cannot be incorporated in the simulations of Section 6.3 due to the short data segments.

### Independent component analysis – Scheme 3

This approach is presented in Chapter 4. The weight vector  $\mathbf{w}'_{\text{opt}}$  is calculated using (6.29). The selection matrix  $\mathbf{U}$  selects the independent component with the largest SNR, computed using (2.11). As a substitute for (6.30), however, the *joint approximate diagonalization of eigenmatrices* (JADE) algorithm has been used for better numerical stability (Cardoso and Soloumiac, 1993). Before application of the algorithm, data is filtered between 77 and 115 Hz and averaged into one sweep (Section 6.4 only) to increase the SNR of the auditory steady-state responses as ICA does not perform well under conditions with low SNR (Hyvärinen et al., 2001). After reducing the eight channels to one channel  $\mathbf{Z}\mathbf{w}'_{\text{opt}}$ , detection is managed using the FFT analysis from (2.11). This scheme is evaluated in Sections 6.3 and 6.4.

### Multichannel Wiener filter – Scheme 4

This approach is presented in Chapter 5. The weight vector  $\mathbf{w}_{\text{opt}}$  is calculated on the entire observation matrix  $\mathbf{Z}$  using (6.19), which is equal to (5.32). After reducing the eight channels to one channel  $\mathbf{Z}\mathbf{w}_{\text{opt}}$ , detection is managed using the FFT analysis from (2.11). This scheme is evaluated in Sections 6.3 and 6.4.

**$\Upsilon_A(\mathbf{z})$ ,  $\mathbf{K}_{\text{spat}}$  fixed,  $\mathbf{K}_{\text{temp}} = \mathbf{I}$  fixed – Scheme 5**

In (6.20),  $\mathbf{K}_{\text{spat}}$  is estimated only once on the entire observation matrix  $\mathbf{Z}$  using (5.26), or equivalently,  $\mathbf{K}_{\text{spat},i} = \mathbf{K}_{\text{spat}}$  in (6.20).  $\mathbf{K}_{\text{temp}}$  is assumed to be equal to the identity matrix  $\mathbf{I}$ , implying that the EEG noise is assumed to be white, uncorrelated and with constant noise power  $\sigma^2 = 1$ . This scheme is evaluated in Sections 6.3 and 6.4.

 **$\Upsilon_A(\mathbf{z})$ ,  $\mathbf{K}_{\text{spat}}$  fixed,  $\mathbf{K}_{\text{temp}} = \sigma\mathbf{I}$  fixed – Scheme 6**

$\mathbf{K}_{\text{spat}}$  is estimated only once on the entire observation matrix  $\mathbf{Z}$  using (5.26).  $\mathbf{K}_{\text{temp}}$  is now scaled by the square root of the noise power  $\sigma^2$ , calculated by applying an FFT to the single channel result  $\mathbf{Z}\mathbf{w}_{\text{opt}}$ , using  $\mathbf{w}_{\text{opt}}$  from (6.18), and applying (2.11) afterwards. This way,  $\mathbf{K}_{\text{temp}} = \sigma\mathbf{I}$  accounts for the noise power that is varying from observation to observation. It is important to note that only the noise power of the local spectrum near to the ASSR frequency is relevant, providing an estimate of the noise power at the reference frequency. Estimation of  $\mathbf{K}_{\text{temp}}$  using (5.26) is inferior, as useless information of other frequency bands is taken into account. Note that when  $\mathbf{K}_{\text{temp}}$  is chosen to be a Toeplitz matrix (and thus sample correlation is incorporated), performance is found to degrade significantly. While in theory, a corresponding adequate estimation of  $\mathbf{K}_{\text{temp}}$  should increase performance, practically this type of estimation is found to be difficult. This scheme is evaluated in Sections 6.3 and 6.4.

 **$\Upsilon_A(\mathbf{z})$ ,  $\mathbf{K}_{\text{spat},i}$  variable,  $\mathbf{K}_{\text{temp}} = \mathbf{I}$  fixed – Scheme 7**

In (6.20),  $\mathbf{K}_{\text{spat},i}$  is estimated for each block  $\mathbf{Z}_i$ , based on (5.26),

$$\hat{\mathbf{K}}_{\text{spat},i} = \hat{\mathbf{N}}_i^T \hat{\mathbf{N}}_i \quad (6.32)$$

This scheme is evaluated in Sections 6.3 and 6.4.

 **$\Upsilon_A(\mathbf{z})$ ,  $\mathbf{K}_{\text{spat},i}$  variable,  $\mathbf{K}_{\text{temp},i} = \sigma_i\mathbf{I}$  variable – Scheme 8**

$\mathbf{K}_{\text{spat},i}$  is estimated for each block  $\mathbf{Z}_i$  using (6.32).  $\mathbf{K}_{\text{temp},i}$  is scaled by the square root of the noise power  $\sigma_i^2$ , calculated by applying (2.11) to the single channel result  $\mathbf{Z}_i\hat{\mathbf{w}}_{\text{opt},i}$ , using  $\hat{\mathbf{w}}_{\text{opt},i} = \mathbf{v}_{\max}\{\text{GEVD}(\mathbf{d}^T\mathbf{d}, \hat{\mathbf{K}}_{\text{spat},i})\}$  from (6.19). The noise region of  $2 \times 3.7$  Hz for the calculation of  $\sigma^2$  is kept constant. The number of noise frequency bins  $M$  used for the estimate of  $\sigma^2$  thus decreases with smaller block lengths of  $\mathbf{Z}_i$ . This scheme is only evaluated in Section 6.4 as the data segments are too short to obtain a reliable assessment of  $\sigma^2$ .

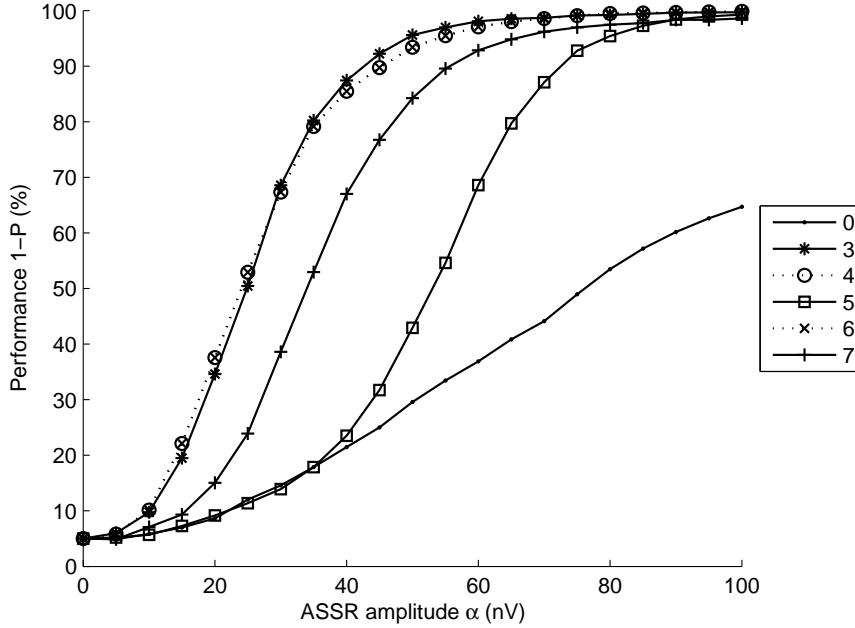


Figure 6.2: Performance (‘1-P’ in Figure 6.1) with relation to ASSR source amplitude  $\alpha$  of the processing schemes 0, 3, 4, 5, 6, and 7 of Section 6.3.2. The performances obtained for schemes 4 and 6 (depicted with dotted lines) coincide.

#### $\Upsilon_A(\mathbf{z})$ , $\mathbf{K}_{\text{spat}}$ fixed, $\mathbf{K}_{\text{temp},i} = P_i P \sigma_n I$ variable – Scheme 9

$\mathbf{K}_{\text{temp},i}$  is calculated based on the rationale of noise-weighted averaging and recalculated for each block  $\mathbf{Z}_i$ .  $P_i$  is the average power between 77 and 115 Hz of  $\mathbf{Z}_i \mathbf{w}_{\text{opt}}$ , with  $\mathbf{w}_{\text{opt}}$  calculated using (6.17).  $P$  is the sum of the reciprocals of  $P_i$ , i.e.  $P = \sum P_i^{-1}$ . A more detailed description of the basics is available in Section 2.2.4. The mean noise power  $\sigma_n^2$  is calculated using (2.11) on a noise-weighted average of  $\mathbf{Z} \mathbf{w}_{\text{opt}}$ . This scheme is only evaluated in Section 6.4 as noise-weighted averaging cannot be incorporated in the simulations of Section 6.3 due to the short data segments.

### 6.3.3 Comparison of the different schemes

When observing Figure 6.2, it is apparent that all multichannel detection schemes are performing better than the single channel approach (scheme 0). This is not surprising, as this scheme does not apply any artifact rejection or noise-weighted averaging. Section 6.4 copes with this problem and describes

a realistic comparison with single channel techniques. The main focus in this section is the performance of the several multichannel techniques described in this research project and their robustness against two types of noise bursts: spatially correlated and uncorrelated noise bursts.

For the construction of the graph of scheme 7 in Figure 6.2, the optimal block size of  $\mathbf{Z}_i$  needs to be determined. Block sizes of 32, 64, 128, 256 and 512 milliseconds are investigated. The ASSR amplitude  $\alpha$  is determined at the  $P = 50\%$  performance point in Figure 6.2 for every block size.  $\alpha$  is equal to respectively 36.4, 33.9, 37.2, 40.0 and 42.6 nV. The optimal block size of  $\mathbf{Z}_i$  in scheme 7 thus is  $T = 64$  milliseconds. It probably corresponds to the time frame in which EEG is still stationary in the 80–120 Hz range. Blocks  $\mathbf{Z}_i$  with larger lengths return a worse estimation of  $\mathbf{K}_{\text{spat}}$  as stationarity is not guaranteed anymore within the same block. Shorter time windows will likely violate the assumption of no correlation between adjacent blocks in (6.21).

When all available channels are used, detection performance is immediately improved. It is important to keep in mind what available information is used however. When  $\mathbf{K}_{\text{temp}}$  is kept equal to the identity matrix (schemes 5 and 7), performance is suboptimal. If the noise amplitude  $\sigma$  is introduced into  $\mathbf{K}_{\text{temp}}$  (scheme 6), the performance increases. Scheme 6 is a special case as the noise amplitude  $\sigma$  used here is identical to  $\sigma$  of (2.11) applied to the multichannel Wiener filter weighted result. As an effect, scheme 6 performs identically compared with multichannel Wiener filtering (scheme 4). A better assessment of  $\mathbf{K}_{\text{temp}}$  in (6.16) and (6.20) could probably improve detection even more.  $\mathbf{K}_{\text{temp}}$  in the simulations is set equal to a noise-scaled identity matrix, as this setting currently seems best. Any other practical estimate of  $\mathbf{K}_{\text{temp}}$  currently degrade performance significantly. Firstly, if the noise power  $\sigma^2$  is calculated on a part of the spectrum that is taken too broad around the reference frequency, performance degrades. Secondly, if correlation between samples (the ‘color’ of the frequency spectrum) can be incorporated,  $\mathbf{K}_{\text{temp}}$  becomes non-diagonal and a better estimate, theoretically. However, all efforts to add non-diagonal non-zero elements to  $\mathbf{K}_{\text{temp}}$  do not improve performance.

The multichannel Wiener filter (scheme 4) performance is similar to the independent component analysis performance (scheme 3). However, as ICA processing does not take into account any a priori information, except the assumption of independent sources, it was expected that the MWF approach, incorporating a known reference frequency, would perform better. This can indicate that the ICA approach is a model sufficient enough for ASSR detection. Results using real subjects already indicated large similarities between MWF and ICA processing in Chapters 4 and 5. This is confirmed in Section 6.4.

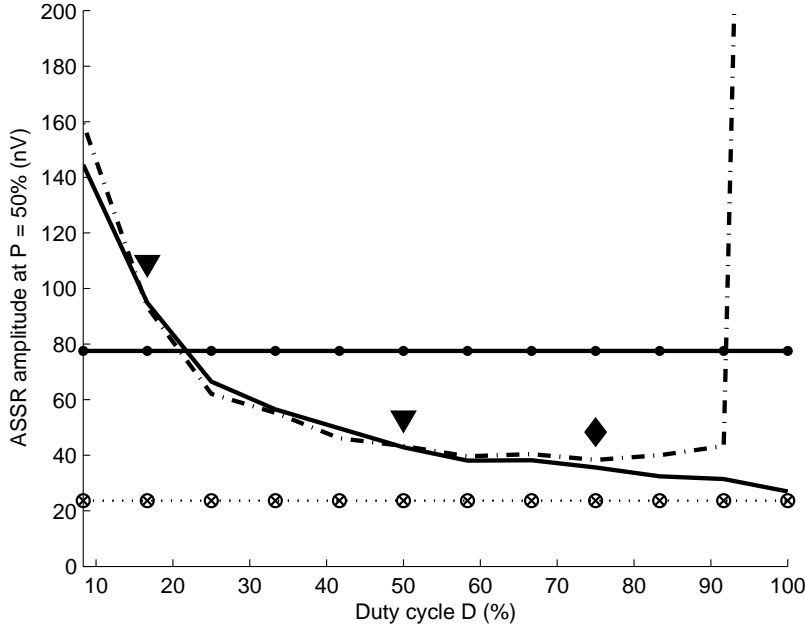


Figure 6.3: ASSR source amplitude  $\alpha$  at  $P = 50\%$  with relation to a varying duty cycle  $D$  for a  $\mathbf{Z}_i$  block size of  $T = 128$  milliseconds.  $D$  increases with steps of  $\frac{1}{12}$ . The horizontal solid line represents the single channel reference with  $D = 100\%$  (scheme 0). The dotted line shows the multichannel Wiener filter approach with  $D = 100\%$  (scheme 4) and  $\Upsilon_A(\mathbf{z})$  processing with  $D = 100\%$  (scheme 6). The solid line indicates the performance of  $\Upsilon_A(\mathbf{z})$  processing with a variable duty cycle  $D$  (scheme 6 with a modified desired signal  $\tilde{\mathbf{S}}$ ). The dash-dotted line shows a variant of scheme 8, with  $\mathbf{K}_{\text{spat},i}$  calculated on noise-only periods and  $\sigma$  assessed on the weighted observation matrix  $\mathbf{Z}$ . The diamond shows the lowest ASSR amplitude achieved for this last approach (38.3 nV at  $D = \frac{9}{12}$ ). The triangles indicate the range ( $\frac{2}{12} \leq D \leq \frac{6}{12}$ ) for which the last approach improves detection compared to  $\Upsilon_A(\mathbf{z})$  processing with a variable duty cycle  $D$  for scheme 6.

### 6.3.4 Varying the duty cycle $D$

Figure 6.3 shows the source amplitude  $\alpha$  of the ASSR that will be detected in 50 % of the experiments, as a function of the duty cycle  $D$  of the desired signal  $\tilde{\mathbf{S}}$  in (6.22). The horizontal lines indicate the  $P = 50\%$  thresholds (in nV) of schemes 0, 4 and 6 from Figure 6.2 for a duty cycle  $D = 100\%$ . Scheme 6 with a variable  $D$  and a variant of scheme 8 are shown as continuous graphs, based on a desired signal  $\tilde{\mathbf{S}}$  with a duty cycle  $D$ . Orthogonality of  $\tilde{\mathbf{S}}$  is



required, implying the use of an integer number of periods of the active desired signal  $\tilde{\mathbf{S}}$ . This makes an increased  $\mathbf{Z}_i$  block size necessary, compared with the previous subsection, to guarantee enough data points for the construction of Figure 6.3. Block sizes of  $T = 128$  milliseconds are now used, which makes 12 data points available for Figure 6.3 instead of 6 data points, as  $T = 128$  milliseconds contains 12 periods of the modulation frequency  $f_m = 93.75$  Hz.

The variant of scheme 8 implements the theory described in Section 6.2.4.  $\mathbf{K}_{n,\text{spat},i}$  is estimated on the noise-only periods, with no desired signal active, as opposed to  $\mathbf{K}_{\text{spat},i} = \mathbf{N}_i^T \mathbf{N}_i$  of the original scheme 8.  $\mathbf{K}_{n,\text{spat},i}$  is practically calculated using the noise-only parts of both blocks  $\mathbf{Z}_i$  and  $\mathbf{Z}_{i-1}$ , or equivalently, referring to (6.24),  $\bar{\mathbf{N}}_k$  is assessed using both  $\bar{\mathbf{N}}_{k-1}$  and  $\bar{\mathbf{N}}_{k+1}$ . This use of adjacent noise-only parts improves the assessment of  $\mathbf{K}_{n,\text{spat},i}$  and is justified as  $\bar{\mathbf{N}}_k$  is assumed to be both stationary ‘with’  $\bar{\mathbf{N}}_{k-1}$  and  $\bar{\mathbf{N}}_{k+1}$  as a result of the small block size  $T$ .

Similar to scheme 8,  $\mathbf{K}_{n,\text{temp},i}$  is taken equal to  $\sigma_i \mathbf{I}$  with  $\sigma_i^2$  the noise power. Unfortunately, the block size  $T$  of  $\mathbf{Z}_i$  is too small to obtain a reliable assessment of  $\sigma_i^2$ . Therefore,  $\sigma_i^2$  is calculated by applying (2.11) on the single channel result  $\mathbf{Z}_i \mathbf{w}_{\text{opt},i}$ , using  $\mathbf{w}_{\text{opt},i} = \mathbf{v}_{\max}\{\text{GEVD}(\mathbf{D}\mathbf{D}^T, \mathbf{K}_{n,\text{spat},i})\}$ . The steering matrix  $\mathbf{D}$  is calculated using (5.4) with  $\tilde{\mathbf{S}}$  from (6.22) as the desired signal.

The best performance using the duty cycle ASSR  $\check{\mathbf{S}}$  is achieved for  $D = 75$  % at  $\alpha = 38.3$  nV. For duty cycles  $D$  between  $\frac{2}{12}$  and  $\frac{6}{12}$ , calculating  $\mathbf{K}_{n,\text{spat},i}$  on the noise-only parts of data blocks  $\mathbf{Z}_i$  and  $\mathbf{Z}_{i-1}$ , improves performance compared to the calculation of  $\mathbf{K}_{\text{spat}}$  on the entire observation matrix  $\mathbf{Z}$ . By introducing less desired signal in (5.1) by means of  $\check{\mathbf{S}}$  from (6.22) with a specific duty cycle  $D$ , less energy of the desired signal  $\tilde{\mathbf{S}}$  is available for detection. Consequently detection performance degrades. This detection performance loss is compensated by the better noise estimates  $\mathbf{K}_{n,\text{spat},i}$  using the noise-only parts that become available. However, this effect is not large enough to compensate fully for the loss of desired signal. Stationarity in one single data block  $\mathbf{Z}_i$  is present, but it is not strong enough.

These results are inferior to those of the multichannel experiments using a  $D = 100$  % duty cycle desired signal  $\mathbf{S}$ . It indicates that full cycle ASSR stimuli are preferred even if the DC-ASSR technique would be physically possible. At these small block sizes  $T$  it is impossible to generate responses with such steep rise and fall times.

### 6.3.5 Robustness against noise bursts

In this section, a series of noise bursts was added to all channels. The noise bursts disrupt the EEG recordings. In the simulations, two types are used: uncorrelated and correlated noise bursts. Uncorrelated noise bursts are added

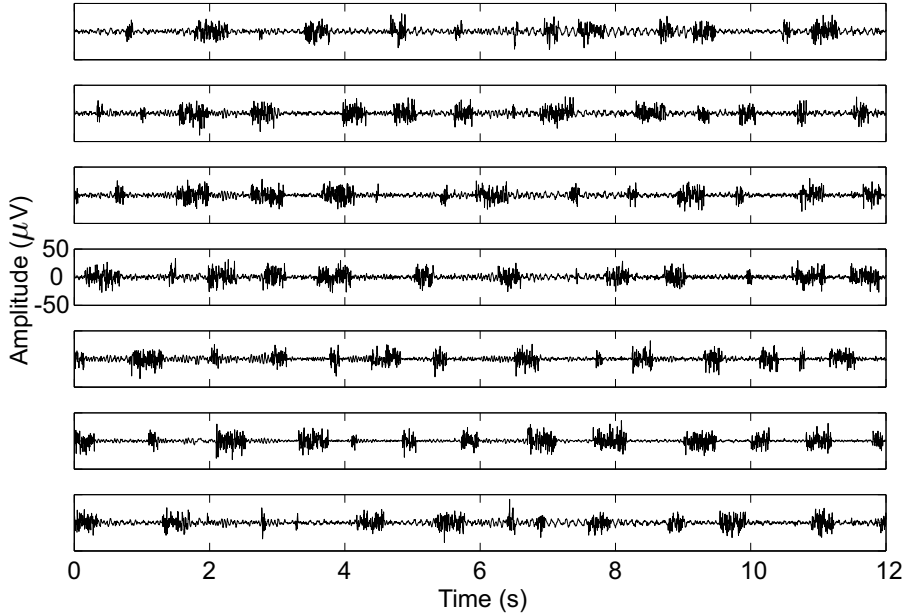


Figure 6.4: Series of spatially uncorrelated noise bursts across 7 EEG channels (first 12 seconds). Each noise burst consists of uniformly distributed white noise. Each noise burst has a maximum length of 512 milliseconds. The distance between two noise bursts is minimally 256 milliseconds and maximally 1024 milliseconds.

that appear spatially uncorrelated on all channels, as shown in Figure 6.4. The noise bursts per channel are maximally 512 milliseconds long, with a jitter period between 256 and 1024 milliseconds. For the correlated noise bursts, the first series of noise bursts is taken and added scaled to the seven EEG channels. Block sizes of  $\mathbf{Z}_i$  are  $T = 64$  milliseconds.

Figure 6.5 shows the simulation results for the seven-channel EEG contaminated with spatially uncorrelated noise bursts from Figure 6.4. Note the difference in ASSR amplitude scale with Figure 6.2. The performance of all schemes except scheme 7 collapses due to the addition of the noise bursts. The impact of spatially uncorrelated noise bursts can thus only be countered when the weight vectors used to combine the available data channels are varied locally within blocks  $\mathbf{Z}_i$ . If this is not the case, the weight vector is not able to select a burst-free channel, and therefore inevitably introduces severe noise corruption in the final result. All schemes that do not introduce a spatial noise covariance matrix  $\mathbf{K}_{\text{spat},i}$ , varying for each block  $\mathbf{Z}_i$ , collapse due to noise influences.

In contrast with the simulations without noise bursts in Section 6.3.3, the multichannel Wiener filter based approach (schemes 4 and 6) now performs

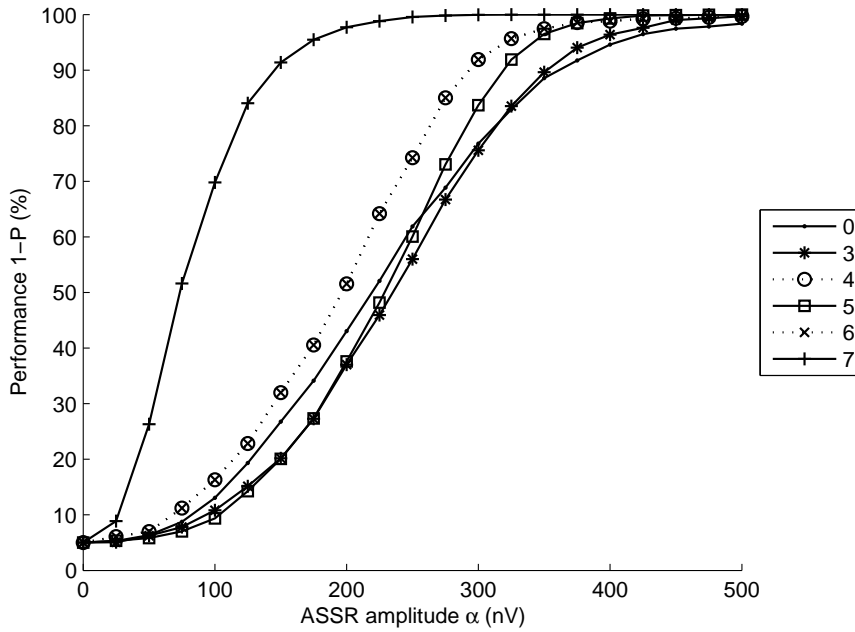


Figure 6.5: Performance (‘1-P’ in Figure 6.1) with relation to ASSR source amplitude  $\alpha$  for schemes 0, 3, 4, 5, 6 and 7 of Section 6.3.2 when imposing spatially uncorrelated noise bursts, as depicted in Figure 6.4. The performances obtained for schemes 4 and 6 (depicted with dotted lines) coincide.

better than the independent component analysis based approach (scheme 3). This shows that ICA does not perform well under conditions with low SNR (Hyvärinen et al., 2001).

Figure 6.6 shows the effects of spatially correlated noise bursts appearing on the different EEG channels. Scaled versions from the first noise burst series in Figure 6.4 are superimposed on all EEG channels. The most important observation is that the benefit of scheme 7 disappears and that the multichannel Wiener filter based approach (schemes 4 and 6) actually performs best. This means that in this case a  $\mathbf{K}_{\text{spat}}$  calculated on the entire observation matrix is sufficient, as the weights are determined such that as much of the noise bursts are eliminated. The performance of the single channel method is obviously poor, but without any artifact rejection or noise-weighted averaging any comparisons are unfair. This issue is covered in Section 6.4. That section will also give an idea which model of noise bursts represents reality best.

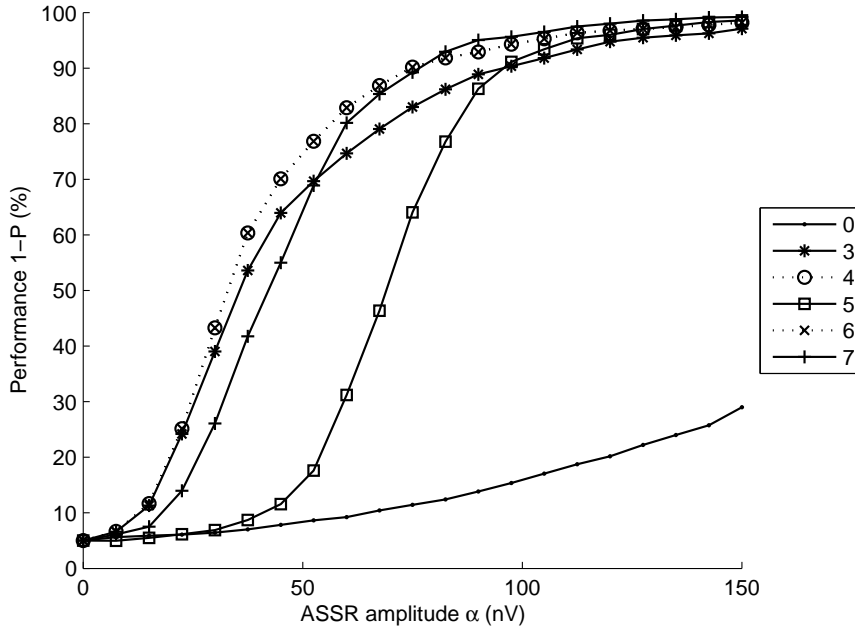


Figure 6.6: Performance (‘1-P’ in Figure 6.1) with relation to ASSR source amplitude  $\alpha$  for schemes 0, 3, 4, 5, 6 and 7 of Section 6.3.2 when imposing spatially correlated noise bursts. The performances obtained for schemes 4 and 6 (depicted with dotted lines) coincide.

## 6.4 Practical evaluation using real-life multi-channel EEG

This section describes an evaluation setup in Section 6.4.1 that is conducted using multichannel data from ten normal-hearing adults. It is a practical extension of the simulations conducted in Section 6.3. Two types of EEG are used: EEG with few artifacts and EEG with a significant number of (controlled) artifacts. The performance of the newly developed sufficient statistic based approach from Section 6.4.2 is evaluated in Section 6.4.3 and compared with an efficient single channel processing technique, noise-weighted averaging from John et al. (2001a), and with multichannel approaches as MWF and ICA from Chapters 4 and 5.

### 6.4.1 Evaluation design

Ten normal hearing subjects (eight male, two female) with mean age 28.2 years (range 22–32 years) were selected. Their hearing thresholds did not exceed 20 dBHL on the octave audiometric frequencies. All experiments were carried out in a double-walled soundproof room with Faraday-cage. Subjects were asked to lie down on a bed and relax or sleep. Lights were switched off. All experiments were identically carried out a second time several days or weeks after the first experiment. This way test-retest comparisons could be done.

Kendall jelly snap electrodes were placed on the positions described in Table 6.1 according to the international 10–20 system from Figure 2.2 (Malmivuo and Plonsey, 1995). This configuration was chosen similar to van der Reijden et al. (2004) with some extra channels added to ensure a symmetrical configuration of all electrodes. They were placed on the subject’s scalp after the skin was abraded with Nuprep abrasive skin prepping gel. A conductive paste was used to keep the electrodes in place and to avoid that inter-electrode impedances exceeded 5 k $\Omega$  at 30 Hz. The electrodes were connected to a low-noise eight-channel Jaeger-Toennies amplifier. Each EEG channel was amplified ( $\times 50,000$ ) and bandpass filtered between 70 and 170 Hz (6 dB/octave). The sampling rate was set equal to 1000 Hz and downsampled later on to 250 Hz. No artifact rejection was applied initially, but a threshold was determined offline that rejected around 10 % of the recorded data blocks (‘epochs’) that exceeded this threshold. All separate acoustic stimuli were calibrated at 70 dB SPL, using a Brüel & Kjær Sound Level Meter 2260 in combination with a 2-cc coupler DB0138 and an artificial ear 4152. All stimuli were presented to the subject and amplified EEG signals recorded using the SOMA program from Chapter 3 and a RME multichannel soundcard.

Two combined stimuli with four 100% amplitude modulated (AM) & 20% frequency modulated (FM) carrier frequencies each, were applied to each ear. The carrier frequencies were the same for both ears, namely 0.5, 1, 2, and 4 kHz. The modulation frequencies were taken close to respectively 82, 90, 98, and 106 Hz for the left ear, and 86, 94, 102, and 110 Hz for the right ear. These modulation frequencies were adjusted to ensure that a non-fractional number of modulation cycles fitted into one data block (‘epoch’) of 1.024 seconds. For example, 82 Hz is converted into  $\frac{\text{round}(82 \cdot 1.024)}{1.024}$  Hz (John and Picton, 2000a). Throughout the rest of the chapter, one will refer to the non-adjusted frequencies for reasons of conciseness.

Stimuli were applied at 30 dB SPL (36 sweeps, each sweep lasting 16.384 seconds). After EEG data collection, each separate channel, or each combination of channels, was reduced to 32 sweeps, using an artifact rejection threshold that removed exactly 4 sweeps per channel (or per combination of channels). Artifact rejection for multichannel data (i.e. a combination of channels) implies the removal of all simultaneous epochs over different channels. Otherwise, any

Table 6.1: Recording electrode positions for a eight-channel setup according to the international 10–20 system (Malmivuo and Plonsey, 1995). All channels are referenced to the reference electrode on top of the head (Cz).

Channel	Position
1	occiput (Oz)
2	P4
3	P3
4	right mastoid (rMa)
5	left mastoid (lMa)
6	F4
7	F3
8	forehead (Fpz)
reference	vertex (Cz)
ground	right clavicle

correlation over simultaneous channels will be removed. This EEG dataset is referred to as *clean EEG* throughout the rest of the chapter.

For the performance analysis of the different processing schemes in Section 6.3.2, the number of ASSR detections is counted. It is assumed that all responses to the applied stimuli are present in the EEG as the stimulus application intensity is 30 dB SPL and well above subject’s hearing thresholds. There are thus 16 responses per subject (test and retest) available for detection. This way, the maximum number of detections is 160. For single channel measurements, the detection threshold is fixed at 4.82 dB SNR using (2.11), which corresponds to 5 % allowed false detections. To calculate a detection threshold for multichannel data, all multichannel processing methods are applied on frequencies without a response (e.g. 1 Hz below the modulation frequencies that could contain a response) of the artifact-free multichannel EEG on the current intensity of 30 dB SPL, extended with extra multichannel EEG from the same subjects. The total EEG data length used here was six times longer than the EEG data length for the detection of the responses. Afterwards, the 95 percentile of this noise distribution is defined as the detection threshold.

For the analysis of robustness against artifacts, the EEG data from the previous paragraph is used together with extra measurements that encouraged the generation of artifacts. Subjects were asked to sit on a chair. They carried out a repeating series of movements in cycles of approximately 6 seconds while the two combined stimuli described above were applied at an intensity of 30 dB SPL. Measurements were 32 sweeps long. No artifact rejection was applied. The movements were carried out in the following order: turn the head up, down, left, and right. This procedure served as a controlled generator of arti-

facts on all channels due to muscle activity and electrode cable movement. This EEG dataset is referred to as *dirty EEG* throughout the rest of the chapter.

### 6.4.2 Processing schemes for evaluation

Processing schemes 1 to 9 from Section 6.3.2 are practically evaluated in this section. Five of them are based on the sufficient statistic  $\Upsilon_A(\mathbf{z})$  of (6.20), with varying restrictions on  $\mathbf{K}_{\text{spat}}$  and  $\mathbf{K}_{\text{temp}}$ . Optimal channels, channel combinations and block sizes are determined in Section 6.4.3.

### 6.4.3 Results

Before the different processing schemes of Section 6.3.2 are evaluated all together for both ‘clean’ and ‘dirty’ EEG, the channel (combination) with the maximum number of ASSR detections is determined so that comparisons are done correctly. This is done for single channel EEG and for the combination of multiple EEG channels.

Table 6.2 shows the number of detections for each of the eight separate channels of Table 6.1. The test–retest statistic for the ‘clean EEG’ (with few artifacts) over all channels is not significant ( $p = 0.129$ , no interactions). The data for the ‘dirty EEG’ (with a significant number of artifacts) is not normally distributed. Based on Table 6.2 it is decided to take channel 1 (vertex–occiput) as the reference channel for the ‘clean EEG’ in the global evaluation further on, as this channel returns the most ASSR detections of all channels. For the ‘dirty EEG’, channel 4 (vertex–right mastoid) is withheld as the reference channel. These optimal EEG derivations are confirmed by van der Reijden et al. (2004), who quote that a small set of 3 derivations (Cz–Oz combined with the right mastoid–Cz and left mastoid–Cz) yield the best SNRs in a larger number of participants than would be expected if all derivations were equally efficient. If one discards channel Cz–P4 from the detection top–four list in Table 6.2, a channel not being used by van der Reijden et al. (2004), the top three from this referenced paper (channels 1, 4 and 5) correspond to their recommended channels. The table finally shows the effect of applying the same artifact rejection level of the ‘clean’ EEG channels to the ‘dirty’ EEG channels. The number of ASSR detections degrades as only few EEG data sweeps are withheld. This shows artifact rejection in the presence of many artifacts is not recommended.

Similarly, Table 6.3 shows the number of detections for a specific set of channels that are combined using (6.16). The ‘clean’ and ‘dirty’ EEG test–retest statistics are not significant ( $p = 0.413$  and  $p = 0.155$  respectively, no interactions). Channels are added incrementally using the order of decreasing number of ASSR detections from the second column of Table 6.2. Due to the statistical multiple testing syndrome, each added channel increases the number of false detections. Therefore, the detection threshold needs to be made more strict,

Table 6.2: Number of ASSR detections for each separate channel from Table 6.1 for three types of EEG. First, ‘clean EEG’ generated by relaxed subjects with about 10 % of epochs removed exceeding a custom rejection threshold. Second, ‘dirty EEG’ generated by controlled head movement without any artifact rejection. Third, ‘dirty EEG’ with the same rejection threshold as for ‘clean EEG’, resulting in a serious length reduction (depicted in the most right column). The channels in bold are the channels with the highest number of ASSR detections for ‘clean EEG’ and ‘dirty EEG’ respectively.

Channel	‘clean EEG’ (w art rej) 36 → 32 sw	‘dirty EEG’ (wo art rej) 32 → 32 sw	‘dirty EEG’ (w art rej) 32 → $X$ sw	Sweeps $X$ after art rej $\hat{\mu}_X \pm \hat{\sigma}_X$
<b>1 (Oz)</b>	<b>116</b>	17	8	$4.9 \pm 8.1$
2 (P4)	104	36	37	$11.0 \pm 10.5$
3 (P3)	90	19	25	$13.2 \pm 10.9$
<b>4 (rMa)</b>	104	<b>52<sup>3</sup></b>	14	$5.2 \pm 7.5$
5 (lMa)	100	27	16	$4.8 \pm 5.9$
6 (F4)	46 <sup>1</sup>	13	4	$3.8 \pm 6.8$
7 (F3)	62 <sup>2</sup>	8	3	$5.8 \pm 8.0$
8 (Fpz)	55 <sup>2</sup>	12	2	$3.0 \pm 5.5$

which is depicted in the fourth column of Table 6.3 for  $p$ -values and SNR. Channel combination 1-3 (121 detections for clean EEG, 23 for dirty EEG) and channel combination 2-3-4-5 (114 detections for clean EEG, 69 for dirty EEG) have the highest number of detections for clean and dirty EEG respectively for all possible combinations (not shown in the table). It is opted for the combination 1-2-3-4-5 however, as this channel combination contains both channels 1 and 4.

For the overall comparison between single channel and multichannel techniques in the rest of this chapter, channels 1 and 4 are used together with channel combination 1-2-3-4-5.

The optimal lengths for partial observation matrices  $\mathbf{Z}_i$  for schemes 7, 8 and 9 are determined in Table 6.4. Block sizes of  $\mathbf{Z}_i$  are varied and the number of ASSR detections are calculated. This way, it is concluded that for schemes 7 to 9, the optimal block length is 8.192 seconds (or 2048 samples) to guarantee the highest number of detections for both ‘clean EEG’ (with few artifacts) and

<sup>1</sup>Significantly different from channel 1 ( $p < 0.01$ ) after pairwise comparison with Bonferroni correction.

<sup>2</sup>Significantly different from channel 1 ( $p < 0.05$ ) after pairwise comparison with Bonferroni correction.

<sup>3</sup>Data not normally distributed. A significant difference ( $p < 0.001$ ) between channels exists.



Table 6.3: Number of ASSR detections in ‘clean’ and ‘dirty’ EEG for individual channels 1 and 4 from Table 6.1 and for specific channel combinations constructed using a decreasing detection order from Table 6.2. Channel combinations are processed using (6.16). The most right column shows the detection thresholds used for each channel (combination). The individual channels in bold are the channels with the highest number of ASSR detections for ‘clean EEG’ and ‘dirty EEG’ respectively. The channel combination in bold returns the most ASSR detections in general.

Channel combination	‘Clean EEG’ (w art rej)	‘Dirty EEG’ (wo art rej)	$p$ -value (%) <SNR (dB)>
<b>1</b>	<b>116</b>	<b>17</b> <sup>5</sup>	5.0 <4.82>
<b>4</b>	<b>104</b>	<b>52</b>	5.0 <4.82>
1-4	117	46 <sup>6</sup>	1.6 <6.28>
1-2-4	111	57	0.97 <6.75>
1-2-4-5	110	56	0.51 <7.32>
<b>1-2-3-4-5</b>	<b>114</b> <sup>4</sup>	<b>66</b>	0.37 <7.58>
1-2-3-4-5-7	111	66	0.23 <7.94>
1-2-3-4-5-7-8	105	63	0.10 <8.52>
1-2-3-4-5-6-7-8	101	65	0.063 <8.81>

‘dirty EEG’ (with a significant number of artifacts). For ‘clean EEG’, this block length could correspond to the stationarity period of the EEG, but 8 seconds seems rather long for EEG stationarity in a frequency window of 70 to 170 Hz (Section 6.4.1). Detection differences between block lengths are not large anyway. Practically the same performance is obtained for data blocks of 256 milliseconds, which could be more plausible (Section 6.3.3). In the case of ‘dirty EEG’ (with the repeated controlled artifact generation), the 8 second period corresponds to the period of repetition. The detection values for this specific block length of 8.192 seconds are used in the overall comparison.

Table 6.5 shows the performance of the schemes, with statistical comparisons, for measurement lengths of 32 sweeps (approximately 9 minutes, one sweep being 16.384 seconds). ‘Clean’ EEG (with few artifacts) and ‘dirty’ EEG (with a significant number of artifacts) as described in Section 6.4.1 are used.

Noise-weighted averaging improves ASSR detection for both single channel as multichannel approaches, especially in the case of ‘dirty EEG’, where a significant number of artifacts is present. This improvement is obtained for channel 1

<sup>4</sup>No significant differences with other channel combinations in the same column.

<sup>5</sup>Significantly different from combination 1-2-3-4-5 ( $p < 0.01$ ) after pairwise comparison with Bonferroni correction.

<sup>6</sup>Significantly different from combination 1-2-3-4-5 ( $p < 0.05$ ) after pairwise comparison with Bonferroni correction.

Table 6.4: Number of ASSR detections in ‘clean’ and ‘dirty’ EEG for schemes 7, 8 and 9 from Section 6.3.2, while varying the block sizes of the partial observation matrices  $\mathbf{Z}_i$ . The block size in bold returns the most ASSR detections in general.

Samples Seconds	32 0.128	64 0.256	128 0.512	256 1.024	512 2.048	1024 4.096	<b>2048</b> <b>8.192</b>	4096 16.384
7 (clean)	69	73	70	74	73	74	<b>74</b>	75
7 (dirty)	54	58	66	71	75	75	<b>82</b>	79
8 (clean)	92	98	96	100	101	104	<b>104</b>	105
8 (dirty)	36	48	49	47	46	46	<b>55</b>	52
9 (clean)	91	117	116	113	116	114	<b>114</b>	115
9 (dirty)	27	61	60	64	69	73	<b>82</b>	78

(17 detections for scheme 1a  $\rightarrow$  35 detections for scheme 2a) and channel 4 (52 detections for scheme 1b  $\rightarrow$  65 detections for scheme 2b), and the multichannel combination 1-2-3-4-5 (66 detections for scheme 6  $\rightarrow$  82 detections for scheme 9). The differences are not always significant.

For the dataset used in this study, the number of ASSR detections is the highest for channel 1 compared with channel 4 for ‘clean EEG’ with few artifacts (116 detections for scheme 1a versus 104 detections for scheme 1b). Channel 1 is the Cz–Oz (vertex–occiput) derivation that is used in several studies (Luts and Wouters, 2005; van der Reijden et al., 2004) and it is similar to the Cz–neck derivation used in other studies (Herdman and Stapells, 2001; John and Picton, 2000b; Lins and Picton, 1995; van der Reijden et al., 2004). However, the number of ASSR detections in channel 1 for ‘dirty EEG’ is lower than the number of ASSR detections in channel 4 (17 detections for scheme 1a versus 52 detections for scheme 1b). Based on this dataset, a difficult choice needs to be made if only three electrodes are available. Channel 1 (vertex–occiput) from literature or channel 4 (vertex–right mastoid) which is more robust against artifacts. If more than three electrodes are available, this choice could be avoided using more channels and applying multichannel signal processing to the recorded data. Moreover, the optimal channels with the highest number of ASSR detections cannot be determined beforehand, together with the fact whether the subject will be relaxed (few artifacts) or stressed (lots of artifacts) during the measurement (or both).

Table 6.5: Number of ASSR detections in ‘clean’ and ‘dirty’ EEG for the schemes from Section 6.3.2. Schemes 7, 8 and 9 use block sizes of 8.192 seconds for  $\mathbf{Z}_i$ , according to Table 6.4. The single channel schemes in bold are the schemes with the highest number of ASSR detections for ‘clean EEG’ and ‘dirty EEG’ respectively. The multichannel scheme in bold returns the most ASSR detections in general.

Scheme	Description	‘Clean EEG’	‘Dirty EEG’
1a	normal averaging (ch 1)	116	17 <sup>9</sup>
1b	normal averaging (ch 4)	104	52 <sup>10</sup>
<b>2a</b>	<b>weighted averaging (ch 1)</b>	<b>120</b>	<b>35</b> <sup>11</sup>
<b>2b</b>	<b>weighted averaging (ch 4)</b>	<b>109</b>	<b>65</b>
3	ICA	124 <sup>8</sup>	53
4 <sup>7</sup>	MWF	114	66
5	$\mathbf{K}_{\text{spat}}$ fixed, $\mathbf{K}_{\text{temp}} = \mathbf{I}$ fixed	78	98
6 <sup>7</sup>	$\mathbf{K}_{\text{spat}}$ fixed, $\mathbf{K}_{\text{temp}} = \sigma\mathbf{I}$ fixed	114	66
7	$\mathbf{K}_{\text{spat},i}$ var, $\mathbf{K}_{\text{temp},i} = \mathbf{I}$ fixed	74	82
8	$\mathbf{K}_{\text{spat},i}$ var, $\mathbf{K}_{\text{temp},i} = \sigma_i\mathbf{I}$ var	104	55 <sup>10</sup>
<b>9</b>	<b><math>\mathbf{K}_{\text{spat}}</math> fixed, <math>\mathbf{K}_{\text{temp},i} = P_i P \sigma_n \mathbf{I}</math> var</b>	<b>114</b>	<b>82</b>

The best single channel results (schemes 2a and 2b) can be compared separately with the best multichannel result (scheme 9), thus discarding the pairwise comparisons and the Bonferroni correction. For the ‘dirty EEG’, scheme 9 produces significantly ( $p < 0.001$ ) more ASSR detections (82 detections) than scheme 2a (35 detections), which corresponds to the noise-weighted averaging of channel 1. Compared with the noise-weighted averaging of channel 4 (65 detections for scheme 2b), the increased number of ASSR detections (82 detections for scheme 9) is close to significance ( $p = 0.063$ ). Thus, when the reference channel is taken as channel 1 (Cz–Oz) for most single channel measurements, the application of scheme 9 to channel combination 1-2-3-4-5 outperforms a noise-weighted version of channel 1 (scheme 2a) significantly for EEG corrupted with artifacts. For EEG with few artifacts, the difference between approaches is not significant.

<sup>7</sup>The results from schemes 4 and 6 are identical.

<sup>8</sup>Significantly different from scheme 8 ( $p < 0.05$ ) after pairwise comparison with Bonferroni correction.

<sup>9</sup>Significantly different from schemes 2b, 7 and 8 ( $p < 0.05$ ), schemes 4, 5 and 6 ( $p < 0.01$ ) and scheme 9 ( $p < 0.001$ ) after pairwise comparison with Bonferroni correction.

<sup>10</sup>Significantly different from scheme 9 ( $p < 0.01$ ) after pairwise comparison with Bonferroni correction.

<sup>11</sup>Significantly different from schemes 5 and 9 ( $p < 0.05$ ) after pairwise comparison with Bonferroni correction.

Independent component analysis (scheme 3) and multichannel Wiener filtering (scheme 4) are similar in performance. However, as ICA processing does not take into account any a priori information, except the strong assumption of independent sources, it was expected that the MWF approach, incorporating a known reference frequency, would perform better. This could indicate that the ICA approach is sufficient for ASSR detection. Results in previous chapters already indicated large similarities between MWF and ICA processing (Chapters 4 and 5). The biggest drawback of the ICA approach is that the available data needs to be filtered and averaged into smaller data blocks first. Otherwise, ASSR amplitudes are too small in the observation matrix and the ICA algorithm cannot separate them from the noise. This shows that ICA does not perform well under conditions with low SNR (Hyvärinen et al., 2001).

The multichannel Wiener filtering approach (scheme 4) and scheme 6 ( $\mathbf{K}_{\text{spat}}$  fixed,  $\mathbf{K}_{\text{temp}} = \sigma \mathbf{I}$ ) are identical in performance. The approaches are characterized entirely by the optimal weight vector  $\mathbf{w}_{\text{opt}}$  and the (square root of the) noise power  $\sigma^2$ .

Varying  $\mathbf{K}_{\text{spat}}$  reduces detection performance for multichannel EEG recordings (scheme 5  $\rightarrow$  scheme 7, scheme 6  $\rightarrow$  scheme 8). This is rather unexpected. Pilot simulations from Section 6.3 using data containing artificial ASSRs showed that spatially uncorrelated high-intensity noise bursts scattered over different EEG channels containing ASSRs could only be processed efficiently using a varying  $\mathbf{K}_{\text{spat},i}$  per data block  $\mathbf{Z}_i$ . Within this data block, an uncorrupted channel could be selected from the several channels of the multichannel data block. Additional pilot simulations in Section 6.3 with spatially correlated noise bursts across channels indicated however that the effect of varying  $\mathbf{K}_{\text{spat}}$  performed similar or even worse than keeping  $\mathbf{K}_{\text{spat}}$  fixed. The results on EEG data from real subjects provided in this chapter show that the second assumption may be correct. High-intensity noise bursts in the form of muscle or movement artifacts emerge spatially correlated on the different EEG channels and can be reduced by choosing an appropriate  $\mathbf{K}_{\text{spat}} \sim \mathbf{w}_{\text{opt}}$  for the entire observation matrix.

$\mathbf{K}_{\text{temp}}$  models the (local) characteristics of the EEG signal. Increasing the precision of the estimation of  $\mathbf{K}_{\text{temp}}$  improves ASSR detection performance. When considering schemes 5, 6 and 9, a gradual increase in the total number of detections ('clean EEG' + 'dirty EEG') is observed. Scheme 5 does not incorporate any noise information at all in  $\mathbf{K}_{\text{temp}}$ . Its total number of detections is the lowest ( $78 + 98 = 176$  detections). The high value ( $98$  detections for scheme 5) of the 'dirty EEG' dataset indicates that EEG with a significant number of artifacts is better off without any noise power estimation as the power of the artifacts (noise bursts) is also present in  $\sigma^2$  ( $98$  detections for scheme 5  $\rightarrow$   $66$  detections for scheme 6). Omitting  $\sigma^2$  has a negative effect on 'clean EEG' however ( $114$  detections for scheme 6  $\rightarrow$   $78$  detections for scheme 5). This opposing effect can be solved by a noise-weighted approach, weighting

each data block  $\mathbf{Z}_i$  with the noise  $P_i$  in that block. The effect of the noise bursts protruding in  $\sigma^2$  is still present (98 detections for scheme 5  $\rightarrow$  82 detections for scheme 9), but better assessed locally by the  $P_i P$  terms (66 detections for scheme 6  $\rightarrow$  82 detections for scheme 9).

A more precise assessment of  $\mathbf{K}_{\text{temp}}$  probably can be achieved by making  $\mathbf{K}_{\text{temp}}$  non-diagonal, incorporating any frequency domain information (the ‘color’ of the noise). Perhaps the fine-structured information of the EEG can be modelled too, taking blocks  $\mathbf{Z}_i$  with a length close to dimensions where EEG stationarity may be assumed (e.g. less than 100 milliseconds). Neither approach succeeds on the EEG dataset used in this study however.

## 6.5 General discussion

Most results have already been discussed in Section 6.4.3. Some general remarks are discussed below.

This is the first time the authors know of that a detection theory approach is investigated for the processing of multichannel EEG data with the main purpose of improving ASSR detection. The presented results support the (simplified) framework and the detection theory approach described in this chapter. When appropriate multichannel processing is applied (multichannel Wiener filtering, independent component analysis or the newly proposed procedural framework based on detection theory), multichannel measurements demonstrate an improvement compared to (noise-weighted) single channel ASSR recordings. When its performance is compared to a noise-weighted version of a standard electrode configuration for single channel measurements (vertex-occiput from Luts et al. (2006) and from van der Reijden et al. (2004)), this improvement is significant. The observations above are only valid when the available EEG is corrupted by artifacts originating from realistic head movements. Otherwise, performance is similar.

When focussing on situations where hearing threshold determination is of clinical relevance, the proposed method  $\Upsilon_A(\mathbf{z})$  implies a serious improvement for its robustness against artifacts. Unless the subject is sedated, ASSR measurements are difficult to conduct in a short period of time because of a large number of artifacts due to movement, distress and agitation, particularly in the patient population where nowadays ASSRs are mostly being applied (in neonates and young children). This type of recording sessions are rarely described in literature, but are unfortunately very current in clinical settings. The proposed procedural framework in this chapter could be highly useful in these cases.

## 6.6 Conclusions

After a small introduction in Section 6.1, a simplified procedural framework is proposed in Section 6.2 that allows the development of a multichannel processing strategy for ASSR detection starting from a detection theory approach. It is shown that a *sufficient statistic* can be calculated that best captures the amount of ASSR in the observation matrix. This sufficient statistic based approach can exploit spatio-temporal stationarity present in the EEG measurements and can be linked with the development and application of the multichannel Wiener filter (MWF) approach and independent component analysis (ICA) based approach to ASSR detection.

Section 6.3 describes simulations that are conducted using EEG data from 10 subjects and artificial ASSRs with varying amplitude and phase. Comparisons between several multichannel approaches are made. The proposed sufficient statistic based approach is identical to the MWF based approach from Chapter 5, and similar to the ICA based approach from Chapter 4 for uncontaminated EEG. A stimulus with a duty cycle of 100 % is preferred. When the EEG is contaminated by spatially uncorrelated artifacts, the proposed technique is performing better than the ICA and MWF based approach. When contamination occurs by spatially correlated artifacts, all multichannel techniques are performing similar.

Section 6.4 works out a practical evaluation using multichannel EEG data from ten normal-hearing adults. Two types of EEG are used: EEG with few artifacts and with a significant number of (controlled) artifacts. It is concluded that most single- and multichannel approaches are similar in performance when applied to uncontaminated EEG. When artifact-rich EEG is used, the proposed detection theory based approach improves the number of ASSR detections compared with the noise-weighted average of the best channel of this dataset (vertex-right mastoid). In general this best electrode configuration cannot be known beforehand. When compared with a noise-weighted common EEG channel derivation (vertex-occiput), the proposed approach improves ASSR detection significantly.

Although the optimal electrode configuration for each individual subject is impossible to know a priori, Chapter 7 determines a minimum set of electrodes that guarantees to deliver a close-to-optimal electrode configuration in terms of response detection for as many subjects as possible.

## Chapter 7

# Optimal electrode selection for multichannel EEG based detection of auditory cortex and brainstem ASSRs

Chapter 6 presented a multichannel processing technique for more than one EEG channel that improved the number of ASSR detections compared with single channel measurements. In the previous chapter it was not addressed however which multiple channels should be used for optimal detection. This chapter evaluates the practical performance with adults of the multichannel processing strategy proposed in Chapter 6. Its results should serve as an intermediate step for application of the multichannel technique to infants.

Section 7.1 starts this chapter with a small introduction. Sections 7.2 and 7.3 describe a setup using eight-channel EEG measurements from ten normal-hearing adults. First, the EEG derivation with the best estimation of the behavioral hearing threshold for ASSRs originating mainly from the brainstem is determined (modulation frequencies between 80 and 110 Hz). For ASSRs originating mainly from the auditory cortex (here, with a modulation frequency of 10 Hz), the minimum EEG derivation with the largest SNR is determined similarly. Afterwards, the minimum EEG channel combination for brainstem

---

The material presented in this chapter has been submitted for publication (Van Dun et al., 2008a).

ASSRs with a close-to-optimal estimation of the behavioral hearing threshold for as many subjects as possible is presented. For auditory cortex ASSRs, the EEG channel combination with the close-to-largest SNR for a maximum number of subjects is presented. Finally, the robustness of the proposed multichannel processing technique against artifacts is evaluated. Section 7.4 discusses the results and Section 7.5 ends this chapter with the conclusions.

## 7.1 Introduction

Many research studies have focused on optimizing detection of ASSRs. First, different stimulus types have been investigated. Amplitude and/or frequency modulated stimuli are most widely accepted (John et al., 2001b). In addition, the application of other stimuli has been studied, such as stimuli with exponential modulation envelopes (John et al., 2002a), broadband and band-limited noise (John et al., 2003), stimuli composed of several carriers modulated with the same modulation frequency (Stürzebecher et al., 2001), and complex stimuli with broader frequency spectra (Riquelme et al., 2006; Stürzebecher et al., 2006). In general, responses are more difficult to detect as the stimuli become more frequency specific. Second, normal averaging has been compared with weighted averaging and artifact rejection (John et al., 2001a). Third, the possibility of increasing the efficiency of ASSR detection by means of multichannel EEG recordings has been evaluated (van der Reijden et al., 2004). Finally, several statistical decision criteria have been studied: tests that use spectral phase information, spectral amplitude information, or both; tests that evaluate only the first harmonic of the response (Cebulla et al., 2001; Dobie and Wilson, 1996) and recently, tests that include higher harmonics for response detection (Cebulla et al., 2006). Most of these techniques have been discussed in detail in Section 2.2.

Most of these studies however are based on EEG obtained from subjects who were instructed to relax or sleep. These experiment conditions can be considered as almost ideal. In practice, however, and especially when e.g. newborns are tested, numerous artifacts prevent a successful measurement unless one takes appropriate measures against these artifacts. This can be achieved by using artifact rejection or weighted averaging (John et al., 2001a). The downside of these techniques is the unacceptably long measurement time in the case only one EEG channel is used for EEG signal recording, as data blocks severely corrupted with artifacts are entirely discarded or are not allowed to contribute much to the final result. Therefore, other techniques have been investigated that are more robust against artifacts. These techniques should allow to detect significantly more responses in artifactual EEG without extending measurement time.



The Cz–Oz or Cz–neck derivations are the most commonly used derivations for adults (John et al., 2001a; Lins and Picton, 1995; Luts and Wouters, 2005). The choice for these derivations lies in the general experience that the mean ASSR thresholds are smaller for these EEG derivations. This experience is confirmed for the Cz–Oz derivation by van der Reijden et al. (2004), who use a multichannel setup to search for the EEG derivations offering the largest signal-to-noise (SNR) values and thus the smallest ASSR thresholds for adults. van der Reijden et al. (2004) also indicate that this Cz–Oz derivation does not always guarantee the largest SNR and that sometimes other derivations (like e.g. Cz–mastoid, Cz–Pz, and Cz–neck) could offer larger SNRs. This is backed by the observation that ASSR thresholds (and their SNRs) show large variations across subjects (John et al., 2001a; Lins and Picton, 1995; Luts and Wouters, 2005). The problem is that the derivation with the largest SNR is subject dependent and cannot be predicted beforehand. Recording EEG using multiple derivations simultaneously, combined with appropriate multichannel processing, is required to select the ‘best’ channel (or combination of channels) for as many subjects as possible.

## 7.2 Methods

Two types of studies will be described here: one based on ASSRs mainly evoked in the brainstem (modulation frequencies between 80 and 110 Hz) and one based on ASSRs mainly evoked in the auditory cortex (10 Hz). The theoretical background for this chapter is entirely described in Section 6.2.

The 80–110 Hz range is the frequency region of interest for hearing threshold determination (Picton et al., 2003). Several studies focus on the sources of these ASSRs. Herdman et al. (2002) suggest that 88 Hz ASSRs are mainly generated in the brainstem. Kuwada et al. (2002) indicate that ASSRs with modulation frequencies below 80 Hz mainly originate in the auditory cortex. Above 80 Hz, there appears to be at least two generators that are likely subcortical. This is confirmed by Purcell et al. (2004) who use response latencies to derive the source locations. They conclude that for modulation frequencies above about 75 Hz, most of the ASSR is generated in the brainstem. Above 95 Hz, the source lies entirely in the brainstem. As a result, responses to stimuli modulated with frequencies between 80 and 110 Hz are referred to as ‘responses mainly evoked in the brainstem’.

The modulation frequency of 10 Hz is chosen because its relevance for speech envelope modulations and speech perception (Shannon et al., 1995). When lowering the modulation frequency of the stimulus, the main source of the auditory steady-state response shifts more to the auditory cortex. According to Herdman et al. (2002), the main source can be found in the auditory cortex,

but a source originating from the brainstem may still be present. This is supported by Purcell et al. (2004), showing that the brainstem still contributes at lower modulation frequencies, but with the main source of the low-frequency ASSR placed in the auditory cortex. Responses to 10 Hz-modulated stimuli are therefore referred to as ‘responses mainly evoked in the auditory cortex’.

For both studies, 10 normal hearing subjects (8 male, 2 female) with mean age 28.2 years (range 22–32 years) were selected. Their hearing thresholds did not exceed 20 dBHL on the octave audiometric frequencies. All experiments were carried out in a double-walled soundproof room with Faraday-cage. Subjects were asked to lie down on a bed and relax or sleep for the brainstem study in Section 7.2.1, and to stay awake by watching a movie for the auditory cortex study in Section 7.2.2. Lights were switched off. This procedure was repeated a number of weeks later to collect both test and retest data.

The used setup is identical to the setup of Section 6.4.1. Kendall jelly snap electrodes were placed on the positions described in Table 7.1 according to the international 10–20 system from Figure 2.2 (Malmivuo and Plonsey, 1995). This configuration was chosen similar to van der Reijden et al. (2004) with some extra channels added to ensure a symmetrical configuration of all electrodes. They were placed on the subject’s scalp after the skin was abraded with Nuprep abrasive skin prepping gel. A conductive paste was used to keep the electrodes in place and to avoid that inter-electrode impedances exceeded 5 k $\Omega$  at 30 Hz. The electrodes were connected to a low-noise eight-channel Jaeger-Toennies amplifier. Each EEG channel was amplified ( $\times 50,000$ ) and bandpass filtered between 70 and 170 Hz (6 dB/octave) for the brainstem study in Section 7.2.1, and between 1 and 30 Hz (6 dB/octave) for the auditory cortex study in Section 7.2.2. The sampling rate was set equal to 1000 Hz, and downsampled afterwards to 250 Hz. No artifact rejection was applied initially, but a threshold was determined offline that rejected around 10 % of the recorded data blocks (‘epochs’) that exceeded this threshold. All separate acoustic stimuli were calibrated at 70 dB SPL, using a Brüel & Kjær Sound Level Meter 2260 in combination with a 2-cc coupler DB0138 and an artificial ear 4152. All stimuli were presented to the subject and amplified EEG were signals recorded using the SOMA program from Chapter 3 and an RME multichannel soundcard.

It is stressed that all possible measures have been taken to avoid false responses not originating from statistical noise effects (the relative number of false responses originating from statistical noise is allowed to be 5 %). Measures to avoid additional false responses include appropriate shielding of stimulus devices and electrodes, filtering prior to downsampling, and adequate separation of stimulus and electrode cables. The absence of false responses was regularly checked by applying the stimuli to a simulated ‘deaf’ subject (by blocking the ear canal of the subject using ear plugs).

Table 7.1: Recording electrode positions for an eight-channel setup according to the international 10–20 system (Malmivuo and Plonsey, 1995). All channels are referenced to the reference electrode on top of the head (Cz). The term ‘Side’ indicates on which hemisphere of the head the electrode is placed in the case it is not positioned on the middle line from front to back. In general, the sides can be ‘Left’ or ‘Right’. However, for ASSRs from the auditory cortex, being described in Section 7.3.2 (with stimuli referenced to the right ear), one refers to the sides as ‘Ipsilateral’ and ‘Contralateral’. The ipsilateral side is the side of stimulation.

Channel	Position	Side
1	occiput (Oz)	
2	P4	R / I
3	P3	L / C
4	right mastoid (rMa)	R / I
5	left mastoid (lMa)	L / C
6	F4	R / I
7	F3	L / C
8	forehead (Fpz)	
reference	vertex (Cz)	
ground	right clavicle	

### 7.2.1 Brainstem stimulation

Two combined stimuli with four 100% amplitude modulated (AM) and 20% frequency modulated (FM) carrier frequencies each, were applied to each ear. The carrier frequencies were the same for both ears, namely 0.5, 1, 2, and 4 kHz. The modulation frequencies were chosen close to respectively 82, 90, 98, and 106 Hz for the left ear, and 86, 94, 102, and 110 Hz for the right ear. These modulation frequencies were adjusted to ensure that a non-fractional number of modulation cycles fitted into one data block (‘epoch’) of 1.024 seconds. For example, 82 Hz is converted into  $\frac{\text{round}(82 \cdot 1.024)}{1.024}$  Hz (John and Picton, 2000a). Throughout the rest of the chapter, one will refer to the non-adjusted frequencies for conciseness.

For Section 7.3.1, stimuli were applied at a stepwise decreasing intensity from 60 to 10 dB SPL, and lowered by 10 dB after a trial of approximately 10 minutes per intensity (36 sweeps, each sweep lasting 16.384 seconds). After EEG data collection, each separate channel, or each combination of channels, was reduced to 32 sweeps, using an artifact rejection threshold that removed exactly 4 sweeps per channel (or per combination of channels). Artifact rejection for multichannel data (i.e. a combination of channels) implies the removal of all simultaneous epochs over different channels to preserve correlation over

simultaneous channels.

For the analysis of robustness against artifacts in Section 7.3.1, the EEG data from the previous paragraph is used together with extra measurements that emphasized the occurrence of artifacts. Subjects were instructed to sit on a chair, and carried out a repeated series of movements in cycles of approximately 6 seconds while the two combined stimuli described above were applied at an intensity of 30 dB SPL. Measurements were 32 sweeps long. No artifact rejection was applied. The movements were carried out in the following order: turn the head up, down, left, and right. This procedure served as a controlled generator of artifacts on all channels due to muscle activity and electrode cable movement.

## 7.2.2 Auditory cortex stimulation

In contrast with the ASSRs generated mainly in the brainstem (Dimitrijevic et al., 2002; Herdman and Stapells, 2001; Luts and Wouters, 2005), the use (and consequently their mutual interactions) of simultaneous stimuli with different modulation frequencies at 40 Hz and below has not yet been thoroughly documented (Armstrong and Stapells, 2007). Therefore, only one carrier was modulated (instead of eight as in Section 7.2.1), which implies that only one ear was stimulated.

White noise was spectrally weighted to conform to the average spectrum of spoken Flemish sentences, which was available as the *Leuven intelligibility sentences test* (LIST) from Van Wieringen and Wouters (2008). This speech-weighted noise was 100 % amplitude modulated with an adjusted modulation frequency close to 10 Hz. The resulting stimulus was applied to the ear with the smallest pure-tone average (PTA), with an intensity of 50 and of 70 dB SPL. Multichannel EEG recordings had a length of 23 sweeps, reduced to 20 sweeps by removing epochs that exceeded the calculated artifact rejection threshold (per channel or per combination of channels), similar to Section 7.2.1.

## 7.3 Results

This section addresses the following questions:

- What is the most appropriate single EEG channel to record brainstem ASSRs and auditory cortex ASSRs respectively?
- What is the most appropriate combination of multiple EEG channels to record brainstem ASSRs and auditory cortex ASSRs respectively?
- Does the side of stimulation influence the conclusions resulting from the questions stated above?

### 7.3.1 Brainstem stimulation

All results in this section are based on the concept of difference scores. A *difference score* is the difference (in dB) between ASSR thresholds (in dBSPL) and behavioral thresholds (in dBSPL). The difference score is used in various ASSR studies (Dimitrijevic et al., 2002; Herdman and Stapells, 2001; Luts and Wouters, 2005). In this study, it allows to combine the available SNR data at six different intensities into one quantitative value. It averages out the variation in behavioral thresholds between subjects and between test sessions within the same subject.

The *ASSR threshold* is defined as the lowest intensity that still produces an auditory steady-state response at the corresponding modulation frequency. In the current study this definition is more strict to avoid false detections. A threshold (in dBSPL) is only accepted as threshold if at more than half of the levels (in steps of 10 dBSPL) up to 60 dBSPL a threshold is obtained. For example, one does not define a threshold at 30 dBSPL in the case stimulus levels at 30 and 40 dBSPL show a response, but those at 50 and 60 dBSPL do not. However, if a response is found at 20, 30, and 40 dBSPL, the ASSR threshold is defined at 20 dBSPL. If no threshold is found at 60 dBSPL, the ASSR threshold in this case is defined at 70 dBSPL.

First, parts of the EEG data are presented using SNRs. Then, the conversion to difference scores is justified. The rest of this section will present the results using difference scores.

#### Separate channels

For each EEG channel, Table 7.2 displays the mean SNRs (and standard deviations) of the eight modulation frequencies, together with the root-mean-square (RMS) values of the noise between 77 and 115 Hz with the modulation frequencies omitted, when stimulating at 30 dBSPL. It shows the means over 10 subjects, with the test and retest data averaged out for each subject (test-retest not significant over all intensities,  $p = 0.404$ , no interactions). For conciseness, the values at other intensities (60 to 40, 20, and 10 dBSPL) are not shown.

When the difference scores over all frequencies and all subjects are compared with the SNRs over all frequencies and all subjects, the following Spearman correlations  $r_s$  are obtained per stimulus intensity ( $p < 0.001$ ):  $r_s = -0.110$  (10 dBSPL),  $r_s = -0.378$  (20 dBSPL),  $r_s = -0.676$  (30 dBSPL),  $r_s = -0.707$  (40 dBSPL),  $r_s = -0.678$  (50 dBSPL), and  $r_s = -0.573$  (60 dBSPL). The strong correlations just above ASSR hearing threshold seem to justify the use of difference scores instead of SNRs per intensity. The test-retest statistic for difference scores is not significant ( $p = 0.439$ , no interactions), similar to the test-retest statistic of the SNRs. Difference scores obtained from the same subject can thus be averaged.

Table 7.2: For each channel from Table 7.1, the mean SNRs (and standard deviations between brackets) of the modulation frequencies modulating the corresponding carrier frequency (500, 1000, 2000, and 4000 Hz, left and right) at a stimulus level of 30 dB SPL are displayed. The mean RMS value (and its standard deviation) of each channel are depicted on the right of the table.

Channel number	Left (dB SNR)				Right (dB SNR)				RMS ( $\mu$ V)
	500 Hz	1000 Hz	2000 Hz	4000 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	
1 (Oz)	5.1 (6.5)	9.3 (2.8)	8.3 (4.0)	5.9 (6.3)	7.1 (5.2)	9.5 (7.0)	10.2 (3.9)	6.8 (5.6)	4.0 (1.3)
2 (P4)	4.2 (5.9)	5.8 (5.4)	4.9 (4.6)	3.1 (5.6)	5.9 (6.7)	9.7 (7.1)	10.1 (3.1)	6.0 (5.1)	2.5 (0.6)
3 (P3)	5.5 (4.7)	8.5 (4.6)	7.9 (4.1)	4.7 (4.9)	4.1 (5.3)	5.6 (6.3)	2.1 (5.4)	-0.4 (4.5)	2.7 (0.6)
4 (rMa)	3.8 (7.6)	7.7 (5.0)	4.6 (6.6)	3.6 (7.0)	7.4 (7.7)	12.0 (5.5)	9.9 (2.5)	6.8 (4.7)	7.1 (3.1)
5 (Ma)	4.5 (7.6)	10.2 (4.8)	8.1 (6.8)	5.1 (5.7)	4.5 (6.1)	7.2 (6.5)	8.7 (4.4)	3.4 (5.1)	6.1 (1.9)
6 (F4)	0.9 (2.2)	1.8 (5.0)	-0.5 (2.5)	-0.3 (4.7)	2.5 (3.1)	4.6 (4.5)	5.4 (3.0)	0.7 (3.2)	4.9 (3.0)
7 (F3)	2.5 (3.0)	5.0 (5.7)	1.9 (8.6)	0.5 (6.4)	3.0 (4.6)	3.9 (5.5)	2.9 (6.8)	0.6 (5.5)	3.7 (1.8)
8 (Fpz)	0.1 (4.8)	2.7 (5.1)	0.9 (5.1)	-1.3 (5.0)	1.8 (6.9)	3.5 (4.2)	3.3 (4.8)	0.7 (3.0)	4.9 (2.6)

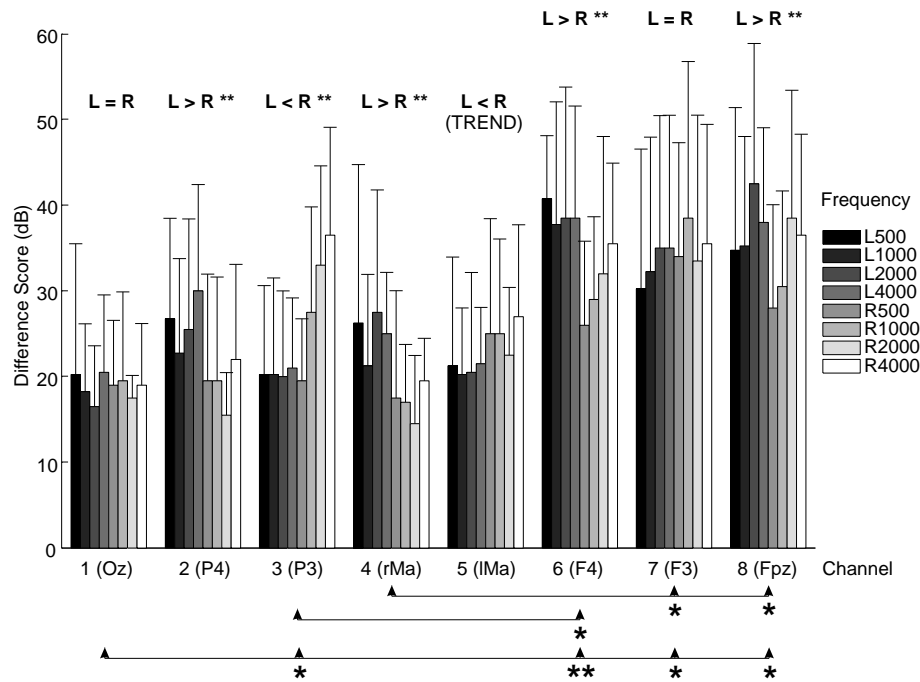


Figure 7.1: For each channel from Table 7.1, the mean difference scores (and standard deviations) are displayed for all modulated carriers (500, 1000, 2000, and 4000 Hz, left and right). Above each cluster of bars, the relation between carriers applied to the left ear (L) and the right ear (R) is shown. Under each cluster of bars, a comparison between channels is displayed. If a statistical difference is present, this is indicated with \* ( $0.05 > p \geq 0.01$ ) and \*\* ( $0.01 > p \geq 0.001$ ). A trend is considered present if  $0.1 > p \geq 0.05$ .

Figure 7.1 shows the mean difference scores (and their standard deviations) per EEG channel for all eight modulated carrier frequencies. Difference scores vary significantly over channels ( $p = 0.004$ ). Comparing pairwise (with Bonferroni correction,  $p < 0.05$ ), channel 1 has significantly better difference scores than channels 3, 6, 7, and 8, channel 3 has significantly better difference scores than channel 6, channel 4 has significantly better difference scores than channels 7 and 8. There is an interaction between subjects and channels ( $p = 0.028$ ). This indicates a significant difference is present between the best channels of one subject compared to the best channels of another subject. Optimal recording results at one particular predefined EEG channel will not be obtained for all subjects.

There is no significant difference ( $p \geq 0.05$ ) between frequencies and between left–right presentation over all channels. Only one other interaction effect is present, namely between channels and left–right application ( $p = 0.006$ ). In Figure 7.1 it is indicated within which channels a significant difference between application on the left ear and application on the right ear is present. Electrodes placed on the right hemisphere will produce difference thresholds that are significantly smaller for carrier frequencies that are applied to the right ear. A similar, but somewhat smaller effect can be observed for the left hemisphere. The effects described here, significantly present on channels 2, 3, 4, and 6, are confirmed on all these channels by six out of ten subjects. Nine out of ten subjects show these effects on at least three of the four significant channels.

If only three electrodes are available for placement, according to Table 7.1, the best positions for the active and the reference electrodes would be the vertex (Cz) and the occiput (Oz), which corresponds to channel 1. This channel returns significantly smaller difference scores than four other channels, including all channels with electrodes on the front side of the head. The side of stimulation is not of significant importance for this channel.

### Combination of channels

The previous section presented the data of all individual channels from Table 7.1 and resulted in an appropriate single channel for smallest difference scores. This section will show the optimal combination of these channels for maximum detection performance (in terms of smallest difference scores), valid for as many subjects as possible and with the restriction that as few electrodes as possible are actually employed. Channels are combined using the detection theory based multichannel processing technique described in Section 6.2.

All 255 ( $2^8 - 1$ ) combinations of the eight EEG channels in Table 7.1 are subject to analysis. Channels are combined using (6.16). Adding extra channels increases the number of detections. The number of extra channels to be added is limited however, as the number of both true and false detections increases. To keep under control the number of false detections, i.e. detections originating from noise only (kept at 5 % in this study), detection decision thresholds have to be tightened accordingly. In theory, a Bonferroni correction should be applied. Practically, this correction is too strict. A detection threshold  $\eta$  is calculated using ‘noise’ frequencies with no response present. These ‘noise’ frequencies are selected close to the modulation frequencies that could contain a response (e.g. 1 Hz smaller). For calculation of the noise estimate, frequency bins with a response are removed. By selecting the 95 percentile of this distribution of  $\Upsilon_A(\mathbf{z})$  values that are calculated on ‘noise’ frequencies, the practical detection threshold  $\eta$  is found. This is done for each of the 255 combinations.



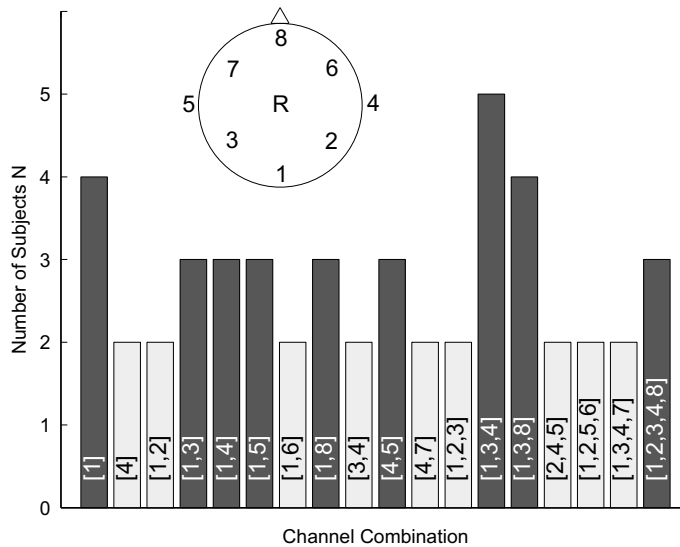


Figure 7.2: Channel combinations appearing in the top-ten list of more than two subjects are represented by the dark bars. Channel combinations appearing only twice in the lists are displayed using light bars. 56 combinations appear only once and are not shown here. Channels are combined using (6.16). Only the combinations with the dark bars are withheld as these exceed the upper level  $U = 2$  of the 95 % confidence interval of the 95 percentile of the distribution of appearances in top-ten lists, determined using Monte Carlo simulations. The inset represents a scalp viewed from above showing the electrode positions from Table 7.1.

For each of the 10 subjects (both test and retest), difference scores of eight modulated carrier frequencies are calculated for all 255 combinations. No significant difference between test and retest is found ( $p = 0.754$ ). Therefore, for each subject a difference score average is available per frequency and per combination ( $10 \times 255 \times 8$  difference scores in total).

The difference scores of most combinations do not have a normal distribution across subjects. Multivariate analysis requires data with a normal distribution. We follow the nonparametric analysis described in van der Reijden et al. (2005), in which outliers are not an issue and equal weights are given to the combinations from each subject regardless of his/her average difference scores.

### Constructing a top-ten list of combinations

For each of ten subjects, a top-ten list is made of ten combinations with the smallest frequency-averaged difference scores. Five ties are included, limited

to the same number of combined channels. This implies that when e.g. channel combination [1,4,8] is ranked 10th, channel combination [3,6,7] with the same difference score is still included, while channel combination [2,3,4,5] (also with the same difference score) is not, as this combination requires one extra electrode which would be less practical. In this particular case, the top-ten list thus contains 105 combinations. Next, the number of times a particular combination appears in the lists is summed over the subjects. The numbers of times the combinations occur are summarized into the histogram in Figure 7.2. This histogram is then subjected to further analysis to identify any high-ranking combinations that are dominant in these subjects and that will therefore suffice to record the ASSR efficiently.

The average chance level  $\bar{d}$  to appear in the top-ten list is  $\bar{d} = \frac{105}{255} = 0.412$ . Monte Carlo simulations are used to find the upper limit  $U$  of the 95 % confidence interval of the 95 percentile of this distribution  $d$  (Press et al., 1992). Therefore, with a reliability of 95 %, a specific combination can only appear more than  $U$  times by chance in less than 5 % of the cases. Practically, 1000 distributions  $d_i$  are constructed. From each distribution  $d_i$ , the 95 percentile  $P_i^{95}$  is determined. This returns 1000 values  $P_i^{95}$ . A 95 % confidence interval (and its upper limit  $U$ ) can be calculated of its mean value  $\bar{P}^{95}$ . Each distribution  $d_i$ , containing 5000 instances  $N_j$ , is constructed as follows. Out of 255 possible combinations (represented by values between 1 and 255, each corresponding to a particular combination), a selection  $S_j$  of 105 random combinations is made. An arbitrary combination  $c_j$  is chosen and the number of appearances  $N_j$  of this combination  $c_j$  among selection  $S_j$  is counted. Instance  $N_j$  is one of the 5000 instances that the distribution  $d_i$  is built up with. As a validation, the mean values  $\bar{d}_i$  are indeed all close to  $\frac{105}{255}$ . The upper limit  $U$  is equal to 2. All values above  $U = 2$  in the histogram of Figure 7.2 (depicted by the dark bars) are thus particular combinations not appearing by chance ( $p = 0.05$ ), with a confidence of 95 %.

The channel combinations in Figure 7.2 that appear more than  $U = 2$  times in the top-ten list with the smallest difference scores over all subjects, are: [1] [1,3] [1,4] [1,5] [1,8] [4,5] [1,3,4] [1,3,8] [1,2,3,4,8]. This series is also displayed in the fourth column of Table 7.3. This result seems promising as these combinations generally appear to contain channels with the smallest difference scores (Figure 7.1). However, a comparison of the test and retest data shows that the test-retest reproducibility of individual combinations in the top-ten list is low. For the test data the withheld combinations are: [2] [1,2] [1,3] [1,5] [1,6] [1,8] [4,7] [1,2,4] [1,3,7] [1,2,3,4]. For the retest data these are: [1] [1,3] [1,4] [4,5] [1,3,4]. This non-reproducibility effect is mainly caused by the fact that a large number (i.e. 255) combinations are candidates for the top-ten list. By focusing on the individual channels appearing in the combinations, this test-retest difference may be reduced.

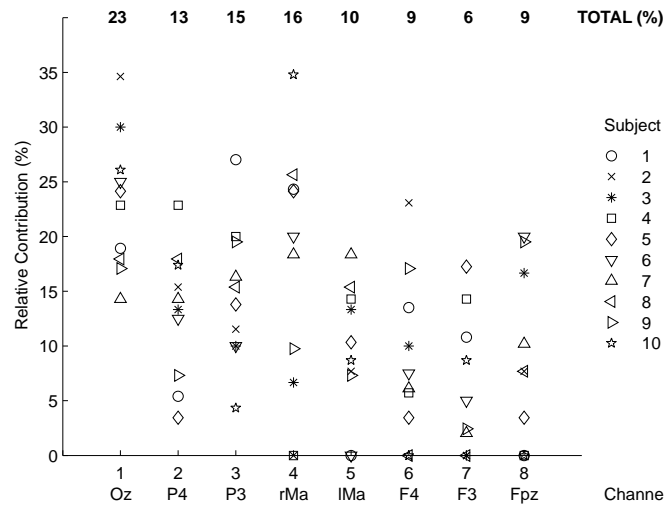


Figure 7.3: Relative contribution of each channel to the top-ten list from each subject (and total relative contribution over all subjects). Brainstem responses from both left and right stimulus presentation are used to construct this figure (cf. fourth column of Table 7.3).

#### Relative contribution of each channel to a top-ten list

The relative contribution of each individual channel to the top-ten lists described in the previous section is an indicator of the importance of each individual channel. For each subject, the number of appearances of a particular channel in the subject's top-ten list is counted and divided by the number of channels in the top-ten list. In this way, the test and retest are not significantly different ( $p = 0.509$ , no interactions). The test-retest combined data are displayed per subject in Figure 7.3, together with the total relative contribution per channel over all subjects. The relative contribution figures can also be found in the fourth column of Table 7.3. An interesting example shows that 'unusable' channels are discarded immediately. During the recording of the data in the first (test) session, channel 4 from both subjects 2 and 4 suffered severely from muscle artifacts at the right mastoid. No contribution is allowed for channel 4 in these subjects.

Based on the total relative contribution over all subjects, a ranking can be produced for the eight channels. For the current derivation the channel ranking is 1 4 3 2 5 6 8 7 (cf. Table 7.3). Guided by this ranking, the minimum number of channels can be determined to cover as many subjects' top-ten lists as possible. The step-by-step construction of this minimum channel set  $\mathcal{M}$  is displayed in Table 7.3. The minimum channel set  $\mathcal{M}_{80}$  covering at least 80% of all subjects, is depicted in bold. In the fourth column of Table 7.3, at least

Table 7.3: Step-by-step derivation of most optimal electrode combination for responses originating mainly from the brainstem and mainly from the auditory cortex.

	BRAINSTEM (80 – 110 Hz)				CORTEX (10 Hz)
	Left	Right	Combined		
Selected combinations ( $U > 2$ )	[1][1,3][1,4] [1,5][1,8][4,5] [1,3,4][1,3,8]	[1][2][4] [1,2][2,4][3,4] [4,7][1,2,4]	[1][1,3][1,4][1,5] [1,8][4,5][1,3,4] [1,3,8][1,2,3,4,8]	[5][1,5][4,6][5,7] [2,5,7][4,5,7][5,6,7] [5,7,8][2,5,7,8][5,6,7,8] [4,5,6,7,8][2,4,5,6,7,8]	
Relative contribution (%)	22 11 17 16 13 6 7 8	19 16 12 18 9 11 7 8	23 13 15 16 10 9 6 9	6 10 6 12 25 12 18 11	
Channel ranking	1 3 4 5 2 8 7 6	1 4 2 3 6 5 8 7	1 4 3 2 5 6 8 7	5 7 6 4 8 2 3 1	
Minimum sets $\mathcal{M}$	{1} – 5 {1,3} – 9 {1,3,8} – 10	{1} – 3 {1,4} – 7 {1,4,2} – 8 {1,4,2,3} – 9 {1,4,2,3,6} – 10	{1} – 4 {1,4} – 7 {1,4,3} – 10	{5} – 5 {5,7} – 6 {5,7,6} – 7 {5,7,6,8} – 9 {5,7,6,8,2} – 10	
Recommended electrodes based on this dataset	{1,3}	{1,4,2}	{1,4,3}	{5,7,6,8}	

one of the seven combinations that can be made from the set  $\mathcal{M}_{80} = \{1, 4, 3\}$  will appear in 10 out of 10 (i.e. all) top-ten lists.

When the minimum channel set  $\mathcal{M}_{80} = \{1, 4, 3\}$  is determined, the list of withheld combinations ( $U > 2$ ) can be reduced. Channel combinations from this list that can be formed by  $\mathcal{M}_{80}$  are [1] [1,3] [1,4] [1,3,4]. Mean difference scores from these channel combinations range between 18.5 and 19.1 dB ( $\pm 1.2$  to 1.5 dB). The difference scores of all these combinations are not significantly different and are significantly ( $p < 0.05$ ) pairwise correlated. As a result, any channel selection out of [1] [1,3] [1,4] or [1,3,4] will result statistically in a minimum mean difference score. To be able to select a channel combination that guarantees a difference score that is always in the top-ten of each subject, the three active electrodes should be placed at position 1 (Oz), position 3 (P3), and position 4 (right mastoid), a selection corresponding to minimum set  $\mathcal{M}_{80}$ . Table 7.4 illustrates this by showing the difference scores of each of these four combinations for each subject, together with their ranking compared with all other combinations within that subject.

### Splitting up the data in left and right stimulus application

To check whether the asymmetric setup of the previous section is asymmetric for an underlying reason, the available data are split up corresponding to the two sides of stimulation (left or right). In this way, the importance of the side of stimulation for electrode placement is addressed.

Responses originating from stimuli applied to the left ear are considered separately from those resulting from stimuli applied to the right ear. The procedures from Section 7.3.1 are repeated identically.

As shown in the second column of Table 7.3, the minimum set  $\mathcal{M}_{80}^L$  for stimuli applied to the left ear is  $\mathcal{M}_{80}^L = \{1, 3\}$ . This corresponds to active electrodes placed at positions Oz and P3.  $\mathcal{M}_{80}^R = \{1, 4, 2\}$  is taken as a minimum set  $\mathcal{M}_{80}^R$  for stimuli applied to the right ear. Active electrodes should be placed at positions Oz, P4, and the right mastoid. In seven of ten subjects, electrodes on the side of stimulation contribute relatively more to the top-ten lists than the contralateral electrodes. In only one subject, this effect is inverted.

The resulting minimum set  $\mathcal{M}_{80} = \{1, 4, 3\}$  of the simultaneous stimuli left and right is a combination of the minimum sets  $\mathcal{M}_{80}^L$  and  $\mathcal{M}_{80}^R$  of stimuli applied to left and right ear separately. No apparent reason for the asymmetry in  $\mathcal{M}_{80}$  can be found however and can be mainly addressed to the used dataset. When compared with symmetric sets  $\{1, 2, 3\}$  and  $\{1, 4, 5\}$ , the sum of the relative contributions of the single channels from the asymmetric set  $\{1, 4, 3\}$  to the top-ten lists is the largest in five of ten subjects. Only in one subject, this sum is the smallest. In the individual subjects, this asymmetry is thus present.

Table 7.4: Difference scores (dB) averaged over all eight frequencies for six relevant combinations, with their ranking (between brackets) compared with all other combinations within the same subject. The six combinations displayed are the best (smallest difference score) and the worst combination, and the four combinations selected in Section 7.3.1. Combinations that were in a subject's top-ten, are highlighted in bold.

Subject	1	2	3	4	5	6	7	8	9	10
Best	14.4 (1)	20.9 (1)	18.1 (1)	13.1 (1)	13.1 (1)	12.5 (1)	16.9 (1)	20.3 (1)	18.4 (1)	10.0 (1)
[1]	21.9 (154)	<b>22.8 (9)</b>	<b>18.1 (1)</b>	17.5 (58)	14.4 (11)	17.5 (44)	19.3 (83)	25.3 (194)	<b>20.9 (10)</b>	<b>10.0 (1)</b>
[1, 3]	<b>14.4 (1)</b>	23.4 (15)	<b>19.4 (3)</b>	<b>13.8 (3)</b>	15.6 (52)	19.4 (114)	20.0 (134)	23.4 (91)	22.2 (35)	13.8 (23)
[1, 4]	20.0 (83)	22.8 (12)	20.6 (13)	20.0 (145)	14.4 (12)	<b>15.0 (6)</b>	<b>18.1 (9)</b>	22.2 (27)	24.1 (108)	<b>11.3 (4)</b>
[1, 3, 4]	<b>16.3 (6)</b>	25.9 (90)	<b>20.0 (6)</b>	18.8 (103)	<b>13.8 (6)</b>	17.5 (47)	18.1 (15)	<b>20.9 (3)</b>	27.2 (188)	<b>12.5 (10)</b>
Worst	41.3 (255)	37.2 (255)	53.8 (255)	30.0 (255)	29.4 (255)	33.8 (255)	41.9 (255)	45.9 (255)	52.2 (255)	42.5 (255)

### Artifact robustness

Four different processing methods are considered. Two processing methods are single-channel methods. Two methods are multichannel. Channel 1 is taken from Table 7.1 as the reference channel for single channel brainstem ASSR measurements with adults (Luts and Wouters, 2005; van der Reijden et al., 2001). Channels 1, 3, and 4 come forward as the minimum set in Section 7.3.1 and are combined using the multichannel processing scheme from Section 6.2.

The processing schemes are applied on two multichannel EEG datasets obtained at an intensity of 30 dB SPL. Firstly, they are administered on an EEG dataset with few artifacts ('clean EEG') that is identical to the EEG in Section 7.3.1. Secondly, they are applied on an EEG dataset with a considerable number of artifacts ('dirty EEG'), generated as described in Section 7.2.1. To calculate a detection threshold  $\eta$ , each of the four processing methods below are applied on frequencies without a response (e.g. 1 Hz below the modulation frequencies that could contain a response) of the artifact-free multichannel EEG from Section 7.3.1 on all intensities. Then, the 95 percentile of this noise distribution is defined as the detection threshold. In this way four different detection thresholds are obtained for each processing method.

1. *Channel 1, normal averaging.* Channel 1 is divided into 32 sweeps (data blocks of 16.384 seconds) and averaged. Detection is managed using (2.11).
2. *Channel 1, noise-weighted averaging.* Each unfiltered epoch (data blocks of 1.024 seconds) is transformed to the frequency domain by using an FFT. The average power  $P_i$  between 77 and 115 Hz is computed after removing the power at the 8 frequencies at which responses occurred. The time domain epoch is then weighted with the average power  $P_i$  and concatenated with the preceding epochs to form sweeps. Each epoch of the final summed sweep is then divided by the sum of the weights  $P = \sum P_i^{-1}$  of the epochs that have been combined to form that particular epoch (adapted from John et al. (2001a)). A more detailed description of the basics is available in Section 2.2.4. Detection is managed using (2.11).
3. Channels 1, 3, and 4 are combined using (6.16) with  $\mathbf{K}_{\text{spat}}$  fixed and  $\mathbf{K}_{\text{temp}} = \sigma I$  fixed. The mean noise power  $\sigma^2$  is calculated using (2.11) on the combined channels 1, 3, and 4.
4. Channels 1, 3, and 4 are combined using (6.21) with  $\mathbf{K}_{\text{spat}}$  fixed.  $\mathbf{K}_{\text{temp},i} = P_i P \sigma_n I$  is variable and is recalculated for each block  $i$  of 8.192 seconds. This choice of the block size returns the best results, similar to Table 6.4.  $\mathbf{K}_{\text{temp},i}$  is calculated based on the rationale of noise-weighted averaging. A more detailed description of the basics is available in Section 2.2.4.  $P_i$

Table 7.5: Number of detections (with a maximum of 160) for each processing method when applied to EEG with few artifacts (‘clean EEG’) and EEG with a significant number of artifacts (‘dirty EEG’). For methods 1, 2, 3, and 4, comparisons have been conducted pairwise with Bonferroni correction. Methods 3’, 3”, 4’, and 4” process two symmetric channel sets, but are not used for the pairwise comparisons. They are only placed in this table for illustrative purposes.

Method	Description	Clean EEG	Dirty EEG
1	Normal averaging channel 1	117	17
2	Weighted averaging channel 1	122	35
3	$\mathbf{K}_{\text{temp}} = \sigma I$ fixed; [1,3,4]	117	57 <sup>1</sup>
4	$\mathbf{K}_{\text{temp},i} = P_i P \sigma_n I$ var; [1,3,4]	<b>122</b>	<b>72<sup>2</sup></b>
3’	$\mathbf{K}_{\text{temp}} = \sigma I$ fixed; [1,2,3]	113	31
4’	$\mathbf{K}_{\text{temp},i} = P_i P \sigma_n I$ var; [1,2,3]	122	50
3”	$\mathbf{K}_{\text{temp}} = \sigma I$ fixed; [1,4,5]	108	44
4”	$\mathbf{K}_{\text{temp},i} = P_i P \sigma_n I$ var; [1,4,5]	112	66

is the average power between 77 and 115 Hz of data block  $i$ .  $P$  is the sum of the reciprocals of  $P_i$ , i.e.  $P = \sum P_i^{-1}$ . The mean noise power  $\sigma_n^2$  is calculated using (2.11) on a noise-weighted average of the combined channels 1, 3, and 4.

It is assumed that all responses to the applied stimuli are present in the EEG as the stimulus intensity is 30 dB SPL and hence well above the subject’s hearing thresholds. There are thus 16 responses for each of the ten subjects (test and retest) available for detection. The maximum number of detections is 160. Table 7.5 shows the number of detections for each approach for EEG with few artifacts (‘clean EEG’) and EEG with a significant number of artifacts (‘dirty EEG’). Test and retest data are not significantly different for both datasets ( $p \geq 0.67$ ).

For EEG with few artifacts, best results are obtained when a noise-weighted averaging approach is used. The improvement relative to normal averaging is not significant however. It does not make a difference if multiple channels are used for processing or just one (reference) channel.

Response detection in EEG with numerous artifacts is improved significantly when using more than one channel for processing. This improvement can increase significantly when using a noise-weighted averaging based approach.

<sup>1</sup>Significantly better than approach 1 ( $p < 0.01$ ).

<sup>2</sup>Significantly better than approach 1 ( $p < 0.001$ ), approach 2 ( $p < 0.01$ ), and approach 3 ( $p < 0.05$ ).



In Section 7.4, it is discussed whether a symmetric minimum set would give a more logical electrode placement compared with an asymmetric minimum set. Therefore, for illustrative purposes, Table 7.5 adds four extra lines showing the multichannel processing results of symmetric minimum sets  $\{1, 2, 3\}$  and  $\{1, 4, 5\}$ .

### 7.3.2 Auditory cortex stimulation

This section will present results from stimulation at only one ear. Six out of ten subjects are stimulated at the right ear. To preserve symmetry, all electrode positions of the other four subjects are mirrored as if stimulation occurred at the right ear. This swapping operation is allowed as no significant difference ( $p = 0.904$ ) in SNR between groups with left ear stimulation and right ear stimulation is found. As a result, Table 7.1 refers for this section to the (corrected) side of stimulation as ‘ipsilateral’ and the opposite side as ‘contralateral’.

#### Separate channels

Figure 7.4 shows the mean SNR (and standard deviation) of the 10 Hz ASSR in all eight channels from Table 7.1 for intensities of 50 and 70 dB SPL. Test and retest data are averaged per subject ( $p = 0.314$ , no interactions). There is no significant difference between both intensities ( $p = 0.113$ , no interactions). A main effect between channels is present ( $p = 0.003$ , no interactions). When comparing pairwise (Bonferroni correction) in Figure 7.4 at 50 dB SPL, the SNRs of channels 4 and 5 are significantly larger than those of channels 6, 7, and 8. Channel 1 is significantly different from channels 6 and 7 and channel 3 only is significantly different from channel 7. At 70 dB SPL, only the SNRs of channel 4 are significantly larger than those of channels 7 and 8. Between subjects an interaction is present with the EEG channels ( $p = 0.044$ ). This indicates that the EEG channel with the largest SNR for one subject is not necessarily the channel with the largest SNR for another subject. The best EEG channel is subject dependent.

Due to the high variability of the SNRs, it is difficult to identify an optimal electrode placement from Table 7.1 in the case of three electrodes available for placement. From the pairwise comparisons above, channels 4 and 5 (both mastoids) appear to be good choices.

#### Combination of channels

Although no significant difference is present between both intensities in the previous subsection, only the results obtained at 50 dB SPL is focused on. The structure of this section is similar to the structure of Section 7.3.1. Results

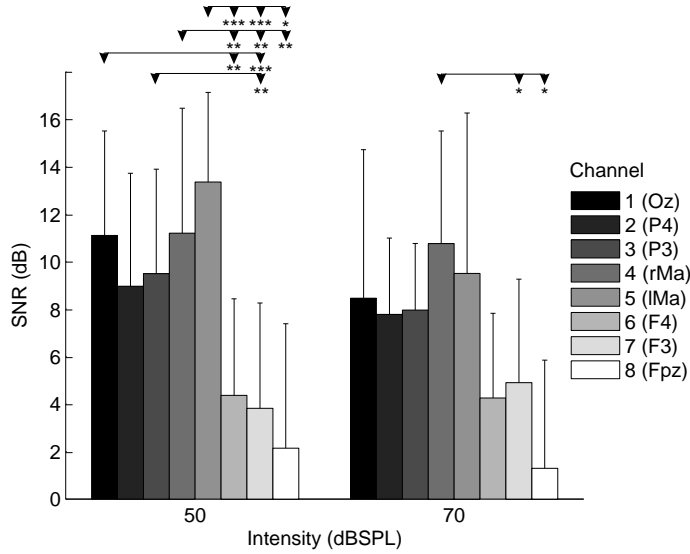


Figure 7.4: For stimulus intensities 50 and 70 dB SPL, the mean SNR (and standard deviation) of the 10 Hz response is displayed per channel from Table 7.1. Stimulus application is at the right ear (next to channel 4). Above each intensity, a comparison between channels is displayed. If a statistical difference is present, this is indicated with \* ( $0.05 > p \geq 0.01$ ), \*\* ( $0.01 > p \geq 0.001$ ), and \*\*\* ( $0.001 > p$ ).

at 70 dB SPL are similar and will only be described briefly whenever different from those at 50 dB SPL.

In Section 7.3.1, difference scores between the ASSR threshold and the behavioral threshold were used to quantify performance. As no behavioral data at the stimuli are available, only signal-to-noise ratios are used in the analysis, with a minor modification. A response of e.g. 12 dB resulting from a five-channel combination does not perform as well as a 12 dB response from a three-channel combination. As a result, the concept of corrected SNR appears to be applicable here. The *corrected SNR* is the difference between the response SNR and a threshold SNR. This latter value increases when more channels are added for combination. The *threshold SNR* is determined as the SNR that allows 5 % false detections when observing a frequency bin without any response present (e.g. 1 Hz below the stimulus frequency of 10 Hz). For calculation of the noise estimate, frequency bins with a response are removed.

Again, all 255 combinations of the eight EEG channels in Table 7.1 are subject to analysis. Channels are combined using (6.16) from Section 6.2. No significant difference between the corrected SNRs of test and retest data is found

( $p = 0.805$ ). Therefore, for each subject an average corrected SNR is available for the 10 Hz response and per combination ( $10 \times 255$  corrected SNRs in total). The corrected SNRs of most combinations do not have a normal distribution across subjects. Therefore, again a top-ten list is made for each of ten subjects. Zero ties are included. The top-ten list thus contains 100 combinations. The number of times a combination appears in this top-ten list is counted (with a mean chance level of  $\bar{d} = \frac{100}{255} = 0.392$ ). All combinations appearing more than  $U = 2$  times in the top-ten list ( $U$  as a result of Monte Carlo simulations), are considered not appearing by chance ( $p = 0.05$ ), with a confidence of 95 %. The withheld combinations for the 10 Hz ASSR are: [5] [1,5] [4,6] [5,7] [2,5,7] [4,5,7] [5,6,7] [5,7,8] [2,5,7,8] [5,6,7,8] [4,5,6,7,8] [2,4,5,6,7,8]. They are also displayed in the fifth column of Table 7.3. Only combinations [5] [5,7] [2,5,7] [5,7,8] can also be found in both test and retest results. To guarantee test-retest reproducibility, the focus is on the individual channels appearing in these combinations instead.

According to the fifth column of Table 7.3, channels 5 and 7 are by far the channels with highest occurrence in the subject's top-ten lists. By adding channels 6 and 8, 90 % of all subjects are covered. The minimum channel set is thus  $\mathcal{M}_{80} = \{5, 7, 6, 8\}$ . However, when following the channel ranking of Table 7.3 exactly, the minimum set  $\{5, 7, 6, 4\}$  should be the first choice. This minimum set offers the same amount of coverage as  $\{5, 7, 6, 8\}$ , but combination [5,7,6,4] does not appear in the list of withheld combinations however, so channel 8 is chosen instead, and consequently for minimum set  $\{5, 7, 6, 8\}$ .

Based on the minimum channel set  $\mathcal{M}_{80}$ , only channel combinations [5] [5,7] [5,6,7] [5,7,8] [5,6,7,8] are considered from the list with withheld combinations. Their mean corrected SNRs are between 8.2 and 9.0 dB ( $\pm 1.1$  and 1.2 dB) and not significantly different. They are all significantly pairwise correlated ( $p < 0.05$ ). As a result, any channel selection out of these five combinations will result statistically in a maximum mean corrected SNR. To be able to select a channel combination that guarantees a corrected SNR that is always in de top-ten of each subject, however, the four active electrodes should be placed at position 5 (contralateral mastoid), position 6 (F4), position 7 (F3), and position 8 (forehead). When the data obtained at 70 dB SPL is analyzed, a coverage of 80 % is possible with  $\mathcal{M}_{80} = \{5, 7, 6, 4\}$ , which is similar to the results at 50 dB SPL.

## 7.4 Discussion

All combinations of EEG channels in this chapter are processed using a sufficient statistic that theoretically captures best the amount of useful ASSR target signal present in these EEG channels. This sufficient statistic originates from

the realm of detection theory (Green and Swets, 1966). It can be linked with the multichannel Wiener filter (MWF), an optimal filtering technique which computes an optimal (minimum mean square error) estimate of a reference signal (Scharf, 1991). In the case of auditory steady-state response detection, this comes down to a maximization of the signal-to-noise ratio of the combined channels (Simmer et al., 2001). This point of view makes it easier to get an idea of what happens when combining EEG data using (6.16).

Adding extra data channels will increase the amount of information that is gathered in the same time period. However, to combine these channels, the use of more advanced signal processing is necessary to truly achieve a performance improvement. If multichannel data is combined using e.g. simple time domain averaging over channels, the statistical multiple testing penalty will increase the number of false detections proportionally to the number of extra correct detections. By intelligently combining these channels, this multiple testing problem can partly be avoided.

This section will discuss the results from the previous section for mainly stimulating the brainstem and the auditory cortex respectively.

#### 7.4.1 Brainstem stimulation

The transition from the clinical approach of difference scores to the practical quantitative measures of signal-to-noise ratios seems to be allowed due to the high correlations between both measures. It is assumed that results for one measure can be used for the interpretation of the results of the other measure.

When observing the difference scores of reference channel 1 (Cz–Oz), two observations can be made. First, difference scores are slightly elevated compared to other studies. Standard deviations are comparable however (Dimitrijevic et al., 2002; Herdman and Stapells, 2001; Luts and Wouters, 2005; Picton et al., 2003). This difference score gap may be caused by the fact that not as much attention has been directed to the placement of all (i.e. more) electrodes. The determination method of ASSR hearing thresholds also appears to be rather strict in this study. However, this determination method is kept identical for the whole study and final conclusions should not be influenced. Second, although not significantly different, carrier frequencies of 500 and 4000 Hz show a trend of having larger difference scores than 1 and 2 kHz at these intensities, which is also observed in other studies (Dimitrijevic et al., 2002; Herdman and Stapells, 2001; Luts and Wouters, 2004; Picton et al., 1998). Both observations suggest the data in this study is representative for general brainstem ASSR measurements.

As already indicated in Section 7.2, several studies focus on the sources of the ASSR with frequencies between 80 and 110 Hz, the frequency region of interest

for hearing threshold determination. They conclude that for modulation frequencies above about 75 Hz, most of the ASSR is generated in the brainstem. Above 95 Hz, the source lies entirely in the brainstem. These assumptions seem to support the observations made in this chapter. With sources mainly in the brainstem the smallest difference scores (and the largest SNRs consequently) are recorded in electrodes close to the brainstem, on the back of the head and the mastoids. The ipsilateral effect of larger SNRs and smaller difference scores (and thus the optimal combination of electrodes) at the side of stimulus application is partly confirmed by Small and Stapells (2008). They indicate that ASSR amplitudes are significantly smaller at the contralateral side. Difference scores with adults do not differ significantly however. For infants, this ipsilateral effect is reported to be strongly present, both for ASSR amplitudes as ASSR thresholds. This asymmetric effect is also confirmed by van der Reijden et al. (2005). This could signify that the orientation of the ASSR sources is stimulus side dependent for adults, and especially dependent in the case of infants. The asymmetrical electrode positioning for binaural stimulation cannot be confirmed in literature. We assume that the results in this chapter are merely due to the used dataset and that a symmetrical electrode placement is preferred. According to Table 7.5, a recommended symmetric minimum set could be  $\{1, 4, 5\}$ , only slightly reducing the number of detections for EEG with few artifacts as well for EEG with many artifacts. To conclude, van der Reijden et al. (2004) report that a small set of three derivations (Cz–Oz combined with the right mastoid–Cz and the left mastoid–Cz, which corresponds to the minimum set  $\{1, 4, 5\}$ ) yields the best SNR in a larger number of adults than would be expected if all derivations were equally efficient. This result is very similar to the one obtained in this study.

Multichannel processing does not significantly improve detection performance when it is applied to EEG that does not contain many artifacts. This shows that most single channel studies using a Cz–Oz or Cz–neck electrode configuration likely get close to the best obtainable result (for an extensive overview, we refer to Picton et al. (2003)). The EEG in these studies is mostly obtained from relaxing adults and sedated children. Studies that specifically incorporate EEG with numerous artifacts are rare however. Unfortunately, these situations reflect the conditions observed mostly in a clinical environment. The proposed technique is significantly more robust against artifacts and may improve measurements in these difficult conditions.

Further research should focus on the confirmation of the results in this study and the manuscript of van der Reijden et al. (2004). Additionally, the methods proposed here should also be applied to multichannel data collected from infants. The only studies that already cover part of this topic are the studies of Small and Stapells (2008) and van der Reijden et al. (2005).

### 7.4.2 Auditory cortex stimulation

As already indicated in Section 7.2, when lowering the modulation frequency of the stimulus, the main source of the auditory steady-state response shifts more to the auditory cortex, while still having a source in the brainstem. This may explain the SNRs being largest on the electrodes close to the brainstem (back of the head) and the auditory cortices (both mastoids).

Although no significant differences in SNR are found in Figure 7.4 between the single electrodes at the ipsilateral and the contralateral mastoid (also reported by Herdman et al. (2002)), the dominant channel (contralateral mastoid, channel 5) of the minimum set  $\{5, 7, 6, 8\}$  clearly is located at the opposite side of stimulation. This could be explained by the crossing of the auditory paths beyond the brainstem (Hall, 2007). The extra channels 6, 7, and 8 of the minimum set  $\{5, 7, 6, 8\}$  lie at the front of the head, in contrast with the channels with the largest SNRs lying at the back of the head as shown in Figure 7.4. This can be explained by the underlying mechanism of the multichannel processing technique from Section 6.2. For responses originating mainly from the auditory cortex, this result is likely obtained by taking the channel with the largest response (channel 5), and by adding channels that have a small response (channels 6, 7, and 8) and that have a high noise correlation with the noise in the channel with the largest response. This observation is in contrast with the conclusions in Section 7.4.1, that locates the minimum set  $\{1, 4, 2\}$  in the back of the head for responses originating mainly from the brainstem. The dominant channel for this minimum set is the occiput (channel 1). Any additional channels with small responses (channels 6, 7, and 8 in front of the head according to Figure 7.1) appear to be too remote as noise correlation with these remote channels is presumably too low in the frequency region of interest (80 – 110 Hz).

The trend in observing smaller SNRs for higher intensities of stimulation is in agreement with Hall (2007) where it is stated that ‘maximum response typically is obtained for moderate (50–60 dB) versus high-intensity stimuli’.

Further research should confirm that contralateral electrode placement is preferred to ipsilateral placement in the case of low-frequency ASSRs. It would be interesting to determine optimal electrode placement for a range of modulations, both at lower (e.g. 4 Hz) as at higher (e.g. 40 or 200 Hz) frequencies.

## 7.5 Conclusions

In this chapter, the practical performance of a multichannel EEG processing strategy for ASSR detection based on a detection theory approach is evaluated. After a short introduction in Sections 7.1, Sections 7.2, 7.3, and 7.4 describe

a setup, its results and the discussion of eight-channel measurements from ten normal-hearing adults. First, auditory steady-state responses mainly originating from the brainstem are considered. It is shown that electrodes should be placed at the back of the head and at the mastoids to obtain the largest mean SNRs and the smallest mean difference scores for individual electrodes. When applying the multichannel processing strategy, the smallest difference scores are found with all subjects when electrodes are placed at five positions: Oz, P3, and right mastoid with Cz as a reference electrode and a ground electrode at e.g. the right clavicle. This combination is significantly more robust against artifacts when compared with a single channel, three electrode, setup. The number of ASSR detections is more than doubled when EEG with artifacts is considered. Second, auditory steady-state responses with a modulation frequency of 10 Hz and mainly originating from the auditory cortex are studied. Again, the largest mean SNRs are obtained at single electrodes located at the back of the head and at both mastoids. When applying multichannel processing, the largest mean SNRs for minimum 80% of the subjects are obtained when placing electrodes at the contralateral mastoid, F4, F3, and the forehead with Cz as a reference electrode and a ground electrode at e.g. the right clavicle.





## Chapter 8

# General Conclusions and Further Research

The auditory steady-state response technique is a reliable way to assess hearing thresholds objectively. Its most important merit compared to other objective hearing threshold estimation techniques, e.g. the auditory brainstem response, is the frequency specificity. For each (audiometric) frequency, an ASSR threshold can be determined that is an objective and reliable estimation of the subject's behavioral hearing thresholds.

The most important target group of the ASSR technique are infants, and more specifically infants that are referred to a specialized clinical center after a failed screening with oto-acoustic emissions or the auditory brainstem response technique. The use of auditory steady-state responses leads to important diagnostic info: based on the acquired and additional frequency specific information, appropriate measures can be taken. These measures include middle ear surgery, the application of a hearing aid or the implantation of a cochlear implant. Based on this frequency specific information, a hearing aid can be fitted in a more optimal way for subjects with limited or no feedback.

Unfortunately, the ASSR technique has some disadvantages. The responses are small (in the order of a couple of nanoVolts when close to the hearing threshold) compared to the present EEG noise. Therefore, detection of ASSRs is more difficult than the detection of e.g. auditory brainstem responses (with amplitudes in the order of hundreds of nanoVolts). Recordings can take a long time due to the nature of the responses and the presence of unwanted noise and artifacts. As a result, even long measurement sessions of 45 to 60 minutes do not always guarantee a reliable hearing threshold estimation, especially when the subject is not relaxed or sleeping.

This chapter is the final chapter of this thesis and presents its conclusions in Section 8.1. It offers recommendations for further research in Section 8.2.

## 8.1 General conclusions

This research project focused on the search for techniques that reduce ASSR recording time and that increase the robustness of the ASSR technique against many inevitable artifacts. The project can be divided in three parts. The first part focused on the development of a robust and multichannel experimental test platform for recording auditory steady-state responses for research and clinical applications (Section 8.1.1). Based on this setup, studies were conducted that evaluated the multichannel processing techniques proposed in the second part of the thesis. This second part focused especially on recording time reduction (Section 8.1.2). The third part presented a simplified procedural framework for ASSR detection that allowed the development of a multichannel processing strategy starting from a detection theory approach (Section 8.1.3). The multichannel techniques introduced in the second part could be fitted into this framework. This final part was oriented more on optimal electrode placement, the robustness against artifacts and increasing the number of ASSR detections explicitly. Of course, this implicitly reduced measurement time too.

The global benefit that is possible using the findings of this research project is summarized in this paragraph. This benefit can be ascribed entirely to the use of multichannel EEG recordings, combined with multichannel EEG signal processing techniques. The global result is described for two types of EEG, namely EEG with few artifacts (optimal recording conditions) and EEG with a significant number of artifacts (avoidable measurement conditions). Most measurements are a combination of both types of EEG. When recording conditions are optimal, a measurement time reduction of up to 60 % is possible for intensities above hearing threshold. When the applied intensities are around hearing threshold, no mean recording time reduction is possible anymore, given the observation that hearing threshold estimations are not significantly different for single channel and multichannel measurements with the same length. Notwithstanding, the use of multiple EEG channels offers the benefit that for each individual subject a close-to-optimal recording of the response of interest is guaranteed. If the EEG is contaminated with artifacts, the measurement time reduction for intensities above hearing threshold is probably equal to or even larger than the reduction obtained for EEG with few artifacts. This assumption can be supported by the significant number of extra response detections using the multichannel technique. At hearing threshold, this reduction still will be present, as artifacts force single channel measurements to discard EEG-data. This immediately implies longer measurement times. In the multichannel case, EEG data can be withheld however.

### 8.1.1 A multichannel platform for ASSR measurements

Commercially available ASSR measurement platforms currently only record up to two channels simultaneously. Multichannel ASSR recordings with more than two channels are not possible using these devices unless one sequentially repositions the available electrodes. This way, no simultaneous measurements can be carried out however, which is necessary for the multichannel techniques proposed in this thesis. Another drawback of the commercially available setups is the non-availability of applying custom made stimuli to subjects. In the study of Chapter 7, this function was required.

A possible critical remark could be why one not just uses a commercial system with e.g. 64 EEG channels, with electrodes applied to the scalp using an EEG-recording head cap. This could solve the problem of the non-availability of enough channels and the requirement to build a setup oneself. First, such a multichannel system could provide an overkill in channels as Chapters 4 and 5 indicate that five channels should suffice, which is confirmed in Chapters 6 and 7. This assumption should be investigated however. It is likely that when measuring with more than eight channels a different maximum number of channels is determined. Second, one has to take care that the quality of the amplifiers is sufficient for auditory steady-state response amplification as these responses are very faint. This implies the use of qualitative and expensive low-noise amplifiers. Using a significant number of such EEG-amplifiers could prove to be costly.

**Chapter 3** presented a research platform for multichannel ASSR measurements, referred to as SOMA (*Setup ORL for Multichannel ASSR*). The setup allows multichannel measurements, the use of own stimuli and independent intensity changes. It can be used to assess ASSR hearing thresholds reliably. It is not restricted by the limitations of commercial software and is thus better suited for research and several clinically diagnostic purposes. The mobile setup is based on an inexpensive multichannel RME soundcard and software is written in C++. Both hardware and software of the setup were described. An evaluation study with nine normal-hearing subjects showed that no significant performance differences exist between the proposed platform and the MASTER setup from John and Picton (2000a). At present this experimental system is being used in UZ Leuven for intra-operative ASSR measurements in the operation room.

### 8.1.2 Development of multichannel signal processing techniques improving ASSR detection

The moment the setup developed in Chapter 3 was up and running, the multichannel processing techniques worked out in this research project could be evaluated using multichannel ASSR measurements from real-life subjects.

The first multichannel processing technique proposed and evaluated in this thesis was independent component analysis. Independent component analysis (ICA) allows finding the underlying factors from multivariate statistical data by looking for components that are both statistically independent, and non-Gaussian. The auditory steady-state response, which is sinusoidal from nature and has a platykurtic distribution, is statistically different from the surrounding EEG noise (mesokurtic, i.e. almost Gaussian). It was shown in **Chapter 4** that this assumption indeed was correct and that ICA applied to multichannel data containing ASSRs increased detection performance and reduced measurement time. ICA applied on single- and multichannel recordings yielded a significantly better performance than the clinically used single channel reference technique for data obtained at intensities above hearing threshold. For single channel measurements a time reduction up to 23 % for a single subject was acquired. For multichannel EEG measurements there was a significant measurement time reduction possible of 52 % for 48-sweep measurements compared to the single channel reference technique. For individual subjects, an improvement between -2 and 63 % in measurement time was recorded. When both single- and multichannel techniques were combined, performance could be improved even more. The huge variability in performance across subjects was confirmed in Chapter 7, where it was shown that a predefined EEG channel performing well statistically over an entire group does not guarantee good performance for a single subject. Five-channel ICA yielded the same performance as the seven-channel ICA for ASSRs generated mainly in the brainstem at modulation frequencies between 80 and 110 Hz. This observation was repeated in the practical evaluation study of Chapter 6 by applying the newly proposed framework to another dataset. Electrodes were best placed on the back of the head. These results were confirmed in Chapter 7.

According to Chapter 6, multichannel Wiener filtering (MWF) and independent component analysis applied to multichannel data containing ASSRs were comparable in performance. Both techniques could be linked together using the framework that was developed in that chapter. **Chapter 5** singled out the multichannel Wiener filter technique (following QR factorization) applied to multichannel data containing ASSRs by evaluating its performance based on the same dataset that was used for the ICA evaluation. The MWF solution was identical to the solution of an output-SNR maximization of the ASSR. Results were highly similar to the ICA-based results obtained in Chapter 4, which was confirmed by the theory, the simulations and the evaluation in Chapter 6. However, several benefits exist with the MWF-based technique compared with the ICA-based technique. No artificially created ‘channels’ need to be added to keep the variability under control. No prior averaging of the data is required. The MWF-based technique also can be implemented with significantly less code lines, which improves the speed of execution.

### 8.1.3 Unifying multichannel techniques into a simplified framework for ASSR detection

**Chapter 6** proposed a simplified procedural framework for ASSR detection that allowed the development of a multichannel processing strategy starting from a detection theory approach. It was shown that a sufficient statistic could be calculated that best captured the amount of useful signal (ASSR) in the observation matrix. This sufficient statistic based approach exploits spatio-temporal stationarity present in the EEG measurements and could be linked with the development and application of the independent component analysis (ICA) based approach in Chapter 4 and the multichannel Wiener filter (MWF) based technique in Chapter 5. First, simulations were conducted using EEG data from 10 subjects and artificial ASSRs with varying amplitude and phase. It was concluded that multichannel techniques are similar in performance when applied to uncontaminated EEG. For this condition, the proposed sufficient statistic based approach was similar to the MWF based approach (and similar to an ICA based approach). It was not recommended to use a stimulus evoking a response that had a duty cycle smaller than 100 %. Local EEG stationarity thus could not be exploited. When the EEG was contaminated by spatially uncorrelated artifacts, the proposed technique (with a varying spatial covariance matrix) was performing better than the multichannel techniques described in Chapters 4 and 5. When spatially correlated noise bursts were superimposed on the multichannel EEG, the multichannel techniques performed similar.

As the artificial simulations in the first part of Chapter 6 did not provide a fully appropriate comparison with real-life recordings and single channel techniques, the second part of that chapter focused on a realistic comparison between single- and multichannel processing techniques using multichannel EEG data from ten normal-hearing adults. Two types of EEG were used: EEG with few artifacts and with a significant number of (controlled) artifacts. It was concluded that most single- and multichannel approaches were similar in performance when applied to uncontaminated EEG. When artifact-rich EEG is used, the proposed detection theory based approach improves the number of ASSR detections compared with the noise-weighted average of the best channel of this dataset (vertex-right mastoid). In general this best electrode configuration can not be known beforehand. When compared with a noise-weighted common EEG channel derivation (vertex-occiput), the proposed approach improves ASSR detection significantly.

After the sufficient statistic based multichannel processing strategy was validated in Chapter 6, it was applied to two datasets for further evaluation in **Chapter 7**. In this chapter, the optimal recording electrode sites were determined based on these datasets, in contrast with the algorithm evaluation of Section 6.4 which used the same datasets for another end. One EEG dataset contained auditory steady-state responses originating mainly from the brain-

stem, another dataset mainly contained an ASSR from the auditory cortex. The former were evoked using stimuli with modulation frequencies between 80 and 110 Hz (identical to the stimuli that are used to assess hearing thresholds). The latter was a response with a modulation frequency of 10 Hz, with a longer latency and thus mainly generated in the auditory cortex. For both types of ASSRs, two research questions were investigated using eight-channel measurements of ten normal-hearing adults. First, the EEG derivation with the best estimation of the behavioral hearing thresholds (brainstem) and the highest SNR (auditory cortex) was searched for. Second, the EEG channel combination with a close-to-optimal estimation of the behavioral hearing thresholds (brainstem) and the close-to-highest SNR (auditory cortex) was determined for as many subjects as possible, subject to as few electrodes as possible.

For brainstem ASSRs, it was shown that separate electrodes should be placed at the back of the head and at the mastoids, with a preference for the Cz–Oz derivation, to obtain the lowest mean difference scores for individual electrodes. When applying the multichannel processing strategy, the close-to-lowest difference scores were obtained for all subjects when electrodes are placed at five positions: Oz, P3, and right mastoid with Cz as a reference electrode and a ground electrode at e.g. the right clavicle. This combination was significantly more robust against artifacts when compared with a noise-weighted single channel, three electrode, reference setup (Cz–Oz). The number of ASSR detections was more than doubled when EEG full of artifacts was considered. Unfortunately the proposed optimal set is asymmetric. If one wants to opt for a symmetric multiple electrode set, it was recommended to use the occiput (Oz) combined with both mastoids as active electrodes. For auditory cortex ASSRs, the close-to-highest mean SNRs were obtained at separate electrodes located at the back of the head and at both mastoids. When applying multichannel processing on a combined multiple electrode set, the highest mean SNRs for minimum 80% of the subjects were obtained when placing six electrodes at the contralateral mastoid, F4, F3, and the forehead with Cz as a reference electrode and a ground electrode at e.g. the right clavicle.

## 8.2 Suggestions for further research

This section describes some ongoing projects and presents several suggestions for further research concerning the recording platform and multichannel ASSR processing.

### 8.2.1 The multichannel measurement platform

All studies in this research project were based on the data collected with the recording setup SOMA. It was possible to record multichannel data (up to

eight channels) and to present binaurally custom made stimuli, two features that were required for the evaluation of the multichannel processing techniques proposed in this thesis. Custom made stimuli are currently used in different research projects. At present modulation frequencies below 40 Hz at intensities above hearing threshold are investigated. It scans for ideal modulation frequency candidates studying the possible link between low-modulation frequency ASSRs, speech understanding, and dyslexia. SOMA can change the intensity of different modulated carriers independently. By decreasing the intensity of a carrier that already evokes a response, while other carriers do not, measurement times can be decreased significantly. The moment an adequate protocol with the right parameters is available, an automatic intensity modification strategy could be implemented in SOMA. This way, frequency specific hearing thresholds could be determined faster and entirely automatically (John et al., 2002b; Mühler et al., 2005). This aspect is currently a topic of research.

Multichannel EEG could be processed real-time by algorithms based on the findings in this thesis.

Electrode impedances are currently monitored before and after a measurement. A continuous impedance monitoring in SOMA would increase the reliability of the ASSR measurements to monitor the electrode contacts during the measurements. This feature already exists in commercially available setups.

Constant monitoring of the stimulus intensity at the eardrum also could increase the measurement reliability. It would be possible to know exactly what the effective stimulus intensity is at the eardrum, which can vary greatly between subjects (especially infants). If the used insert phones are not placed correctly, or move during measurement, this can be detected and compensated for. The best practical solution for this would be the use of real-ear-to-coupler differences (RECDs). The RECD is the difference between the sound pressure level measured at the eardrum (the real-ear SPL or RESPL) and the sound pressure level measured in a coupler for the same signal (Munro and Salisbury, 2002).

SOMA can be extended for eABR and eASSR measurements for CI subjects. This is subject of current research. Measuring auditory steady-state responses in CI subjects proves to be very challenging as the electrical artifacts generated by the cochlear implant obscures the faint responses. Advanced multichannel processing is certainly required for this purpose.

### **8.2.2 Multichannel EEG signal processing**

Although the multichannel ASSR framework already improves detection significantly for artifactual EEG compared to single channel measurements, there is still room for improvement. The presented framework currently is a simpli-

fied version of reality. Most auditory steady-state responses have at least two sources at different places in the brain. This results in an ASSR with varying amplitude and phase. Moreover, ASSR amplitudes vary over time during the measurement due to different levels of attention, especially with responses consisting mainly of generators in the auditory cortex. The varying EEG noise conditions may be better modelled by estimating the temporal covariance matrix  $\mathbf{K}_{\text{temp}}$  more precisely.

Next to the optimal electrode positions for recordings with modulation frequencies already studied in this research project (10 Hz and frequencies between 80 and 110 Hz), the study for optimal electrode positioning should be repeated for other modulation frequencies, varying from a few Hertz (generators in and beyond the auditory cortex), over 40 Hz (mainly auditory cortex) to frequencies above 110 Hz (brainstem only). This could increase the knowledge about the position and the magnitude of the different ASSR sources, and could indicate which electrode positions are preferred above other derivations.

It would be interesting to focus on the minimization of the *variance* of the difference scores in Chapter 7, instead of the *mean* of these scores. The minimization of the difference score means is widely used as a major criterion for ASSR detection. However, except with functional hearing loss and particular stimulus conditions (brief tones in patients with steep high frequency losses), the physiological threshold will always be higher than the behavioral threshold. One therefore usually just subtracts out the expected mean difference in order to come up with an estimated behavioral threshold. If this is so, then the physiological-behavioral difference is less important than the inter-subject variability of that difference. It would therefore be worthwhile to compare this variability across the different channels and channel combinations. The minimization of difference scores is strongly correlated with the SNR of the ASSR. As a result, the SNR-maximization criterion described in this research project can be used for difference score minimization. Unfortunately, a small pilot study showed that this criterion cannot be applied to difference score variance minimization. Moreover, there even was only a small correlation between SNR variance minimization and the minimization of difference score variances. A totally new signal processing approach should be worked out that focuses on reducing the variance of the difference score, or on a combined approach that assigns predefinable weights to the importance of difference score means and variances.

The determination of optimal electrode positions for three specific clinical applications should be highlighted too. Firstly, and highly important, the optimal electrode positions for infants vary with age. van der Reijden et al. (2005) already suggested some preferred scalp locations for infants, but it would be of clinical value to combine the proposed multichannel techniques in this thesis with the EEG data collected from these electrodes. Especially the robustness against artifacts and the benefits hereof should be investigated thoroughly. Sec-



ondly, the previous results should be checked with bone conduction ASSR. Auditory steady-state responses evoked by bone conductors suffer greatly from stimulus artifacts, and optimal electrode placement combined with the proposed multichannel techniques could reduce these artifacts. Finally, the holy grail of cochlear implant fitting would be the objective determination of comfortable and threshold levels using ASSRs. Currently an adequate fitting of a cochlear implant with (very) young children is extremely difficult, one of the reasons why one relies on hearing aids at first in the early months of life. The possibility to reliably determine the necessary thresholds for CI fitting objectively would signify an extra argument for early implantation. Additionally, the decision for early implantation depends on a hearing threshold estimation using the ABR- and/or ASSR-technique, and this decision should be trustworthy as cochlear implantation is currently not reversible. The ASSR-technique applied to CI fitting supposedly could be made possible by removing electric artifacts that a CI inevitably produces using smart electrode placement and appropriate signal processing.



# Bibliography

- Alaerts, J., Luts, H., Van Dun, B., Desloovere, C., and Wouters, J. (2008), “Latencies of auditory steady-state responses recorded in early infancy,” Submitted for publication.
- Alaerts, J., Luts, H., Van Dun, B., and Wouters, J. (2007a), Evaluating speech understanding by means of suprathreshold ASSR, in *Proc. of the 20th International Evoked Response Audiometry Study Group (IERASG)*, Bled, Slovenia, p. 54.
- Alaerts, J., Luts, H., and Wouters, J. (2007b), “Evaluation of middle ear function in young children: clinical guidelines for the use of 226-Hz and 1000-Hz tympanometry,” *Otol. Neurotol.*, 28, 727–732.
- Aoyagi, M., Kiren, T., Furuse, H., Fuse, T., Suzuki, Y., and Yokoto, M. (1994), “Pure-tone threshold prediction by 80-Hz amplitude-modulation following response,” *Acta Otolaryngol.*, suppl 511, 7–14.
- Aoyagi, M., Kiren, T., Kim, Y., Suzuki, Y., Fuse, T., and Koike, Y. (1993), “Optimal modulation frequency for amplitude-modulation following response in young children during sleep,” *Hear. Res.*, 65, 253–261.
- Armstrong, M., and Stapells, D. R. (2007), Multiple-stimulus interactions in the brainstem (80 Hz) and cortical (14 & 40 Hz) auditory steady-state responses, in *Proc. of the 20th International Evoked Response Audiometry Study Group (IERASG)*, Bled, Slovenia, p. 148.
- Back, A. D., and Weigend, A. S. (1997), “A first application of independent component analysis to extracting structure from stock returns,” *Int. J. Neural Syst.*, 8(5), 473–484.
- Batra, R., Kuwada, S., and Maher, V. L. (1986), “The frequency-following response to continuous tones in humans,” *Hear. Res.*, 21, 167–177.
- Beattie, R. C., and Kennedy, K. M. (1992), “Auditory brainstem response to tone bursts in quiet, notch noise, highpass noise, and broadband noise,” *J. Am. Ac. Audiol.*, 3, 349–360.

- Beattie, R. C., and Torre, P. (1997), "Effects of rise-fall time and repetition rate on the auditory brainstem response to 0.5 and 1 kHz tone bursts using normal-hearing and hearing-impaired subjects," *Scand. Audiol.*, 26, 23–32.
- Berg, A. L., Spitzer, J. B., Towers, H. M., Bartosiewicz, C., and Diamond, B. E. (2005), "Newborn hearing screening in the NICU: profile of failed auditory brainstem response/passed otoacoustic emission," *Pediatrics*, 116(4), 933–938.
- Bland, J. M., and Altman, D. G. (1995), "Multiple significant tests: the Bonferroni method," *Br. Med. J.*, 310, 170.
- Bohórquez, J., and Ozdamar, O. (2008), "Generation of the 40-Hz auditory steady-state response (ASSR) explained using convolution," *Clin. Neurophysiol.*, Epub ahead of print.
- Boukong, S., Toch, B., Saad, D., and Lowe, D. (2003), "ICA for watermarking digital images," *J. Mach. Learn. Res.*, 4, 1471–1498.
- Bradley, A. P., and Wilson, W. J. (2004), "On wavelet analysis of auditory evoked potentials," *Clin. Neurophysiol.*, 115(5), 1114–1128.
- Brillinger, D. R. (2001), *Time series: data analysis and theory* Philadelphia: SIAM.
- Capua, B. D., Costantini, D., Martufi, C., Latini, G., Gentile, M., and Felice, C. D. (2007), "Universal neonatal hearing screening: the Siena (Italy) experience on 19,700 newborns," *Early Hum. Dev.*, 83(9), 601–606.
- Cardoso, J., and Soloumiac, A. (1993), "Blind beamforming for non-Gaussian signals," *IEE Proc.-F*, 140(6), 362–370.
- Cebulla, M., Stürzebecher, E., and Elbering, E. (2006), "Objective detection of auditory steady-state responses: comparison of one-sample and q-sample tests," *J. Am. Ac. Audiol.*, 17, 93–103.
- Cebulla, M., Stürzebecher, E., and Wernecke, K. D. (2001), "Objective detection of the amplitude modulation following response (AMFR)," *Audiology*, 40, 245–252.
- Clark, G. (2003), *Cochlear implants, fundamentals & applications* New York: Springer-Verlag.
- Cohen, L. T., Rickards, F. W., and Clark, G. M. (1991), "A comparison of steady state evoked potentials to modulated tones in awake and sleeping humans," *J. Acoust. Soc. Am.*, 90, 2467–2479.
- Comon, P. (1994), "Independent component analysis, a new concept?," *Signal Process.*, 36, 287–314.

- Comon, P. (2004), "Blind identification and source separation in  $2 \times 3$  underdetermined mixtures," *IEEE Trans. Signal Process.*, 52(1), 11–22.
- Cone-Wesson, B., Rickards, F. W., Poulis, C., Parker, J., Tan, L., and Pollard, J. (2002), "The auditory steady-state response, III: clinical applications in infants and children," *J. Am. Ac. Audiol.*, 13, 270–282.
- 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, 1530–1540.
- Davies, M. E., and James, C. J. (2007), "Source separation using single channel ICA," *Signal Process.*, 87, 1819–1832.
- Davila, C. E., and Mobin, M. S. (1992), "Weighted averaging of evoked potentials," *IEEE Trans. Biomed. Eng.*, 39, 338–345.
- Davis, P. A. (1939), "Effects of acoustic stimuli on the waking human brain," *J. Neurophysiol.*, 2, 494–499.
- De Lathauwer, L., De Moor, B., and Vandewalle, J. (2000), "Fetal electrocardiogram extraction by blind source subspace separation," *IEEE Trans. Biomed. Eng.*, 47(5), 567–572.
- Deville, Y., Benali, M., and Abrard, F. (2004), "Differential source separation for underdetermined instantaneous or convolutive mixtures: concept and algorithms," *Signal Process.*, 84, 1759–1776.
- Dillon, H. (2001), *Hearing aids* New York: Thieme Publishers.
- Dimitrijevic, A., John, M. S., and Picton, T. W. (2004), "Auditory steady-state responses and word recognition scores in normal-hearing and hearing-impaired adults," *Ear Hear.*, 25(1), 68–84.
- Dimitrijevic, A., John, M. S., Van Roon, P., and Picton, T. W. (2001), "Human auditory steady-state responses to tones independently modulated in both frequency and amplitude," *Ear Hear.*, 22(2), 100–111.
- Dimitrijevic, A., John, M. S., Van Roon, P., Purcell, D. W., Adamonis, J., Ostroff, J., Nedzelski, J. M., and Picton, T. W. (2002), "Estimating the audiogram using auditory steady-state responses," *J. Am. Ac. Audiol.*, 13, 205–224.
- Dobie, R. A., and Wilson, M. J. (1989), "Analysis of auditory evoked potentials by magnitude-squared coherence," *Ear Hear.*, 10(1), 2–13.
- Dobie, R. A., and Wilson, M. J. (1994), "Objective detection of 40-Hz auditory evoked potentials: phase coherence vs. magnitude-squared coherence," *Electroencephalogr. Clin. Neurophysiol.*, 92, 405–413.

- Dobie, R. A., and Wilson, M. J. (1996), "A comparison of t test, F test, and coherence methods of detecting steady-state auditory-evoked potentials, distortion-product otoacoustic emissions, or other sinusoids," *J. Acoust. Soc. Am.*, 100(4, pt.1), 2236–2246.
- Dolphin, W. F., and Mountain, D. C. (1992), "The envelope following response: scalp potentials elicited in the Mongolian gerbil using sinusoidally AM acoustic signals," *Hear. Res.*, 58, 70–78.
- Elberling, C., Don, M., Cebulla, M., and Stürzebecher, E. (2007), "Auditory steady-state responses to chirp stimuli based on cochlear traveling wave delay," *J. Acoust. Soc. Am.*, 122(5), 2772–2785.
- Everest, F. A. (2001), *Master Handbook of Acoustics*, 4th edn New York: McGraw-Hill Professional.
- Fang, H., and De-Shuang, H. (2005), "Extracting mode components in laser intensity distribution by independent component analysis," *Appl. Opt.*, 44(18), 3646–3653.
- Fawcett, T. (2004), *ROC Graphs: notes and practical considerations for researchers. Technical Report* Palo Alto: HP Laboratories.
- Fobel, O., and Dau, T. (2004), "Searching for the optimal stimulus eliciting auditory brainstem responses in humans," *J. Acoust. Soc. Am.*, 116(4 Pt 1), 2213–2222.
- Fontaine, C. M., and Stapells, D. R. (2007), Efficiency of single versus multiple stimuli for 40 Hz auditory steady-state responses, in *Proc. of the 20th International Evoked Response Audiometry Study Group (IERASG)*, Bled, Slovenia, p. 147.
- Galambos, R., Makeig, S., and Talmachoff, P. J. (1981), "A 40-Hz auditory potential recorded from the human scalp," *Proc. Natl. Acad. Sci. USA*, 78(4), 2643–2647.
- Geisler, C. D. (1960), *Average responses to clicks in man recorded by scalp electrodes. Technical Report 380* Cambridge: MIT Research Laboratories of Electronics.
- Gorga, M. P., Johnson, T. A., Kaminski, J. R., Beauchaine, K. L., Garner, C. A., and Neely, S. T. (2006), "Using a combination of click- and tone burst-evoked auditory brain stem response measurements to estimate pure-tone thresholds," *Ear Hear.*, 27(1), 60–74.
- Gray, H. (1918), *Anatomy of the human body*, 20th edn Philadelphia: Lea & Febiger.
- Green, D., and Swets, J. A. (1966), *Signal detection theory and psychophysics* New York: John Wiley & Sons Inc.

- Hall, J. W. (2007), *New handbook of auditory evoked responses* Boston: Pearson.
- Hanley, J. A., and McNeil, B. J. (1982), "The meaning and use of the area under the Receiver Operating Characteristic (ROC) curve," *Radiology*, 143, 29–36.
- Hanley, J. A., and McNeil, B. J. (1983), "A method of comparing the areas under receiver operating characteristic curves derived from the same cases," *Radiology*, 148, 839–843.
- Herdman, A. T., Lins, O., Van Roon, P., Stapells, D. R., Scherg, M., and Picton, T. W. (2002), "Intracerebral sources of human auditory steady-state responses," *Brain Topogr.*, 15, 69–86.
- Herdman, A. T., and Stapells, D. R. (2001), "Thresholds determined using the monotic and dichotic multiple auditory steady-state response technique in normal-hearing subjects," *Scand. Audiol.*, 30, 41–49.
- Hoke, M., Ross, B., Wickesberg, R., and Lütkenhöner, B. (1984), "Weighted averaging - theory and application to electric response audiometry," *Electroencephalogr. Clin. Neurophysiol.*, 57, 484–489.
- Hyvärinen, A. (1999), "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Networks*, 10(3), 626–634.
- Hyvärinen, A., Karhunen, J., and Oja, E. (2001), *Independent component analysis* New York: John Wiley & Sons Inc.
- ISO 389 (1998), "Acoustics – Reference zero for the calibration of audiometric equipment," *International Organization for Standardization*, <http://www.iso.org>.
- James, C. J., and Lowe, D. (2003), "Extracting multisource brain activity from a single electromagnetic channel," *Artif. Intell. Med.*, 28(1), 89–104.
- Jewett, D. L., and Williston, J. S. (1971), "Auditory-evoked far fields averaged from the scalp of humans," *Brain*, 94(4), 681–696.
- John, M. S., Brown, D. K., Muir, P. J., and Picton, T. W. (2004), "Recording auditory steady-state responses in young infants," *Ear Hear.*, 25(6), 539–553.
- John, M. S., Dimitrijevic, A., and Picton, T. W. (2001a), "Weighted averaging of steady state responses," *Clin. Neurophysiol.*, 122, 555–562.
- John, M. S., Dimitrijevic, A., and Picton, T. W. (2002a), "Auditory steady-state responses to exponential modulation envelopes," *Ear Hear.*, 23, 106–117.

- John, M. S., Dimitrijevic, A., and Picton, T. W. (2003), "Efficient stimuli for evoking auditory steady-state responses," *Ear Hear.*, 24, 406–423.
- John, M. S., Dimitrijevic, A., Van Roon, P., and Picton, T. W. (1998), "Multiple auditory steady-state responses (MASTER): stimulus and recording parameters," *Audiology*, 37, 59–82.
- John, M. S., Dimitrijevic, A., Van Roon, P., and Picton, T. W. (2001b), "Multiple auditory steady-state responses to AM and FM stimuli," *Audiol. Neuro-Otol.*, 6, 12–27.
- John, M. S., and Picton, T. W. (2000a), "MASTER: a Windows program for recording multiple auditory steady-state responses," *Comput. Methods Programs Biomed.*, 61, 125–150.
- John, M. S., and Picton, T. W. (2000b), "Human auditory steady-state responses to amplitude-modulated tones: phase and latency measurements," *Hear. Res.*, 141(1–2), 57–79.
- John, M. S., Purcell, D. W., Dimitrijevic, A., and Picton, T. W. (2002b), "Advantages and caveats when recording steady-state responses to multiple simultaneous stimuli," *J. Am. Ac. Audiol.*, 13, 246–259.
- Johnson, D. H., and Dudgeon, D. E. (1993), *Array Signal Processing: Concepts and Techniques* London: Prentice-Hall.
- Johnson, T. A., and Brown, C. J. (2005), "Threshold prediction using the auditory steady-state response and the tone burst auditory brain stem response: a within-subject comparison," *Ear Hear.*, 26(6), 559–576.
- Joint Committee on Infant Hearing (2000), "Year 2000 position statement: Principles and guidelines for early hearing detection and intervention programs," *Pediatrics*, 106, 798–817.
- Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M. J., Iragui, V., and Sejnowski, T. J. (2000), "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, 37, 163–178.
- Junius, D., and Dau, T. (2005), "Influence of cochlear traveling wave and neural adaptation on auditory brainstem responses," *Hear. Res.*, 205(1–2), 53–67.
- Kemp, D. T. (1978), "Stimulated acoustic emissions from the human auditory system," *J. Acoust. Soc. Am.*, 64, 1386–1391.
- Kemp, D. T., Bray, P., Alexander, L., and Brown, A. M. (1986), "Acoustic emission cochleography – practical aspects," *Scand. Audiol.*, 25, 71–95.
- Knuth, K. H. (1999), A Bayesian approach to source separation, in *ICA '99 Proceedings*, Aussois, France, pp. 283–288.



- Kuwada, S., Anderson, J. S., Batra, R., Fitzpatrick, D. C., Teissier, N., and D'Angelo, W. R. (2002), "Sources of the scalp recorded amplitude-modulation following response," *J. Am. Ac. Audiol.*, 13, 188–204.
- Kuwada, S., Batra, R., and Maher, V. L. (1986), "Scalp potentials of normal and hearing-impaired subjects in response to sinusoidally amplitude-modulated tones," *Hear. Res.*, 21, 179–192.
- Lins, O. G., Picton, P. E., Picton, T. W., Champagne, S. C., and Durieux-Smith, A. (1995), "Auditory steady-state responses to tones amplitude-modulated at 80–110 Hz," *J. Am. Ac. Audiol.*, 97, 3051–3063.
- Lins, O. G., and Picton, T. W. (1995), "Auditory steady-state responses to multiple simultaneous stimuli," *Electroencephalogr. Clin. Neurophysiol.*, 96, 420–432.
- Lins, O. G., Picton, T. W., Boucher, B. L., Durieux-Smith, A., Champagne, S. C., Moran, L. M., Perez-Abalo, M. C., Martin, V., and Savio, G. (1996), "Frequency-specific audiometry using steady-state responses," *Ear Hear.*, 17(2), 81–96.
- Lütkenhöner, B., Hoke, M., and Pantev, C. (1985), "Possibilities and limitations of weighted averaging," *Biol. Cybern.*, 52, 409–416.
- Luts, H. (2005), Diagnosis of hearing loss in newborns: Clinical application of auditory steady-state responses, PhD thesis, Katholieke Universiteit Leuven, Leuven, Belgium.
- Luts, H., Desloovere, C., and Wouters, J. (2006), "Clinical application of dichotic multiple-stimulus auditory steady-state responses in high-risk newborns and young children," *Audiol. Neuro-Otol.*, 11, 24–37.
- Luts, H., Van Dun, B., Alaerts, J., and Wouters, J. (2008), "The influence of the detection paradigm in recording auditory steady-state responses," *Ear Hear.*, 29(4), 638–650.
- Luts, H., and Wouters, J. (2004), "Hearing assessment by recording multiple auditory steady-state responses: the influence of test duration," *Int. J. Audiol.*, 43(8), 471–478.
- Luts, H., and Wouters, J. (2005), "Comparison of MASTER and AUDERA for measurement of auditory steady-state responses," *Int. J. Audiol.*, 44, 244–253.
- Macmillan, N. A., and Creelman, C. D. (1991), *Detection theory: a user's guide* Cambridge: Cambridge University Press.
- Maiste, A. C., and Picton, T. W. (1989), "Human auditory evoked potentials to frequency-modulated tones," *Ear Hear.*, 10, 153–160.

- Malmivuo, J., and Plonsey, R. (1995), *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields* New York: Oxford University Press.
- Mason, J. A., and Herrmann, K. R. (1998), "Universal infant hearing screening by automated auditory brainstem response measurement," *Pediatrics*, 101, 221–228.
- Maurizi, M., Almadori, G., Paludetti, G., Ottaviani, F., Rosignoli, M., and Luciano, R. (1990), "40-Hz steady-state responses in newborns and in children," *Audiology*, 29, 322–328.
- McCormick, B. (1993), *Paediatric audiology 0-5 years*, 2nd edn London: Whurr Publishers.
- Mehl, A. L., and Thomson, V. (2002), "The Colorado newborn hearing screening project, 1992–1999: on the threshold of effective population-based universal newborn hearing screening," *Pediatrics*, 109, E7.
- Moller, A. R. (1974), "Responses of units in the cochlear nucleus to sinusoidally amplitude-modulated tones," *Exp. Neurol.*, 45, 105–117.
- Moller, A. R. (1994), *Neural generators of auditory evoked potentials* Boston: Allyn and Bacon.
- Monzingo, R. A., and Miller, T. W. (1980), *Introduction to Adaptive Arrays* New York: John Wiley and Sons.
- Moore, B. C. J. (2003), *An introduction to the psychology of hearing*, 4th edn Amsterdam: Academic Press.
- Morgan-Jones, R. A. (2001), *Hearing differently, the impact of hearing impairment on family life* London: Whurr Publishers.
- Mühler, R., von Döbenek, F., Ziese, M., and von Specht, H. (2005), Recording of multiple frequency ASSR with adaptive stimuli, in *Proc. of the 19th International Evoked Response Audiometry Study Group (IERASG)*, Havana, Cuba.
- Munro, K. J., and Salisbury, V. A. (2002), "Is the real-ear to coupler difference independent of the measurement earphone?," *Int. J. Audiol.*, 41, 408–413.
- Northern, J. L., and Downs, M. P. (2001), *Hearing in children*, 5th edn Baltimore: Williams & Wilkins.
- Norton, S. J., Gorga, M. P., Widen, J. E., Folsom, R. C., Sininger, Y., Cone-Wesson, B., Vohr, B. R., Mascher, K., and Fletcher, K. (2000), "Identification of neonatal hearing impairment: evaluation of transient evoked otoacoustic emission, distortion product otoacoustic emission, and auditory brainstem response test performance," *Ear Hear.*, 21(5), 508–528.

- Oates, P. A., and Purdy, S. C. (2001), "Frequency specificity of the human auditory brainstem and middle latency responses using notched noise masking," *J. Acoust. Soc. Am.*, 110(2), 995–1009.
- Oates, P. A., and Stapells, D. R. (1997), "Frequency specificity of the human auditory brainstem and middle latency responses to brief tones. II. Derived response analyses," *J. Acoust. Soc. Am.*, 102, 3609–3619.
- Pantev, C., and Khvoles, R. (1984), "Comparison of the efficiency of various criteria for artifact rejection in the recording of auditory brain-stem responses (ABR)," *Scand. Audiol.*, 13, 103–108.
- Parra, L., and Sajda, P. (2003), "Blind source separation via generalized eigenvalue decomposition," *J. Mach. Learn. Res.*, 4, 1261–1269.
- Perez-Abalo, M. C., Savio, G., Torres, A., Martín, V., Rodríguez, E., and Galán, L. (2001), "Steady state responses to multiple amplitude modulated tones: an optimized method to test frequency-specific thresholds in hearing-impaired children and normal-hearing subjects," *Ear Hear.*, 22, 200–211.
- Petitot, C., Collet, L., and Durrant, J. D. (2005), "Auditory steady-state responses (ASSR): effects of modulation and carrier frequencies," *Int. J. Audiol.*, 44(10), 567–573.
- Picton, T. W., Dimitrijevic, A., John, M. S., and Van Roon, P. (2001), "The use of phase in the detection of auditory steady-state responses," *Clin. Neurophysiol.*, 112, 1698–1711.
- Picton, T. W., Durieux-Smith, A., Champagne, S. C., Whittingham, J., Moran, L. M., Giguere, C., and Beauregard, Y. (1998), "Objective evaluation of aided thresholds using auditory steady-state responses," *J. Am. Ac. Audiol.*, 9, 315–331.
- Picton, T. W., Durieux-Smith, A., and Moran, L. M. (1994), "Recording auditory brainstem responses from infants," *Int. J. Pediatr. Otorhinolaryngol.*, 28, 93–110.
- Picton, T. W., John, M. S., Dimitrijevic, A., and Purcell, D. (2003), "Human auditory steady-state responses," *Int. J. Audiol.*, 42(4), 177–219.
- Picton, T. W., Linden, R. D., Hamel, G., and Maru, J. T. (1983), "Aspects of averaging," *Semin. Hear.*, 4, 327–341.
- Picton, T. W., Skinner, C. R., Champagne, S. C., Kellett, A. J., and Maiste, A. C. (1987), "Potentials evoked by the sinusoidal modulation of the amplitude or frequency of a tone," *J. Acoust. Soc. Am.*, 82, 165–178.
- Picton, T. W., Van Roon, P., and John, M. S. (2007), "Human auditory steady-state responses during sweeps of intensity," *Ear Hear.*, 28(4), 542–547.

- Plourde, G., and Picton, T. W. (1990), "Human auditory steady-state responses during general anesthesia," *Anesth. Analg.*, 71, 460–468.
- Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1992), "Confidence Limits on Estimated Model Parameters in *Numerical Recipes in C*," Cambridge: Cambridge University Press, chapter 15.6, pp. 689–698.
- Purcell, D. W., John, M. S., Schneider, B. A., and Picton, T. W. (2004), "Human temporal auditory acuity as assessed by envelope following responses," *J. Acoust. Soc. Am.*, 116(6), 3581–3593.
- Purdy, S. C., Houghton, J. M., Keith, W. J., and Greville, K. A. (1989), "Frequency-specific auditory brainstem responses. Effective masking levels and relationship to behavioural thresholds in normal hearing adults," *Audiology*, 28, 82–91.
- Rance, G., Rickards, F. W., Cohen, L. T., Vidi, S. D., and Clark, G. M. (1995), "The automated prediction of hearing thresholds in sleeping subjects using auditory steady-state evoked potentials," *Ear Hear.*, 16, 499–507.
- Rance, G., and Tomlin, D. (2006), "Maturation of auditory steady-state responses in normal babies," *Ear Hear.*, 27, 20–29.
- Rance, G., Tomlin, D., and Rickards, F. W. (2006), "Comparison of auditory steady-state responses and tone-burst auditory brainstem responses in normal babies," *Ear Hear.*, 27(6), 751–762.
- Rees, A., Green, G. G., and Kay, R. H. (1986), "Steady-state evoked responses to sinusoidally amplitude-modulated sounds recorded in man," *Hear. Res.*, 23, 123–133.
- Regan, D. (1989), *Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine* Amsterdam: Elsevier.
- Reyes, S. A., Lockwood, A. H., Salvi, R. J., Coad, M. L., Wack, D. S., and Burkard, R. F. (2005), "Mapping the 40-Hz auditory steady-state response using current density reconstructions," *Hear. Res.*, 204(1–2), 1–15.
- Rickards, F. W., Tan, L. E., Cohen, L. T., Wilson, O. J., Drew, J. H., and Clark, G. M. (1994), "Auditory steady state evoked potentials in newborns," *Br. J. Audiol.*, 28(6), 327–337.
- Riquelme, R., Kuwada, S., Filipovic, B., Hartung, K., and Leonard, G. (2006), "Optimizing the stimuli to evoke the amplitude modulation following response (AMFR) in neonates," *Ear Hear.*, 27, 104–119.
- Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (1967), "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey," *J. Neurophysiol.*, 30, 769–793.

- Ross, B., and Pantev, C. (2004), "Auditory steady-state responses reveal amplitude modulation gap detection thresholds," *J. Acoust. Soc. Am.*, 115(5(1)), 2193–2206.
- Savio, G., Cardenas, J., Perez-Abalo, M., Gonzalez, A., and Valdes, J. (2001), "The low and high frequency auditory steady-state responses mature at different rates," *Audiol. Neuro-Otol.*, 6, 279–287.
- Scharf, L. L. (1991), *Statistical signal processing : Detection, estimation and time series analysis*, 1st edn New York: Addison Wesley.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995), "Speech recognition with primarily temporal cues," , 270(5234), 303–304.
- Simmer, K. U., Bitzer, J., and Marro, C. (2001), "Post-filtering techniques in *Microphone Arrays: Signal Processing Techniques and Applications*," M. Brandstein, D. Ward (eds.), Berlin: Springer-Verlag, chapter 3, pp. 39–60.
- Small, S. A., and Stapells, D. R. (2006), "Multiple auditory steady-state response thresholds to bone-conduction stimuli in young infants with normal hearing," *Ear Hear.*, 27, 219–228.
- Small, S. A., and Stapells, D. R. (2008), "Normal ipsilateral/contralateral asymmetries in infant multiple auditory steady-state responses to air- and bone-conduction stimuli," *Ear Hear.*, 29, 185–198.
- Stapells, D. R., Galambos, R., Costello, J. A., and Makeig, S. (1988), "Inconsistency of auditory middle latency and steady-state responses in infants," *Electroencephalogr. Clin. Neurophysiol.*, 71, 289–295.
- Stapells, D. R., Galambos, R., and Makeig, S. (1987), "Auditory steady-state responses: threshold prediction using phase coherence," *Electroencephalogr. Clin. Neurophysiol.*, 67, 260–270.
- Stapells, D. R., and Oates, P. (1997), "Estimation of the pure-tone audiogram by the auditory brainstem response: a review," *Audiol. Neuro-Otol.*, 2, 257–280.
- Stapells, D. R., Picton, T. W., and Smith, A. D. (1982), "Normal hearing thresholds for clicks," *J. Acoust. Soc. Am.*, 72, 74–79.
- Stögbauer, H., Kraskov, A., Astakhov, S. A., and Grassberger, P. (2004), "Least-dependent-component analysis based on mutual information," *Phys. Rev. E*, 70, 066123.
- Stürzebecher, E., Cebulla, M., and Elberling, C. (2005), "Automated auditory response detection: statistical problems with repeated testing," *Int. J. Audiol.*, 44, 110–117.

- Stürzebecher, E., Cebulla, M., Elberling, C., and Berger, T. (2006), “New efficient stimuli for evoking frequency-specific auditory steady-state responses,” *J. Am. Ac. Audiol.*, 17(6), 448–461.
- Stürzebecher, E., Cebulla, M., and Pschirrer, U. (2001), “Efficient stimuli for recording of the amplitude modulation following response,” *Audiology*, 40, 63–68.
- Tang, Y., and Norcia, A. M. (1995), “An adaptive filter for steady-state evoked responses,” *Electroencephalogr. Clin. Neurophysiol.*, 96, 268–277.
- Thumak, A. I., Durrant, J. D., and Collet, L. (2007), “80 Hz auditory steady-state responses (ASSR) at 250 Hz and 12,000 Hz,” *Int. J. Audiol.*, 46(1), 26–30.
- Trotter, H. F. (1959), “An elementary proof of the central limit theorem,” *Arch. Math.*, 10, 226–234.
- Valdes, J. L., Perez-Abalo, M. C., Martin, V., Savio, G., Sierra, C., Rodriguez, E., and Lins, O. (1997), “Comparison of statistical indicators for the automatic detection of 80 Hz auditory steady state responses,” *Ear Hear.*, 18, 420–429.
- van der Reijden, C. S., Mens, L. H. M., and Snik, A. F. M. (2001), “Comparing signal-to-noise ratios of amplitude modulation following responses from four EEG derivations in awake normally hearing adults,” *Audiology*, 40, 202–207.
- van der Reijden, C. S., Mens, L. H. M., and Snik, A. F. M. (2004), “Signal-to-noise ratios of the auditory steady-state response from fifty-five EEG derivations in adults,” *J. Am. Ac. Audiol.*, 15(10), 692–701.
- van der Reijden, C. S., Mens, L. H. M., and Snik, A. F. M. (2005), “EEG derivations providing auditory steady-state responses with high signal-to-noise ratios in infants,” *Ear Hear.*, 26, 299–309.
- Van Dun, B., Moonen, M., and Wouters, J. (2008a), “Optimal electrode selection for multichannel electroencephalogram based detection of auditory cortex and brainstem auditory steady-state responses,” Submitted for publication.
- Van Dun, B., Rombouts, G., Wouters, J., and Moonen, M. (2008b), “A procedural framework for auditory steady-state response detection,” *IEEE Trans. Biomed. Eng.*, Accepted for publication.
- Van Dun, B., Verstraeten, S., Alaerts, J., Luts, H., Moonen, M., and Wouters, J. (2008c), “A flexible research platform for multi-channel auditory steady-state response measurements,” *J. Neurosci. Meth.*, 169, 239–248.

- Van Dun, B., Wouters, J., and Moonen, M. (2006), Independent component analysis applied on multi-channel auditory steady-state response measurements, in *Proc. of the Independent Component Analysis Research Network (ICArn) International Workshop*, Liverpool, UK, pp. 89–92.
- Van Dun, B., Wouters, J., and Moonen, M. (2007a), “Improving auditory steady-state response detection using independent component analysis on multi-channel EEG data,” *IEEE Trans. Biomed. Eng.*, 54(7), 1220–1230.
- Van Dun, B., Wouters, J., and Moonen, M. (2007b), Multi-channel Wiener filtering based auditory steady-state response detection, in *Proc. of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vol. II, Honolulu, HI, USA, pp. 929–932.
- Van Kerschaver, E., and Stappaerts, L. (2004), *De ALGO gehoorscreening, rapport van de werkjaren 2001 en 2002* Brussels: Kind & Gezin.
- Van Trees, H. L. (2001), *Detection, estimation, and modulation theory* New York: John Wiley & Sons.
- Van Wieringen, A., and Wouters, J. (2008), “LIST and LINT: Sentences and numbers for quantifying speech understanding in severely impaired listeners for Flanders and the Netherlands,” *Int. J. Audiol.*, 47, 348–355.
- Warner, E. S., and Proudler, I. K. (2003), “Single-channel blind signal separation of filtered MPSK signals,” *IEE Proc. Radar Sonar Navig.*, 150(6), 396–402.
- Wegner, O., and Dau, T. (2002), “Frequency specificity of chirp-evoked auditory brainstem responses,” *J. Acoust. Soc. Am.*, 111(3), 1318–1329.
- Yoshinaga-Itano, C., Sedey, A. L., Coulter, D. K., and Mehl, A. L. (1998), “Language of early- and later-identified children with hearing loss,” *Pediatrics*, 102, 1161–1171.
- Yost, W. A. (2000), *Fundamentals of hearing: An introduction*, 4th edn San Diego: Academic Press.
- Zhang, R., McAllister, G., Scotney, B., McClean, S., and Houston, G. (2006), “Combining wavelet analysis and Bayesian networks for the classification of auditory brainstem response,” *IEEE Trans. Inf. Technol. Biomed.*, 10(3), 458–67.





# List of Publications

## International journal papers

- Van Dun, B., Wouters, J., and Moonen, M. (2007), “Improving auditory steady-state response detection using independent component analysis on multi-channel EEG data,” *IEEE Trans. Biomed. Eng.*, 54(7), 1220–1230.
- Van Dun, B., Verstraeten, S., Alaerts, J., Luts, H., Moonen, M., and Wouters, J. (2008), “A flexible research platform for multi-channel auditory steady-state response measurements,” *J. Neurosci. Meth.*, 169, 239–248.
- Van Dun, B., Rombouts, G., Wouters, J., and Moonen, M. (2008), “A procedural framework for auditory steady-state response detection,” *IEEE Trans. Biomed. Eng.*, Accepted for publication.
- Van Dun, B., Moonen, M., and Wouters, J. (2008), “Optimal electrode selection for multichannel electroencephalogram based detection of auditory cortex and brainstem auditory steady-state responses,” Submitted for publication.
- Luts, H., Van Dun, B., Alaerts, J., and Wouters, J. (2008), “The influence of the detection paradigm in recording auditory steady-state responses,” *Ear. Hear.*, 29(4), 638–650.
- Alaerts, J., Luts, H., Van Dun, B., Desloovere, C., and Wouters, J. (2008), “Latencies of auditory steady-state responses recorded in early infancy,” Submitted for publication.
- Alaerts, J., Luts, H., Hofmann, M., Van Dun, B., and Wouters, J. (2008), “Auditory steady-state responses to low modulation rates and their relation to speech understanding,” Submitted for publication.

## International conference papers

- Van Dun, B., Wouters, J., and Moonen, M. (2006), “Independent component analysis applied on multi-channel auditory steady-state response measurements,” *Proc. of the ICA Research Network International Workshop*, Liverpool, United Kingdom, 89–92.
- Van Dun, B., Wouters, J., and Moonen, M. (2007), “Multi-channel Wiener filtering based auditory steady-state response detection,” *Proc. of the 32nd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Honolulu, HI, USA, vol. 2, 929–932.

## Abstracts without conference proceedings

- Van Dun, B., Wouters, J., Moonen, M., Vergult, A., Welkenhuyzen, J., Laneau, J., and Luts, H., “Improvement of response detection with SSEP after comparison of several statistical tests,” *4th Belgian Society of Audiology (BSA)*, Liège, Belgium, 13/12/2003.
- Van Dun, B., Moonen, M., and Wouters, J., “The quest for babies and their hearing thresholds,” *PhD Symposium*, Leuven, Belgium, 20/04/2005.
- Van Dun, B., Wouters, J., and Moonen, M., “Improving ASSR detection using independent component analysis,” *19th Biennial International Evoked Response Audiometry Study Group (IERASG)*, Havana, Cuba, 12–16/06/2005.
- Alaerts, J., Luts, H., Van Dun, B., Desloovere, C., and Wouters, J., “Objective hearing assessment in young children,” *40 jaar LAW Symposium*, Leuven, Belgium, 30/09/2006.
- Luts, H., Alaerts, J., Van Dun, B., and Wouters, J., “De ASSR-techniek als opvolging van de neonatale gehoorscreening,” *40 jaar LAW Symposium*, Leuven, Belgium, 30/09/2006.
- Luts, H., Alaerts, J., Van Dun, B., and Wouters, J., “Clinical application of auditory steady-state responses,” *Annual congress of the Royal Belgian Society of Otorhinolaryngology*, Head & Neck Surgery, Leuven, Belgium, 17–18/11/2006.
- Alaerts, J., Luts, H., Van Dun, B., and Wouters, J., “Evaluating speech understanding by means of suprathreshold ASSR,” *20th International Evoked Response Audiometry Study Group (IERASG)*, Bled, Slovenia, 10–14/06/2007.

- Luts, H., Van Dun, B., Alaerts, J., and Wouters, J., “Objective detection of ASSR: do’s and don’ts,” *20th International Evoked Response Audiometry Study Group (IERASG)*, Bled, Slovenia, 10–14/06/2007.
- Van Dun, B., Wouters, J., and Moonen, M., “Improving ASSR detection using multi-channel Wiener filtering,” *20th International Evoked Response Audiometry Study Group (IERASG)*, Bled, Slovenia, 10–14/06/2007.
- Luts, H., Alaerts, J., Van Dun, B., and Wouters, J., “Building a bridge between detection of hearing loss and rehabilitation with the auditory steady-state response,” *Conference of the International Collegium for Rehabilitative Audiology (ICRA)*, Leuven, Belgium, 17–20/06/2007.
- Luts, H., Van Dun, B., Alaerts, J., and Wouters, J., “Do’s and don’ts in the recording of ASSRs,” *9th International Tinnitus Seminars*, Göteborg, Sweden, 15–18/06/2008.
- Alaerts, J., Luts, H., Van Dun, B., Royackers, L., Desloovere, C., and Wouters, J., “Obtaining a reliable audiogram in neonates,” *Intensive programme speech language therapy*, Barcelona, Spain, 28–30/08/2008.
- Alaerts, J., Luts, H., Van Dun, B., Royackers, L., Desloovere, C., and Wouters, J., “ASSR als audiologische follow-up na neonatale gehoorscreening: evaluatie na 5 jaar,” *Symposium ‘Hoe vroeger, hoe beter?!’*, Leuven, Belgium, 11/10/2008.



# Curriculum Vitae



Bram Van Dun was born in the lovely town of Zoersel, Belgium, on the 6th of December 1979. In 2003 he received the electrical engineering degree from the Katholieke Universiteit Leuven, Belgium. From 2003 he has been employed as a research assistant in a joint research project of the group of Experimental Oto-Rhino-Laryngology (Dept. Neurosciences) and the Signals, Identification, System Theory and Automation group (Dept. Electrical Engineering), both

of the K.U.Leuven, under supervision of Prof. dr. J. Wouters and Prof. dr. ir. M. Moonen. His research interests are in the area of applied signal processing techniques for EEG and auditory evoked potentials. One of his main research activities aimed at studying the impact of signal processing strategies on hearing threshold determination. In 2003 and 2004, he received funding from the 'Research Foundation - Flanders' (FWO-Vlaanderen). From 2005 till 2008 he was funded by the 'Innovation through Science and Technology in Flanders' (IWT-Vlaanderen).

Bram Van Dun is happily married to Ilse Cornelis.



