

# **Automatic Detection and Assessment of Swallowing Based on Bioimpedance and Electromyography Measurements**

**Enabling Control of Functional Electrical Stimulation Synchronously to Volitional Swallowing in Dysphagic Patients**

vorgelegt von  
Dipl.-Ing.  
Holger Nahrstaedt  
geb. in Berlin

von der Fakultät IV - Elektrotechnik und Informatik  
der Technischen Universität Berlin  
zur Erlangung des akademischen Grades  
Doktor der Ingenieurwissenschaften  
-Dr.-Ing.-  
genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr.-Ing. Reinhold Orglmeister

Gutachter: Prof. Dr.-Ing. Jörg Raisch

Gutachterin: Prof. Dr. Simona Ferrante

Gutachter: Prof. Dr.-Ing. Clemens Gühmann

Tag der wissenschaftlichen Aussprache: 09. Mai 2017

Berlin 2017



## Zusammenfassung

Schluckstörungen komplizieren für die Betroffenen jegliche Nahrungs- und Flüssigkeitsaufnahme und können zu folgenschweren Erkrankungen führen. Durch den Verlust eines normalerweise perfekt koordinierten Zusammenspiels der Muskulatur wird der Kehlkopf und das Zungenbein nicht weit genug oder im falschen Augenblick nach oben und vorne gezogen. Dadurch wird der Zugang zu den unteren Atemwegen nicht verschlossen und/oder der Zugang zur Speiseröhre nicht ausreichend geöffnet. Es kommt zum Übertritt von Teilen des Speichels und/oder der Nahrung in die Atemwege und folgenschwere Entzündungen in den unteren Atemwegen können hervorgerufen werden. Durch das verborgene Zusammenwirken vieler Komponenten (26 Muskelpaare, fünf Hirnnerven und drei Zervikalnerven) gestaltet sich das Erstellen einer exakten Diagnose schwierig. Obwohl zuverlässige video-basierte Systeme zur Diagnose bereitstehen, kann durch die Entwicklung von kleinen, preiswerten Sensoren, welche den Schluckvorgang in Echtzeit erfassen können, sowohl die Therapie als auch Diagnose von Schluckstörungen entscheidend verbessert werden. Solche Sensoren könnten längere Untersuchungen ermöglichen und Feedback zum Patienten und Therapeuten geben. Eine zum Schluck synchronisierte, elektrische Muskelstimulation könnte ebenfalls durch solche Sensoren ausgelöst werden.

Im Rahmen dieser Arbeit wurde ein Messsystem entwickelt, mit dem ein Elektromyogramm (EMG) und eine Bioimpedanz (BI) am Hals gemessen wird. Es wurde eine automatisierte Schluckerkennung entwickelt und erfolgreich an Daten sowohl von gesunden Probanden als auch von Patienten mit Schluckstörungen evaluiert. Es konnte gezeigt werden, dass die BI Messsignale mit der Bewegung von Zungenbein und Kehlkopf korrelieren und somit die pharyngeale Schluckphase abbilden. Mittels Messung von BI und EMG konnte die submentale Muskulatur zur Schluckunterstützung rechtzeitig zum Anfang der Hehebewegung von Zungenbein und Kehlkopf über elektrische Stimulation unterstützt werden und somit der Schutz der Atemwege während des Schlucks verbessert werden. An einem Patienten wurde gezeigt, dass die Muskelstimulation rechtzeitig zum Schluckbeginn erfolgte und die Hebung verbessert wurde.

Die Ergebnisse aus dieser Arbeit legen den Grundstein für ein System zur automatisierten Auswertung von erfassten Schlucken, welches als Biofeedback-Therapie eingesetzt werden könnte.

Es wurde gezeigt, dass die zeitgerechte Aktivierung einer elektrischen Muskelstimulation zur Unterstützung des Schlucks möglich ist. Ein solches System würde Patienten helfen und könnte auch außerhalb von Kliniken eingesetzt werden.



## Abstract

Dysphagia (swallowing disorders) complicates any food and fluid intake for patients or makes them impossible. Due to the loss of a normally coordinated interplay of muscles, the larynx, and the hyoid bone are not moved properly. Thus, the bolus might not be prevented from entering the lower airways and/or the entrance to the esophagus might not be opened sufficiently. The transfer of parts of the saliva and food into the lower airway can cause aspiration pneumonia. Since swallowing is triggered by a hidden interplay of numerous components (26 pairs of muscles, five cranial nerves, and three cervical nerves), an accurate diagnosis is difficult. Although powerful video-based systems, such as endoscopy and fluoroscopy, provide enough information for diagnosis, the development of a small, inexpensive, and non-invasive sensor system, which could detect the swallowing process in real-time, could drastically improve therapy and diagnosis of dysphagia. Such sensors could allow longer examinations and could give feedback to the patient and therapist. Electrical stimulation synchronous to swallowing could be triggered by such sensors.

In this work, a measurement system was developed which is able to measure electromyography (EMG) and bioimpedance (BI) at the neck. An automated swallowing detection was developed and successfully evaluated on data from healthy subjects and patients. It could be shown that the BI measurement signal is correlated to the displacement of the hyoid bone and the larynx. In order to support swallowing by functional electrical stimulation, algorithms have been developed which are able to detect the onset of a swallow in the very beginning. This allows supporting the airway protective displacement of hyoid bone and larynx by electrical muscle stimulation. It has been shown on one patient that the developed methods based on BI and EMG measurements are able to control functional electrical stimulation and that swallowing could be significantly improved by this.

Based on the achieved results in this work, a system able to automatically analyze swallows was developed that can be used for diagnosis and biofeedback therapy. It has been shown that triggering of functional electric stimulation is possible in order to support swallowing. This system will help patients suffering from dysphagia and can also be used outside a clinical environment.



Dedicated to my beloved wife Claudia and my sweet daughter Tamara Arite.



## Acknowledgements

I would like to express my gratitude to my first supervisor Prof. Dr.-Ing. Jörg Raisch for giving me the possibility to perform my research and writing this thesis and to my second supervisor Dr. Thomas Schauer for supporting me and discussing every part of my thesis in detail. Furthermore, I wish to express my gratitude to my dissertation evaluators Prof. Dr. Simona Ferrante and Prof. Dr.-Ing. Clemens Gühmann and to Prof. Dr.-Ing. Reinhold Orglmeister for chairing the doctoral committee.

Sincere thanks go to Dr. phil. Corinna Schultheiss and PD Dr. med. Rainer O. Seidl for recording several hours of measurement data containing swallows from healthy subjects and patients and explaining to me the complicated process of swallowing. I am really thankful that Corinna manually prepared all datasets by marking all swallows within. My personal thanks go to all subjects and patients who supported this research by allowing to record their swallows.

I would like to thank Dr. phil. Corinna Schultheiss, Benjamin Riebold, Christina Salchow and Markus Valtin for reviewing parts of my thesis and helping me to improve it.

I have to thank my wife Claudia for pushing me forward, helping me improving my English and giving me the required time for letting me finish my thesis.

Finally, I would like to acknowledge the thousands of individuals who have coded for open-source projects. It is due to their efforts that scientific work can be conducted with powerful tools.



# Table of Contents

<b>List of Figures</b>	<b>xv</b>
<b>List of Tables</b>	<b>xix</b>
<b>Abbreviations</b>	<b>xxi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The Swallowing Process . . . . .	1
1.2 Dysphagia . . . . .	4
1.3 Electromyography . . . . .	6
1.4 Bioimpedance . . . . .	7
1.5 Functional Electrical Stimulation . . . . .	8
1.6 Main Contributions of this Thesis . . . . .	9
1.7 Outline of this Thesis . . . . .	10
1.8 Publications . . . . .	12
<b>2 State of the Art in Diagnosis and Treatment of Swallowing Disorders (Dysphagia)</b>	<b>13</b>
2.1 Diagnosis of Swallowing Disorders . . . . .	13
2.1.1 Clinical Swallowing Tests . . . . .	14
2.1.2 Videofluoroscopic Swallowing Study . . . . .	15
2.1.3 Fiberoptic Endoscopic Evaluation of Swallowing . . . . .	15
2.1.4 Cervical Auscultation . . . . .	16
2.1.5 Sonography . . . . .	16
2.1.6 Real-time Capable Measurement Methods . . . . .	16
2.1.6.1 Piezoelectric Sensors . . . . .	17
2.1.6.2 Swallowing Sounds . . . . .	17
2.1.6.3 Accelerometry . . . . .	18
2.1.6.4 Electromyography . . . . .	20
2.1.6.5 Electroglottography . . . . .	21
2.1.6.6 Impedance Tomography . . . . .	21
2.1.6.7 Bioimpedance Measurement . . . . .	22
2.2 Therapy of Swallowing Disorders . . . . .	22
2.2.1 Strengthening Exercises and Compensatory Methods . . . . .	23
2.2.2 Sensory Stimulation . . . . .	24

## TABLE OF CONTENTS

---

2.2.3	Electrical Nerve Stimulation . . . . .	25
2.2.4	Intramuscular Stimulation . . . . .	26
2.2.5	Neuromuscular surface stimulation . . . . .	26
2.3	Conclusions . . . . .	28
<b>3</b>	<b>BI- and EMG-Measurement System</b>	<b>31</b>
3.1	Summary . . . . .	31
3.2	Motivation . . . . .	32
3.3	Measurement Device PHYSIOSENSE . . . . .	33
3.4	Current Sources . . . . .	37
3.5	Voltage Measurement . . . . .	40
3.6	Analog Signal Processing . . . . .	42
3.6.1	EMG Filtering . . . . .	42
3.6.2	BI Filtering for Measurement at 50 kHz . . . . .	44
3.6.3	BI Filtering for Measurement at 100 kHz . . . . .	44
3.6.4	Amplitude Demodulation . . . . .	45
3.7	Fault Detection and Automatic BI Configuration . . . . .	47
3.8	Digital Signal Processing . . . . .	49
3.9	Calibration . . . . .	53
3.9.1	Voltage Calibration for EMG Measurement . . . . .	53
3.9.2	Impedance Calibration for BI Measurement . . . . .	54
3.9.3	BI and EMG Response on High Voltage Pulses . . . . .	56
3.10	Video Reference System . . . . .	57
3.11	Conclusions . . . . .	58
<b>4</b>	<b>Automatic Segmentation and Classification of Swallowing</b>	<b>61</b>
4.1	Summary . . . . .	61
4.2	Motivation . . . . .	62
4.3	Methods . . . . .	63
4.3.1	Signal Acquisition . . . . .	63
4.3.2	Study Procedure . . . . .	63
4.3.3	Signal Preprocessing of EMG and BI . . . . .	64
4.3.4	Segmentation of EMG Activity Periods . . . . .	67
4.3.4.1	Introduction to the Double-Threshold Detector . . . . .	67
4.3.4.2	Statistical Properties of the Double-Threshold Detector . . . . .	71
4.3.4.3	Method for Tuning the Double-Threshold Detector . . . . .	75
4.3.4.4	Optimization of the Parameters $r_0$ and $m$ . . . . .	75
4.3.4.5	Determination of an Optimal Threshold $\hat{\zeta}$ . . . . .	76
4.3.4.6	Performance Benchmark Tests . . . . .	79
4.3.5	Piece-wise Linear Approximation of BI . . . . .	82
4.3.6	Segmentation of Swallowing by a Heuristic Valley Search . . . . .	87
4.3.7	Feature Extraction . . . . .	89

4.3.8	Support Vector Machine . . . . .	95
4.3.9	Validation of the Support Vector Machine . . . . .	97
4.4	Results . . . . .	98
4.4.1	Data Collection . . . . .	98
4.4.2	Detection of EMG Activity Periods . . . . .	99
4.4.3	Piece-wise Linear Approximation of BI . . . . .	100
4.4.4	Segmentation of BI Data by a Heuristic Valley Search . . . . .	100
4.4.5	Swallowing Segmentation . . . . .	102
4.4.6	SVM and Feature Parameter Optimization . . . . .	103
4.4.7	Classification Performance . . . . .	109
4.5	Conclusions . . . . .	113
<b>5</b>	<b>Assessment of Pharyngeal Swallowing by Bioimpedance and EMG</b>	<b>117</b>
5.1	Summary . . . . .	117
5.2	Motivation . . . . .	118
5.3	Methods . . . . .	119
5.3.1	Subjects . . . . .	119
5.3.2	Measurement Setup . . . . .	119
5.3.3	Extraction of Hyoid and Laryngeal Displacement . . . . .	119
5.3.4	Signal Pre-processing of EMG and BI . . . . .	121
5.3.5	Swallow Segmentation . . . . .	121
5.3.6	Data Processing . . . . .	122
5.3.7	Statistical Analysis across all Patients . . . . .	123
5.3.8	Individual Statistical Analysis for each Patient . . . . .	124
5.4	Results . . . . .	124
5.4.1	Test Subjects . . . . .	124
5.4.2	Correlation and Regression Analysis . . . . .	124
5.5	Conclusions . . . . .	130
<b>6</b>	<b>BI- and EMG-Based Control of FES For Supporting Swallowing</b>	<b>133</b>
6.1	Summary . . . . .	133
6.2	Motivation . . . . .	134
6.3	Methods . . . . .	136
6.3.1	Experimental Setup . . . . .	136
6.3.2	Stimulation Artifact Removing and Preprocessing of EMG and BI . . . . .	137
6.3.3	Online Detection of EMG Activity Periods . . . . .	139
6.3.4	Online Segmentation of BI Measurements . . . . .	140
6.3.5	Control of FES . . . . .	143
6.4	Results . . . . .	145
6.5	Conclusions . . . . .	148
<b>7</b>	<b>Conclusion and Future Work</b>	<b>149</b>

## TABLE OF CONTENTS

---

<b>References</b>	<b>151</b>
<b>Appendix A Component Lists for PHYSIOSENSE</b>	<b>169</b>
A.1 Current Source for BI Measurement . . . . .	169
A.2 Voltage Measurement . . . . .	171
A.3 Analog Signal Processing . . . . .	172
A.3.1 EMG Filtering . . . . .	172
A.3.2 BI Filtering for Measurement at 50 kHz . . . . .	173
A.3.3 BI Filtering for Measurement at 100 kHz . . . . .	175
A.3.4 Amplitude Demodulation . . . . .	177

# List of Figures

1.1	Sagittal section of the head and neck. Anatomical illustration is from [2] . . . . .	2
1.2	Infrahyoid and suprahyoid muscles. Anatomical illustration is from [2] . . . . .	3
1.3	Muscles of the pharynx. Anatomical illustration is from [2] . . . . .	4
3.1	Measurement system PHYSIOSENSE. The picture is also used in [135] and [136] . . . . .	34
3.2	Structure of the PHYSIOSENSE measurement system . . . . .	35
3.3	BI Measurement using the two- and four-electrode measurement method . . . . .	36
3.4	Schematic of the floating current source with CMFB . . . . .	38
3.5	Bioimpedance measurement ranges . . . . .	39
3.6	Circuit design of a bipolar voltage measurement channel . . . . .	40
3.7	CMRR measurement for the EMG channel . . . . .	41
3.8	Structure of the filter and amplifier block FiltEMG . . . . .	42
3.9	Simulated frequency responses of the filter characteristics of EMG1, BI1, and BI2 . . . . .	43
3.10	Structure of FiltBI1 for extracting BI at 50 kHz and FaultBI1 for fault detection . . . . .	44
3.11	Structure of FiltBI2 for extracting BI at 100 kHz and FaultBI2 for fault detection . . . . .	45
3.12	Schematics of the precision full wave rectifier . . . . .	45
3.13	Simulated output of the full wave rectifier B4 with a sine wave of 50 kHz as input . . . . .	47
3.14	Automatic configuration of current amplitude and gain setting . . . . .	48
3.15	FFT of an EMG and BI measurement . . . . .	50
3.16	Frequency response of the FIR filter for EMG and BI measurements . . . . .	52
3.17	Measured frequency response of the PHYSIOSENSE device . . . . .	53
3.18	Impedance calibration of the BI measurement . . . . .	54
3.19	Electrode position for measuring BI/EMG on the forearm while FES is active. Picture is from [146] . . . . .	56
3.20	EMG and BI measurement during FES . . . . .	57
4.1	Electrode positions for measuring BI/EMG at the neck. Anatomical illustrations are from [2] . . . . .	64
4.2	Histogram of the discrete difference sequence from an EMG recording . . . . .	65
4.3	EMG signal with and without removal of spikes . . . . .	66
4.4	EMG and BI before and after the preprocessing. Based on [132] . . . . .	67
4.5	Frequency response of the whitening filter . . . . .	68
4.6	Demonstration of the double-threshold detector on 40 measurement samples . . . . .	70

LIST OF FIGURES

---

4.7	ROC curves of the double-threshold detector . . . . .	71
4.8	ROC of the double-threshold for different values of $r_0$ . . . . .	76
4.9	Estimation of the disturbance variance $\sigma_d^2$ . . . . .	78
4.10	Exemplary variance profile of a trial . . . . .	80
4.11	Signals of the double-threshold detector benchmark . . . . .	81
4.12	Comparison of different approximation methods . . . . .	83
4.13	Implementation of the bottom-up algorithm . . . . .	85
4.14	Numbered line segments of a BI measurement around a swallow . . . . .	88
4.15	Adjustment of minimum and end point of a swallow. Based on [132] . . . . .	89
4.16	Definition of area based features. Based on [132] . . . . .	91
4.17	Features generated by the SAX method from a normalized BI sequence. Based on [132]	92
4.18	Features generated by the SAX method from a normalized and squared EMG sequence	93
4.19	Influence of $VS_{onset}$ and $VS_{emg}$ on segmenting datasets I, II, and III . . . . .	101
4.20	Influence of $VS_{onset}$ and $VS_{emg}$ on segmenting dataset IV . . . . .	102
4.21	Example of a ROC curve . . . . .	104
4.22	Grid search for finding the best learning parameters for the datasets I, II, III, and IV	106
4.23	SAX-features for EMG and BI of data from datasets I, II, and III . . . . .	107
4.24	SAX-features for EMG and BI of data from dataset IV . . . . .	108
4.25	Sensitivity, specificity, and accuracy of all subjects and patients . . . . .	110
5.1	Definition of the tracking points and their coordination system. The VF image is also used in [147] and [172] . . . . .	120
5.2	Pre-processed BI measurement data during a swallow together with the extracted displacement trajectories of hyoid and larynx . . . . .	123
5.3	Correlation of the largest drop in BI to the maximal displacement of hyoid and larynx	128
5.4	Correlation of the change in BI to the displacement of the hyoid and larynx . . . .	129
6.1	Electrode positions for applying electrical stimulation within several studies. Anatomical illustrations are from [2] . . . . .	135
6.2	Electrode positions for electrical stimulation and BI/EMG measurements. Anatomical illustrations are from [2] . . . . .	136
6.3	Signal processing of BI/EMG signals during active FES . . . . .	138
6.4	Step response of the online and the offline EMG filter . . . . .	139
6.5	Implementation of the BEST_LINE function . . . . .	142
6.6	Implementation of the modified SWAB algorithm . . . . .	143
6.7	Influence of the $max\_diff$ parameter on the online slope detection . . . . .	144
6.8	Comparison of a swallow with and without stimulation support . . . . .	146
A.1	Circuit layout with components of the floating current source . . . . .	169
A.2	Circuit layout with components of the voltage measurement input stage . . . . .	171
A.3	Structure of FiltEMG with components . . . . .	172
A.4	Structure of FiltBI1 and FaultBI1 with components . . . . .	173
A.5	Structure of FiltBI2 and FaultBI2 with components . . . . .	175

A.6 Circuit layout with components of the high-precision full-wave rectifier . . . . . 177



# List of Tables

4.1	Results of the double-threshold detector benchmark . . . . .	82
4.2	Example for the matrix $T$ and vector $M$ . . . . .	87
4.3	Definition of accuracy, sensitivity, and specificity based on the confusion matrix . .	96
4.4	Datasets I, II, III, and IV . . . . .	98
4.5	Influence of $max\_error$ on the piece-wise linear approximation of BI . . . . .	100
4.6	List of the parameters for the swallowing segmentation of datasets I, II, III, and IV .	103
4.7	Cross-validation AUC for the ten best SAX parameter combinations . . . . .	105
4.8	Cross-validation AUC for different features . . . . .	105
4.9	SAX and training parameters used for swallowing classification . . . . .	106
4.10	Performance of the classifier on the datasets I, II, III, and IV . . . . .	109
4.11	Performance of the classifier on the datasets I, II, and III (healthy subjects) . . . . .	111
4.12	Achieved performance of the trained classifier for different bolus types of dataset IV	112
4.13	Achieved performance of the classifier trained only on the dataset IV for different bolus types of dataset IV . . . . .	112
5.1	Time difference between the swallow related events and the onset of the trans-UES flow . . . . .	125
5.2	Pearson correlation coefficients of the correlation between the maximum drop in BI and the maximal displacement . . . . .	126
5.3	Pearson correlation coefficients of the correlation between the amplitude change in BI and the displacement trajectories of hyoid and larynx . . . . .	126
5.4	Correlation coefficients of the correlation between BI and displacement trajectory of hyoid bone and larynx for each patient . . . . .	127
6.1	List of parameters for the online swallowing segmentation used in the pilot study .	147
6.2	Effect of different stimulation currents during swallowing . . . . .	147
A.1	Component list for the floating current source . . . . .	170
A.2	Component list for the voltage measurement input stage . . . . .	172
A.3	Component list for FiltEMG . . . . .	173
A.4	Component list for FiltBI1 and FaultBI1 . . . . .	174
A.5	Component list for FiltBI2 and FaultBI2 . . . . .	176
A.6	Component list for the high precision full-wave rectifier . . . . .	177



# Abbreviations

- AC alternating current 7, 136
- ADC analog to digital converter 31
- Ag/AgCl silver/silver-chloride 6, 33, 136
- APCA adaptive piece-wise constant approximation 83
- AS active shielding 35
- AUC area under the ROC curve 103
- BI bioimpedance 1, 22, 31, 61, 117, 133, 149
- BIA bioimpedance analysis 7
- BPF band-pass filter 45
- C-SVM C-support vector machine 95
- CMFB common-mode feedback 37
- CMRC common mode rejection circuit 41
- CMRR common-mode rejection ratio 37
- DC direct current 31
- DDS direct digital frequency synthesis 38
- DFT discrete Fourier transform 82
- DWT discrete wavelet transform 82
- ECG electrocardiography 32
- EGG electroglottography 21, 32
- EIT electrical impedance tomography 32
- EMG electromyography 1, 17, 31, 61, 117, 133, 149
- ENT ear, nose, and throat 99
- FEES fiber-optic endoscopic evaluation of swallowing 5, 14, 62, 131, 149
- FES functional electrical stimulation 1, 26, 31, 133, 149
- FFDA feedback fully differential amplifier 43
- FFT fast Fourier transform 50
- FIR finite impulse response 43
- HPF high-pass filter 38
- IA instrumentation amplifier 35
- IIR infinite impulse response 65, 137
- LMN lower motor neuron 6
- LPF low-pass filter 38
- MFB multiple feedback 42
- MOSFET metal oxide semiconductor field-effect transistor 42
- MU motor unit 6
- MUAP motor unit action potential 6
- NMES neuromuscular electrical stimulation 134
- OP operational amplifier 40
- PAA piece-wise aggregate approximation 83, 138
- PC personal computer 10, 31, 119, 136
- PDF probability density function 65
- PEG percutaneous endoscopic gastrostomy 22
- PLA piece-wise linear approximation 82, 121, 140
- RBF radial basis function 95
- RD reference driver 35
- RMS root mean square 37
- ROC receiver operating characteristic 71
- SAX symbolic aggregate approximation 91
- SNR signal-to-noise ratio 50, 69
- SPI serial peripheral interface 51
- SQNR signal-to-quantization-noise ratio 50
- SVD singular value decomposition 83
- SVM support vector machine 10, 18, 61, 149
- SWAB Sliding Window And Bottom-up 141
- TMS transcranial magnetic stimulation 24
- TTL transistor-transistor logic 34, 137
- UES upper esophagus sphincter 118
- USART universal asynchronous receiver transmitter 51
- USB universal serial bus 31, 119
- VF videofluoroscopy 1, 15, 62, 117, 149
- VFSS videofluoroscopic swallowing study 5, 13, 62, 119



# 1

## Introduction

This thesis is about the development and testing of new methods for diagnosis and treatment of swallowing disorders. Measurement of electromyography (EMG) and bioimpedance (BI) at the neck were made possible by the development of a measurement device. Automatic swallowing detection algorithms were developed and tested on the recorded measurement data of healthy people and patients. The relationship between swallowing and EMG and BI measurements was analyzed by performing videofluoroscopy (VF) on patients. It was possible to trigger functional electrical stimulation (FES) of the muscles involved in swallowing in real-time such that the swallowing could be supported.

This chapter introduces the basics of swallowing, dysphagia, EMG, BI, and FES. The structure of this thesis is outlined and the main contributions by the author are given.

### 1.1 The Swallowing Process

Swallowing is a vital and complex process that is generated by complex movements of the jaw, tongue, hyoid bone, larynx and esophagus. Swallowing is associated with ingestion and therefore volitional initiated, but can also be reflexive triggered in order to protect the lower airways. Swallowing can be divided into three sequential phases [1]. The first phase is referred to as the oral phase and is sometimes divided in literature into an oral preparation phase and an oral absorption phase. The entire oral phase is controlled voluntarily. The preparation phase includes the food or fluid intake, size reduction and insalivation of the bolus. In the oral absorption phase, parts of the food or fluid are separated by the tongue in order to form a bolus. Undulating movements of the tongue transport the bolus into the posterior pharynx. In the transition to the pharyngeal phase of swallowing, the bolus is pushed from the tongue dorsally towards the soft palate and the posterior pharyngeal wall.

The pharyngeal phase is reflex controlled and the initial triggering of the swallowing reflex is caused by irritation of the posterior oral cavity and pharynx. Although it is reflex controlled, adaption to environmental conditions (e.g. drinking in headstand position) and compensating of

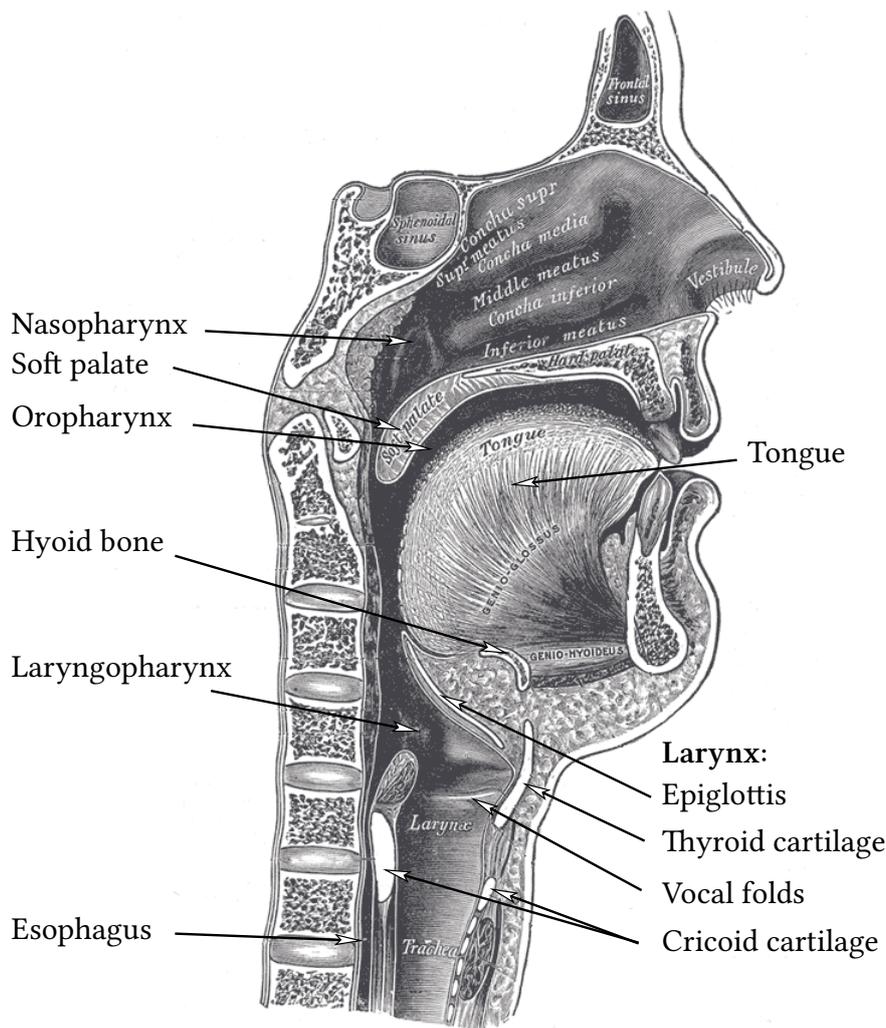


Figure 1.1: Sagittal section of the head and neck. Anatomical illustration is from [2].

impairment is possible in the pharyngeal swallowing phase [3]. The swallowing reflex inhibits respiration. The pharyngeal phase begins with the closing and the stabilization of the jaw by the jaw muscles. This supports the suprahyoid muscles<sup>1</sup> by pulling the hyoid bone up- and forward in direction of the jaw. Subsequently, the larynx (cf. Fig. 1.1) is raised by the infrahyoid muscles<sup>2</sup>. The suprahyoid and infrahyoid muscles are depicted in Figure 1.2. The access to the nasopharynx is obstructed by muscular lifting and stretching of the soft palate in order to prevent a transfer of bolus parts into the upper respiratory tract. By the up- and forward movement of the larynx and the hyoid bone, the epiglottis passively tilts downwards over the entrance to the larynx and seals the lower respiratory tract. Should parts of the bolus nevertheless reach the entrance to the lower airways in height of the vestibular folds or below, a reflexive cough is triggered in order to clear the lower airways. As additional protection, the vestibular folds and vocal cords are rapidly closed when a small amount of liquid triggers the pharyngoglottal closure reflex [4]. The tongue is pushed back- and downwards into the pharynx, and the posterior pharyngeal wall contracts towards the tongue. Beginning from the nasopharynx down to the esophagus, peristalsis

<sup>1</sup>Suprahyoid muscles: mylohyoid, stylohyoid, digastric, and geniohyoid.

<sup>2</sup>Infrahyoid muscles: sternohyoid, sternothyroid, thyrohyoid, and omohyoid.

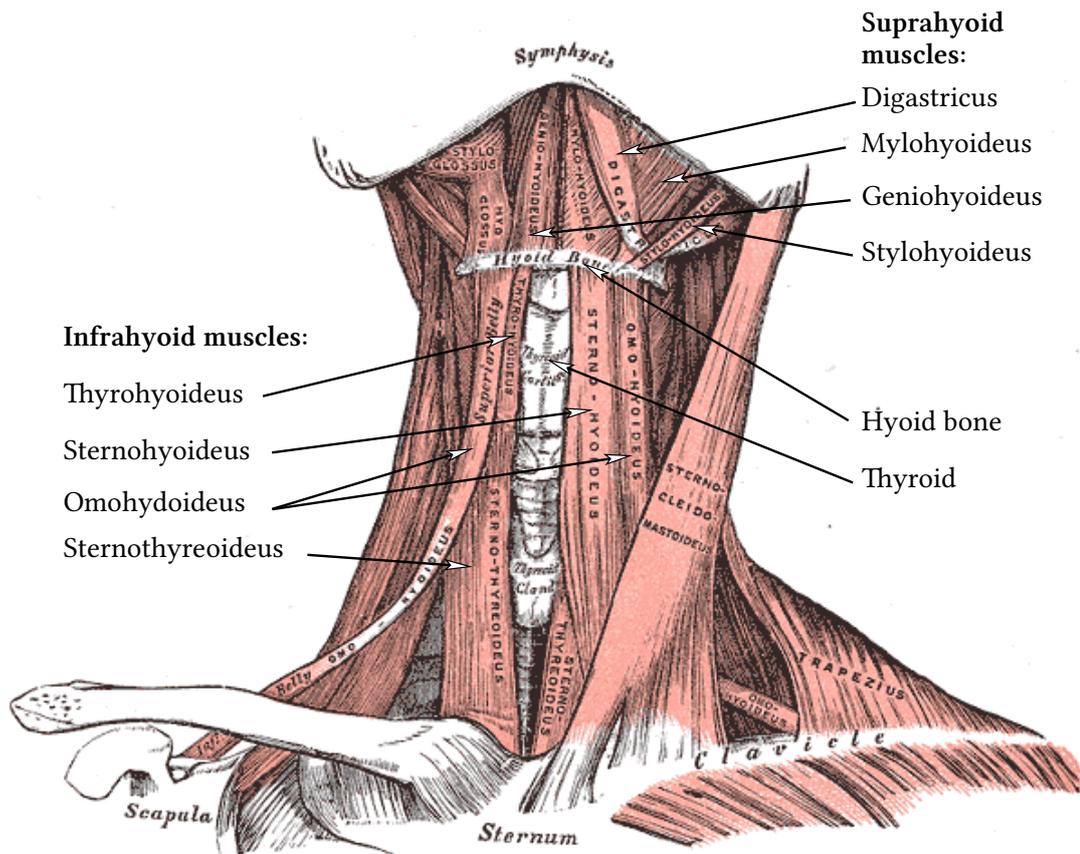


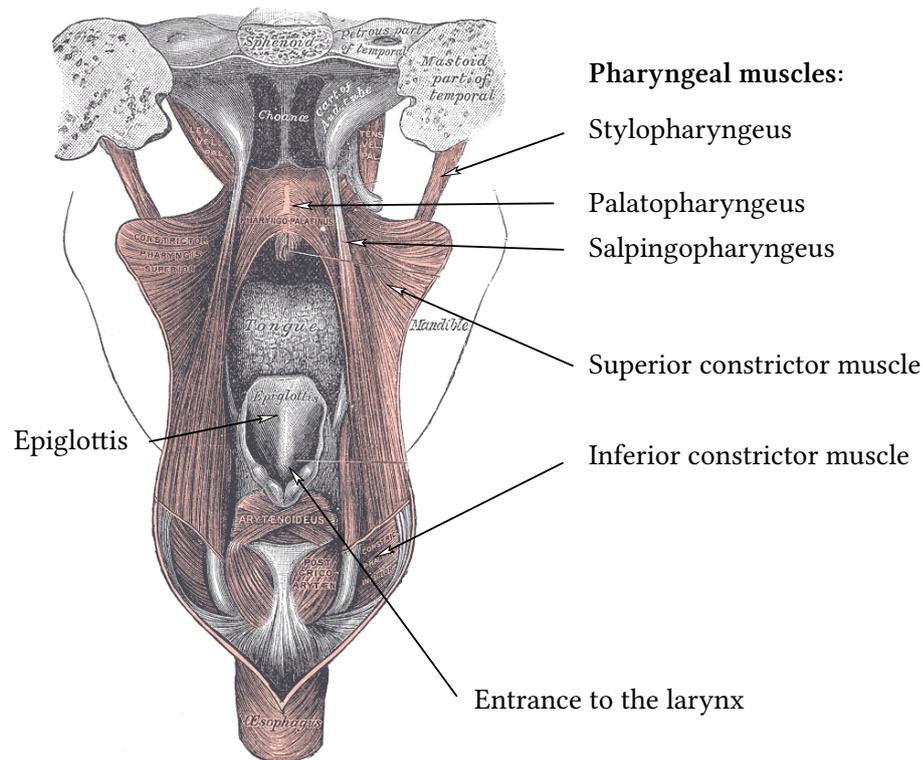
Figure 1.2: Infrahyoid and suprahyoid muscles. Anatomical illustration is from [2].

of the pharyngeal muscles (cf. Fig. 1.3) propels the bolus into the esophagus by wave-like muscle contractions. The pharyngoesophageal segment is opened by relaxation of the upper esophageal sphincter which consists mainly of the lower part of the inferior constrictor. Counting from onset of lingual peristalsis, the upper esophageal sphincter opens after 0.35 s to 0.85 s and closes after 0.9 s to 1.2 s. The duration of upper esophageal sphincter opening is between 0.3 s to 0.6 s [5].

In the esophageal phase, which is mainly controlled by the enteric nervous systems, the bolus enters the esophagus and is propelled downwards by peristalsis of the striated and smooth muscles. Finally, the lower esophageal sphincter relaxes and the bolus is pushed into the stomach.

The swallowing phases are coordinated in an adult through the brain stem and cortical paths. Four motor cranial nerve nuclei<sup>3</sup> and their cranial nerves are involved in the complete swallowing process in order to fire the respective motor neurons of the muscles involved. In the pharyngeal phase, additionally the motor neurons arising from vertebrae C1 to C3 are involved. Sensory feedback is triggered during swallowing via the spinal trigeminal nucleus and the nucleus of the solitary tract in the brain stem. The wave-like muscle activation pattern in the esophagus, which is used in the esophageal phase in order to transport the bolus into the stomach, is controlled by the enteric nervous system.

<sup>3</sup>Cranial nerve nuclei: spinal trigeminal nucleus, facial nerve nucleus, nucleus ambiguus, and the nucleus nervi hypoglossi.



**Figure 1.3:** Muscles of the pharynx. The middle constrictor muscle is not shown. Anatomical illustration is from [2].

## 1.2 Dysphagia

Dysphagia (swallowing disorder) can be classified into oropharyngeal dysphagia and esophageal dysphagia. In the following, only the oropharyngeal dysphagia is considered.

The complete closure of the larynx and its timing take a central role in safe swallowing. Besides swallowing, the pharynx is used in respiration and speech. The air flow path and the bolus transport path cross each other in the laryngopharynx. The lower airway and the nasopharynx must be securely sealed during the pharyngeal phase of swallowing. In case of closure failure, which can be caused by pharyngeal dysphagia, a transfer of saliva, liquid or food into the lower airway (aspiration) takes place. Aggregated in the lung, bacteria from the bolus can cause acute aspiration pneumonia which can be fatal. In general, swallowing disorders include difficulty to transport food or liquid from the mouth to the stomach and can lead to dehydration, malnutrition, chronic lung disease and acute aspiration pneumonia. Symptoms besides aspiration that are related to dysphagia are drooling, leaking, nasal penetration, laryngeal penetration, retention and pharyngeal regurgitation [6].

Dysphagia may result from changes in the swallowing-related structures, caused by surgery, cancer, infectious diseases, burns or trauma. Swallowing problems are common in patients who had a stroke which is caused by a blockage of blood flow (ischemia) or internal bleeding (hemorrhage) within the brain. Dysphagia may also be caused by any neurological diseases that effect swallowing as cranial nerve dysfunction, myasthenia gravis, Guillain-Barré syndrome, multiple sclerosis, motor neurone disease, Parkinson's disease, or Alzheimer's disease [7].

Depending on the examination method, several studies have shown that 37 % to 78 % of acute stroke patients show symptoms of dysphagia [8]. According to the World Health Organization<sup>4</sup>, 15 million people suffer a stroke each year. A pneumonia occurred in around 7 % to 33 % of stroke patients with dysphagia [8]. Especially older people are at a greater risk for aspiration pneumonia when they suffer from dysphagia [9]. Six month after an acute stroke, 17 of 67 patients (25 %) with initial swallowing abnormalities showed aspiration [10].

These numbers show that diagnosis of dysphagia, especially clinical dysphagia assessment, is an important tool in order to decrease mortality and complications for patients. Based on diagnosis of cause and development of each specific case of dysphagia, appropriate intervention techniques have to be selected and evaluated on each patient. By this, the bolus transportation and airway protection during swallowing should be improved. Otherwise, the airway has to be blocked by a tracheotomy tube in order to prevent saliva aspiration [11].

Because of the complex anatomy, the overlapping muscular processes and the complex control, the diagnosis of swallowing disorders is complex and significant differences are possible in the injury patterns. Especially the detection of an existent aspiration is crucial for the patient's prognosis and the decision of further treatment steps in the underlying disease. In approximately 50 % of the patients who aspirate during swallowing, the aspiration triggers no physiological signs such as coughing, face turning red or uncoordinated breathing [12]. This is also called silent aspiration and means a significant risk of complications for the patient as the aspiration may remain undetected.

Both gold standards, videofluoroscopic swallowing study (VFSS) and fiber-optic endoscopic evaluation of swallowing (FEES), for examination of the swallowing process are not real-time capable. They can only be used for the assessment of very few swallows and have only a moderate inter- and intra-rater reliability. A comparison of scores given by different raters for the same video recording results in an inter- and intra-rater agreement of 50 % to 75 % [13], [14]. An automated evaluation of the swallowing process and its phases is not possible with neither of the two methods.

Treatments of dysphagia are selected by a therapist based on the initial diagnosis. Due to the high diagnosis effort of the gold standards VFSS and FEES, the therapeutic impact may not be frequently reevaluated. Real-time measurement methods that are able to automatically evaluate the swallowing process, and that utilize simply attachable, inexpensive and non-invasive sensors, could improve the diagnosis and therapy of swallowing disorders. Such a system should robustly detect swallows in the measurement signals even in presence of signal artifacts which could be caused by noise, speech, coughing, or movement of tongue, neck, or head. The detected swallow in the time series should then be classified into normal and abnormal swallows. Systems based on such methods could also be used outside the clinical environment for diagnosis of swallowing disorders, e.g. at home, at a retirement home or at a doctors office. Continuous long term measurements would become possible that could be used for examination of the swallowing behavior during sleep or how the swallowing behavior changes without supervision of therapists or doctors.

Besides the development and validation of new diagnostic tools, new therapies that utilize real-time measurement systems could significantly improve the treatment of dysphagia. Standard treatment of swallowing disorders uses dietary adjustments, compensatory techniques, sensory stimulation or strengthening techniques including continuous electrical muscle stimulation. A real-

---

<sup>4</sup><http://www.who.int/whr/2002/en/>

time measurement method could be used for new therapeutic measures that are synchronized with swallowing. Thus, instantaneous biofeedback would become feasible and would help the teaching of compensatory swallowing techniques. Instead of an undirected and constant stimulation of sensory or muscle fibers, stimulation could be precisely activated during swallowing. Such a measurement system could be used for controlling functional electrical stimulation of swallowing related muscles in order to increase elevation of hyoid bone and larynx as well as lower airway protection.

### 1.3 Electromyography

EMG is the recording of the electrical activity of muscle tissue and can be used to study activity patterns. Aberrations produced by various diseases, especially upper motor neuron disorders, are measurable with EMG [15].

A muscle is composed of many fibers which are activated by lower motor neurons (LMNs). A motor unit (MU) consists of a single LMN cell body, its axon which innervates the corresponding muscle fibers, neuromuscular junctions and all innervated muscle fibers. How many muscle fibers belong to an MU depends on the specific muscle.

Every skeletal muscle is innervated by LMNs. The cell body of an LMN lays within the spinal cord or in a somatic motor nucleus of the brain stem. The connection to the brain is accomplished by the upper motor neurons which connect the motor area of the cortex with the appropriate level in the spinal cord.

When a LMN is activated, a wave of depolarization is going down the axon to the neuromuscular junctions and then to each connected muscle fiber causing depolarization which propagates along the muscle fibers. The electrical potential which is associated with the activation of a single MU is called motor unit action potential (MUAP). The muscle contraction depends on the number of activated MUs and their activation rate. Muscle activity can be recorded by surface silver/silver-chloride (Ag/AgCl) electrodes, hooked wire electrodes and needle electrodes.

Surface EMG measurements can be used to indicate which muscle or muscle group contracts at the moment and provide an approximate measure of the contraction strength in comparison to the maximal possible contraction [16]. Surface EMG is normally measured bipolarly with two identical Ag/AgCl gel electrodes in direction of the muscle fibers. The measured signal is a composite of all MUAPs occurring in the muscles underlying the skin between the electrodes.

Hooked wire electrodes are inserted into a muscle and are more selective. The activity of deep muscles (e.g. the geniohyoid) can only be recorded with hooked wire electrodes.

Needle electrodes have only a very small contact area and are highly selective. Therefore, needle electrodes can be used to analyze action potentials of a single or few motor units. Needle EMG is mainly used for assessment of motor units in order to diagnose neuromuscular diseases[17].

As swallowing is not possible without muscle activity, EMG can be used as an indicator for swallowing activity. During the oral phase of swallowing, the masseter muscle is active and remains active during the pharyngeal phase in order to stabilize the mandible. Its activation can be recorded with surface EMG electrodes. The suprahyoid muscle group<sup>5</sup> whose activity can also be recorded with surface EMG electrodes is then activated to initiate the pharyngeal swallowing phase by

---

<sup>5</sup>Mylohyoid, geniohyoid, stylohyoid, and anterior digastric muscles form the suprahyoid muscle group.

elevating the hyoid and larynx. Vaiman et al. [18] showed in their work that surface EMG can be used as a screening method for dysphagia. An overview over the usage of EMG in the field of dysphagia can be found in Section 2.1.6.4.

## 1.4 Bioimpedance

The passive electrical properties of biological tissue can be grouped together under the term of bioimpedance. Impedance is defined as the relation of voltage to current in terms of amplitude and phase. In case that the load contains capacitive and/or inductive parts (e.g. tissue), both are frequency-dependent. Bioimpedance depends on the type of tissue (e.g. bone or muscle tissue) and is the result of the conductivity of the intra- and extracellular space and the behavior of cell membranes. Depending on the frequency range, tissue behaves more as a volume conductor (below 100 kHz) or as a dielectric (over 50 kHz) [19]. As biological tissue has normally an inhomogeneous resistivity across its volume, BI measurement can only collect the average impedance distribution in a certain area.

Bioimpedance analysis (BIA) is used to determine the fat-free mass and total body water by measuring BI of the whole body [20]. It is assumed that the tissue itself is time invariant during the measurement period.

If bioimpedance is measured at a specific frequency, it can be used to measure the impedance changes over time. For example, the stroke volume of the heart can be estimated by the maximum rate of change of BI within the thorax in comparison to the measured average BI [21]. In the following, the term BI refers to the measurement of a time dependent BI signal at a specific measurement frequency.

Bioimpedance can be measured using two, three or four electrodes [19]. An alternating current (AC) is applied between two electrodes and the resulting electrical potential is measured between another pair of electrodes (tetrapolar), between one additional electrode and one current electrode (tripolar) or between the current electrodes itself (dipolar).

In biological tissue, the impedance of the electrodes and the connection between electrodes and tissue can be greater than the impedance of the tissue itself. In case of dipolar measurement, the electrode impedance affects the measurement and cannot be separated from the tissue impedance.

Tetrapolar measurements allow to measure the potential from a different pair of electrodes with a high resistance amplifier which reduces strongly the influence of the electrode impedance to measurement results. Using four electrodes, one pair of electrodes is used for current injection into the tissue and the other pair is measuring the resulting voltage. The measured voltage divided by the applied current is called transfer impedance [19]. The term refers to the fact that the current flow between the current electrodes is transferred through the tissue into a voltage across the voltage measurement electrodes. A more detailed discussion about BI measurement methods will be given in Chapter 3 of this thesis.

It was demonstrated by Kusuhara et al. [22] that a four-electrode transcutaneous BI measurement at the neck area can capture swallowing induced movements. They used a measurement frequency of 50 kHz. BI measurements seems to be convenient for swallowing assessment and

should be further investigated. More details about the state of art in BI measurement can be found in Section 2.1.6.7.

### 1.5 Functional Electrical Stimulation

Electrical stimulation of tissue is a technique by which an electrical field is created across cells in order to release their action potential. In contrast to therapeutic electrical stimulation or transcutaneous nerve stimulation, FES is applied on one or more muscles in order to generate functional movements which partly compensate lost or impaired functions [23]. The stimulation electrodes are attached near or in muscle tissue which is sufficiently innervated by LMNs such that the applied electrical field triggers the action potential of several motor neurons.

Contrary to the natural motor nerve recruitment pattern, in which firstly smaller diameter fibers that are weaker but more fatigue-resistant are contracting [24], [25], electrical stimulation activates firstly the fast fatiguing and large diameter fibers at low stimulation intensities [26]. This inverse recruitment order leads to faster fatigue in electrically activated muscles. The second reason why FES leads to faster fatigue is that during FES all nearby motor neurons are synchronously stimulated. In voluntary contraction, different motor neurons are activated asynchronously, which allows fatigued fibers to recover and leads to a lower effective stimulation frequency for each fiber.

Different stimulus waveforms are possible. Mostly biphasic current-controlled waveforms are used which consist of a positive and a negative square pulse. The current amplitudes of both squares are equal and the negative square pulse follows after a delay of around 100  $\mu\text{s}$ . A biphasic, charge balanced stimulation pattern prevents migration of ions from the electrodes into the tissue and allows muscle stimulation at both electrodes. The current amplitude is defined as the height of the square waves. The pulse width is defined as the width of one square pulse and is normally in a range from 10  $\mu\text{s}$  to 500  $\mu\text{s}$ . The stimulation frequency, by which the square pulse pattern is repeated, lays in a range of 20 Hz to 100 Hz, and is usually constant. Instead of using one pair of stimulation pulses,  $N$  closely spaced stimulation pulse pairs which are called  $N$ -lets can be used for enhancing muscle force produced by FES.

The relation between produced force and applied stimulation is nonlinear and time variant due to the recruitment function and fatigue of the stimulated muscle fibers [27]. The recruitment level and thus the produced muscle force depend nonlinear on the stimulus current amplitude and the pulse width at a fixed stimulation frequency [28]. In order to produce a muscle contraction of desired intensity, the stimulation intensity has to be tuned for each patient and can be controlled in order to compensate nonlinearity and fatigue. Normally, either the current amplitude or the pulse width is kept constant.

Most studies which are using electrical stimulation in therapy of dysphagia apply therapeutic electrical stimulation with surface electrodes. During several sessions, the submental and laryngeal muscles of a patient are stimulated for a defined duration. The stimulation does not produce functional movements. The stimulation intensity is set manually to a maximum tolerated level at which muscle contraction occurs. The swallowing related muscles are strained by this therapy, which results in a better swallowing due to muscle strengthening. Some studies showed that the outcome of this therapy is similar to the outcome of standard swallowing therapy [29]. It was shown

that intramuscular stimulation is able to elevate the larynx [30] which is an important function for swallowing and airway protection during swallowing. Using surface electrodes, the achieved effect is lower and selective stimulation of the thyrohyoid muscle is impossible [31]. A study by Humbert et al. [32] investigated ten different electrode positions in comparison with VF on healthy adults. Only the submental stimulation did not decrease hyoid or larynx during swallowing in comparison to swallowing without stimulation. However, the stimulation did not significantly increase the elevation. The study was repeated on chronic dysphagic patients with therapeutic electrical stimulation on the submental and laryngeal muscles [33]. The results showed that an increase of elevation of the larynx and hyoid triggered by surface muscle stimulation could only be observed in two patients whereas for the other eight patients the elevation of larynx and hyoid decreased which led to an increased aspiration risk in these patients. Thus, therapeutic electrical stimulation on the laryngeal muscles with surface electrodes may strengthen swallowing involved muscles but generally increases the aspiration risk and does not elevate the hyoid bone or the larynx. Therapeutic electrical stimulation on the submental muscles does not seem to have any negative effects.

Besides surface stimulation, functional movement can be triggered by electrical nerve stimulation and intramuscular stimulation. Hadley et al. [34] were able to show that selective stimulation of the hypoglossal nerve could achieve highly laryngeal elevation in canines. By stimulating the left recurrent laryngeal nerve, Broniatowski et al. [35] could achieve closure of the true vocal folds which helps preventing aspiration. Burnett et al. [36] showed the feasibility of intramuscular stimulation in order to achieve laryngeal elevation on healthy subjects. While these studies look very promising for the future, there exists no satisfactory solution for automatically triggering such supporting stimulation in the correct phase of swallowing.

In a study by Leelamanit et al. [37], the region around the thyrohyoid muscles is stimulated by surface electrodes for one second whenever EMG activity of the posterior tongue could be measured. The results are promising, but other studies showed that it is difficult to stimulate the thyrohyoid muscle with surface electrodes [30], [31]. Secondly, EMG measurements of the suprahyoid muscles are not truly swallowing specific and are timely inconsistent due to the varying relationship between EMG threshold and swallowing onset. Healthy subjects seem to be able to self-trigger FES to support their swallow [36]. To the best knowledge of the author, no other solutions for triggering FES in order to improve swallowing have been proposed until now. Within this thesis, methods are proposed, which utilize measurements of EMG and BI for triggering transcutaneous FES of submental muscles during swallowing.

## 1.6 Main Contributions of this Thesis

In the presented thesis, a developed measurement device *PHYSIOSENSE* is described which allows the measurement of BI through the neck by transcutaneous Ag/AgCl electrodes in order to gain information about the swallowing process. The device additionally measures EMG over the same electrodes. The presented measurement device consists of a protected amplifier circuit. It has the ability to quickly recover after a stimulation pulse in order to allow measurements during active FES which is applied to muscles near the measurement electrodes. The measured data samples are sent to

a personal computer (PC) in real-time. The measurement system is able to trigger FES synchronously with the onset of swallowing in order to assist the movement of hyoid bone and larynx.

The measurement signals are used to automatically detect swallowing events in recorded time series of EMG and BI. An EMG activity detector was developed and implemented. The detector is based on a double-threshold detector and extended with robust noise and disturbances estimation in order to find the optimum detection threshold. Possible segments in the BI time series, which are likely to contain a swallow, are found by a valley search algorithm. The heuristic algorithm uses a piece-wise linear approximation of the BI signal. Only valleys that coincide with EMG activity are then selected for feature extraction. Besides time, value and area based features, symbolic approximations are generated for each possible swallow and directly used as features. The classification of the extracted valleys is done by a support vector machine (SVM). The classifier has been trained and tested on data from healthy subjects and patients.

The measurement principle is evaluated by correlating the EMG and BI time series to the displacement of the hyoid bone and the laryngeal cartilage which were determined by VF. The movement trajectory of hyoid bone and the laryngeal cartilage is extracted from 2828 VF images which include 92 swallows from 17 patients.

Parts of the algorithms from the swallow detection are used in the swallow trigger algorithm which is able to trigger functional electrical stimulation for improving elevation of hyoid bone and larynx and lowering the risk of aspiration and penetration. Timely muscle activation by FES requires an swallowing effort detection algorithm which is able to trigger the stimulation within some hundred milliseconds. The stimulation system, which automatically triggers FES at the onset of swallowing, was successfully tested on one patient.

In summary, it can be said that within this thesis a measurement device was developed and evaluated that is able to capture the pharyngeal swallowing phase, delivers parameters which are able to characterize the swallow, and allows the control of FES in order to support swallowing.

### 1.7 Outline of this Thesis

Besides the introduction, this thesis consists of five further chapters and is concluded by a discussion and outlook.

In the second chapter, the state of art in swallowing diagnosis and therapy is presented. It is pointed out that swallowing and swallowing disorders are complex and swallowing diagnosis as well as therapy are challenging problems.

The developed measurement system PHYSIOSENSE, which utilizes EMG and BI for capturing swallowing is described in the third chapter. Important parts of this measurement systems are explained in detail.

The measurement system is used in a swallow detection algorithm which is presented in the fourth chapter. The algorithms for EMG detection, BI approximation, heuristic valley searching, and feature extraction are explained in detail. The influence of important parameters is evaluated. Finally, a classifier based on an SVM is trained and tested on the extracted features from data of healthy subjects and patients. The sensitivity and specificity of the complete approach is determined.

The correlation between displacement of hyoid bone and larynx and the measured BI signal is investigated within the fifth chapter. Based on VF videos of swallows from 17 patients, correlation coefficients are calculated and interpreted.

In the sixth chapter, the algorithms are described which allow detection of voluntary swallowing effort in real-time and are able to trigger FES for improving elevation of hyoid bone and larynx and lowering the risk of aspiration and penetration. The system was evaluated on one patient. FES was triggered based on the measured EMG and BI signals such that the patient swallows could be successfully supported.

Finally, a conclusion about the results within this thesis is drawn and recommendations about future work are given.

### 1.8 Publications

Parts of this thesis were published in the following articles:

- . H. Nahrstaedt and T. Schauer. **A bioimpedance measurement device for sensing force and position in neuroprosthetic systems.** In *4th European Conference of the International Federation for Medical and Biological Engineering*, 22:1642–1645, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. DOI: 10.1007/978-3-540-89208-3\_390
- . R. O. Seidl, H. Nahrstaedt, and T. Schauer. **Electric stimulation in dysphagia therapy – a review (Article in German).** *Laryngo-Rhino-Otologie*, 88(12):768–774, 2009. DOI: 10.1055/s-0029-1237350
- . H. Nahrstaedt, T. Schauer, and R. O. Seidl. **Messsystem für eine Bioimpedanz-geregelte Schluckneuroprothese.** In *Biomed Tech 2010*, 55(Suppl. 1):1–4, Walter de Gruyter, 2010.
- . H. Nahrstaedt, T. Schauer, and R. O. Seidl. **Bioimpedance based measurement system for a controlled swallowing neuro-prosthesis.** In *Proc. of 15th Annual International FES Society Conference and 10th Vienna Int. Workshop on FES*, 49–51, 2010.
- . H. Nahrstaedt, C. Schultheiss, T. Schauer, and R. O. Seidl. **Swallowing detection based on combined EMG and bioimpedance measurements.** In *45. Jahrestagung der DGBMT*, Freiburg, Germany, 2011.
- . H. Nahrstaedt, C. Schultheiss, T. Schauer, and R. O. Seidl. **Automatische Erkennung und Beurteilung von Schluckvorgängen mittels kombinierter EMG- und Bioimpedanzmessung.** In *AUTSYM - 6th International Symposium on Automatic Control*, Wismar, Germany, 2011.
- . H. Nahrstaedt, C. Schultheiss, R. O. Seidl, and T. Schauer. **Swallow Detection Algorithm Based on Bioimpedance and EMG Measurements.** In *8th IFAC Symposium on Biological and Medical Systems*, 91–96, Budapest, Hungary, 2012. DOI: 10.3182/20120829-3-HU-2029.00067
- . H. Nahrstaedt, C. Schultheiss, T. Schauer, and R. O. Seidl. **Bioimpedance- and EMG-triggered FES for improved protection of the airway during swallowing.** In *Biomedical Engineering / Biomedizinische Technik*, 2013. DOI: 10.1515/bmt-2013-4025
- . R. O. Seidl, H. Nahrstaedt, C. Schultheiss, and T. Schauer. **Entwicklung einer Neuroprothese zur Therapie von Schluckstörungen - Erste Ergebnisse.** In *84. Jahresversammlung der Deutschen Gesellschaft für Hals-Nasen-Ohren-Heilkunde*, 2013. DOI: 10.3205/13hnod031
- . C. Schultheiss, T. Schauer, H. Nahrstaedt, and R. O. Seidl. **Evaluation of an EMG bioimpedance measurement system for recording and analysing the pharyngeal phase of swallowing.** *European Archives of Oto-Rhino-Laryngology*, 270(7):2149–2156, 2013. DOI: 10.1007/s00405-013-2406-3
- . C. Schultheiss, T. Schauer, H. Nahrstaedt, and R. O. Seidl. **Automated Detection and Evaluation of Swallowing Using a Combined EMG/Bioimpedance Measurement System.** *The Scientific World Journal*, 2014:405471, 2014. DOI: 10.1155/2014/405471

# 2

## **State of the Art in Diagnosis and Treatment of Swallowing Disorders (Dysphagia)**

### **2.1 Diagnosis of Swallowing Disorders**

An analysis of a questionnaire regarding clinical examination of dysphagia by Mathers-Schmidt et al. [38] showed that most clinicians perform a clinical (e.g. bedside) examination of dysphagia prior to an instrumental diagnostic procedure. This study was limited to clinicians from the western Washington state. For almost all clinicians, clinical examination contains the following points: history, assessment of vocal quality, assessment of lip seal and dentition for chewing, ability to perform a volitional cough, observation of oral movement and motor function, assessment of pharyngeal delay and excursion, the patient's perception and mental status. For specific hypothetical cases in which the clinical examination did not lead to clear results, more than 80 % of the clinicians agreed that an instrumental evaluation is warranted. videofluoroscopic swallowing study (VFSS) was the most recommended instrumental evaluation method and almost all clinicians had access to VFSS in their local area.

The existence of accurate instrumental evaluation methods is important for the subsequent diagnosis of patients who showed presence of clinically indicators that predict a risk of aspiration and for monitoring of swallowing therapy.

There are various methods for assessing the swallowing process and still there is a need for improving these methods. In literature, measurement methods are described which gain objective information about timing, pressure, range, and strength of structural movements, bolus flow pattern, bolus clearance and efficiency, airway protection, and sensation [39].

### 2.1.1 Clinical Swallowing Tests

Clinical swallowing tests try to classify dysphagia and aspiration based on the response of the patient on standardized food intake. The clinical swallowing tests utilize either water swallows, swallows of different viscosity, measurement of oxygen saturation or water swallows combined with oxygen saturation measurements [40]. The tests are intended to be a screening tool which is conducted by therapists, is non-invasive and produces reliable results. A questionnaire is used in order to categorize the severeness of a swallowing disorder into a numeric scale which can be used for dietary recommendation. Especially nurses and therapists should be able to perform these tests as they play an important role in observation and treatment of patient with dysphagia. As acute stroke patients have a high risk to develop dysphagia in the first days after stroke and their risk to aspirate is high [8], the swallowing functions of acute stroke patients should be checked as early as possible. This can be done with bedside swallowing tests.

Ramsey et al. [41] compared bed side assessment with videoflouroscopy and found a sensitivity of 47 % and a specificity of 72 %. Pulse oximetry was measured simultaneously and the combination with oxygen saturation showed a sensitivity of 60 % and 53 % and a specificity of 41 % and 67 % depending on the desaturation level. Ramsey et al. [41] concluded that a bedside swallowing assessment cannot be used for identifying the presence or absence of aspiration as its sensitivity and specificity is not high enough to be reliable.

In a study by Daniels et al. [42], six clinical indicators were used to identify patients with moderate to severe dysphagia. 56 acute stroke patients were evaluated. 38 patients were identified with at least two clinical markers which indicate a risk for dysphagia. In only 24 of these 38 (63 %) patients, moderate to severe dysphagia was confirmed by VFSS. None of the patients developed pneumonia.

The Burke Dysphagia Screening Test (BDST) was used by DePippo et al. [43] for examination of 139 patients. 82 patients failed the BDST. 12 patients, from which 11 (92 %) failed the BDST, developed pneumonia, upper airway obstruction or died.

The Toronto Bedside Swallowing Screening Tool (TOR-BSST) was evaluated by Martino et al. [44]. 311 stroke patients were evaluated and in one part of patients a subsequent VFSS was performed. In 4 of 32 (12.5 %) patients which passed the TOR-BSST and received VFSS, dysphagia was found. In 22 from 36 (61 %) patients which failed the TOR-BSST and received VFSS, dysphagia was diagnosed.

The Gugging Swallowing Screen (GUSS) was evaluated by Trapl et al. [45]. 49 stroke patients were separated into two groups and evaluated by GUSS and fiber-optic endoscopic evaluation of swallowing (FEES). In direct comparison to FEES, sensitivity of 100 % and a specificity of 50 % and 69 %, respectively, could be achieved.

62 patients were evaluated by using the Bolus Swallow Test (BST) in combination with Saliva Swallow Test (SST) by Schultheiss et al. [46]. In comparison to FEES, a sensitivity of 89.6 % and specificity of 72.7 % was achieved.

These studies show that clinical swallowing tests cannot reach the accuracy of the gold standards VFSS and FEES. Clinical swallowing tests are an alternative for assessing dysphagia in patients when VFSS or FEES are not available.

### 2.1.2 Videofluoroscopic Swallowing Study

The videofluoroscopic swallowing study (VFSS) utilizes videofluoroscopy (VF) which is a radiological method (X-ray) for analysis of the location and severeness of swallowing dysfunctions in the oropharyngeal or esophageal region. VFSS is also called “modified barium swallowing examination” as a bolus containing barium is used for visualization of the swallowing process. The swallowing process itself can be estimated by the movements of the bolus, bony, and cartilaginous structures which are visible in the X-ray examination and which are recorded in real-time by videoradiographic images. The swallowing process can be documented in two planes (frontal and lateral). The slow-motion representation allows an accurate assessment of individual swallowing phases and interference. In these images, the transportation of the bolus through oral cavity, pharyngeal cavity, and esophagus can be observed. Presence and timing of aspiration, in which ingested material enters through the level of the true vocal folds into the trachea, is visible [39]. Effects of various bolus volumes, consistencies, and compensatory strategies can be observed [39].

The VFSS can be used to evaluate the severeness of aspiration or penetration by mapping the occurrence of specific symptoms into numeric scales. The 8-point Penetration-Aspiration scale [47] combines the level of airway invasion with information about ejection of already entered bolus parts. Each recorded swallow is manually mapped into the scale by experts. Rosenbek et al. [47] showed that the inter- and intra-rater agreement for their 8-point Penetration-Aspiration scale lies in a range between 57 % and 84 %. A different VFSS examination, by which the pharyngeal residue severity was rated from none to severe by different raters, found an inter- and intra-rater agreement in a range between 50 % and 75 % [13]. The Videofluoroscopic Dysphagia Scale (VDS) consists of 14 items that can be observed by VFSS. Kim et al. [48] evaluated the inter-rater reliability on 100 patients. The VDS score inter-rater reliability was found out to be low to moderate with an intra-class correlation of 0.556.

VFSS is often considered as gold standard in diagnosis of dysphagia [49] and often the performance of other swallowing diagnosis techniques is measured against the outcome of VFSS. The advantage of videofluoroscopy is the possibility to assess of the entire swallowing process. A description of the different stages, that can be observed in the VFSS, can be found with sample images in [39]. The disadvantages are the significant technical complexity and radiation exposure during the investigation, which makes frequent follow-up examinations inappropriate. The average radiation exposure duration is around 3 min to 5 min [39]. Evaluation of saliva swallows is only possible when they occur directly after a bolus swallow containing sufficient contrast agent.

### 2.1.3 Fiberoptic Endoscopic Evaluation of Swallowing

Fiber-optic endoscopic evaluation of swallowing (FEES) is performed by a transnasally introduced flexible endoscope by which swallowing of saliva or food of different consistencies can be observed. In comparison to the VFSS method, FEES is a safe nonradiologic alternative without any side effects [50].

The disadvantage of the FEES method is that only a limited area can be observed by the endoscopist. In order to detect penetration and aspiration, the endoscope has to be inserted inferiorly into the larynx after each swallow. During the approaching of the tongue and posterior

pharyngeal wall, the camera will only transmit a white image (“white out”-effect). This effect can be used for determine whether the aspiration occurred before swallow (bolus disappears in the pharynx before swallow), during swallow (residue visible) or after swallow (patient aspirates as the airway reopens).

The sensitivity of FEES is comparable to VFSS or even better in the following swallowing parameters: assessment of the delay in swallowing initiation, penetration, aspiration, and pharyngeal residue [51]. FEES can also be used for patients who are unable to sit or stand during the examination. The inter- and intra-rater agreement is in the same range as VFSS with an inter-rater agreement of 0.51 [13] or 0.35 to 0.46 [14] and an intra-rater agreement of 0.72 [13] or 0.53 to 0.78 [14].

### 2.1.4 Cervical Auscultation

Cervical auscultation is a method in which trained experts try to classify dysphagia in patients by listening to the swallowing sounds during the pharyngeal phase using an amplifying instrument. Zenner et al. [52] showed that the agreement between VFSS and auscultation has been moderate for tracheal aspiration (76 %) and oral delay (72 %), whereas oral residuals (62 %), pharyngeal delay (66 %) and pharyngeal residuals (42 %) have been a lower agreement with VFSS. The inter-rater agreement was found out to be at fair level (0.46) by Borr et al. [53]. Similar results were achieved by Leslie et al. [54], who compared the experts ratings together with VFSS. The experts achieved a sensitivity of 62 % and a specificity of 66 %. If only the consensus of all experts for each swallow was used for classification, then the sensitivity was improved to 80 % and the specificity was improved to 90 %. Cervical auscultation has been seldom used by clinicians [38].

### 2.1.5 Sonography

Sonography is of limited use for the assessment of swallowing disorders. The following analyses are possible: observation of the tongue function within the oral phase of swallowing [55] and evaluation of the movements of hyoid bone [56], [57], thyroid cartilage [58] as well as the distance between hyoid and thyroid cartilage [59]. Caused by the various overlapping tissue types, movements in the throat are not easy to differentiate and an automatic extraction of movement trajectories is not possible. Detection of aspiration or penetration using sonography is hardly feasible as the position of the bolus is not clearly visible.

### 2.1.6 Real-time Capable Measurement Methods

VFSS and FEES are currently the most important instrumental methods for diagnosis and monitoring of swallowing disorders. Both methods can only be carried out in a clinical environment and need a subsequent evaluation of the recorded video by skilled experts. Repeated examinations with VFSS should be avoided as the patient is exposed to radiation. Both methods are not always available due to logistical problems.

An automated evaluation of swallowing and detection of dysphagia is not possible using VFSS or FEES. This means that these methods are not usable for biofeedback applications, home monitoring or as sensors in a neuro-prosthetic implant. To overcome these disadvantages, several alternative methods have been investigated and are currently being researched. In the following sections,

several methods are introduced which do not need bulky and expensive devices, as they use only some easily attachable sensors. It is also possible to link these methods with digital signal processing and artificial intelligence as discrimination tools. Currently, these methods are only rarely used in the assessment of swallowing disorders in clinical routine.

#### 2.1.6.1 Piezoelectric Sensors

In the study from Ertekin et al. [60], a piezoelectric sensor, which is able to detect laryngeal movement under the sensor due to pressure changes, was used together with submental muscle electromyography (EMG) to evaluate oropharyngeal dysphagia. Only timing information about the laryngeal movement can be measured using a piezoelectric sensor.

#### 2.1.6.2 Swallowing Sounds

Swallowing sounds can be measured using microphones. The method is similar to cervical auscultation but uses digital signal processing and artificial intelligence / machine learning instead of discrimination skills from an expert. The swallowing sound as a sensor signal is used in combination with a trained classifier in some studies for automatic swallow detection and separation of normal and abnormal swallows.

In order to monitor food intake, Sazonov et al. [61] used a microphone located over the laryngopharynx for automatic swallow counting. The recorded microphone signal contained 9966 swallows from 20 healthy subjects. In addition, resting periods with noise and talking as well as the intake of a meal were recorded. The data were divided into 70 visits containing measurements of 20 minutes resting, one meal and another 20 minutes of resting. These datasets were used for generating 70 different intra-visit models. Each model was trained and tested by applying three folded cross-validation on 55 segments of a visit. The classifiers, which were trained to detect presence of a swallow in time periods of 1.5 s, achieved an average weighted accuracy of 84.7%.

Amft et al. [62] recorded the signal of a microphone which was placed near the cricoid cartilage together with the EMG of the infrahyoidal and submental muscles for classification of swallowing, bolus volumes, and consistencies. Using the data of five healthy subjects, swallows were automatically detected with a sensitivity of 65% and a specificity of 31%. The accuracy for discrimination of low and high bolus volume and viscosity was around 70%. The same authors repeated the study on six healthy subjects using a different sensor setup (EMG was recorded directly at the infra-hyoid throat position) [63]. By using sensor fusion between EMG and the microphone, a sensitivity rate of 68% and a precision rate of 20% were achieved. A lot of false positive events were received by the classification.

In the work by Aboofazeli et al. [64], [65], an automated classification of breathing sounds from swallowing sounds was performed on manually segmented data of 15 healthy subjects (including children) and nine cerebral palsy patients. Wavelet-based detection [64], recurrence plot features [64], and hidden Markov models [65] were used. The latter achieved the best results of swallowing sound detection with a false positive rate of 6.7% and 4.8% of missed swallows.

Lazareck et al. [66] used a microphone for classification of swallows from control subjects (12 children (3 to 16 years) and three adults (ages 35, 38, and 54 years)) and swallows without aspiration

from patients with swallowing disorders (ages 16 to 25 years). 350 swallows of different viscosity were recorded. The swallowing sounds were manually extracted from the recorded data based on repeated listening and monitoring of the signal. The achieved accuracy of the trained classifier for the test data, which were produced by the leave-one-subject-out method, depends on the viscosity of the bolus. Normal swallows of thick liquid bolus was classified with an accuracy of 76.6 % (semisolid: 58.0 % and thin liquid: 84.6 %) and an abnormal swallow of thick liquid bolus was classified with an accuracy of 88.1 % (semisolid: 84.9 % and thin liquid: 63.3 %). The classification results were improved by excluding the two adult subjects in the control group, which leads to the question whether the age of the subjects is more important than the presence of swallowing disorder.

On ten patients with dysphagia who showed silent aspiration during a FEES examination, a microphone was used to record the swallowing and breathing sound, which was later manually extracted from the recorded signal by Shirazi et al. [67]. 186 swallows without aspiration and 46 swallows with aspiration were recorded. An unsupervised classifier separated swallowing with aspiration from normal swallows with an average accuracy of 82.3 % (sensitivity of 84.8 % and specificity of 81.4 %). By the same authors [68], a larger study with 50 dysphagic adult patients was performed. Swallowing sound and breathing sounds were recorded with the same microphone together with either VFSS or FEES which were used for manual swallowing sound separation. The breathing sound of up to three breath phases was also selected for analysis. A support vector machine (SVM) was used to distinguish swallowing with severe aspiration from swallowing with mild or without aspiration. The classifier was trained and tested by the leave-one-subject-out method and resulted in 91 % sensitivity and 85 % specificity.

A wearable acoustic sensor for recognition of swallowing water, eating and other activities such as breathing, coughing, speaking, laughing, sighing, whispering, and whistling was developed by Yatani et al. [69]. Features for classification were extracted from non-overlapping frames with a length of 186 ms. The leave-one-subject-out and leave-one-sample-out methods were used for training and testing on the data of ten healthy subjects. Swallowing was classified with a sensitivity of 35 % (78 % for the leave-one-sample-out method) and a precision of 27.8 % (66.1 % for the leave-one-sample-out method) among twelve different activity classes.

Olubanjo et al. [70] recorded swallowing sounds from dry and water swallowing together with other events such as cough, speech, chewing, clearing the throat, head tilts, and head turns from four healthy subjects in two sessions. The first session was used for training and the data from the second session were used for testing. The data were separated into non-overlapping 500 ms window frames from which features were calculated. In the test data set, the correctly detected swallows were counted. A sensitivity rate of 79.9 % and a precision rate of 67.6 % were achieved.

### 2.1.6.3 Accelerometry

Swallowing accelerometry measures the movement and vibration which is caused by swallowing and partially captures sound as it can be seen as an acoustic transducer [71]. An accelerometer has the advantage that it can measure with two or three degrees of freedom and its signal to noise ratio is slightly better in a noisy environment compared to a microphone [72].

According to Reddy et al. [73], the maximum amount of acceleration correlates strongly with the maximum amount of laryngeal elevation. 31 patients with a history of swallowing disorders took part in this study. During each swallow, the maximum elevation of the larynx and the maximum amplitude of the band-pass filtered acceleration signal were extracted and correlated. The product-moment correlation coefficient between both amplitudes was determined to be 0.807 for all swallows.

Lee et al. [74] used a radial basis classifier to distinguish aspiration from normal swallows in children with dysphagia, which were 64 males and 53 females (aged  $6.0 \pm 3.9$ ). The single-axis acceleration signal was extracted manually for 100 swallows and 94 aspirations which were identified using VF. The results for the best two features were a sensitivity of 79.4 % and a specificity of 80.3 %.

Using dual-axis swallowing accelerometry, Sejdić et al. [75] was able to automatically segment swallows in the measurement signals on 20 healthy participants. In comparison to manual segmentation of the same signal by human experts, the segmentation algorithm was able to correctly segment 94.6 % of 295 recorded swallows. The measurement signals were completely free of any disturbances as cough, speech, or excessive head movement. Swallowing duration was extracted from the accelerometry signal and additionally analyzed. In [76], the same authors presented an approach for removing vocalization and coughing on dual-axis swallowing accelerometry data.

Similar results were achieved by Lee et al. [77], who used dual-axis swallowing accelerometry, a nasal airflow sensor, and a submental mechanomyography, which consists of a microphone. Manual segmentation by two raters were conducted for 1292 swallows from 17 healthy subjects. 80 % of the data were used for training of a neural network, while the remaining 20 % were used for testing. The combination of all sensors resulted in an adjusted accuracy of 89.6 % for swallow segmentation, whereas the contribution from the accelerometer was most important.

Damouras et al. [78] developed an automatic online segmentation algorithm that uses dual-axis accelerometry signals in order to segment swallows. 295 swallows from 20 healthy subjects were automatically segmented with 90 % sensitivity and 93 % specificity in comparison to a manual segmentation by two human raters. The performance of the segmentation algorithm was also evaluated on 266 swallows from 37 dysphagic patients. VF was used as reference. The sensitivity of the segmentation algorithm was 87 %, whereas the specificity could not be measured as not all swallows could be segmented through VF.

Combined measurement of the dual-axis accelerometry and the nasal airflow signals was used from Lee et al. [79] on 24 patients with dysphagia (aged  $64.8 \pm 18.62$  years, 22 male) in comparison to VF. Onset and offset of each swallow were marked by hand based on the VF images. Three different classification problems were considered: 1) airway invasion in which material entered the airway below the vocal folds, 2) residuals at the valleculae, 3) residuals at the pyriform sinuses. The training data were randomly separated from the test data set. The performance of the classifier had the best results in classifying problem 2) and 3). The adjusted accuracy of all achieved results laid in a range around 74.7 % to 84.2 %.

29 pediatric patients (aged  $6.8 \pm 4.8$  years, 20 male) were examined by Merey et al. [80] using dual-axis accelerometry in comparison to VF. Periods with vocalization were removed and start and endpoint of each swallow were manually adjusted based on VF images. Classification results

were obtained using an 8-fold cross-validation. Each swallow was assigned to a safe swallow (no aspiration / no penetration) or to an unsafe swallow (aspiration occurs) whereas swallows with penetration were discarded. The classification of 638 swallows resulted in a sensitivity of 89.6 % and a specificity of 92.2 %. Weaknesses of this study are that the training and testing set contained swallows from the same patients and that swallows with penetration were removed.

Steele et al. [81] applied dual-axis accelerometry on 40 patients (mean age 67, 20 male subjects) from which each swallow was classified using VF either as safe swallow (8-point Penetration-Aspiration scale score less than 3) or as unsafe swallow (8-point Penetration-Aspiration scale score greater than 2, with deeper entry of material into the airway without clearance). The swallows were manually segmented using the images from VF. The testing of the classifier was done by the leave-one-out-approach. The results for 154 swallows were good (sensitivity of 90 % and specificity of 77 %). Using the exact same procedure for dual-axis accelerometry, Sejdić et al. [82] developed an algorithm based on wavelets which resulted in even better results on 40 patients (sensitivity of 92.5 % and specificity of 95.6 %).

### 2.1.6.4 Electromyography

EMG can be used to measure the electrical activity of muscles and muscle groups. The EMG signal provides information about the onset and the level of muscle activity. The EMG is suitable for the diagnosis of neuromuscular diseases and can be used in biofeedback procedures. Normally, EMG is normalized on the maximum voluntary contraction of the corresponding muscle. As swallowing related muscles cannot easily be contracted voluntarily, it is normally not possible to normalize swallowing related EMG to maximum voluntary contraction. Therefore, timing information as onset and offset and the measured voltage amplitude of swallowing related muscles can be used for assessment of dysphagia.

Gupta et al. [83] recorded surface EMG at the left side near the thyroid cartilage on 35 healthy subjects during dry and wet swallowing. The mean power differences of the EMG signal between dry and wet swallowings were significant, but it could also be seen that the mean power varied strongly from subject to subject.

Perlman et al. [84] used bipolar hooked wire electrodes which were placed in the following muscles: submental region, superior pharyngeal constrictor, cricopharyngeus, and thyroarytenoid. The onset, offset and duration of muscle activity was measured for saliva, a 5 ml water and a 10 ml water swallow. He concluded that the pattern of muscle activity within subjects is highly consistent and that the bolus volume leads to changes in the muscle activation pattern.

Vaiman et al. [18] used a large reference dataset of EMG recordings to identify abnormal swallowing. Surface EMG was recorded from the masseter, submental muscle group, infrahyoid muscle group, and from trapezius muscle. A test protocol was applied in which the patient has to swallow water of different volume and saliva several times. The EMG recordings were then compared to normal swallows from the database. In [85], this database was used to examine psychogenic swallowing disorders. For 42 patients with suspected psychogenic dysphagia, EMG was recorded as described above. It was found that the inter-rater reliability was high (88.2 %).

EMG has been combined with swallowing sounds (cf. Section 2.1.6.2) and accelerometry (cf. Section 2.1.6.3).

#### 2.1.6.5 Electroglottography

Electroglottography (EGG) is a two-electrode bioimpedance measurement method, in which primarily the impedance changes in the vocal cord area during speech are examined [86], [87]. The electrodes are placed on both sides next to the thyroid cartilage. Therefore, movements of the larynx during the EGG also lead to changes in impedance in the frequency range below 10 Hz. Sorin et al. [88] and Perlman et al. [89], [90] investigated the suitability of an EGG in order to describe the timing of laryngeal elevation during swallowing. Drawbacks of this method for swallowing evaluation are that it does not work well with female subjects and that incorrect placement of the electrodes leads to wrong measurements of the timing aspect of laryngeal movements [89]. However, the EGG was used as a reference measurement for the beginning and end of laryngeal movement in some studies [91], [92].

Kob et al. [93] developed an EGG device with a 2x3 electrode array that had a 2 MHz carrier frequency and allowed to measure EGG signals from a 36-channel combination with a frequency of 44.1 kHz. The authors described a possible way to measure the larynx position under the assumption that the current path with the maximum conductivity is closest to the glottis. It is possible to reconstruct the channel and the position of the maximum conductivity with a frequency of 1000 Hz. Studies with patients have not been carried out yet.

#### 2.1.6.6 Impedance Tomography

Impedance tomography consists of numerous electrodes from which one pair is used as current source and the impedance is measured by all other electrodes. The distribution of conductivity can be determined by a rotary impedance measurement by which the current electrode pair is continuously rotated. Hughes et al. [94], [95] examined the possibility for using an impedance tomography at the neck for the examination of the bolus passage in the throat during swallowing. Videofluoroscopy was used in both studies as a reference method. 16 electrodes were attached in a transverse plane around the neck at the level of the thyroid cartilage and the third cervical vertebra.

Hughes et al. [94] correlated the timing of typical swallowing events, e.g. larynx closed, to the time of peak in conductivity change. They concluded that the swallowing event with the best correlation is the last apposition of anterior and posterior pharyngeal walls. Seven patients were examined who had swallowing problems but could swallow 5 ml in volume without problems.

In the second study [95], 13 patients with swallowing problems were examined. Only timing information was extracted from the measurements and correlated with timings of swallowing events which were measured using videofluoroscopy. Significant correlations were only found for clearance time, duration of hyoid displacement, and duration of apposition of pharyngeal walls. The low sampling rate of the impedance tomography and its low signal to noise ratio were identified as problems of impedance tomography.

### 2.1.6.7 Bioimpedance Measurement

A Japanese group studied the possibilities of a transcutaneous bioimpedance (BI) measurement at the neck area in swallowing [22], [96], [97]. They used a four-electrode bioimpedance measurement at a frequency of 50 kHz by which a current is applied through two electrodes and the resulting voltage is measured at the other two electrodes. The authors described the measurement method as an impedance pharyngography.

In a pilot study by Yamamoto et al. [98], the waveform of a bioimpedance measurement during swallowing was compared for different consistencies and electrode positions. For shifts of the voltage measurement electrode in a range of 1 cm, the bioimpedance waveform remained similar. Different boluses did not lead to strong changes in the bioimpedance waveform. The resulting trace was interpreted as a reflection of the entire swallowing process (oral, pharyngeal, esophageal phase) caused by movement of the larynx, pharynx, throat, and esophagus. The end of the oral phase was located at the beginning of the drop-off in the bioimpedance signal. The pharyngeal phase was assigned to the entire drop-off. Finally, it was assumed that the bioimpedance signal was going back to its baseline value during in the esophageal phase of swallowing.

The same measurement device was used by Morimoto et al. [99], in order to measure the pharyngeal phase of swallowing in 38 patients suffering from dysphagia in comparison to ten healthy control subjects. Besides the impedance pharyngography, the maximum tongue pressure and swallowing sounds were recorded. Videofluoroscopy was used as reference measurement. It was shown that specific features allowed classification of abnormal swallows from normal swallows. Best results were achieved by maximum tongue pressure (sensitivity of 69 % and specificity of 100 %), followed by the duration from begin of the drop-off to the minimum point of the impedance signal (sensitivity of 76 % and specificity of 78 %) and the number of valleys in the impedance signal (sensitivity of 76 % and specificity of 89 %).

Ward et al. [100] used an impedance analyzer to measure BI at the neck during swallowing. Electrode positions and measurement setup were similar to [98]. Shown BI time sequences during swallowing, which were measured on a healthy subject, were similar to the presented measurement curves from [98].

A two channel BI measurement at the neck was combined with a manometric measurement inside the pharynx by Chester et al. [101]. BI was measured at two places using the four-electrode measurement method with a frequency of 40 kHz and 70 kHz, respectively. Measurement signals and some related features from a healthy and a dysphagic patient were shown, but the authors admitted that the device needs further refinement and validation.

## 2.2 Therapy of Swallowing Disorders

The goal of any therapy of swallowing disorders is the rehabilitation of the disturbed swallowing sequences and processes. Meanwhile, airway protection and nutrition have to be assured for the patient. In cases of severe dysphagia, surgical procedures such as a tracheotomy or a percutaneous endoscopic gastrostomy (PEG) system are necessary in order to prevent or minimize complications. The swallowing process may be improved by therapeutic measures and the patient can carefully

begin with food intake. Appropriate treatment options have to be individually selected based on the severity and history of dysphagia for each patient.

Conservative treatments, such as dietary adaptation of food consistencies, are firstly used for improving the swallowing process [102]. Supportive measures, such as posture changes during swallowing, are a way to assist swallowing by enhancing the remaining capabilities of the patient [103].

An important factor for restoring swallowing functions especially after cerebral injury is neural plasticity [104] which describes the ability of the neural substrate to change their behavior and relearn important functions such as swallowing. Neural plasticity can be stimulated by sensory stimulation. Different stimuli (cold stimuli, electrical stimulation, etc.) are used to improve sensitivity for triggering the swallowing reflex [105]. Methods of swallowing intervention utilize repeated exercises with and without swallowing and try to strengthen swallowing involved muscles and excite neural plasticity.

Sensors which autonomously access swallowing in real-time could be part of a biofeedback based therapy. Furthermore, such measurement methods could be used as a sensor for a controlled swallowing neuro-prosthesis which is essential for the treatment of chronic patients.

### 2.2.1 Strengthening Exercises and Compensatory Methods

Exercises, such as the Shaker exercise, Masako maneuver or the Mendelsohn-Maneuver, can be used to strengthen swallowing related muscles and should be performed on regular basis over a longer period.

In the study by Shaker et al. [106], 27 patients practiced the Shaker exercise three times per day. In the Shaker exercise, the patient should lie flat and firstly perform three sustained head raisings for one minute followed by 1 minute rest period and secondly raise his/her head 30 times. In contrast to sham exercises, the real exercise showed significant improvements after six weeks. The Shaker exercise can help to improve swallowing on patients with abnormal upper esophagus sphincter opening.

Kang et al. [107] tested for two months the effectiveness of 30 minutes conventional swallowing therapy combined with additional bedside exercise training of one hour per day against conventional swallowing therapy of 30 minutes per day. Both groups contained 25 comparable patients. The bedside exercise program consisted of oral, pharyngeal, laryngeal, and respiration exercises, including the Shaker exercise. The treatment outcome indicated that the additional exercises improved the oral phase in the exercise patient group. Improvements at the pharyngeal phase were comparable between both groups.

In a study by Wheeler-Hegland et al. [108], submental surface EMG and hyoid bone movements were recorded on 25 healthy adults during Mendelsohn maneuver, effortful swallow, and expiratory muscle strength training. EMG activity and the angle of hyoid elevation, which were measured using videofluoroscopy, were increased in comparison to control swallows. They concluded that these exercises induced additional muscle contractions which were able to strengthen the submental muscles and were important for rehabilitation.

McCullough et al. [109] utilized surface EMG biofeedback for assuring proper execution of the Mendelsohn maneuver by 18 stroke patients who suffered from pharyngeal dysphagia. After performing two weeks of two 45 to 60 minutes exercise sessions per day, significant improvements in “duration of hyoid maximum anterior excursion” and “duration of hyoid maximum elevation” were observed.

Bülow et al. [110] analyzed changes of the hypopharyngeal intrabolus pressure, during supra-glottic swallowing, effortful swallowing or chin tuck swallowing in eight patients with pharyngeal dysfunctions. Although no significant differences in intrabolus pressure using the different techniques could be found, a tendency for an increased intrabolus pressure during the maneuvers was observed. The authors believed that laryngeal elevation is the most important factor for a successful pharyngeal swallow sequence.

Postural changes that have to be performed by the patient itself, as chin tuck, head turn, head tilt, head back or chin tuck with head turn can improve swallowing and reduce or prevent aspiration. Such techniques are especially used in isolated disorders which may be caused by surgeries. Rasley et al. [103] showed that using postures such as chin tuck or head turn could avoid aspiration for small bolus volumes of 1 ml to 3 ml in around 70 % of all examined patients. Higher bolus volumes of 10 ml could successfully be swallowed by 25 % of all patients. The effect of aspiration prevention was examined on 165 patients.

### 2.2.2 Sensory Stimulation

The stimulation of oral and pharyngeal areas seems to be a promising method in order to help patients with dysphagia. Mainly mechanical, thermal, gustatory, and electrical stimuli [111] are used. Sensory stimuli can lower the threshold which evoke pharyngeal swallowing. As the excitability of the central pathway can be enhanced by stimuli, swallowing can be improved by sensory stimulation in patients. Sensory stimulation is used in neurological diseases associated with changes in perception. In case of severe neurological diseases, increasingly complex therapies are used (e.g. F.O.T.T. [112]).

In a pilot study with four stroke patients with dysphagia, Park et al. [113] applied an electrical stimulation at 1 Hz on the soft palate during swallowing. The electrical stimulation electrodes were attached on a palate prosthesis. A clinically significant improvement in the swallowing process was observed on two out of four patients, whereas the total transit time was improved for all patients.

The recovery of swallowing functions in stroke patients is associated with the reorganization of swallowing motor regions in the cortex of the intact hemisphere [114]. Therefore, pharyngeal stimulation which increases the cortical representation of the pharynx could improve swallowing recovery. Electrical sensory stimulation on the pharynx was applied by Hamdy et al. [115] on eight healthy subjects for a duration of 10 minutes. For stimulation of the pharynx at 10 Hz, a bipolar ring electrode was attached on a transnasally inserted catheter. It was shown by transcranial magnetic stimulation (TMS) that pharyngeal stimulation increases motor cortex excitability and cortical representation of the pharynx which lasted for at least 30 minutes.

Fraser et al. [116] performed an initial experiment on eight healthy subjects for finding the best parameters on pharyngeal sensory stimulation for changes in the excitability of the cortex.

They suggested that pharyngeal stimulation at 5 Hz and 75 % of the maximum tolerated intensity for 10 minutes produces the largest effects. These stimulation parameters for electrical pharynx stimulation were then applied on 16 acute stroke patients. After one hour, significant changes in size and excitability of the cortical representation of the pharynx in the unaffected hemisphere were measured. Simultaneously pharyngeal transit time, swallowing response time and aspiration were improved.

The same stimulation parameters were used for pharyngeal sensory stimulation by Jayasekeran et al. [117] for a placebo controlled trial on 67 acute dysphasic stroke patients. The active group showed a higher post-intervention reduction in aspiration than the placebo-stimulation group. The actively stimulated patient group stayed a shorter period of days (median 21 days) in hospital post intervention compared to the placebo-stimulation group (median 26 days).

Power et al. [118] examined the effect of unilateral stimulation of the faucial pillar on ten healthy adults. Significant changes were measured in the pharyngeal response amplitude. This indicates changes in the excitability of the cortical representation and the swallow response time. In a further study by Power et al. [119] with sixteen stroke patients who suffered from dysphagia, electrical stimulation of the faucial pillar was compared with sham stimulation. No significant differences between stimulation and sham stimulation were observed.

### 2.2.3 Electrical Nerve Stimulation

Some of the nerves which are involved in swallowing can be electrically stimulated. Developments towards implantable devices for improving swallowing and preventing aspiration are presented in the following.

Miller et al. [120] applied electrical stimulation by which swallowing could be triggered by the internal laryngeal nerve in 39 cats. The applied stimulation frequency correlated to the number of achieved swallows.

Chi-Fishman et al. [121] investigated the effect of mechanical and thermal stimuli on swallowing induced by electrical stimulation. In four cats, swallowing was triggered by electrical stimulation on the internal laryngeal nerve. Thermomechanical stimuli seemed to be able to modulate the initiation of oropharyngeal swallows induced by electrical stimulation.

In a pilot study on two patients with dysphagia, Broniatowski et al. [35] implanted a laryngeal pacing device which stimulated the left recurrent laryngeal nerve with one channel to close the true vocal folds which helps preventing aspiration during swallowing. The stimulation was manually controlled. On both patients, aspiration was significantly reduced. In a further study [122], five patients with aspiration were implanted the laryngeal pacing device. Statistically significant vocal fold adduction was achieved in all patients. On four patients, pneumonia were prevented during a period of six to twelve months.

Hyoid and laryngeal elevation were achieved in five canines by Hadley et al. [34] through selective hypoglossal nerve stimulation. Direct stimulation of the nerve complex caused significant higher laryngeal elevation than intramuscular stimulation. The elevation was comparable to the elevation within normal swallowing. Selective stimulation of the geniohyoid and thyrohyoid without triggering undesired tongue movement was possible.

### 2.2.4 Intramuscular Stimulation

Burnett et al. [30] examined the possible larynx elevation generated by intramuscular stimulation of the mylohyoid, geniohyoid, and thyrohyoid muscles in 15 healthy men (mean age 42 years). In comparison to a 2 ml water swallow, 50 % of elevation and 80 % of elevation velocity could be achieved by a two channel stimulation. The stimulation was also applied to each muscle individually which resulted in an elevation of around 30 % and a velocity of 50 %.

In order to improve airway protection during swallowing, such a stimulation must be synchronized with swallowing. On nine healthy adults, it was examined by Burnett et al. [36] if the stimulation could be triggered synchronized to swallowing. Functional electrical stimulation (FES) was applied intramuscularly on one site of the neck and EMG was measured intramuscularly on the other site. The volunteers were able to trigger the stimulation manually with a thumb switch synchronized to their own swallowing. In comparison to EMG measurements, the thumb switch occurred 239 ms after an onset of mylohyoid activity. The measured EMG activity was not altered by FES stimulation in comparison to foil stimulation.

Kagaya et al. [31] identified the motor points of the laryngeal elevation muscles. On two healthy subjects, the motor points were stimulated by surface electrodes and implanted electrodes. The achieved movements were greater with implanted electrodes.

### 2.2.5 Neuromuscular surface stimulation

Electrical stimulation can be applied by surface electrodes. The electrical field below the electrodes, which is created by the voltage across the electrodes, releases action potentials of nearby motor neurons. Depending on the electrode positions, more than one muscle may be activated during stimulation.

Freed et al. [123] compared electrical surface stimulation at the neck on 63 patient with dysphagia to thermal-tactile stimulation (36 patient with dysphagia). Both therapies were given to the patients for 60 minutes daily. The electrical stimulation with a frequency of 80 Hz and a pulse width of 300  $\mu$ s was increased until the participants felt a strong vibration around the electrodes. Two surface electrodes were placed above the hyoid and the other two electrodes were placed on the larynx to activate muscles in the submental and laryngeal region simultaneously. The electrical stimulator Staodyn EMS +2 (Staodyn Inc, Longmont, Colorado) was used. A new, not widely accepted, swallow function score (0 - aspirates saliva; 6 - normal swallow) was used for evaluating the outcome. The mean initial swallow function score for both groups was 0.75. In the end, the patients who received electrical stimulation therapy reached an increased swallow function score of 4.52, whereas the other group that received thermal-tactile stimulation could only achieve a swallow function score of 1.39. Afterwards, a commercial stimulator named VitalStim® with the same specification was developed and suggested for the described therapy.

Blumenfeld et al. [124] tried to repeat the positive results from Freed et al. [123] using electrical stimulation. The results of electrical stimulation were compared with standard dysphagia therapy on two groups of 40 patients suffering from dysphagia. The daily given therapies lasted 30 minutes. Both groups showed significant improvements, whereas the electrical stimulation group achieved a higher score of 3.23 in comparison to 1.48 which was achieved by the standard therapy group.

In a study by Oh et al. [125], eight patients with neurogenic dysphagia were stimulated similarly to the study by Freed et al. [123]. Statistically significant improvements in swallowing were detected using VF and FEES. In four patients, the cortical representations of the cricothyroid muscle were expanded after electrical stimulation, which was measured using TMS.

The VitalStim® device was used in a study by Bülow et al. [29]. On 25 patients with dysphagia, electrical stimulation was compared with standard therapy. The results after three weeks of therapy, which was given for 60 minutes daily, showed no statistically significant differences between both groups.

A similar study protocol as in [123] was applied by Lim et al. [126]. 36 patients with dysphagia after stroke were divided into two groups from which one was treated with electrical stimulation and the other group was given thermal-tactile stimulation. The treatment of one hour was repeated for four weeks. The final penetration-aspiration scale showed significant improvement for the electrical stimulation group, whereas in the other group the score did not improve.

The effect of electrical stimulation on the submental muscle activity was examined by Suiter et al. [127]. Muscle activity was measured by surface EMG. Eight subjects were stimulated by the VitalStim® device for one hour daily within two weeks. The change in EMG level was not statistically significant.

The resulting movements and effects on swallowing triggered by the VitalStim® device were investigated by Ludlow et al. [33]. The submental and laryngeal region were simultaneously stimulated with two channels as described by Freed et al. [123]. On eleven patients, no significant results on the movement of hyoid or larynx could be found, whereas on two participants the hyoid bone was descending by 5 mm to 10 mm. The position of hyoid and larynx were assessed by VF. On three participants, the stimulation resulted in a movement of the hyoid bone of around 5 mm in anterior-posterior direction. The larynx was moved 2 mm to 3 mm upwards in two participants by the stimulation. The 8-point Penetration-Aspiration (Pen-Asp) scale and the NIH-swallowing safety scale (NIH-SSS) for swallowing with and without stimulation were determined. The results were not significant. Then, the change in Pen-Asp scale due to the electrical stimulation were correlated to the movement of the hyoid bone achieved by stimulation at rest. The improvement in the Pen-Asp scale was significantly inversely related to the depression of the hyoid bone. It was assumed that the stimulation of the sternohyoid pulled the hyoid downwards and therefore increased the aspiration risk. The electrical stimulation was also applied on a sensory level which resulted in an improvement in the NIH-SSS along the whole group of patients.

VitalStim® was used by Baijens et al. [128] for stimulating ten patients with the Parkinson's disease during swallowing with three different electrode setups. VF was used for measuring the effect of stimulation during swallowing. In the first setup, the submental muscles were stimulated. In the second setup, two electrodes were placed below the hyoid bone and finally in the third setup, four electrodes from the first and second setup were used together. Statistically significant differences between the three setups were observed only in the time duration from laryngeal vestibule closing to its opening during swallowing and in the time duration from beginning of hyoid motion to the maximum horizontal displacement. The observed changes could also be repeated in healthy control subjects without any statistically significant differences.

A study by Humbert et al. [32] investigated ten different electrode positions in comparison with VF on healthy adults. From these ten setups in comparison to swallowing without stimulation, only the submental stimulation did not descent hyoid or larynx during swallowing, which increases the risk for aspiration. However, the submental stimulation did not significantly increase the elevation either.

Soon et al. [129] used four channels for measuring EMG and applying electrical stimulation to the masseter muscle and to the digastric muscle at both sides on eleven patients with chronic dysphagia. At the beginning, EMG on both muscles at both sides was recorded during swallowing of water (2, 5, 10 and 150 ml). Then, each patient was stimulated for 32 minutes three times a week with a fixed recurrent stimulation pattern. Each pattern consisted of stimulating the masseter muscle for 4 seconds. The submental stimulation channel was activated 2 seconds after the masseter stimulation has been started for 4 seconds. After the stimulation was concluded, the patient had 10 seconds for resting. The patient was asked to keep the mouth closed and swallowed together with the active stimulation of the submental muscles. After the treatment, EMG was recorded during swallowing and the improvement ratio was calculated. Significant reduction in EMG and swallowing time were observed for drinking 150 ml water. The authors assumed that due to the treatment the muscle coordination in the patients was improved.

Aspiration can be prevented by closure of the true and false vocal folds. Therefore, Humbert et al. [130] tried to achieve a change in the vocal fold angle by surface electrical stimulation on healthy adults. The VitalStim® with ten different electrodes placements on the neck was used. The measured changes in the fold angle were not significant and the authors concluded that it is not possible to cause movements in the true and false vocal folds using surface electrodes.

Most published studies on surface stimulation used a fixed electrical stimulation pattern independent from swallowing. Leelamanit et al. [37] tried to improve swallowing in patients with reduced laryngeal elevation by synchronized electrical stimulation. Surface EMG of the posterior tongue muscle was used for triggering the electrical stimulation of thyrohyoid muscle on both sides by two pairs of surface electrodes for a duration of one second. Stimulation was applied four hours daily until the criteria for an improved swallow were fulfilled or other intervention seemed necessary. 20 of 23 patients showed sufficient improvements in swallowing after a duration from two to 30 days. After successful treatment, patients were followed-up for a period of 3 to 33 months. In this follow-up period, a second successful treatment was given to six patients. After therapy, improved laryngeal elevation and a gain in body weight due to resumed oral diet were observed during the latest follow-up visit. The therapeutic outcome was not compared to a control group.

### 2.3 Conclusions

A lot of studies deal with dysphagia and proposed solutions either in the field of diagnosis or therapy. Instrumental gold standards for diagnosis are VFSS and FEES. Both are mainly used in the clinical routine. Both methods need trained medical staff for retaining reliable results. Clinical evaluation methods, such as bedside swallowing assessment, play an important role in the clinical environment due to its low instrumental and training requirements and due to the fact that dysphagia screening is important for a wide range of patients with neurological disorders (e.g. acute stroke, Parkinson's

disease, or multiple sclerosis). Bedside swallowing assessment has a lower accuracy than VFSS or FEES for diagnosis of dysphagia and especially aspiration detection.

Alternative measurement methods which map the swallowing process into a time series, e.g. swallowing sounds, accelerometry, or bioimpedance, could be used for accessing the swallowing process. Such easily attachable sensors could be combined with other measurement signals and could improve swallowing diagnosis and therapy. These systems could be used outside a clinical environment and could improve long-term monitoring of dysphagia.

Swallowing sounds and accelerometry are used in many studies and are combined with automatic classification and online segmentation algorithms. Most of these studies use carefully manually segmented data without any disturbances for swallowing classification by which in most cases swallowing with aspiration is separated from normal swallowing. These classifiers may only be useful for clinical practice when the manual segmentation by experts can be replaced by precise automatic segmentation algorithms which are robust against disturbances in the measurement signal.

Automatic segmentation algorithms using swallowing sounds, which took also disturbances into account, were only considered in a few studies [63], [68], [69]. Sensitivity rates of around 80 % could be achieved and the measurement data were divided into non-overlapping frames (width of 125 ms to 500 ms) which were then classified for a swallowing event inside. In order to be used for precise segmentation for automatic swallowing classification, better and more accurate methods in terms of sensitivity and time resolution have to be developed.

BI measurement of the swallowing process is only used in a small number of publications. To the best knowledge of the author, automatic classification and segmentation algorithms using BI have not been published until now. There exist no clues in the literature that the performance of BI for accessing swallowing and swallowing disorders is inferior to other methods. Advantages of the BI measurement method are its robustness against acoustical noise and that EMG measurement can be performed with the same electrodes. Studies in the field of BI measurement during swallowing give the evidence that the time series can be directly interpreted and mapped to a specific swallowing phase.

Therapeutic sessions, which are repeated over a long period, try to utilize sensory stimulation of areas involved in swallowing in order to excite neuronal plasticity. For increasing strength and coordination of swallowing related musculature, repeated activation of these muscles is used in other kind of therapies. Several studies tried to utilize electrical stimulation. In daily therapeutic sessions, electrical stimulation is either applied on a sensory level or on a stronger level by which muscle contraction occurs. From the latter, the VitalStim® is the most commonly used device. Several studies tried to find out if simple repeated hourly sessions by which electrical stimulation is applied to the area around the hyoid and the larynx could be more effective than conventional therapy. Contradictory results were published and there is a strong evidence that the applied stimulation increases the risk of aspiration during the session.

Assessment of swallowing functions is normally applied before and after the treatment in order to select appropriate treatments and then to evaluate the success of the therapeutic sessions. Besides muscle strengthening, new swallowing techniques are taught to the patients, e.g. supraglottic swallow, super-supraglottic swallow, Mendelsohn-Maneuver, effortful swallow, or swallowing with

postural changes. Such techniques could help those patients who are able to learn new swallowing techniques in order to improve the airway closure and the swallowing process. The most helpful technique for the patient is selected by visual control through FEES or VFSS. Biofeedback based on EMG measurement or tactile feedback of the therapists is partly used for supporting the patient in adapting his swallowing act to these techniques.

Dysphagia may lead to penetration and aspiration due to reduced air way protection during swallowing. FES may help to improve the airway protection by improved hyoid and larynx elevation or by closure of the vocal folds. Electrical nerve stimulation or intramuscular stimulation could help patients in recovering swallowing function, but need the developed of an integrated device. In those patients, who partly recovered their swallowing functions, also surface electrical stimulation may support swallowing. The results of some studies imply that the only safe positions for surface stimulation are the submental muscles and the masseter muscles, which can be stimulated in order to improve swallowing. Pilot studies that applied such systems to patients were successful but had triggered the stimulation manually or by unspecific EMG measurements which are sensitive to tongue, jaw, or head movements. As the pharyngeal phase of swallowing is triggered by reflexes, it is not a satisfactory solution to force the patient to manually trigger stimulation support at the correct timing for each swallow he/she wants to make.

In this thesis, it is investigated how a combined BI/EMG measurement at the neck could be used for precise and robust automatic swallowing segmentation as well as detecting an onset of swallowing in real-time for applying FES. Such a system can be used in biofeedback for assisting strengthening and compensatory swallowing therapies. Besides therapeutic sessions, the patient can be monitored over a longer period in his/her normal environment which gives a better understanding of his/her swallowing performance. Assisted swallowing by FES support of submental muscles could be realized by triggering the FES synchronized to the swallowing start, which is determined by such a measurement system. As a robust triggering of FES during swallowing could not be achieved until now, such a system could open up new possibilities for swallowing therapy.

# 3

## BI- and EMG-Measurement System

### 3.1 Summary

**Aim:** The development of a measurement system which is able to continuously measure electromyography (EMG) and up to two bioimpedances (BIs) using transcutaneous as well as needle electrodes at the neck is necessary for researching the ability to assess the swallowing process and is described in this chapter. The device should also operate during active functional electrical stimulation (FES) at the neck.

**Methods:** BI measurement is implemented by applying a sinusoidal current flow through the tissue and measuring the resulting voltage. Two accurate current sources were developed which are able to generate a stable sinusoidal current with a tunable current amplitude and a frequency of either 50 kHz or 100 kHz. The current sources were extended with safety elements and active voltage monitoring for patient protection. Four bipolar voltage measurement channels were integrated. All four channels contain protection elements such that a patient is protected against direct current (DC) and that the measurement device can withstand high-voltage impulses which occur during FES. Two channels are able to measure BI and EMG simultaneously. Analog filters and circuits were designed for separating both signal contents. Extraction of the BI amplitude is performed by accurate envelope detectors. All processed analog voltages are sampled with an accurate 24-bit analog to digital converter (ADC). A micro-controller is used for configuration, supervision and real-time transmission of the measurement data to a personal computer (PC) over a universal serial bus (USB).

**Results:** A combined measurement system, named *PHYSIOSENSE*, was realized. The device fulfills the following standards: IEC 60601-1:1998+A1:1991+A2:1995, IEC 60601-2-40, and IEC 60601-1-2:2007. Therefore, the device can be used for studies with patients and healthy subjects. Using this device, the two-electrode, three-electrode, and four-electrode measurement method for accessing BI can be successfully established using both current sources at 50 kHz and 100 kHz, respectively. It was

experimentally demonstrated that a BI/EMG measurement is possible between stimulation pulses during active FES. An artifact recovery of less than 3 ms was observed.

**Conclusion:** The developed measurement system *PHYSIOSENSE* can be used for assessing swallowing. It is a safe, reliable, and complete integrated measurement device which complies with the requirements in the clinical routine and can be used for scientific research. The presented measurement device *PHYSIOSENSE* can be used in order to investigate the possibility of automatic swallow detection in offline data (Chapter 4), to assess the correlation between hyoid / larynx movement and BI (Chapter 5), and to trigger FES in the very beginning of a swallow by BI/EMG measurement (Chapter 6).

**Contribution:** The author's contribution consists in the conception and development of the presented device as well as in the development of the algorithms for device calibration, failure detection, and automatic device parameterization. Parts of this Chapter were published in [131], [132].

## 3.2 Motivation

The difficulties in measuring BI are well known and many measurement devices were presented in the literature. A challenge is still to develop a system that allows the synchronous measurement of other bio-signals like EMG or electrocardiography (ECG) from the BI voltage recording electrodes at a higher frequency for precise bio-signal analysis and digital signal processing for artifact reduction. For the application treated in this work, i.a. EMG and BI processing, a sampling frequency of 4 kHz is required.

A first measurement system which is able to measure BI and electroglottography (EGG) over the same electrodes was presented by Li et al. [133]. However, the sampling rate of the BI signal was only 5 Hz, which is too low for assessing swallowing. A second system, which is able to measure BI and ECG via the same measurement electrodes, was introduced by Vuorela et al. [134]. The sampling rate of this system was limited to 100 Hz due to wireless data transmission. The bioimpedance measurement device which was used by Kusuhara et al. [22] for evaluating swallowing had a cut-off frequency of 25 Hz. Therefore, this system was not able to assess the signal content for higher frequencies than 25 Hz. Systems which are able to measure at a higher sampling frequency can be found in the field of electrical impedance tomography (EIT). But such devices are complex, expensive, and designed for measuring exclusively BI circular on many electrodes.

A versatile measurement system, which is able to measure BI and EMG by means of surface or needle electrodes also during active FES, is necessary in order to investigate the applicability of combined EMG and BI measurements at the neck for swallowing assessment and FES-assisted swallowing. In a first pilot study with the developed system *PHYSIOSENSE*, it was investigated if passages of fluid through the larynx can be assessed by a BI measurement system. It was shown on a prepared animal larynx, that BI, which was measured with needle electrodes at the larynx, reacts to fluid passage through the larynx. Such an aspiration detection might only be possible by using needle electrodes. Therefore, the possibility to use a combination of needle and surface

electrodes is provided in the **PHYSIOSENSE** which requires an adjustable current and amplifier gain, as the contact resistances are in different ranges for both types of electrodes. Although only one BI and one EMG channel are used in this thesis to assess swallowing with surface electrodes, the measurement system was developed to be able to measure BI at two different positions even with mixed types of electrodes. Further investigations in this direction are planned but outside the scope of this thesis.

The demand of performing combined EMG and BI measurements during active FES has to be considered in the design of the measurement device. Due to the high voltage pulses, which are used in the application of FES, and the immediate proximity of stimulation and measurement electrodes, a measurement device must be developed which is protected against high voltages at all input and output terminals. Since low-frequency high-pass filter would disturb the signals by its oscillating impulse response, they were not used in **PHYSIOSENSE**. Furthermore, a high-resolution ADC with a high sampling rate is provided for an accurate elimination of stimulation artifacts. By using such high-resolution ADC, an instrumentation amplifier with low gain for voltage measurement can be used which minimizes the time in which the amplifier is saturated due to the high voltage stimulation pulses.

### 3.3 Measurement Device **PHYSIOSENSE**

The measuring system **PHYSIOSENSE** allows two independent BI measurements and provides up to four channels of EMG measurement. The device fulfills the following standards: IEC 60601-1:1998+A1:1991+A2:1995, IEC 60601-2-40, and IEC 60601-1-2:2007. It includes protection under a single fault condition and has a proven isolation of the patient from the mains electric supply. The device was designed to be stimulation safe, which means that BI and/or EMG can be measured while FES is active in the same area. It is possible to use either silver/silver-chloride (Ag/AgCl) surface electrodes or monopolar needle electrodes. The developed device is able to be used in a wide range of BI measurement setups.

The structure of the **PHYSIOSENSE** system is shown in Figure 3.2. The device consists of two current sources CS1 and CS2. The sinusoidal current at 50 kHz and 100 kHz can be varied in a range from 3.4  $\mu$ A to 137  $\mu$ A. The optimal currents are found by means of the fault detection circuits FaultCS1 and FaultCS2 which monitor the output voltage of the current sources. Voltage measurement can be performed by four bipolar input channels. Each channel consists of an active shielding circuit AS, a switchable common mode rejection circuit CMRC, an instrumentation amplifier IA, and a filter circuit FiltEMG for EMG measurement. The first bipolar measurement channel CH1 additionally contains a reference driver RD, a demodulation circuit DemodBI1, by which the amplitude of a measured voltage at 50 kHz can be extracted and the fault detection circuit FaultBI1, which monitors the amplitude output of DemodBI1. The second channel CH2 contains the demodulation circuit DemodBI2 which is switchable between 50 kHz and 100 kHz. This circuit extracts the amplitude of the measured voltage at a frequency of 50 kHz or 100 kHz, respectively. The voltage output is monitored by FaultBI2. Both circuits DemodBI1 and DemodBI2 have switchable gains. All filtered and processed voltages are converted into digital values by the ADC. These digital values are filtered at the micro-controller and send together with the state of a galvanically

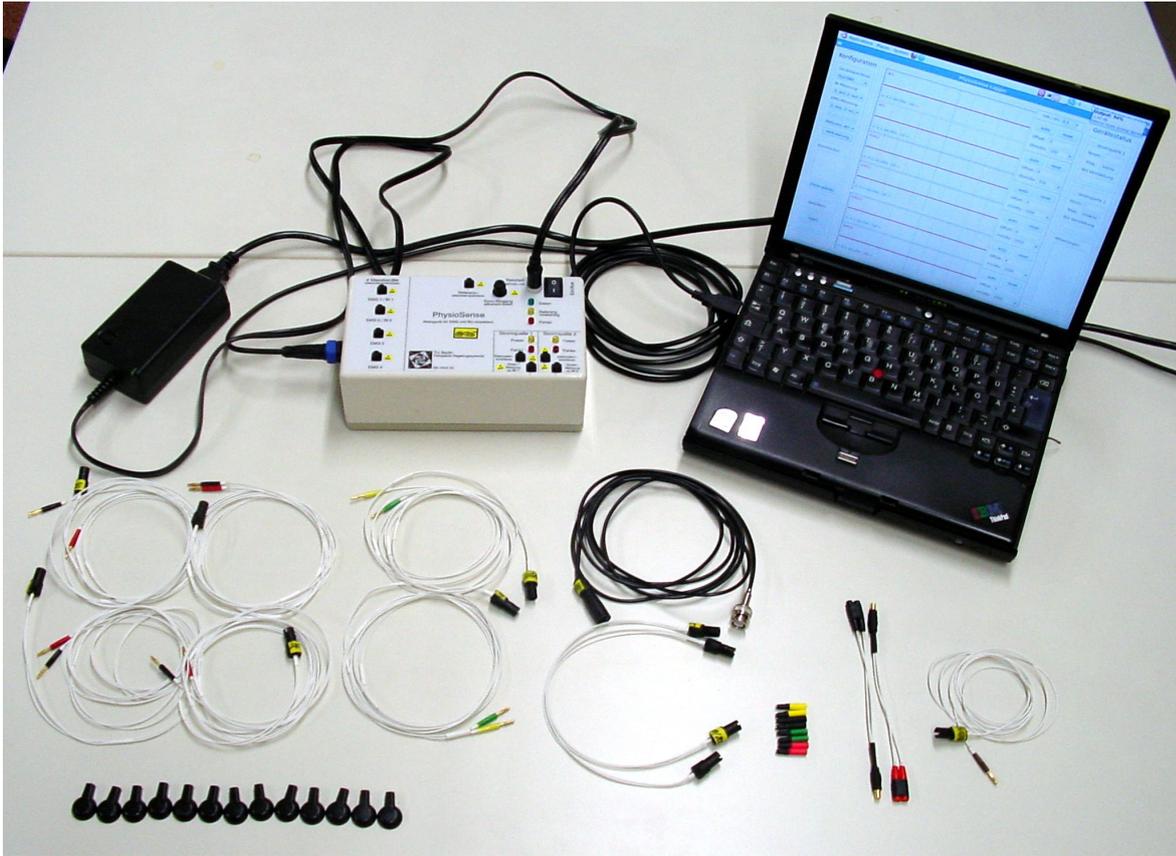


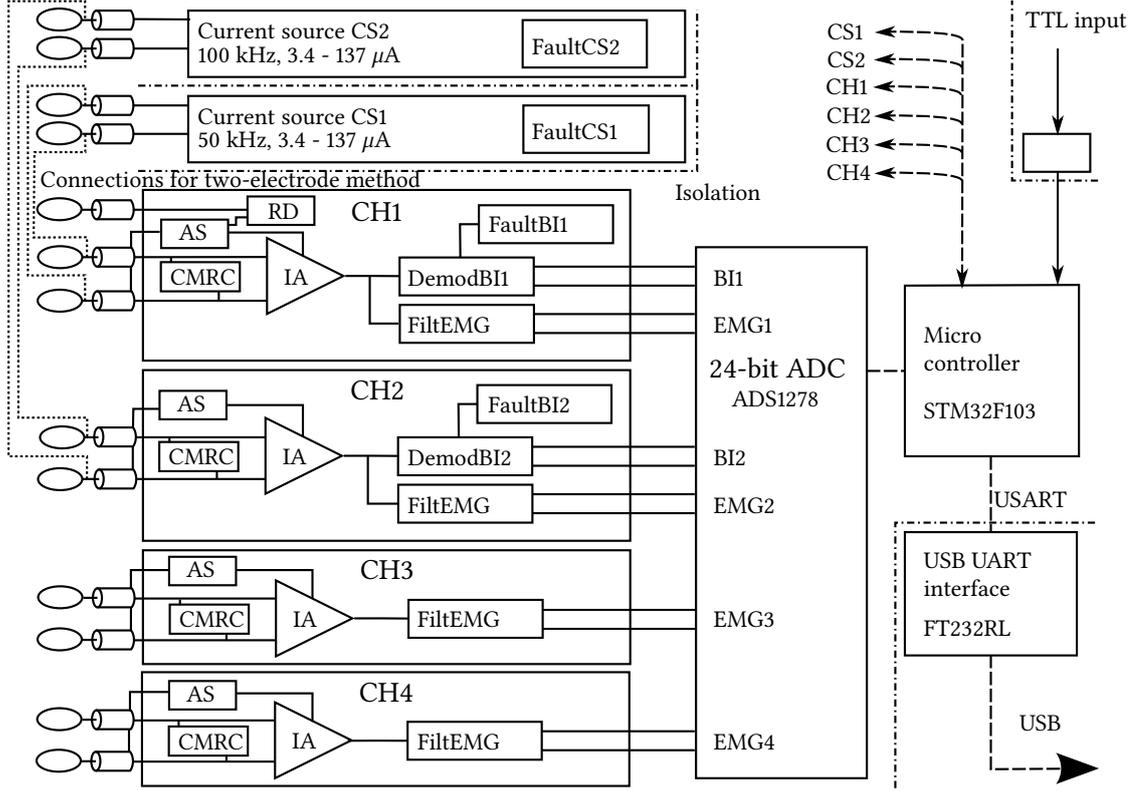
Figure 3.1: Measurement system PHYSIOSENSE with all measurement cables. The system is connected through USB to a notebook on which the measurement signals can be displayed and continuously saved. The picture is also shown in [135] and [136].

isolated transistor-transistor logic (TTL) input to a PC over USB. The individual components of the PHYSIOSENSE device will be explained in detail in the following sections.

A BI measurement consists of an insertion of current into the investigated tissue and a measurement of the resulting voltage drop which corresponds to the impedance according to Ohm's law. Primarily, two different setups are used for measuring BI (cf. Fig. 3.3):

1) In the **two-electrode measurement method**, the voltage is measured directly across the current electrodes. The current, which is induced into the tissue through the electrodes, causes a voltage drop across the electrode-skin contact. As this resistance is time-variant, it will lead to a measurement error. A second error source is the electrode polarization impedance which may influence the measured BI [137].

2) Both undesirable effects of the two-electrode measurement method can be almost avoided by using the **four-electrode measurement method**. Here, the voltage is recorded separately over additional measurement electrodes by an instrumentation amplifier with high input impedance. Since almost no current flows through the voltage measurement electrodes, no disturbing time-varying voltage drop across the electrode-skin contact will be present and the electrode polarization impedance is largely reduced. The measured impedance is a **transfer impedance** which is defined as the ratio of applied current to the resulting measured voltage. The following equation (3.1) describes how an inhomogeneous volume conductor with a resistivity distribution  $\rho$  leads to the transfer



**Figure 3.2:** Structure of the PHYSIOSENSE measurement system. The dash-dotted lines symbolize isolation measures. The dotted connection lines at the electrodes are only present when measuring BI with the two-electrode method. Only one pair of electrodes is used then by each BI measurement channel. When measuring BI with the four-electrode method, the electrodes of the current source are not connected with the electrodes from the measurement unit and two pair of electrodes are used by each BI measurement channel. Connections for digital data transmission are symbolized by a dashed line. (RD - reference driver; IA - instrumentation amplifier; CMRC - common mode rejection circuit; AS - active shielding; CH - channel)

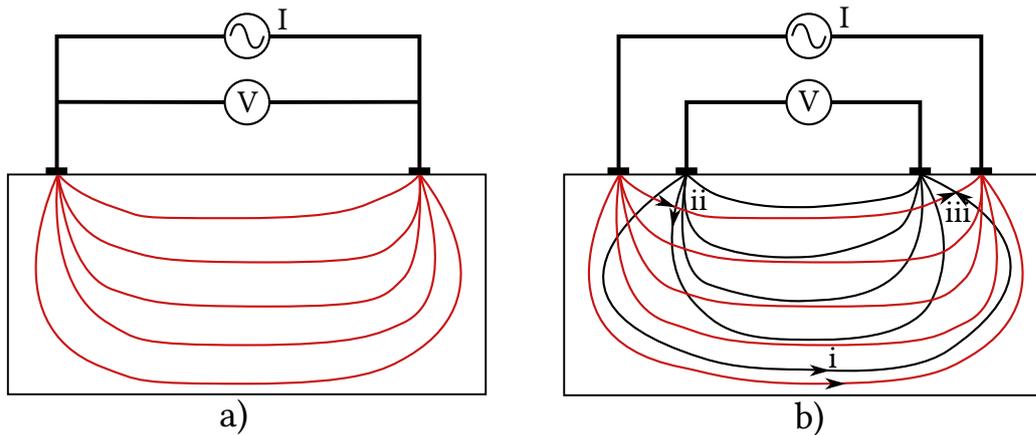
impedance  $Z_t$  [19]:

$$Z_t = \int_v \rho \mathbf{J}'_{cc} \cdot \mathbf{J}'_{reci} dV. \quad (3.1)$$

The resistivity distribution  $\rho$  is weighted by the dot product of the lead field of the voltage measurement electrodes for unit reciprocal current  $\mathbf{J}'_{reci}$  and the lead field of the current feeding electrodes for unit current  $\mathbf{J}'_{cc}$  and finally, integrated over  $v$ , responsible for the measured transfer impedance  $Z_t$ .

As described in detail by Grimnes and Martinsen [137], the transfer impedance between two pairs of electrodes placed on a volume conductor depends on its sensitivity to react to the conductivity change of all small volume elements in the volume. The sensitivity to a particular small volume element depends on its position to the electrodes and can be seen as the scalar product of the current density field from the current electrodes and the reciprocal current density field of the voltage measurement electrodes that would be achieved if the current was injected into the voltage measurement electrodes.

If the transfer impedance is almost zero, it does not necessarily mean that the tissue is a good conductor; it can mean that the injected current results in a very low voltage drop across the



**Figure 3.3:** BI measurement using the two- and four-electrode measurement method. On subplot a) the two-electrode measurement method is shown by which a lead field through the current injection electrodes arises. The sensitivity to small volume conductivity changes depends on the current density in the same area.

For the four-electrode measurement method, which is shown in subplot b), two different lead fields from each pair of electrodes are formed. The measured transfer impedance depends on the volume integral of the dot product from the lead field of the voltage measurement electrodes and the lead field of the current feeding electrodes across the complete volume (cf. Eq. (3.1)). Depending on the densities and the angle between both fields, local resistivity changes contribute differently to the measured transfer impedance. At position i) positive changes in the resistivity distribution in this area increase the measured transfer impedance. Whereas at position ii) local changes in resistivity distribution do not influence the measurement result and at position iii) positive local changes in the resistivity distribution decrease the measured transfer impedance.

measurement electrodes (this could be caused by a bad electrode-skin contact of the measurement electrodes). A second problem is that the sensitivity distribution is negative near the measurement electrodes. This means that a positive change in conductivity in this volume would lead to a higher transfer impedance. With different electrode positions, the transfer function is changing and therefore the transfer impedance changes. This means that the transfer impedance is a function which depends on the electrode position and on the tissue impedance itself. Therefore, BI measurements can only be repeated if the electrodes are placed on the same anatomical landmarks.

As shown in Figure 3.2, the measurement system PHYSIOSENSE consists of four bipolar input channels and two stable bipolar current outputs which operate at a frequency of 50 kHz and 100 kHz respectively. EMG measurement is possible on all four bipolar channels. Active shielding is applied on all input measurement cables in order to reduce noise. The measured common mode voltage from the first channel is used as input for the reference driver, by which common mode interferences at all channels, which are connected to the skin through electrodes, are reduced. Whenever a measurement channel is used for a different purpose, e.g., measurement of a piezoelectric breathing sensor, a common mode rejection circuit CMRC can be connected to both terminals of the corresponding bipolar input channel for reducing common mode interferences. On the first channel, voltage amplitudes at a frequency of 50 kHz can be measured in parallel to EMG by the means of a demodulation circuit. The second channel contains a switchable demodulation circuit which allows for measurement of an amplitude at 50 kHz or at 100 kHz. Together with the current sources, the two-electrode and four-electrode measurement method at 50 kHz and 100 kHz can be realized. Besides measuring BI independently at 50 kHz or/and at 100 kHz, it is possible to measure two transfer

impedances at two different positions using a frequency of 50 kHz (using only one current source). As BI measurement with Ag/AgCl surface electrodes and monopolar needle electrodes should be possible, the current amplitude was made switchable in a range of root mean square (RMS) values from  $3.4 \mu\text{A}$  to  $137 \mu\text{A}$  in eight steps (cf. Fig. 3.5). In order to increase the measurement resolution, different gain settings were implemented into the demodulation circuit. The optimal current and gain setting is chosen by monitoring the outputs of the fault detection circuits at the current source and the demodulation circuit. Different current and gain settings are tested automatically for each measurement until the current source and the demodulation circuit are in their preferred operating range.

After 24 bit analog to digital conversion, digital signal processing is used for further filtering. Together with the state of a galvanically isolated TTL input, all data are sent in real-time to the PC, on which the digital words can be converted into physical units.

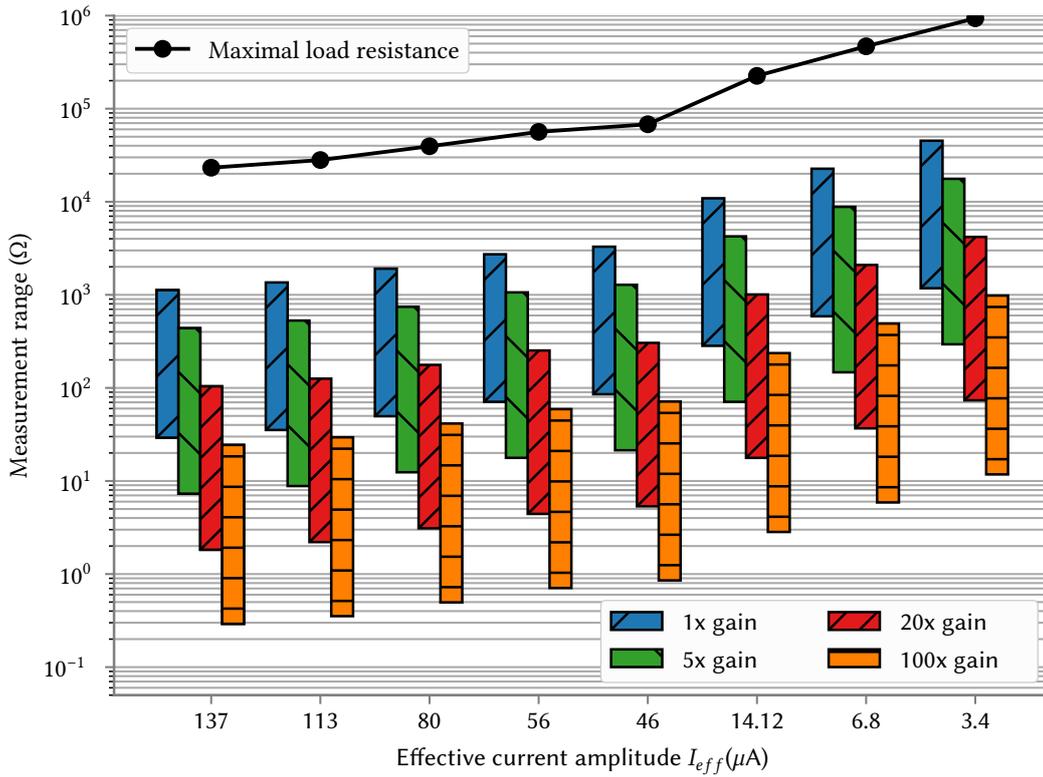
### 3.4 Current Sources

Different kinds of current sources are described in the literature which varies greatly in complexity and performance. Generally, they can be divided into two groups by the type of load. The first group is formed by single-ended current sources that drive a grounded load. The often used Howland current source belongs to this group [138]. A floating load is driven by the second group of current sources. The corresponding current sources are called floating current source and can be built by two single-ended current sources at which the current is injected by one single-ended current source and drained by a second inverted single-ended current source. An undesirable common mode voltage at the output of a floating current source can be caused by an imbalance of both internal single-ended current sources and can be reduced by a common-mode feedback (CMFB) circuit. This reduces the later BI measurement error as the common-mode rejection ratio (CMRR) of the differential amplifier for voltage measuring is limited (see next Section). Identical patient protection elements can be added in series to both output terminals of the floating current source.

Two floating current sources with CMFB are used in the developed measurement device. Their design is based on the patent US6501255B2 [139] and shown in Figure 3.4. The circuit design was enhanced with additional DC blocking Y1-capacitors ( $C_y$ ) which are connected in series to each output terminal. A Y1-capacitor is a fault-save capacitor whose dielectric can withstand very high voltages. In a case of a single fault condition, dangerous currents (especially long-term DC components are dangerous) are blocked by both Y1-capacitors. The output current  $I$  is continuously monitored for fault detection.

The output voltages of the CMFB ( $V_{cmr1}$  and  $V_{cmr2}$ ) are slowly changing voltages which counteract to a common mode voltage on the output terminals. A common mode voltage is mainly caused by external interferences.





**Figure 3.5:** Maximum measurement ranges for measuring BI (absolute values) at 50 kHz. The range is limited by the input range of the amplifier which depends on the gain. The measured voltage can be amplified with a gain of 1x, 5x, 20x, or 100x. The input ranges (peak-to-peak voltages) are 4 mV to 310 mV (1x), 1 mV to 59 mV (5x), 0.25 mV to 14 mV (20x), and 0.04 mV to 3.1 mV (100x). The lower limit of the measurement range is due to the precision (required minimally signal amplitude) of the later used full wave rectification inside the device. The maximum load resistance curve shows the maximum load which can be handled by the current source without going into saturation.

$$Z_{s50\text{kHz}} \approx \Delta R_L \frac{U_1}{\Delta U_L} = R_2 \frac{U_1}{U_2 - U_1} = 500 \Omega \frac{42.155 \text{ mV}}{0.111 \text{ mV}} = 189.9 \text{ k}\Omega, \quad (3.3)$$

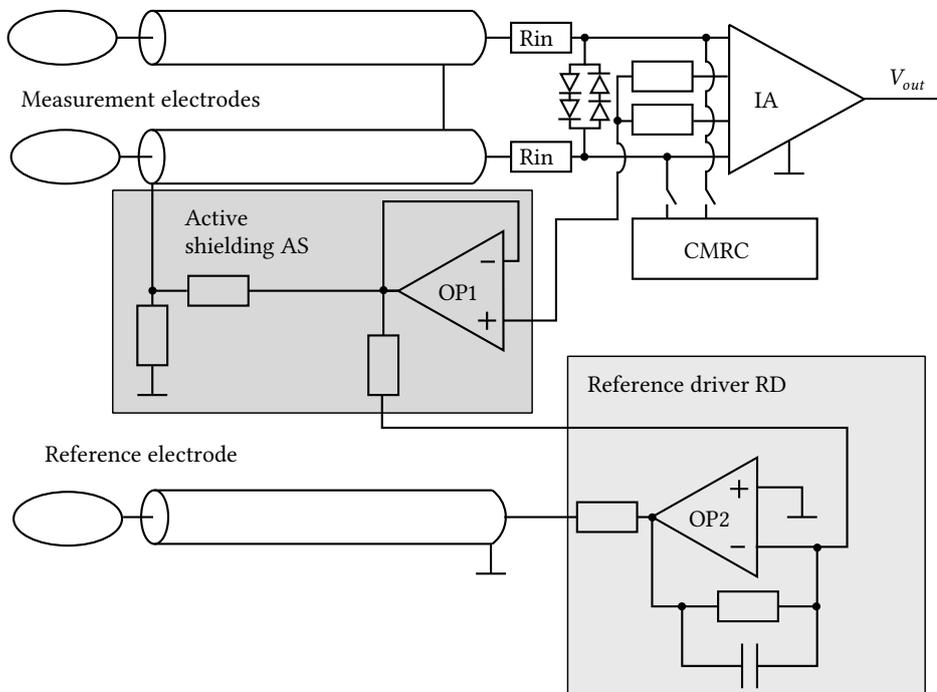
$$Z_{s100\text{kHz}} \approx R_2 \frac{U_1}{U_2 - U_1} = 500 \Omega \frac{42.95 \text{ mV}}{0.309 \text{ mV}} = 69.5 \text{ k}\Omega. \quad (3.4)$$

The electrode-skin impedance of a wet ECG electrode at a measurement frequency of 50 kHz and 100 kHz lays around 100 k $\Omega$  [19]. A single stainless steel monopolar needle electrode has an impedance of around 1 k $\Omega$  at 50 kHz [141]. The expected load impedance is much lower than the measured output impedance of 190 k $\Omega$  and 70 k $\Omega$  and should therefore be sufficient for accurate BI measurements [140].

The BI measurement range which is achieved with a RMS current of 137  $\mu\text{A}$  lays around 1 k $\Omega$  and has not enough safety margins for measuring with needle electrodes (see Figure 3.5). Lowering the current amplitude would increase the maximally possible load but would also reduce the measurement accuracy for wet electrodes. A configurable current amplitude allows the usage of

surface and needle electrodes with the optimal current amplitude and consequently with optimal resolution. The current amplitude  $I_{eff}$  can be set by means of the resistor  $R_k$  to one of eight available settings in the range  $3.4 \mu\text{A}$  to  $137 \mu\text{A}$  which results theoretically in a maximal load resistance between  $23 \text{ k}\Omega$  ( $I_{eff} = 137 \mu\text{A}$ ) to  $935 \text{ k}\Omega$  ( $I_{eff} = 3.4 \mu\text{A}$ ). This wide range enables the applicability of the measurement system for different electrode positions and types. As the measurement range of the instrumentation amplifier is limited, the achievable measurement range is smaller than the maximally possible one. Figure 3.5 shows the measurement ranges for all current amplitudes and gain settings of the voltage measurement (cf. Section 3.6) which are realizable at a measurement frequency of  $50 \text{ kHz}$ . The measurement ranges at  $100 \text{ kHz}$  are slightly different, as the voltage measurement part has different input ranges and gain factors at this frequency.

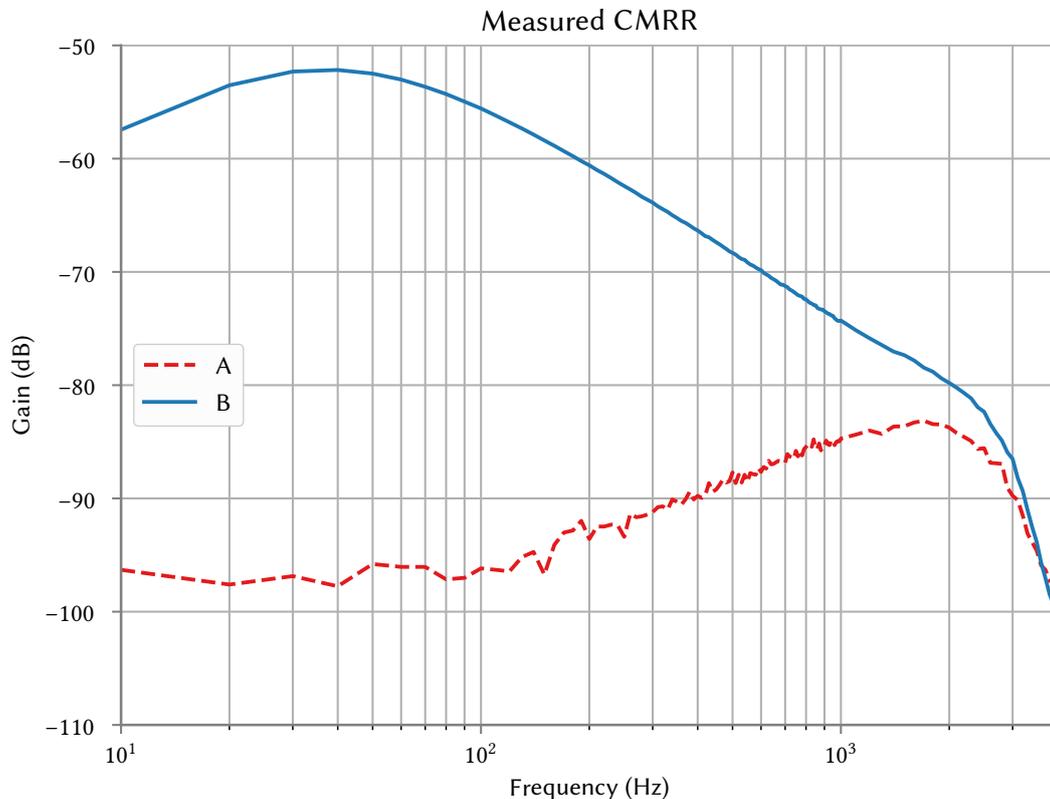
### 3.5 Voltage Measurement



**Figure 3.6:** The circuit design is used in all four bipolar voltage measurement channels. Input elements consist of  $R_{in}$  and diodes ensure protection of the instrumentation amplifier IA and protection of the patient in cases of a failure. The cables are actively shielded by AS in order to ensure good signal quality. Common mode disturbances in the measurement area are reduced by a reference driver RD which is connected to the first channel. In order to reduce power supply interferences by voltage measurements of external sensors, the common mode rejection circuit CMRC can be individually activated for each channel. (OP - operational amplifier)

Figure 3.6 shows the amplifier design of the bipolar voltage measurement channels. The instrumentation amplifier IA (INA128, Texas Instruments, USA) is connected through  $100 \text{ k}\Omega$  high precision resistors ( $R_{in}$ ) to the electrodes in order to achieve patient safety. In a case of a failure, the maximal current is limited by these resistors to  $50 \mu\text{A}$ . The instrumentation amplifier is protected against high voltages by four diodes (BAV199, Infineon Technologies AG, Germany). This allows applying FES in the vicinity of the BI/EMG measurement electrodes.

The measured input impedance of the instrumentation amplifier is  $3\text{ M}\Omega$ . The amplifier gain of IA is 10.26. This low gain makes it possible to measure EMG without high-pass filtering (input range is  $[-150\text{ mV}, 150\text{ mV}]$ ). The sinusoidal carrier wave at 50 kHz or 100 kHz which is used for assessing BI can be measured simultaneously and the measurement range of BI is increased due to the large input range.



**Figure 3.7:** The CMRR of the first EMG channel was measured by applying a sinusoidal voltage with an amplitude of 500 mV across the amplifier ground to both mutually connected input terminals. A - the reference driver was connected to the input terminals; B - the reference driver was disconnected, while the common mode rejection circuit (CMRC) on the amplifier inputs was active.

The measurement cables, which connect the amplifier with the electrodes, are shielded in order to reduce the effect of external noise and electrostatic interference. As carbon coated shielded cables (Mind Media BV, Roermond-Herten, Netherlands) have been used, cable movement artifacts are strongly reduced. The additional parasitic capacitance between shield and amplifier input is reduced by using active shielding in which the shield is driven to the average potential of both input terminals using the unity gain buffer OP1. One common reference signal generator (consisting of operational amplifier OP2) is used for all channels in order to reduce common mode disturbances in the measurement area. The reference signal is equal to the low-pass filtered average signal of the first channel. Thus, the first channel must always be used for measurement with wet electrodes on the skin.

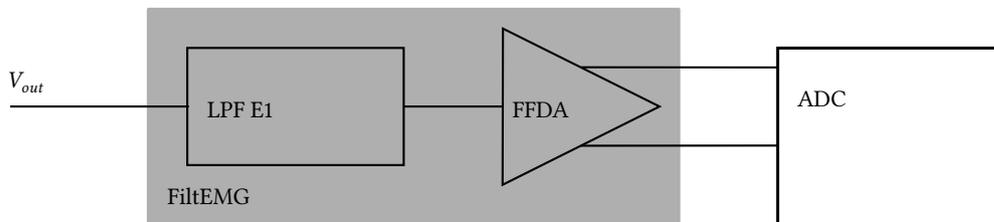
While using needle electrodes, the reference generator may drive the amplifier into saturation due to the voltage difference between the potential on the needles and the reference electrode on

the skin. In order to be able to measure using needle electrodes, the reference generator can be replaced by the CMRC (see Fig. 3.6) which is connected through solid state relays with MOSFET output (AQW210, Panasonic EW, Japan) to both instrumentation amplifier inputs behind  $R_{in}$ . The relays are optically isolated (PhotoMOS). When both switches were active, CMRC reduces common mode voltages at the amplifier input. Normally, the switches are in an off-state and the reference generator from the first bipolar input channel is used for reducing common mode voltages. It is possible to activate CMRC individually for each channel except the first channel, which should be connected to the skin for achieving a working reference generator. By using CMRC, it is also possible to connect sensors with a high output impedance to the measurement system and reduce common mode voltages, especially at 50 Hz. Such sensors can be piezo or thermocouple sensors for measuring respiration. The CMRC is based on the work of Dobrev et al. [142]. Disadvantages of this circuit in comparison to the reference generator, are the not flat amplitude response (drop of 3 dB from 10 Hz to 1 kHz, see Fig. 3.17) and its reduced common mode rejection ratio (52 dB at 50 Hz, see Fig. 3.7).

The CMRR was measured by mutually connecting both input terminals and applying a sinusoidal voltage with an amplitude of 500 mV against amplifier ground. The frequency of the sinusoidal voltage was varied from 10 Hz to 4000 Hz. The results are displayed in Figure 3.7. The best CMRR was achieved for case A, by which the reference driver was connected to the short-circuited input terminals. The measured CMRR at the important main supply frequency of 50 Hz is then 96 dB.

## 3.6 Analog Signal Processing

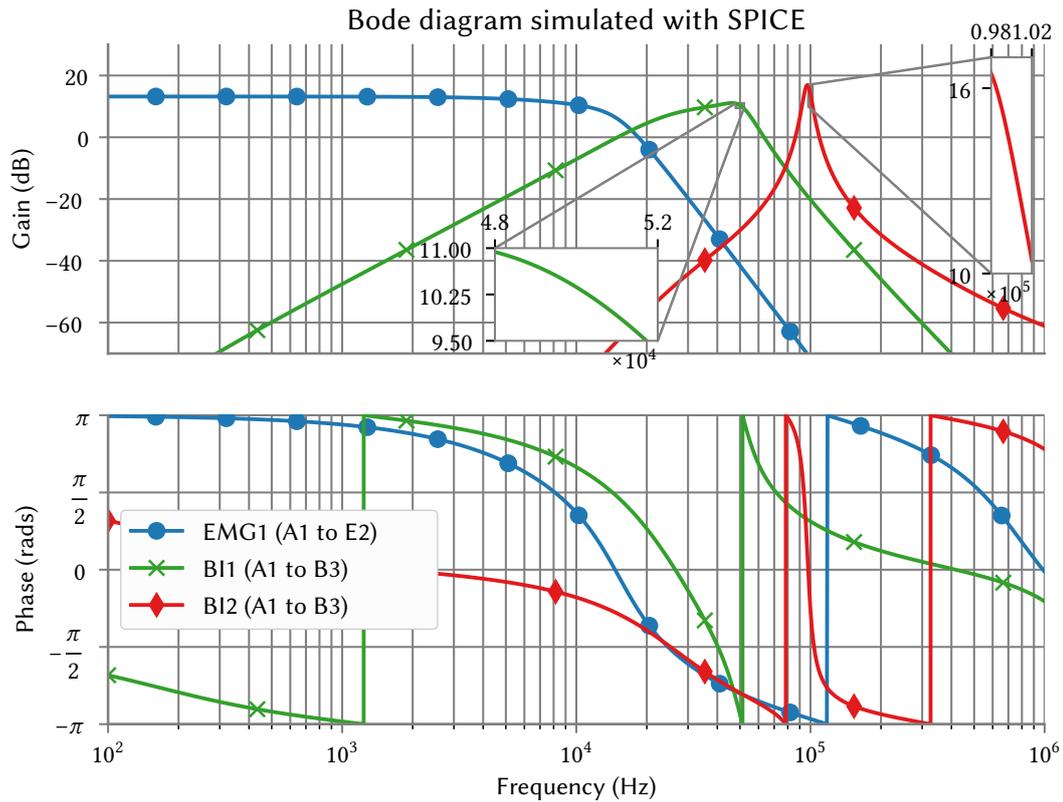
### 3.6.1 EMG Filtering



**Figure 3.8:** The filter and amplifier block FiltEMG is used in all four input channels for EMG signal processing.  $V_{out}$  is the measured voltage from the instrumentation amplifier shown in Figure 3.6. (LPF - low-pass filter; FFDA - feedback fully differential amplifier; ADC - analog to digital converter)

Whenever the device simultaneously measures EMG and BI, the output signal of the instrumentation amplifier is a mixture of the amplitude modulated BI signal, EMG, and noise. The EMG signal has normally a frequency range of 10 Hz to 450 Hz [143] while the amplitude modulated BI signal has a different range around 50 kHz or 100 kHz respectively. Analog filter circuits consisting of operational amplifiers are used for separating both sinusoidal carrier waves and EMG from each other. The multiple feedback (MFB) filter structure is mostly used in the following as this filter structure has a low sensitivity to component variations [144].

The EMG signal is separated from the BI frequency content using a fourth-order MFB low-pass filter E1 with Butterworth characteristics as such a filter has the flattest possible pass-band

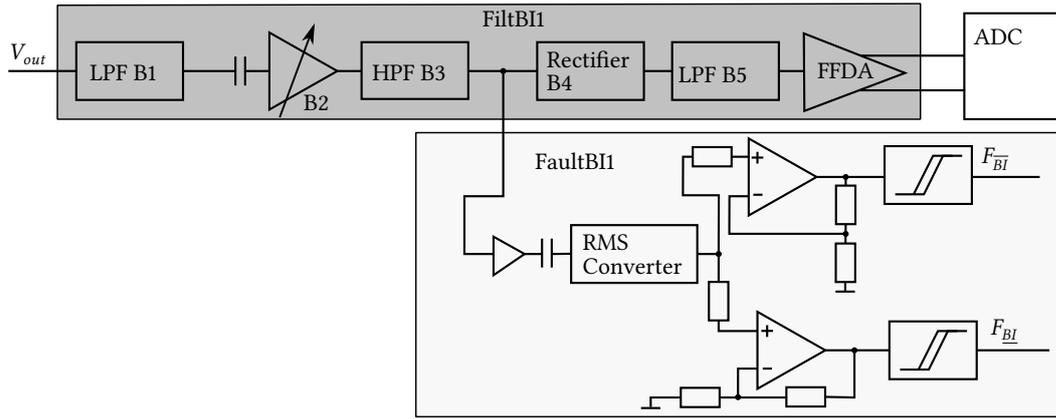


**Figure 3.9:** Simulated frequency responses of the filter characteristics of EMG1, BI1, and BI2. The signal characteristics of the EMG signal are shown without the low-pass filter of the ADC and the finite impulse response (FIR) filter within the micro-controller. The carrier frequency of 50 kHz / 100 kHz for BI measurement is damped with 54.8 dB / 84.7 dB.

EMG signal components in a range up to 1 kHz are damped by at least 40 dB for the BI1 measurement input. The separation between both carrier frequencies could be achieved. The frequency response of BI1 around 50 kHz is shown from 48 kHz to 52 kHz. Around 100 kHz, the frequency response of BI2 is shown separately from 98 kHz to 102 kHz. In both areas around the carrier frequencies, an almost constant descent could be observed which is necessary for an accurate amplitude demodulation. Frequencies higher than 2 kHz will be removed after demodulation by the anti-aliasing low-pass filters.

magnitude response. The cut-off frequency was set to 13 kHz<sup>1</sup>, which was the highest possible cut-off frequency in order to damp the carrier frequency of 50 kHz or 100 kHz sufficiently. The feedback fully differential amplifier FFDA (THS4521, Texas Instruments, USA) is used in order to convert the analog signal to a differential signal, which reduces the analog to digital conversion error in the ADC. The FFDA has a first-order low-pass characteristics with a cut-off frequency of 13 kHz.

The FiltEMG signal processing block (cf. Fig. 3.8) reduces the amplitude modulated BI signal (at a carrier frequency of 50 kHz or 100 kHz) by 54.8 dB and 84.7 dB respectively (see Fig. 3.9).



**Figure 3.10:** Structure of FiltBI1 for filtering and extracting the amplitude of the measured BI signal.  $V_{out}$  is the measured voltage from the instrumentation amplifier shown in Figure 3.6. B1 to B3 form a band-pass filter with switchable gain for extracting the BI carrier wave at 50 kHz. B4 and B5 are used as an envelope detector for amplitude demodulation. The FFDA converts the signal to a differential signal for accurate analog to digital conversion. FaultBI1 is used for monitoring the RMS voltage at the output of B3. (LPF - low-pass filter; HPF - high-pass filter; ADC - analog to digital converter)

#### 3.6.2 BI Filtering for Measurement at 50 kHz

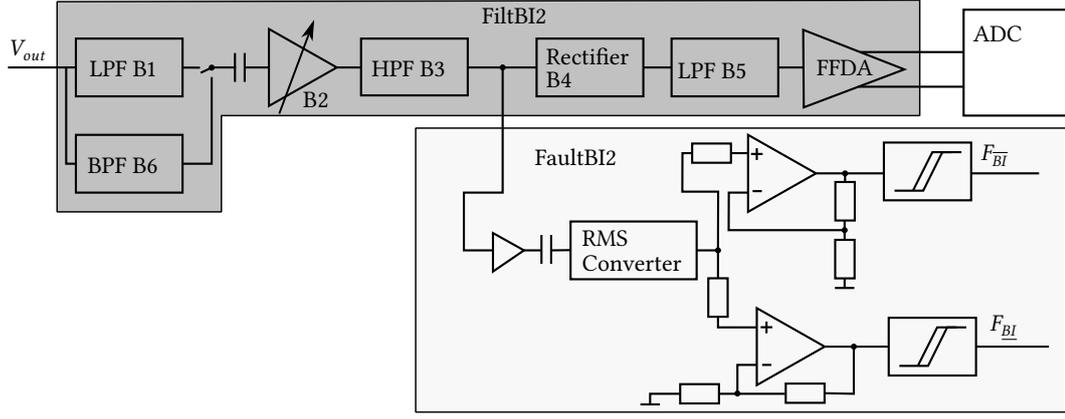
In order to extract the sinusoidal carrier wave, the filters B1 and B3 is used in FiltBI1 (cf. Fig. 3.10). A fourth-order MFB Chebychev (ripple of 0.5 dB) low-pass filter B1 ( $f_{cut-off} = 50$  kHz) is used to separate the carrier wave of 50 kHz from the second BI measurement frequency of 100 kHz. The frequency content at 100 kHz is damped by 30 dB, which was only possible by using Chebychev filter parameters.

In the next step, the signal is amplified (1x, 5x, 20x, or 100x) using a programmable amplifier B2 (LTC6910, Linear Technology, USA). Then, a second-order Butterworth high-pass filter B3 with a Sallen-Key topology ( $f_{cut-off} = 26$  kHz) is used to suppress noise and the EMG signal content. This filter prepares the signal for a full wave rectifier, which can only produce a rectified voltage when all DC components are removed from the signal. The high-pass filter B3 suppresses signal content below 1 kHz by 57 dB, which separates the BI signal content at 50 kHz sufficiently from the EMG signal. A bode diagram, which displays the filter characteristics at the output of B3, is shown in Figure 3.9. The envelope detector containing B4 and B5 is described in Section 3.6.4. The amplifier FFDA is described in Section 3.6.1.

#### 3.6.3 BI Filtering for Measurement at 100 kHz

The second channel is able to measure BI either at a frequency of 50 kHz or at a frequency of 100 kHz. An electrical switch is used in order to select either the B1 or the B6 filter structure at the beginning of FiltBI2 (cf. Fig. 3.11). For measurements at 50 kHz, at which a second transfer impedance can be measured in parallel to the first channel using the same current source, the low-pass filter B1 is used. FiltBI2 has the same filter characteristics as FiltBI1 (see Section 3.6.2).

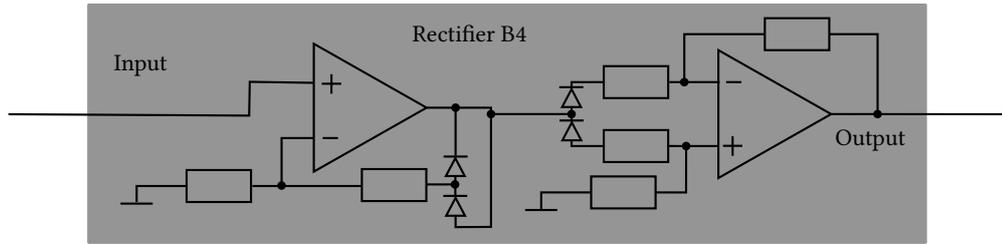
<sup>1</sup>Due to the over-sampling ratio of 64 and digitally low-pass filtering within the ADC, the selected cut-off frequency is sufficient for sampling with 12 kHz later in the ADC. A high sampling rate is necessary for accurately blanking of stimulation artifacts when measuring during active FES.



**Figure 3.11:** Structure of FiltBI2 for filtering and extracting the amplitude of the measured BI signal at the second measurement channel.  $V_{out}$  is the measured voltage from the instrumentation amplifier shown in Figure 3.6. B1 + B2 + B3 and B6 + B2 + B3 are used for extracting the BI carrier wave at 50 kHz and 100 kHz respectively. B4 and B5 are used as an envelope detector for amplitude demodulation. The FFDA converts the signal to a differential signal for accurate analog to digital conversion. FaultBI2 is used for monitoring the RMS voltage at the output of B3. (LPF - low-pass filter; BPF - band-pass filter; HPF - high-pass filter; FFDA - feedback fully differential amplifier; ADC - analog to digital converter)

For measuring BI at 100 kHz, a second-order MFB band-pass filter B6 with a resulting center frequency of 97 kHz and a Q factor of 16 is used. This band-pass filter dampens the carrier wave at 50 kHz of the first BI measurement by 45.4 dB. Thus, two independent BI measurements can be performed by the PHYSIOSENSE device. The resulting amplitude and phase diagram is shown in Figure 3.9. The components B2 and B3 are described in Section 3.6.2, while details about the envelope detector (B4 and B5) are given in Section 3.6.4.

### 3.6.4 Amplitude Demodulation



**Figure 3.12:** Circuit layout of the high precision full-wave rectifier B4 which is used in the envelope detector.

The measured and filtered BI carrier signal  $C(t)$  at time instance  $t$  behind the high-pass filter B3 can be written as

$$C(t) = BI(t) \sin(\omega_c t + \phi_c), \quad (3.5)$$

where  $BI(t)$  is a time-variant amplitude,  $\omega_c$  is a fixed frequency at either  $2\pi 50 \cdot 1 \times 10^3$  rad/s or  $2\pi 100 \cdot 1 \times 10^3$  rad/s, and  $\phi_c$  is an initial phase, which can be neglected. It is assumed that  $BI(t)$  does only contain frequency components up to frequency  $\bar{\omega}_m < 0.5\omega_c$  where  $\bar{\omega}_m = 2\pi 2000$  rad/s is assumed as a maximum frequency content of 2 kHz within the BI measurement signal. In case that only one frequency component is present,  $BI(t)$  can be written as cosine with a constant amplitude

$M$ , a phase  $\phi$ , and an offset  $M_0$ :

$$BI(t) = M \cos(\omega_m t + \phi) + M_0. \quad (3.6)$$

The measured voltage  $C(t)$  at the carrier frequency  $\omega_c$  can then be written as

$$C(t) = (M \cos(\omega_m t + \phi) + M_0) \sin(\omega_c t). \quad (3.7)$$

The equation can be expanded to the following equation:

$$C(t) = M_0 \sin(\omega_c t) + \frac{M}{2} [\sin((\omega_c + \omega_m)t + \phi) + \sin((\omega_c - \omega_m)t - \phi)]. \quad (3.8)$$

The rectification is obtained by a multiplication with the carrier wave  $\sin(\omega_c t)$ :

$$\begin{aligned} C(t) &= \frac{M_0}{2} (1 - \cos(2\omega_c t)) + \frac{M}{2} \cos(\omega_m t + \phi) \\ &\quad - \frac{M}{4} [\cos((2\omega_c + \omega_m)t + \phi) + \cos((2\omega_c - \omega_m)t - \phi)]. \end{aligned} \quad (3.9)$$

Ideal low-pass filtering results in (all frequencies greater than  $\omega_c$  are removed):

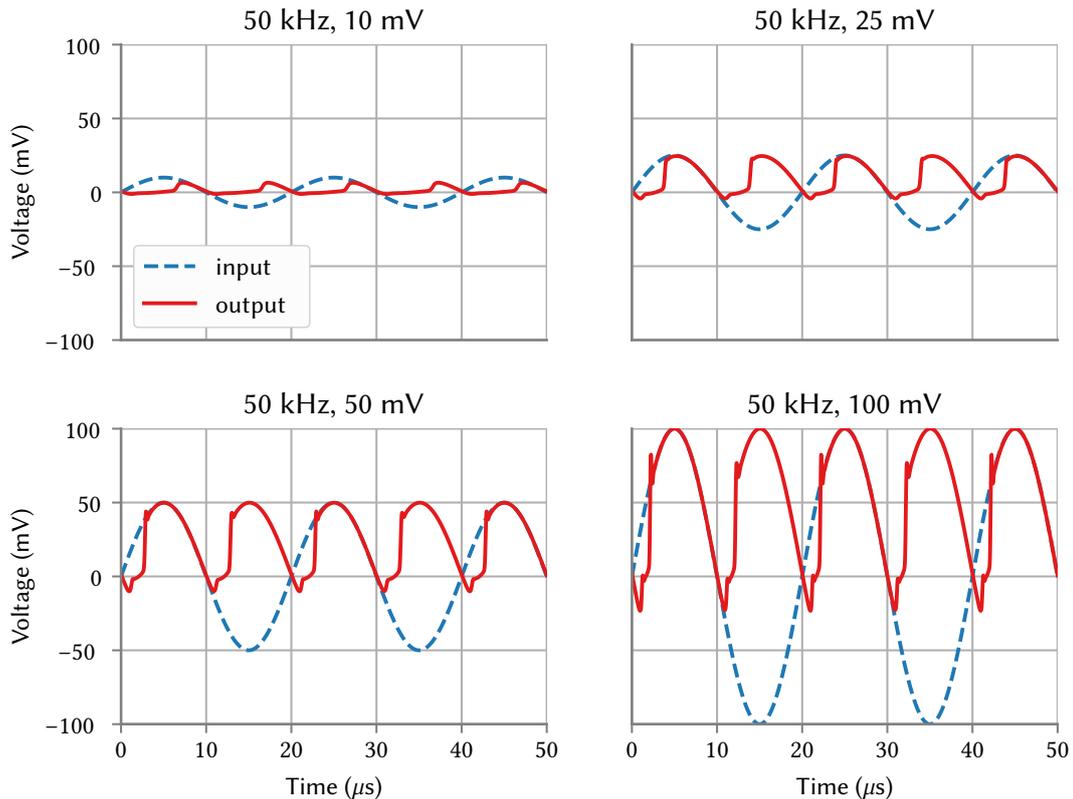
$$BI^*(t) = \frac{M_0}{2} + \frac{M}{2} \cos(\omega_m t + \phi) = \frac{1}{2} BI(t). \quad (3.10)$$

The amplitude modulated signal is transmitted in a frequency range from  $\omega_c - \omega_m$  to  $\omega_c + \omega_m$ . The frequency response  $H_{BI}$  for a BI measurement should fulfill the condition

$$\frac{|H_{BI}(j(\omega_c - \omega_m))| + |H_{BI}(j(\omega_c + \omega_m))|}{2} = |H_{BI}(j\omega_c)| = 1, \quad 0 < \omega_m < \bar{\omega}_m \quad (3.11)$$

in order to allow the amplitude accurate demodulation. In other words  $|H_{BI}|$  should be constant or linear and symmetrically around  $\omega_c$  in the range  $\omega_c - \omega_m$  to  $\omega_c + \omega_m$ . It can be verified in Figure 3.9 that the latter is true of the applied BI filter characteristics in the PHYSIOSENSE device.

As shown in Equations (3.9) and (3.10), an ideal amplitude demodulation can be achieved with an ideal rectification followed by low-pass filtering. This approach has been used inside the PHYSIOSENSE device. The precision full wave rectifier B4 and the fourth-order MFB Butterworth low-pass filter B5 ( $f_{\text{cut-off}} = 13$  kHz) form an envelope detector by which  $BI^*(t)$  (a scaled version of the bioimpedance  $BI(t)$ ) can be determined. This analog demodulation requires a precision full wave rectifier that is able to rectify sine waves with amplitudes in the range from about 50 mV up to the supply voltage. Any precision rectifier which uses diodes should compensate the forward voltage drop across the diodes as well as possible. The used rectifier (see Fig. 3.12 for a detailed circuit diagram) utilizes diodes (BAV199, Infineon Technologies AG, Germany) with very low leakage current (max. 5 nA) and is able to rectify the peak voltages of a sine wave at a frequency of 50 kHz down to an amplitude of 25 mV with an appropriate precision. At a frequency of 100 kHz, the rectifier is able to process amplitudes higher than 50 mV. Figure 3.13 shows the input and output of the rectifier for different amplitudes of a 50 kHz sine wave at the input. The remaining minor errors in the rectification for amplitudes above 25 mV (or 50 mV for 100 kHz) can be reduced by a non-linear



**Figure 3.13:** Simulated output of the full wave rectifier B4 with a sine wave of 50 kHz as input. It can be seen, that the rectifier is not able to fully rectify the input signal for amplitudes of 10 mV. A non-linear behavior can be seen, as the first part of the rising sine wave is not correctly rectified.

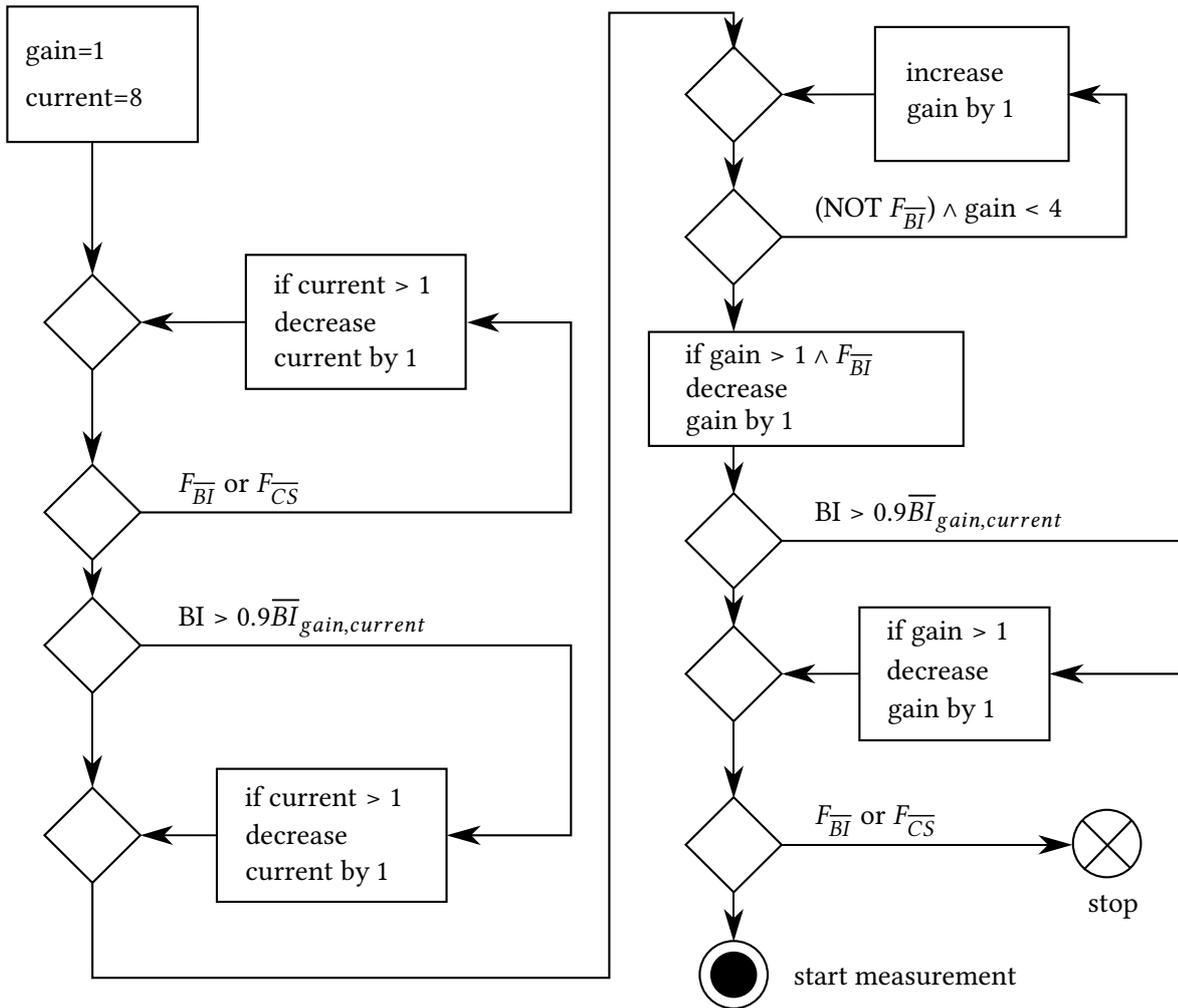
amplitude dependent correction (calibration) of the rectifier output. Details on this will be given in Section 3.9.

An initial automatic BI measurement configuration (see Section 3.7) ensures that the amplitude of the sinusoidal rectifier input does not go below the threshold of 25 mV for 50 kHz (50 mV for 100 kHz), assuming no large changes in the assessed BI measurement range. Unforeseen changes in the BI measurement range or electrode contact problems, that may lead to amplitudes below the reported threshold, will be reported by a fault detection (see Section 3.7). A detected fault will automatically lead to a stop of a running measurement.

### 3.7 Fault Detection and Automatic BI Configuration

The device is able to detect whenever the available measurement ranges (see Fig. 3.5) are left. Therefore, the RMS output voltage of the filter B3, which is part of the signal processing chain `filtBI1` and `filtBI2`, and the RMS voltage across the output terminals of each current source during BI measurement are continuously monitored. A fault is detected if any of these voltages exceed or fall below pre-defined thresholds. After a fault is detected, the running BI measurement will be stopped and afterward restarted with adjusted current and gain settings. If it is not possible to find suitable

settings, the measurement will be permanently stopped. The latter is the case if an electrode has fallen off or if the electrode contact is too bad, because the electrode becomes lose.



**Figure 3.14:** Automatic configuration for finding the optimal current amplitude and gain setting. The highest current amplitude setting (current=8) corresponds to 137  $\mu\text{A}$ , while current=1 means that a current of 3.4  $\mu\text{A}$  should flow between the electrodes. The lowest gain setting (gain=1) corresponds to a gain of 1x, while the highest gain (gain=4) means that the signal is amplified by a factor of 100. See Figure 3.5 for all available gain and current settings.

Measuring smaller impedances than the BI measurement range allows, results in too small voltage amplitudes which cannot sufficiently be processed and demodulated. Too small voltage amplitudes at the current sources are denoted as  $F_{CS}$  and are identified by monitoring the RMS voltage across the output terminals of each current source.  $F_{CS}$  is triggered when this RMS voltage goes below a pre-defined threshold. A second related error  $F_{BI}$  is triggered when the RMS output voltage of the filter B3 falls below a pre-defined threshold. Small voltage amplitudes might be the result of small transfer impedances which can be caused by too narrow placed current electrodes (leading to  $F_{CS}$ ) or too narrow placed voltage measurement electrodes (leading to  $F_{BI}$ ).

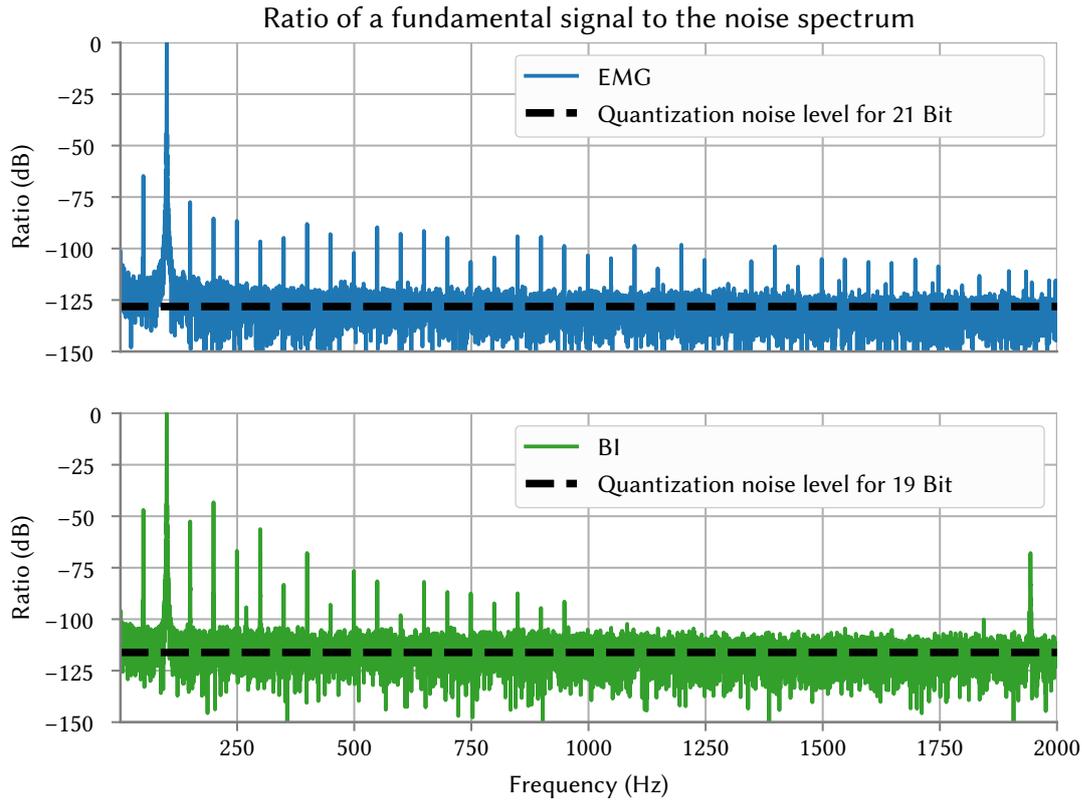
On the other hand, measuring impedances above the predefined BI measurement range leads to too high voltages within the current source and voltage measurement circuits which are too close to the supply voltages. This leads to saturation effects which falsify the measured BI values. Too

high voltages at the current source are monitored by assessing the RMS voltage across the output terminals. An exceeding will lead to an error  $F_{CS}$  which indicates that the load impedance is too high. This error might be triggered during a measurement when one of the current electrodes falls off. Too high voltages at the measurement circuits are prevented by monitoring the RMS output voltage at the filter B3. The related error  $F_{BI}$  is triggered whenever the transfer impedance is too high which could be caused during a measurement by a bad electrode-skin contact. The errors  $F_{CS}$ ,  $F_{BI}$ ,  $F_{CS}$ , and  $F_{BI}$  are transmitted immediately to the PC. The measurement is halted and the current source is switched off.

The fault detection is also used in the beginning of each measurement in order to find the best combination of current and gain for the current measurement setup. As the device is mainly used for the four-electrode BI measurement with separate current sources for each BI measurement, the current amplitude and the amplifier gain are optimized for this type of setup. The current amplitudes for CS1/2 are set initially to the maximally possible value and the amplifier gain for filtBI1/2 are set to the lowest value. By setting the current amplitude to the highest value, it can be assured that the highest possible current amplitude is used in order to have the best possible signal quality. The amplifier gain is set to the lowest value in order to optimize the gain independent from the current amplitude as the lowest gain leads to the highest measurement range. After finding a current amplitude by decreasing the current step by step and stopping at this point at which  $F_{CS}$  and  $F_{BI}$  are not triggered (including the highest current), the gain is increased until  $F_{BI}$  is triggered or the maximally possible gain is reached. If  $F_{BI}$  was triggered and the gain was increased, the gain is set to the next lower value. As soon as a current or gain setting has been selected, it is checked whether the measured BI value is too close to the highest possible measurable value  $\overline{BI}_{gain,current}$  for the selected current and gain. Thus, the current amplitude and also the gain value is decreased when  $BI > 0.9\overline{BI}_{gain,current}$ . When no sufficient current and gain value could be found, the measurement is stopped with an informative error message. The optimizing procedure for finding the best current amplitude and gain setting is shown in detail in Figure 3.14.

### 3.8 Digital Signal Processing

A 24-bit ADC (ADS1278, Texas Instruments, USA) is used to simultaneously sample up to four EMG and two BI inputs at a sampling frequency of 12 kHz. An FFDA (THS4521, Texas Instruments, USA) is used in order to convert each analog signal to a differential signal before sampling. The amplifier is part of FiltEMG, FiltBI1, and FiltBI2. The differential signals from all inputs are fed to the ADC. By using such signal conversion, the signal quality improves as a differential signal is robust against electromagnetic interference and the non-linear conversion error of the ADC is smaller for such signals. The FFDA has a first-order low-pass characteristics with a cut-off frequency of 13 kHz. The chosen cut-off frequency is sufficient due to the over-sampling and anti-aliasing filtering within the ADC. The over-sampling ratio is 64 and the signals are then sampled with 768 kHz. The digitally low-pass filter within the ADC damps all frequency above the Nyquist frequency of 6 kHz with more than 100 dB. Finally, the down-sampled data are send with the output sampling frequency of 12 kHz to the microcontroller.



**Figure 3.15:** A fast Fourier transform (FFT) is performed and shown for the recorded measurement data of a sine-wave signal near full scale at a frequency of 100 Hz. For obtaining an FFT for the BI channel, an amplitude modulated sine wave with a modulation degree of 80 % at 50 kHz with a frequency of 100 Hz is used. The signal-to-noise ratio (SNR) is now defined as the distance between the fundamental signal to the non-harmonic part of the remaining spectrum. For comparison, the signal-to-quantization-noise ratio (SQNR) for 21 bit and 19 bit are plotted.

In order to check the minimally necessary sampling resolution, the SNR was once measured by an FFT analysis (see Fig. 3.15). A sine wave with a frequency of 100 Hz at the maximum input level was applied to a bipolar input channel, filtered through the EMG filters and sampled by the ADC with 24 bit. Harmonic distortions of multiple of 100 Hz were mixed with power supply interferences at 50 Hz and multiple of it. As the level of harmonics at the multiple of 100 Hz did not differ from the harmonics at the multiple of 50 Hz, it can be assumed that almost all harmonic distortions were caused by the power supply interferences which are reduced in real measurements by the reference driver.

Figure 3.15 shows the ratio of the fundamental signal to the noise spectrum in dB. The SQNR of a sinusoidal signal excludes all harmonic frequencies and is defined as

$$SQNR_w = 6.02w + 1.76 \text{ [dB]}, \quad (3.12)$$

where  $w$  is the resolution of the ADC. The SQNR of an effective resolution of 21-bit is 128.18 dB. It can be clearly seen that the SNR within the measured data almost corresponds to the SQNR for an

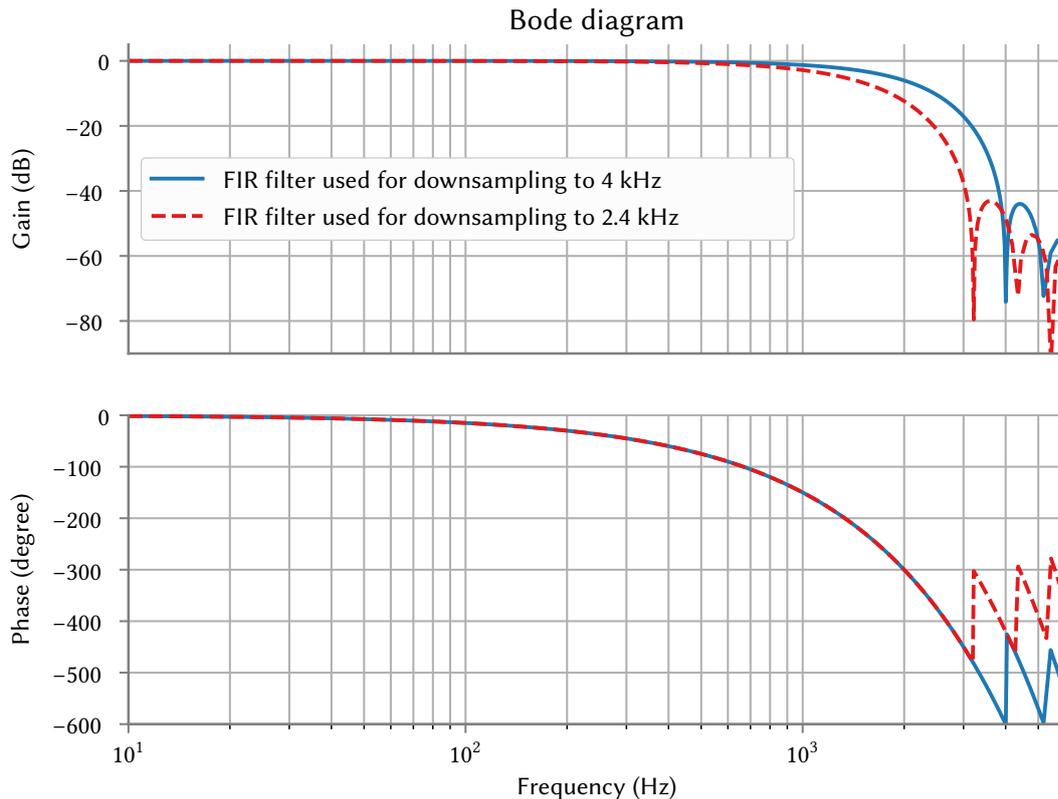
ideal 21-bit sampling. Thus, it is sufficient to transmit only the first 21 bits of EMG measurement data. The EMG measurement resolution is then  $0.15 \mu\text{V}$  for an input range of  $\pm 150 \text{ mV}$ .

In a second investigation, a sine wave with a frequency of 100 Hz was amplitude modulated by a carrier frequency of 50 kHz and applied to the first bipolar input channel. After analog demodulation by the BI filter and processing stage, the voltage is sampled with the 24-bit ADC. The measured SNR of the BI channel corresponded to an SQNR of an ideal 19-bit ADC, which is 116.24 dB (cf. Eq. (3.8)). The effective resolution of 19 bits for the BI measurement results for a two-electrode measurement ( $I_{eff} = 137 \mu\text{A}$ , gain = 1x) on a load resistor of  $100 \Omega$  in an effective measurement resolution of  $0.2 \text{ m}\Omega$ . Harmonic distortions at the multiple of 50 Hz which are caused by power supply interferences and 100 Hz are visible in a range up to  $-40 \text{ dB}$  (cf. Fig. 3.15). Harmonic distortions at the multiple of 100 Hz show a higher level than the harmonics between them, which indicates a non-linear behavior of the envelope detector. In real measurements with electrodes, harmonic distortions which are caused by power supply interferences will be reduced by the reference driver.

The micro-controller (STM32F103RET6, ST microelectronics, Switzerland) collects the data via a serial peripheral interface (SPI) from the ADC and reduces the bit-resolution for all measurement data to a resolution of 21 bits by keeping the first 21 most significant bits. The EMG and BI measurement data are reduced to the same bit-resolution in order to simplify the transmission protocol. Additionally to the data of the input signals, a galvanically isolated TTL input is available and its state is transmitted to the PC together with the other measurement data. In order to send the data over a galvanically isolated serial USB converter (FT232RL, FTDI) to the PC, a universal asynchronous receiver transmitter (USART) bus is used. The transmission bit rate of the USART bus is limited; maximal 16000 Samples per second with a resolution of 21 bits can be transmitted to the PC. Therefore, the sampling rate is reduced to 4 kHz for each signal, when using up to 4 channels, and reduced to 2.4 kHz, when using 5 or 6 channels before passing the data from the microcontroller to the PC. As the sampling frequency of the ADC cannot be changed digitally, down-sampling has to be implemented in the microcontroller. As the used sampling frequency is 12 kHz which is a multiple of 4 kHz and 2.4 kHz only proper digital anti-aliasing filter are necessary for down-sampling.

In order to avoid aliasing by the down-sampling, by which higher frequent signal content would be mapped to lower frequency bands, two different FIR filters for both output sampling frequencies of 2.4 kHz and 4 kHz were designed, respectively. Both FIR filters have an impulse response length of 10 time steps and are implemented on the microcontroller using fix-point calculation. The FIR filter coefficients are symmetrical and are chosen in such a way that the resulting rounding error is minimized. The coefficients for the FIR filter which is used at 4 kHz are calculated from a Hanning window function at a cut-off frequency of 2 kHz. The coefficients are approximated by fractions with 256. The root mean square error to the original coefficients is 0.0258729. If five or six channels should be transmitted, a similar FIR filter is used with a cut-off frequency of 1.2 kHz. The root mean square error for approximated coefficients is then 0.0165650. An amplitude and phase plot in a frequency range from 10 Hz to the internal Nyquist frequency of the ADC of 6 kHz is shown for both FIR filters in Figure 3.16. The coefficients and the output of the digitally implemented FIR filters are determined by using Scilab 5.4.0<sup>2</sup>.

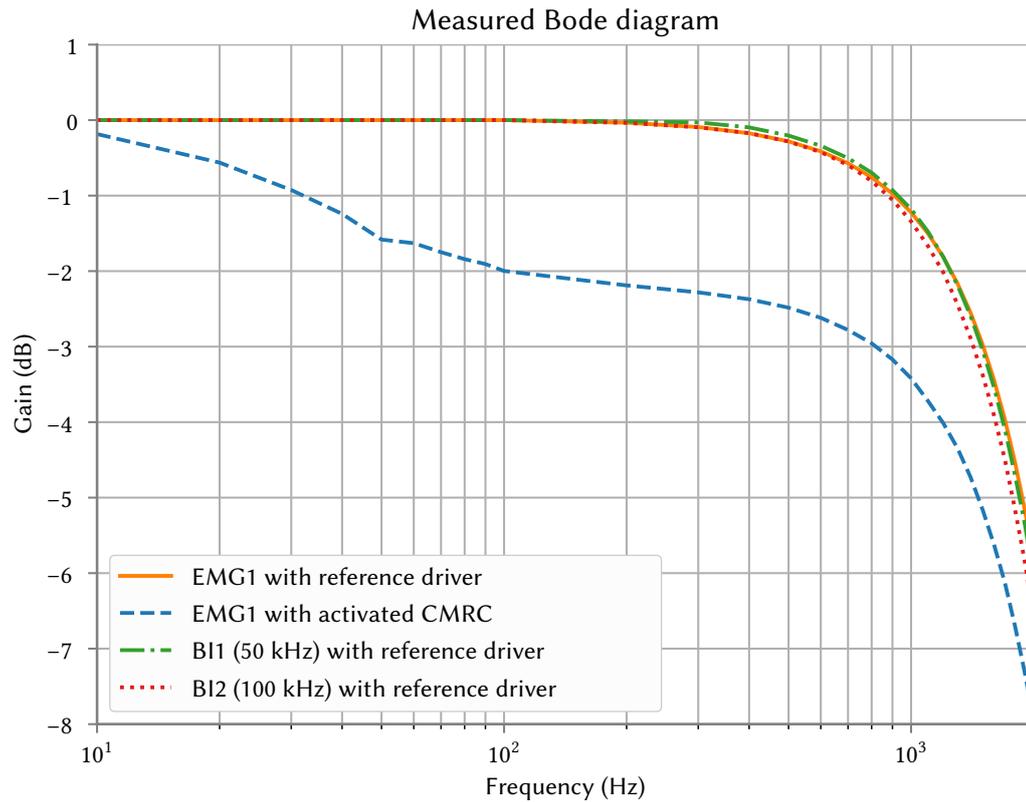
<sup>2</sup><https://www.scilab.org>



**Figure 3.16:** Frequency response of the digitally implemented anti-aliasing FIR filters which are used for down-sampling from 12 kHz to 4 kHz and 2.4 kHz, respectively. The bode plot goes from 10 Hz to the Nyquist frequency of 6 kHz. The FIR filters have a impulse response length of 10. Although the coefficients are rounded to multiples of  $\frac{1}{256}$ , proper damping for avoiding aliasing effects could be achieved.

A transmitting protocol was developed. It consists of a header, a body and a seven-bit cyclic redundancy check. The leading bit of each byte is used for marking the beginning of a package. The header byte characterizes the package. The measurement data from all channels are packed into the body of the package. Other package types are failure report, configuration, and actions as reset, start and stop. In order to check the status of the PC, the PC has to continuously send a package to the device. In the case of any PC failure, the device will stop to measure and goes into a secure mode. All transmitted measurement data are saved in the BDF+<sup>3</sup> format on the PC side. BDF+ is the 24-bit version of the EDF+ [145] file format. The measured frequency response of the PHYSIOSENSE device including analog filters and digital FIR filters is shown in Figure 3.17. EMG and BI measurement channels show a flat amplitude response with small damping near the Nyquist frequency. The CMRR circuit which is used as an alternative to the reference driver for e.g. respiratory measurements has a slightly damped amplitude response (2.5 dB at 500 Hz).

<sup>3</sup>[http://www.biosemi.com/faq/file\\_format.htm](http://www.biosemi.com/faq/file_format.htm)



**Figure 3.17:** Measured frequency response of the PHYSIOSENSE device including digital FIR filters using an Agilent 33522A frequency generator. The measurement data were transmitted to the PC with a sampling frequency of 4 kHz. The frequency response of both BI channels was measured by applying a modulated sine wave with the corresponding carrier frequency of 50 kHz and 100 kHz. EMG and BI measurements have a flat amplitude response while using the reference driver which was connected to one input terminal (3 dB drop at 1.5 kHz). The activated CMRC leads to slightly decreasing amplitude response while the reference driver was disconnected (2.5 dB at 500 Hz).

## 3.9 Calibration

All measured data samples which are arriving at the PC are continuously transformed into physical units.

### 3.9.1 Voltage Calibration for EMG Measurement

All four EMG measurement channels are calibrated separately using an external reference voltage. For each channel  $Ch = \{1, 2, 3, 4\}$ , the ADC output  $out_{Ch}$  is stored in  $\overline{out}_{Ch}$  after applying a DC voltage of 145 mV. Then a DC voltage of -145 mV is applied and the ADC output is saved in  $\underline{out}_{Ch}$ .

The calibration is performed once and the stored parameters  $\overline{out}_{Ch}$  and  $\underline{out}_{Ch}$  are loaded in the beginning of each measurement. During a measurement, each sample  $out_{Ch}$  is converted to an EMG voltage signal  $EMG_{Ch}$  by

$$EMG_{Ch}^{norm} = (out_{Ch} - \underline{out}_{Ch}) / (\overline{out}_{Ch} - \underline{out}_{Ch}), \quad (3.13)$$

$$EMG_{Ch} = (EMG_{Ch}^{norm}(145 \text{ mV} + 145 \text{ mV})) - 145 \text{ mV}. \quad (3.14)$$

### 3.9.2 Impedance Calibration for BI Measurement

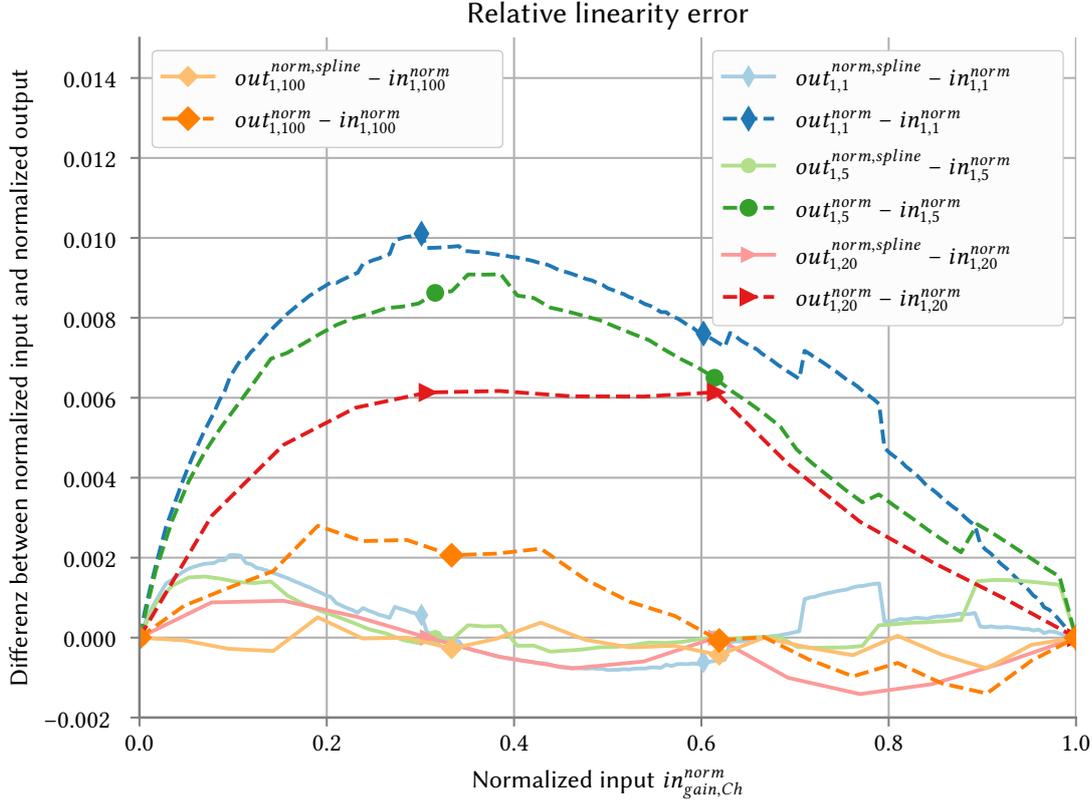


Figure 3.18: Impedance calibration of the BI measurement. Shown is the error between normalized input voltage  $in_{Ch,gain}^{norm}$  and normalized output voltage. The normalized output voltages  $out_{Ch,gain}^{norm}$  for the channel  $Ch$  and the  $gain$  is obtained by linear mapping. Using spline interpolation, the normalized output voltage  $out_{Ch,gain}^{norm,spline}$  can be obtained. The linearity error is smaller for the latter.

The current sources and the BI measurement channels are calibrated separately. First, both BI measurement channels  $Ch = \{1, 2\}$  are calibrated at each available gain setting  $gain = \{1, 5, 20, 100\}$  using an external reference sinusoidal signal with a frequency of 50 kHz for channel 1 and 100 kHz for channel 2. Please note that the second channel could also be calibrated for measurements at 50 kHz following the same procedure; for simplicity, this case is not described in the following.

The measured output  $out_{Ch,gain}$  from each BI channel  $Ch$  and available gain setting  $gain$  is stored while the sinusoidal voltage input with amplitude  $in_{Ch,gain}$  is applied. First, the minimally possible amplitude  $\underline{in}_{Ch,gain}$  and the maximal possible amplitude  $\overline{in}_{Ch,gain}$  is fed to the device input and the resulting outputs  $\underline{out}_{Ch,gain}$  and  $\overline{out}_{Ch,gain}$  are stored.

Based on these two measurements, a normalized output  $out_{Ch,gain}^{norm}$ , which is always in the interval between zero and one, can be calculated for the output  $out_{Ch,gain}$

$$out_{Ch,gain}^{norm} = (out_{Ch,gain} - \underline{out}_{Ch,gain}) / (\overline{out}_{Ch,gain} - \underline{out}_{Ch,gain}). \quad (3.15)$$

Additionally, the input  $in_{Ch,gain}$  is normalized by

$$in_{Ch,gain}^{norm} = (in_{Ch,gain} - \underline{in}_{Ch,gain}) / (\overline{in}_{Ch,gain} - \underline{in}_{Ch,gain}). \quad (3.16)$$

Due to the non-linear behavior of the full-wave rectifier (cf. Fig. 3.13), a maximum linearity error of 0.3 % to 1 % can be observed between  $in_{Ch,gain}^{norm}$  and  $out_{Ch,gain}^{norm}$  depending on the gain. This non-linearity yields a maximum absolute error of 10  $\Omega$  to 30  $\Omega$ , depending on the selected gain and current amplitude. As the measurement ranges for different gain settings are overlapping, such that the next range begins in the center of the previous range, this non-linearity behavior would also lead to jumps in the BI signal when the gain is changed although the BI of the tissue remains constant. E.g., a change of the gain from 1x to 20x at a load impedance of 30  $\Omega$  and a measurement current of 137  $\mu\text{A}$  would lead to a jump in the measured BI of 0.2  $\Omega$ . As the expected changes during a swallow lie in a range of 0.5  $\Omega$  to 3  $\Omega$ , the previously described non-linearity is not tolerable and must, therefore, be reduced, e.g., by using cubic spline interpolation. The linearity error for measuring 30  $\Omega$  with a gain of 20x and a current of 137  $\mu\text{A}$  is reduced from 0.2  $\Omega$  to almost zero ( $in_{20,1}^{norm} = 0.3$ ) when cubic spline interpolation is used. Figure 3.18 shows the linearity error of the first channel for all gain settings and the improvement by cubic spline interpolation.

For calculation of the cubic spline interpolation, measurements at only four specific input voltages  $in_{Ch,gain}$  for each gain and channel are needed. Together with the already measured outputs  $out_{Ch,gain}$  and  $\overline{out}_{Ch,gain}$  at the minimum and maximum input voltage amplitudes  $\underline{in}_{Ch,gain}$  and  $\overline{in}_{Ch,gain}$ , two measurements at input voltage amplitudes of  $1/3(\overline{in}_{Ch,gain} - \underline{in}_{Ch,gain})$  and  $2/3(\overline{in}_{Ch,gain} - \underline{in}_{Ch,gain})$  have to be performed. The results of these four measurements are collected in a vector  $\mathbf{out}_{Ch,gain} \in \mathbb{R}^4$  and the applied input amplitudes are stored into the vector  $\mathbf{in}_{Ch,gain} \in \mathbb{R}^4$ . Both vectors will be normalized by the following equations:

$$\mathbf{out}_{Ch,gain}^{norm} = \frac{\mathbf{out}_{Ch,gain} - \underline{out}_{Ch,gain}}{\overline{out}_{Ch,gain} - \underline{out}_{Ch,gain}}, \quad (3.17)$$

$$\mathbf{in}_{Ch,gain}^{norm} = \frac{\mathbf{in}_{Ch,gain} - \underline{in}_{Ch,gain}}{\overline{in}_{Ch,gain} - \underline{in}_{Ch,gain}}. \quad (3.18)$$

Four cubic polynomials are fitted such that  $\mathbf{out}_{Ch,gain}^{norm}$  is mapped to  $\mathbf{in}_{Ch,gain}^{norm}$  for each channel and gain separately:

$$f_{cubicspline}(\mathbf{out}_{Ch,gain}^{norm}, \boldsymbol{\theta}_{Ch,gain}) = \mathbf{in}_{Ch,gain}^{norm}, \quad (3.19)$$

where  $\boldsymbol{\theta}_{Ch,gain}$  is the parameter vector of the cubic spline interpolation which maps  $\mathbf{out}_{Ch,gain}^{norm}$  to  $\mathbf{in}_{Ch,gain}^{norm}$  such that Equation (3.19) is fulfilled.

Using cubic spline interpolation, the calibrated voltage amplitude  $v_{Ch,gain}^{calib}$  is obtained from the normalized measured output  $out_{Ch,gain}^{norm}$  as follows

$$out_{Ch,gain}^{norm,spline} = f_{cubicspline} \left( out_{Ch,gain}^{norm}, \theta_{Ch,gain} \right), \quad (3.20)$$

$$v_{Ch,gain}^{calib} = out_{Ch,gain}^{norm,spline} \left( \overline{in}_{Ch,gain} - \underline{in}_{gain,Ch} \right) + \underline{in}_{Ch,gain}, \quad (3.21)$$

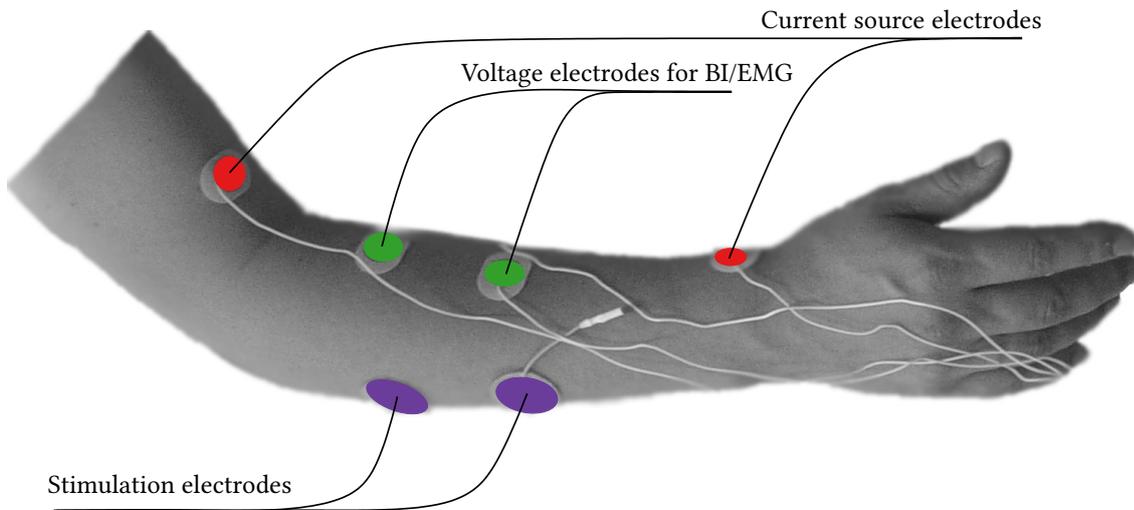
where  $out_{Ch,gain}^{norm,spline}$  is the corrected normalized output voltage using spline interpolation. It can be seen in Figure 3.18, that the difference between  $out_{Ch,gain}^{norm,spline}$  and  $in_{Ch,gain}^{norm}$  is zero at the supporting points 0, 1/3, 2/3, and 1. Between these points, the difference is almost zero which shows that the relation between  $out_{Ch,gain}^{norm,spline}$  and  $in_{Ch,gain}^{norm}$  is almost linear due to the cubic spline interpolation.

The resulting current  $I_{l,CS}$  of the current source  $CS = \{1, 2\}$  is measured across reference resistors using the two-electrode measurement method for each programmable current level  $l = 1, \dots, 8$ .

Finally, the BI amplitude  $BI_{Ch}$  can be calculated by dividing the calibrated voltage amplitude  $v_{Ch,gain}^{calib}$  by the applied current  $I_{l,CS}$ :

$$BI_{Ch} = \frac{v_{Ch,gain}^{calib}}{I_{l,CS}}. \quad (3.22)$$

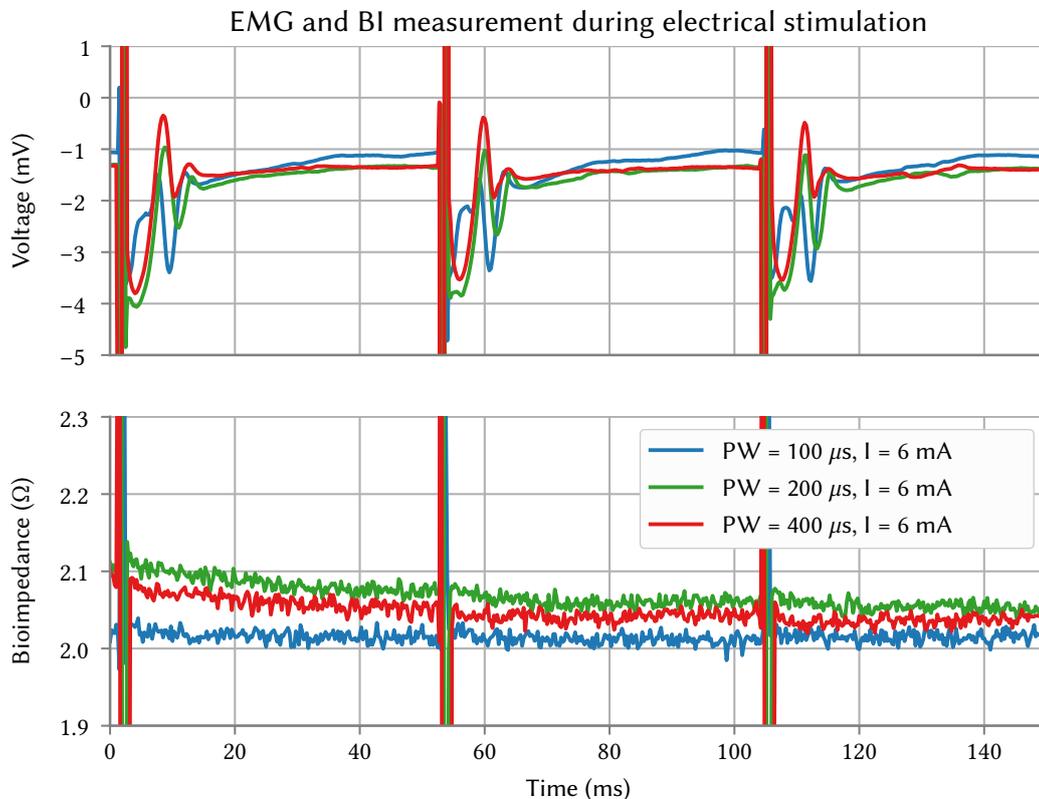
#### 3.9.3 BI and EMG Response on High Voltage Pulses



**Figure 3.19:** Electrode position for measuring BI/EMG on the forearm while FES is active. The arm picture is from [146].

The ability of the measurement system to recover after stimulation artifacts was investigated at the forearm of a healthy subject, on which higher stimulation intensities than on the neck can be applied. The electrode positions are similar to the experiment described in Nahrstaedt et al. [146], where BI was measured by the author with a predecessor measurement device while FES was applied to the forearm muscles. As shown in Figure 3.19, the current BI electrodes were attached to the skin near the wrist and above the elbow joint on the upper arm. The BI/EMG measurement electrodes were placed in a line between the BI-current electrodes. The hydro-gel electrodes for FES ( $\varnothing$  32 mm,

KRAUTH+TIMMERMANN GmbH, Germany) were placed on the musculus extensor carpi radialis longus/brevis parallel to the measurement electrodes.



**Figure 3.20:** EMG and BI measurement at the left forearm while a biphasic electrical stimulation with a frequency of 20 Hz was active. FES electrodes were placed at a distance of 3 cm to the BI/EMG measurement electrodes. In the EMG signal, the M-wave is visible shortly after the stimulation artifact disappears.

Figure 3.20 shows the BI/EMG response after stimulation pulses with a stimulation current amplitude of 6 mA were applied. The pulse width was varied from 100  $\mu$ s to 400  $\mu$ s. Due to fast recovery from stimulation artifacts, measurement of BI and EMG is possible after 2 ms to 3 ms. The EMG recordings show the M-wave (FES-evoked muscle activity). Volitional muscle activity was not present in this test, but could certainly be assessed from the EMG recording between the simulation pulses

### 3.10 Video Reference System

In order to compare the BI/EMG-measurements with endoscopic and videofluoroscopic recordings, a synchronous video stream needs to be saved. A USB video grabber (Grabby, TerraTec, Germany) is used to convert the analog video output into a digital video stream in real-time. A time stamp is immediately written into each frame which arrives on the PC. This time stamp is also saved together with the current time from the PHYSIOSENSE system.

The delay of the video frame conversion was measured. A square wave with a frequency of 1 Hz was generated and displayed on an oscilloscope and at the same time recorded with the `PHYSIOSENSE` device. Simultaneously, an analog camera was filming the screen of the oscilloscope. An analog camera was used, as the endoscopic and videofluoroscopic camera systems in the clinic are analog ones or have an analog video signal output. Slopes in the square wave were marked in the video and in the measured signal. The video stream was saved with a frame rate of 10 frames per seconds. An average delay of 500 ms was measured between the captured video stream and the electrically measured square wave. A jitter error of around 100 ms was measured for the time points in which the slope should be visible in the video frames. This jitter corresponds to the video frame rate of 10 Hz. By compensating the average delay of 500 ms for the recorded data, the video stream can be used as a reference system.

## 3.11 Conclusions

A measurement device was developed, which can be used for measurement of EMG and BI at the neck for the examination of swallowing. The device layout is highly versatile for answering several questions regarding BI measurements in the field of swallowing disorders. It also allows the investigation of controlled FES for supporting swallowing, as measurement is possible while FES is actively used.

Bipolar surface and needle EMG can be independently recorded at four channels in a range of  $\pm 150$  mV with a 21-bit effective resolution up to a frequency of 4 kHz. The device has two BI measurement channels by which BI can be measured either with the two-, three-, or four-electrode measurement method. Two controlled current sources are used for generating a sinusoidal current at a frequency of 50 kHz and 100 kHz, respectively. Both current sources are implemented as dual single-ended current source with guaranteed common mode reduction for the means of patient protection from DC components. As the current amplitude and the measurement gain are switchable, a measurement range from  $0.3 \Omega$  to  $30 \text{ k}\Omega$  could be realized with 19-bit effective ADC resolution. This wide range allows the usage of needle and transcutaneous electrodes. Separation of the EMG signal content from the BI signal could be achieved with analog filters and therefore simultaneous measurement of BI and EMG over the same voltage measurement electrodes is possible.

The implemented fault detection senses a low electrode-skin contact at the BI measurement and current electrodes. In case that a fault has been triggered, a better measurement range is automatically searched. The fault detection circuit is also used for finding the best current amplitude automatically and measurement gain settings individually for a subject at the beginning of each measurement. It has been shown in practice, that sensing bad electrode-skin contact is an important feature for avoiding faulty measurements. Possible reasons are that old electrodes with gel dried out, the skin was not prepared optimally, the skin becomes to sweaty, or measurement cables are broken.

An optional common mode reduction circuit can be activated individually for each voltage measurement channel. This circuit causes a slightly falling amplitude response but is useful for the case that the output voltage of sensors should be measured. Such sensors could be a hand switch or

sensors for measuring respiration (e.g. respiration belt). As they are not connected to the skin, the reference circuit is, contrary to the common mode reduction circuit, not able to reduce the CMRR for these measurement channels.

All measurement data were sent continuously in real-time to the PC and a synchronization with a USB based video-grabber is possible. The video-stream of a videofluoroscopy and of an endoscopy could successfully be synchronized with BI and EMG measurements.

The device is protected against high voltages. It is possible to measure EMG and BI shortly after a stimulation impulse was applied through transcutaneous electrodes in the same area. The measurement device can therefore be used for triggering supportive electrical stimulation.



# 4

## Automatic Segmentation and Classification of Swallowing

### 4.1 Summary

**Aim:** Investigation of a bioimpedance (BI)/electromyography (EMG) segmentation algorithm with subsequent classification for automatic detection of swallowing events in recorded data.

**Methods:** The BI signal is assumed to correlating to the movement of anatomical structures during the pharyngeal phase of swallowing. Muscle activity is a prerequisite for the laryngeal movement. A swallow results in a typical reproducible temporary drop in the BI amplitude with EMG activity in the begin of the valley. Algorithms for robust detection of EMG activity and finding valley like shapes in the approximated BI signal are presented. Only BI valleys that coincide with EMG activity are potential candidates for swallowing events and selected by this combined BI/EMG segmentation algorithm. Specific features are calculated from the BI and EMG signal for each valley. A support vector machine (SVM) classifier is used to decide if a valley represents a swallow or a different event, e.g., tongue, head, or neck movement artifacts.

**Results:** The SVM classifier has been trained and tested on data from 31 healthy subjects and 41 patients with oropharyngeal swallowing disorder. The dataset contained 3661 manually marked swallows of different bolus sizes and consistency and was affected by movements and speech. The combined BI/EMG segmentation algorithm detected 97.9 % of all swallowing events but included also 12656 non-swallowing related events. The subsequently applied classifier was tested by cross-validation and showed a sensitivity of 89.9 % and a specificity of 91.5 %.

**Conclusion:** The results show that EMG and BI recording at the neck can be used for swallowing segmentation. Even in the presence of disturbances such as speaking, head, tongue, or jaw movements, good sensitivity and specificity values could be achieved. Thus, long-term monitoring

of swallowing and swallowing assessment in an uncontrolled environment seem to be possible. The beginning and end point of the pharyngeal swallowing phase can be accurately determined.

**Contribution:** The author contributed the presented signal processing methods, the improvement of the EMG activity detector, the algorithms for piece-wise linear approximation of the BI signal, the BI valley detection algorithm, and the methods for feature extraction.

Parts of this chapter were published by the author in [132], [147], and [148].

## 4.2 Motivation

Systems with small and inexpensive sensors for automatic detection of swallowing and evaluation of swallowing performance in clinical and non-clinical settings would greatly improve diagnosis and therapy of swallowing disorders.

Two gold standards (videofluoroscopic swallowing study (VFSS) and fiber-optic endoscopic evaluation of swallowing (FEES)) are used for evaluation of dysphagia and mapping of dysphagia severeness into different scales is performed by the examiners. The outcome depends strongly on the experience of the examiner which is reflected in a moderate inter- and intra-rater agreement for both methods [13], [14]. An automatic swallowing evaluating tool could estimate several swallow related parameters, such as the level and speed of laryngeal elevation, and could enhance established clinical swallowing scales. The automatic detection and evaluation of swallowing events could play an important role in assisting inexperienced examiners during a VFSS or FEES.

First approaches used dual-axis swallowing accelerometry signals. The work by Lee et al. [74], [79], Merey et al. [80], and Steele et al. [81] utilized a classifier to determine the severeness of dysphagia and compared the results to the outcome of VFSS. Using an aspiration scale, an agreement of up to 90% with manual examination was achieved. However, in all studies a manual segmentation of clearly recognizable swallows was performed based on videofluoroscopy (VF) recordings.

Other accelerometry studies had put their focus on swallow segmentation. The studies from Sejdić et al. [75] and Damouras et al. [78] used accelerometry recordings of 295 swallows from healthy subjects. All disturbances as vocalizations, coughing, and excessive head movements were completely excluded from the data. In comparison to manual segmentation, the segmentation algorithm by Sejdić et al. [75] was able to recognize swallows with a sensitivity of 94.5% and a precision of 94%. Damouras et al. [78] developed an automatic real-time segmentation algorithm, which was able to detect swallows from the same dataset with 90% sensitivity and 93% precision. The algorithm used a window length of one second which limited the time resolution of the swallowing detector. The performances of both algorithms were also evaluated on a second dataset which contains 266 swallows and vocalization from 37 dysphagia patients [78]. The swallow segmentation algorithm from Sejdić et al. [75] resulted in a sensitivity of only 39% and a precision of 54%. A sensitivity of 87% and a precision of only 40% was achieved by the algorithm proposed by Damouras et al. [78].

In a study by Sazonov et al. [61], swallowing sounds were used as measurement signals. 70 measurement session from 20 healthy subjects containing resting periods and talking as well as meal intake were recorded. The measurement signal was divided into small fragments which were

called epochs of lengths from 0.375 s to 3 s. Optimal results were obtained by epoch lengths of 1.5 s. Swallowing was segmented with an average accuracy of 84.7 %. The weakness of this study is that the classifier was individually trained for each session and subject. In addition, the classifier had only a time resolution of 1.25 s. A higher time resolution of 0.25 s was achieved by Amft et al. [63] but with a lower accuracy (sensitivity of 68 % and a precision of 20 %).

Aboofazeli et al. [65] used hidden Markov models to detect the onset and end point of swallowing with a delay of less than 100 ms. The dataset contained swallowing measurements of 15 healthy children and nine young adult cerebral palsy patients. A false positive rate of only 6.7 % was achieved, whereas 4.8 % of swallows were missed by the detector. The effects of disturbances caused by head turning, speaking, or coughing were not investigated.

To the best knowledge of the author, nobody tried to apply automatic swallowing detection based on BI and EMG in order to determine onset and duration of the pharyngeal phase of swallow. Such a solution could be able to robustly detect swallows even in the presence of disturbances. Methods for signal preprocessing, swallowing segmentation, and finally swallowing classification based on the combined EMG and BI measurement at the neck are investigated in this chapter.

## 4.3 Methods

### 4.3.1 Signal Acquisition

The developed measuring system *PHYSIOSENSE*, which is described in detail in Chapter 3, is used for measuring BI and EMG at the neck. The four-electrode measurement method<sup>1</sup> at a frequency of 50 kHz is used for BI measurement. EMG is recorded at the BI voltage measuring electrodes while a reference electrode is also attached to the skin. A battery powered hand switch is connected to the second bipolar voltage measurement input. Its state is saved together with the BI and EMG signal during each session. The sampling frequency is set to 4 kHz for all measurement channels.

### 4.3.2 Study Procedure

The measurement electrodes are placed carefully at the neck by the examiner and BI and EMG are recorded during the examination. The electrodes of the current source are placed bilaterally on the upper onset of the sternocleidomastoid muscle below the mastoid insertion. The voltage measurement electrodes are placed laterally on the gap between hyoid bone and the thyroid cartilage symmetrically on both sides. Lastly, the reference electrode is attached to the cheek. The electrode positions are schematically visualized in Figure 4.1. Blue Sensor N-00-S ECG electrodes (Ambu A/S, Denmark) are used for all measurements.

<sup>1</sup>The four-electrode measurement method consists of a bipolar current source and a bipolar voltage measurement. The current source applies a constant sinusoidal current at a fixed frequency through one pair of electrodes (current electrodes) to the tissue. The resulting voltage is measured by additional voltage measurement electrodes and the resulting transfer impedance can be calculated.

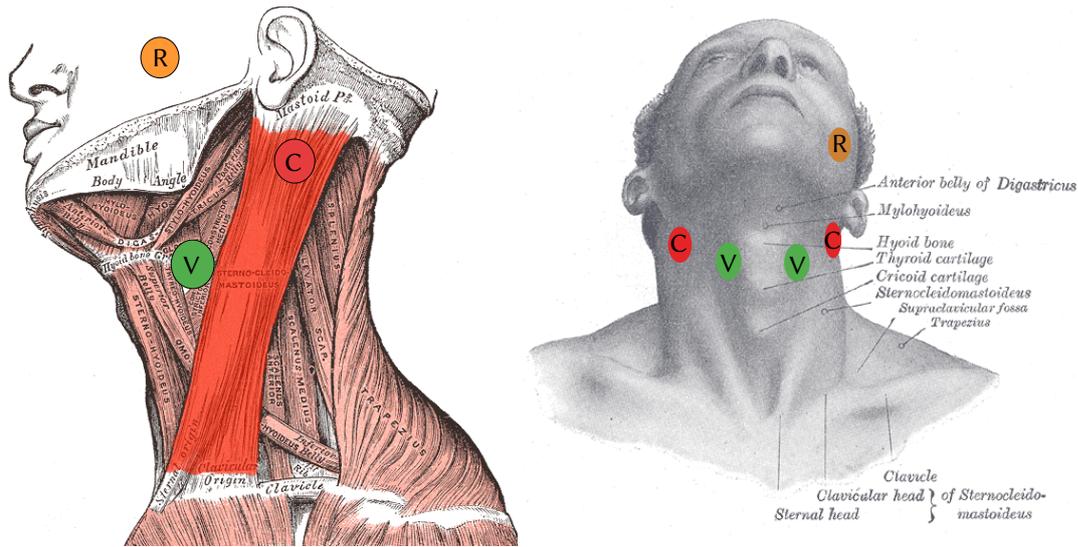


Figure 4.1: Electrode positions for measuring BI/EMG at the neck (C - current electrodes, V - voltage measurement electrodes, R - reference electrode). Anatomical illustrations are from [2].

### 4.3.3 Signal Preprocessing of EMG and BI

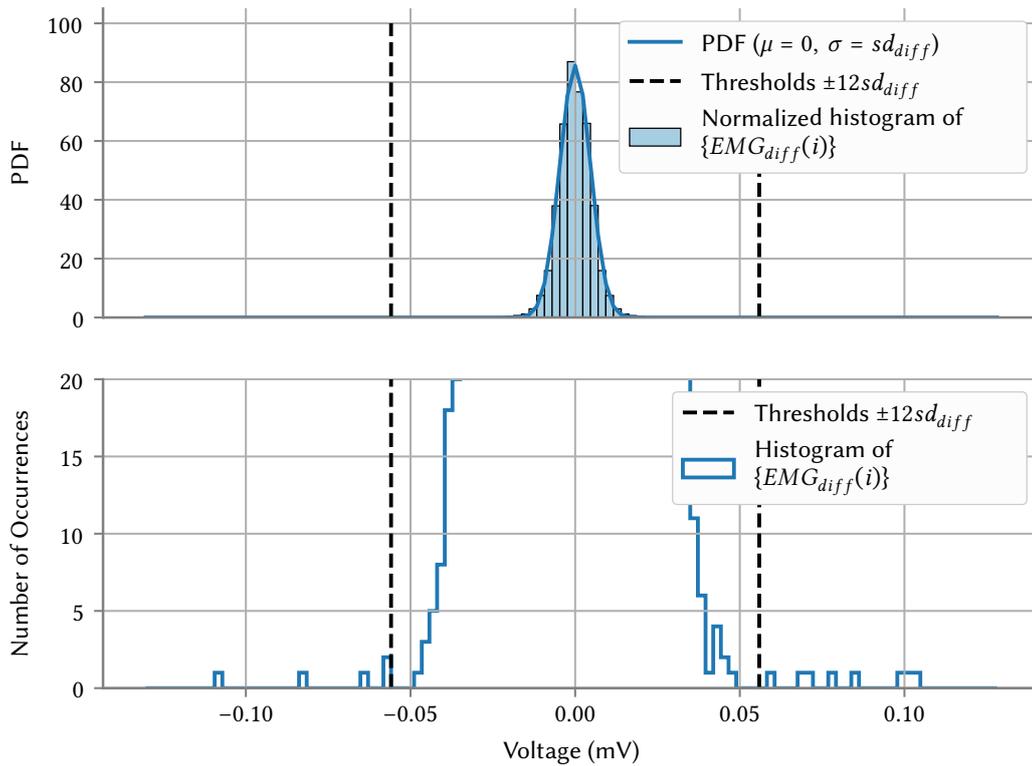
First, jumps and spikes in the EMG signal are detected and removed. This is performed by calculating the discrete difference  $\{EMG_{diff}(i)\}$  of the EMG time series by:

$$EMG_{diff}(i) = EMG(i) - EMG(i - 1), \quad i = 1, \dots, N, \quad (4.1)$$

where  $EMG(i)$  is an EMG sample at time instance  $i$  and  $N$  is the number of available samples. It is assumed that the sequence  $EMG_{diff}(i)$  is an almost stationary Gaussian process with zero mean. The sample standard deviation  $sd_{diff}$  is calculated from the data of the first five seconds from the sequence  $EMG_{diff}(i)$ . Figure 4.2 shows the histogram of an exemplary measurement together with the estimated Gaussian distribution.

A spike or jump in the original signal is detected if the corresponding value in the absolute difference sequence  $\{|EMG_{diff}(i)|\}$  is at least 12 times greater than the previously defined standard deviation  $sd_{diff}$ . The corresponding thresholds are visualized in Figure 4.2 as black dashed vertical lines. It is assumed that the sequence  $\{EMG_{diff}(i)\}$  follows a normal distribution with standard deviation  $sd_{diff}$ . The probability that a difference of two subsequent EMG samples exceeds the threshold is then  $2 \cdot 10^{-32}$ .

In order to decide if a spike or jump is present, the EMG value before the sequence  $\{EMG_{diff}(i)\}$  exceeds the threshold is stored. The length of the disturbance is estimated by finding the time position after the disturbance in which its EMG value passes the stored EMG value. If the EMG signal is going up or down after the disturbance such that it crosses the stored value within one second, a spike is assumed and all samples of the EMG signal within the spike are set to the stored value before the spike. Otherwise, a jump is assumed, and the height of the jump is subtracted from all subsequent EMG samples. The height of the jump is estimated by subtracting the mean EMG value before and after the disturbance. Removing of a spike is exemplarily shown in Figure 4.3.

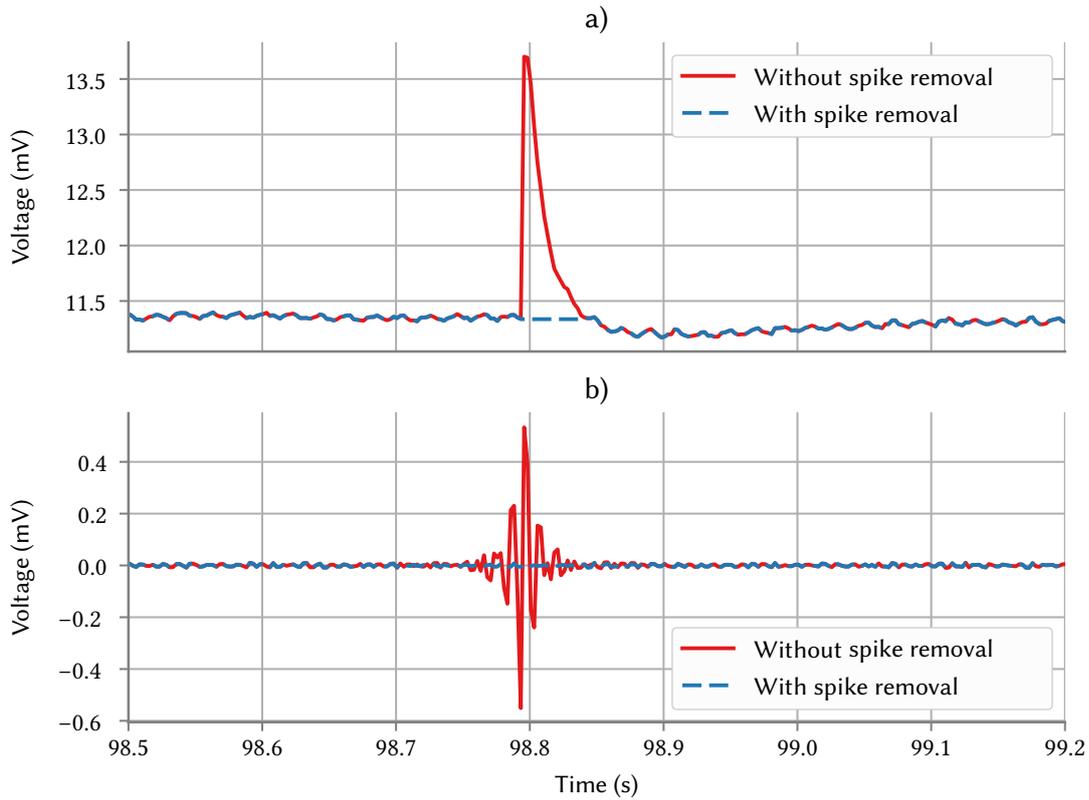


**Figure 4.2:** The histogram of the discrete difference sequence  $\{EMG_{diff}(i)\}$  from an EMG recording is shown together with the estimated probability density function (PDF) of a Gaussian distribution in the first sub-plot. The thresholds which separate spikes or jumps from normal EMG data are shown as black dashed lines. In the second sub-plot, spikes which cause occurrences greater than zero outside the thresholds can be seen in the unnormalized histogram.

In the next step, a non-causal elliptic infinite impulse response (IIR) high-pass filter of third order with the cut-off frequency of 10 Hz is applied to the EMG signal. This filter removes low frequent movement artifacts. The supply voltage causes disturbances at the first and third harmonics around 50 Hz and 150 Hz, respectively. Therefore, a non-causal elliptic IIR band-stop filter of third order with cut-off frequencies at 47 Hz and 53 Hz and a second similar band-stop filter with cut-off frequencies at 140 Hz and 160 Hz were additionally applied. All three filter are implemented as zero-phase filter realization, by which the signal is filtered forward and then backward.

The BI signal is mainly disturbed by noise. A heuristic comparison between low-pass filtering and wavelet denoising has shown that wavelet denoising can almost completely reduce noise without changing the signal and is therefore preferable. Wavelet denoising applies a signal transformation into wavelet coefficients. Adaptive shrinkage is performed on these coefficients [149]. The denoised signal is then found by reconstructing the signal from the wavelet coefficients.

A db4 wavelet is used for discrete wavelet transformation of the BI signal. Wavelets coefficients were calculated up to a level of eight. The noise variance is estimated from the first level of wavelet coefficients. The “minimax” threshold [149] multiplied with the estimated noise is then used as threshold for soft thresholding on all wavelet coefficients.

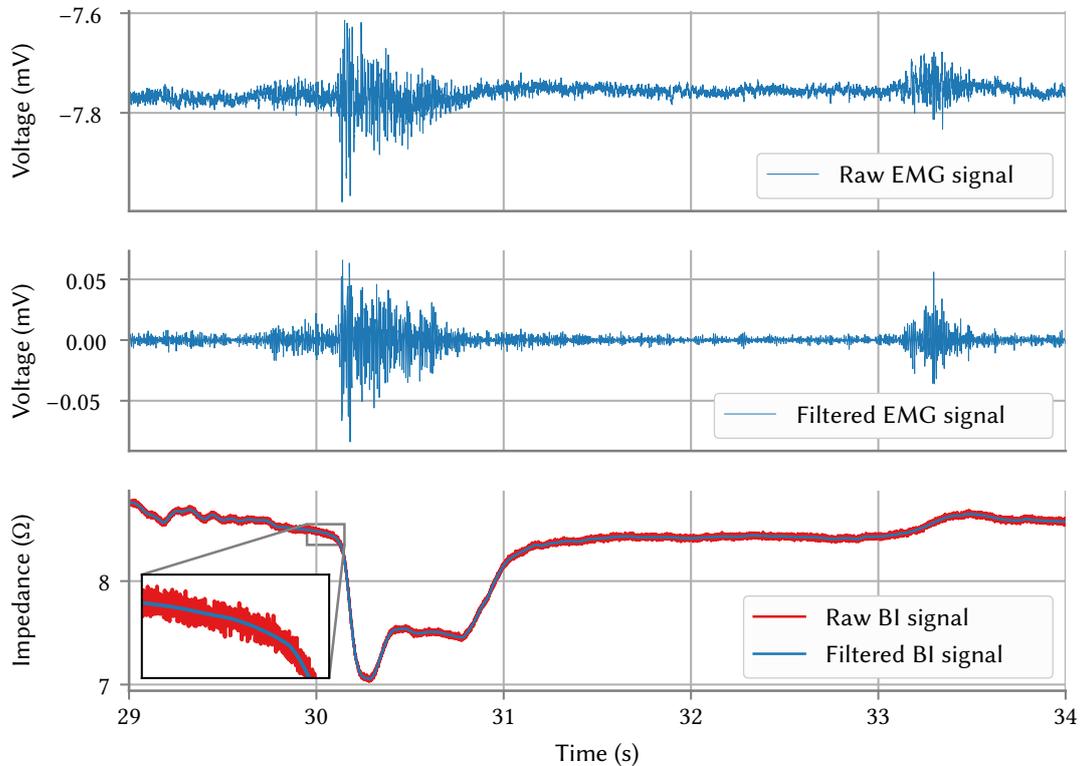


**Figure 4.3:** Unfiltered and filtered EMG signal with and without removal of spikes. In each subplot, two versions of the same signal are compared. The blue line shows the signal without applying a spike-removing algorithm whereas the red line is showing the signal with applied spike removal. In subplot a) the unfiltered EMG signals and in b) the filtered EMG signals are shown.

The coefficients of the wavelet transform are saved in a filter bank structure [150]. Each level represents a high-pass and a low-pass filtering step, where high-frequency coefficients are called detail coefficients and low-frequency coefficients are called approximation coefficients. Each filter output is down-sampled by two. Only the highest level of coefficients in the filter bank contains approximation and detail coefficients. All lower levels contain only approximation coefficients at the specific level. The approximation coefficients at level eight represent the signal content with frequencies up to  $f_n/2^8 = 7.81\text{ Hz}$ , where  $f_n = 2\text{ kHz}$  is the Nyquist frequency of the signal. The detail coefficients at level eight contain the signal content from  $f_n/2^8$  to  $f_n/2^7 = 15.62\text{ Hz}$ . The approximation coefficients of the next level ( $level = 7$ ) represent the frequency content of the signal from  $f_n/2^7$  to  $f_n/2^6 = 31.25\text{ Hz}$ . Finally, the approximation coefficients at level one represent the signal frequency content from  $f_n/2$  to  $f_n$ .

In order to be able to down-sample the BI signal to 250 Hz, all coefficients representing frequencies greater than 125 Hz will be set to zero. As the filter bank consists of the eight levels and the Nyquist frequency is 2 kHz, all approximation coefficients corresponding to a level of four or lower will be set to zero. Finally, the BI signal is reconstructed from all wavelet coefficients and down-sampled to 250 Hz by keeping only each eighth sample.

Exemplary, Figure 4.4 shows the unprocessed EMG and BI signals in comparison to the preprocessed signals.



**Figure 4.4:** EMG and BI recording of a water swallow from a healthy subject before and after the preprocessing. The figure is based on Fig. 2 from [132].

#### 4.3.4 Segmentation of EMG Activity Periods

##### 4.3.4.1 Introduction to the Double-Threshold Detector

Periods of muscle activation are detected by using a double-threshold detector with automatic threshold adaption based on the estimated noise and disturbance levels within the EMG signal. This algorithm uses an auxiliary sequence  $\{z(i)\}$  in which the EMG activity is detected. The auxiliary sequence is calculated by

$$z(i) = x(i)^2, \quad (4.2)$$

where  $x(i)$  is the preprocessed and whitened measurement sample at sampling instance  $i$  with a sampling time  $T_s^{EMG}$ . Whitening is a signal processing step in which non-stochastic components were removed. The Stulen-De Luca filter [151] is used as model for myoelectric activity:

$$H_{sf}(s) = \frac{ks(2\pi f_h)^2}{(s + \rho 2\pi f_h)(s + 2\pi f_h)^2}, \quad (4.3)$$

where the scaling factor is  $k = 7$ , the high cut-off frequency is set to  $f_h = 120$  and the low cut-off frequency  $f_l = \rho f_h$  is defined by  $\rho = 2/3$ . The Stulen-De Luca filter is z-transformed and is inverted in order to be used as whitening filter for EMG. A further pole at zero is added in order to make the filter causal.

The resulting amplitude response of the inverted filter is damped for low frequencies by a 2-order IIR high-pass filter  $H_{HP}$  with a cut-off frequency of 40 Hz as the inverted filter has a strongly increasing amplitude response for low and high frequencies. The large gain for frequencies near the Nyquist frequency is reduced by a 2-order IIR low-pass filter  $H_{LP}$  with a cut-off frequency of 700 Hz. The resulting whitening filter is

$$H_{whitening}(z) = g \frac{H_{sf}^{-1}(z)}{z} H_{LP}(z) H_{HP}(z), \quad (4.4)$$

where the gain  $g$  of  $H_{whitening}(z)$  is set to  $g = 1.8$  such that a variance of Gaussian noise is preserved, after applying the Stulen-De Luca filter and the whitening filter.

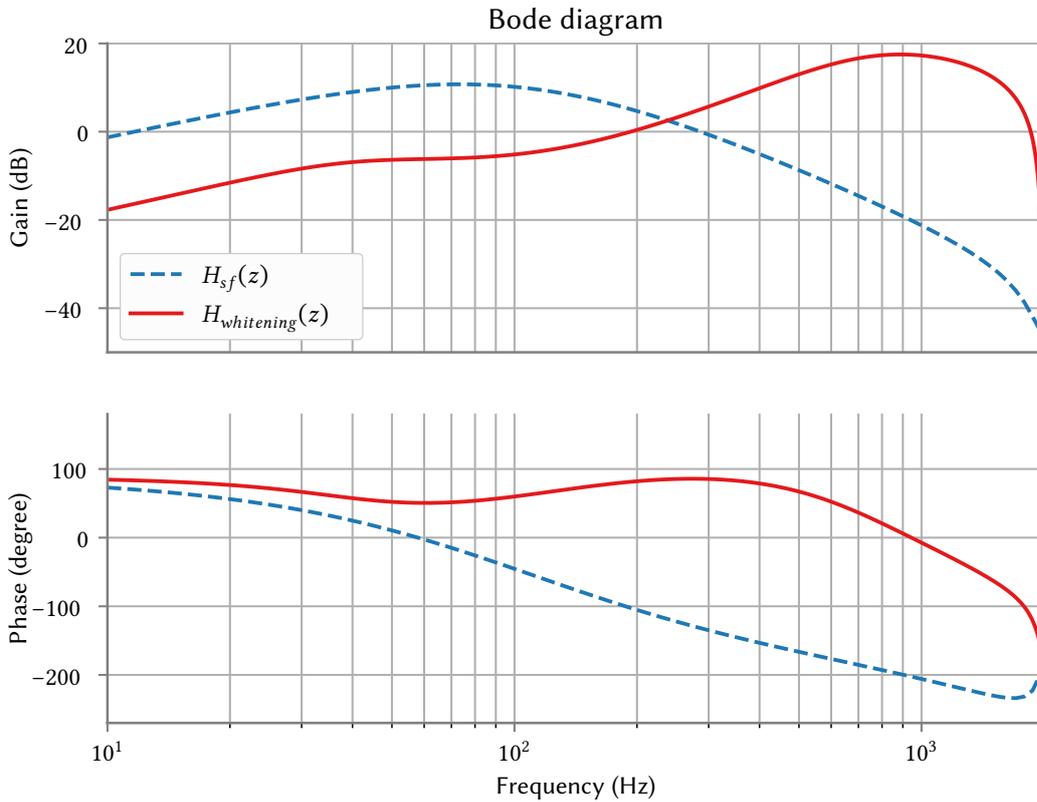


Figure 4.5: Frequency response of the whitening filter  $H_{whitening}(z)$  in comparison to the Stulen-De Luca filter  $H_{sf}(z)$ .

It is assumed that preprocessed and whitened surface EMG can be modeled as Gaussian process. Thus, elements of the auxiliary sequence  $\{z(i)\}$  are equal or greater than zero and the sequence can be modeled by a  $\chi^2$  (chi-squared) distribution with one degree of freedom. The knowledge about the distribution of  $\{z(i)\}$  is used for optimally estimating parameters of the double-threshold detector.

EMG activity is detected for sampling instance  $i$  for all  $N$  samples by

$$act_{\zeta}(i) = \sum_{k=i-m+1}^i \begin{cases} 1, & \text{if } z(i) > \zeta \\ 0, & \text{otherwise} \end{cases}, \quad \zeta \in \mathbb{R}^+, \quad m \in \mathbb{N}^*, \quad i = m, \dots, N, \quad (4.5)$$

$$act_{r_0}(i) = \begin{cases} 1, & \text{if } act_{\zeta}(i) \geq r_0 \\ 0, & \text{otherwise} \end{cases}, \quad r_0 \in \mathbb{N}^*, \quad r_0 \leq m, \quad i = m, \dots, N. \quad (4.6)$$

The number of samples from the auxiliary sequence  $\{z(i)\}$  which exceed  $\zeta$  (the first threshold) within the sliding window of  $m$  samples are counted and assigned to  $act_{\zeta}$  at the sample index  $i$  which is located at the endpoint of the sliding window (cf. Eq. (4.6)). When  $r_0$  (denoted as second threshold) samples from a sliding window of the auxiliary sequence  $\{z(i)\}$  were above  $\zeta$ , EMG activity is assumed and the variable  $act_{r_0}(i)$  is set to one instead of zero.

After onset of a period of continuous EMG activity in which all samples of  $\{z(i)\}$  are above threshold  $\zeta$ ,  $act_{r_0}$  toggles to one within a delay of  $r_0$  samples. At the end of such a continuous EMG activity period,  $act_{r_0}$  goes back to zero with a delay of  $m - r_0 + 1$  samples. As an exact onset time identification is important for the subsequent swallow segmentation, the detected activity will be shifted  $r_0 - 1$  samples to the left such that the detected onset is exactly at the begin of the activity:

$$act(i) = act_{r_0}(i - r_0 + 1), \quad i = m + r_0 - 1, \dots, N + r_0 - 1. \quad (4.7)$$

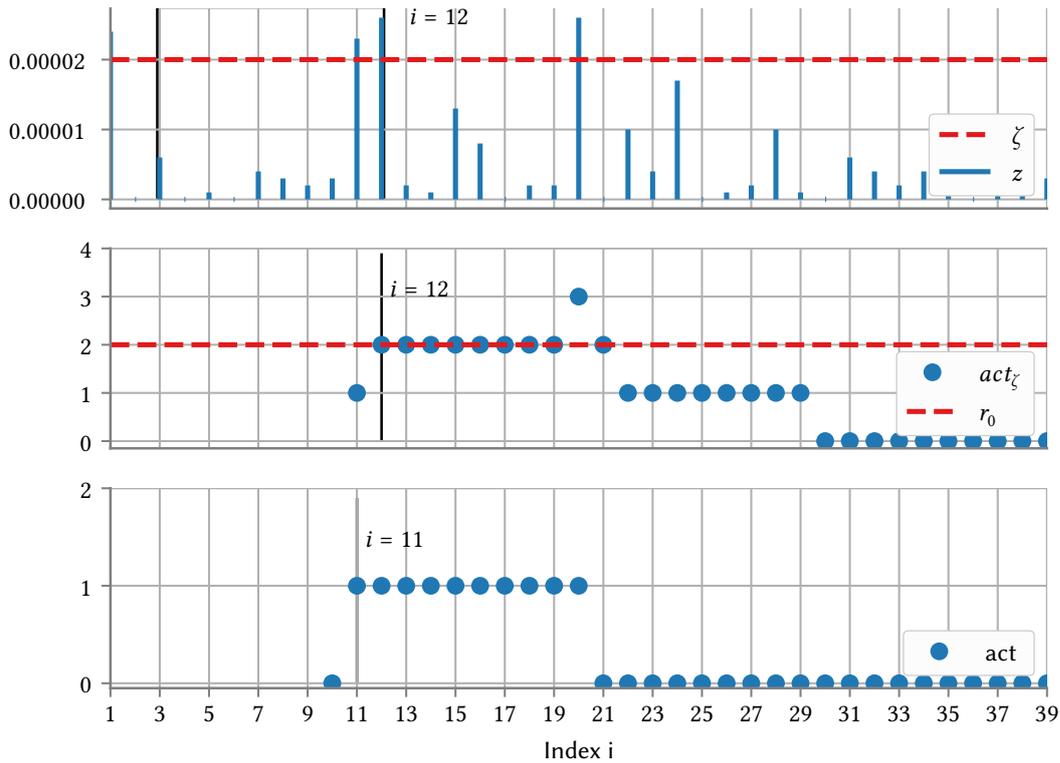
An example for detecting EMG activity is presented in Figure 4.6. The window length of the double-threshold detector is exemplarily set to  $m = 10$  and the second threshold is set to  $r_0 = 2$ .

The performance of the detector is defined by the false-alarm probability  $P_{fa}$ , which defines how likely Gaussian noise and disturbances are misclassified as EMG activity. The detection probability  $P_d$  describes how likely real EMG activity will be correctly classified as EMG activity. The ratio between  $P_{fa}$  and  $P_d$  for a given signal-to-noise ratio (SNR) can be improved by increasing the window length  $m$  or by increasing the degree of freedom of the  $\chi^2$  distribution by building the auxiliary sequence with more than one EMG sample.

In literature, two different approaches to create the auxiliary sequence  $\{z(i)\}$  were proposed. Bonato et al. [152] tried to improve the relation between  $P_{fa}$  and  $P_d$  by forming the auxiliary sequence from the sum of two squared successive, whitened EMG samples by which the degree of freedom of the  $\chi^2$  distribution is increased to two. This detector was compared with other EMG onset detectors by Staude et al. [153] with good results (99.9 % detected onsets and a mean detection delay of 4 ms).

Bonato et al. [152] had calculated the threshold  $\zeta$  depending on the estimated noise variance from a manually selected signal window and a given false-alarm probability. They found out, that the sliding window with length  $m$  should be as large as possible in order to increase the detection probability. However, a larger value decreases the time resolution of the detector. Bonato et al. [152] limited the window length to  $m = 5$  for having a time resolution of 10 ms. They had shown that  $r_0 = 1$  for  $m = 5$  gives the best relation between detection probability and false-alarm probability.

An improved version of the double-threshold detector from Bonato et al. was presented by Severini et al. [154]. The noise variance was estimated by sample variance calculation from the

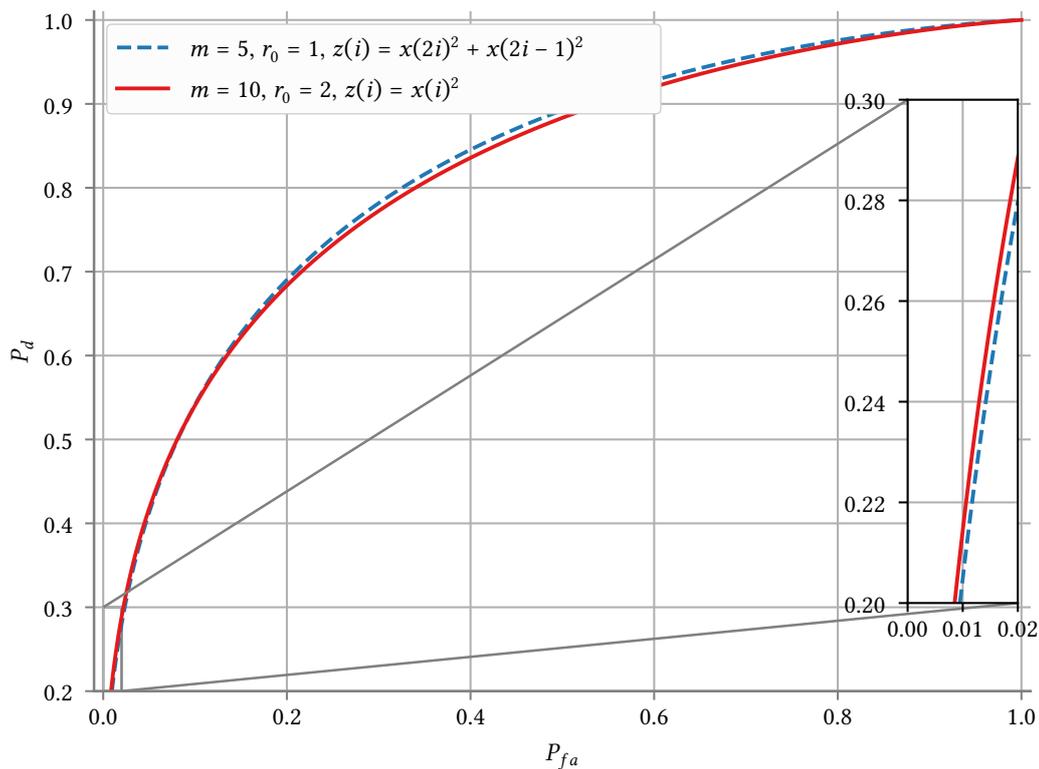


**Figure 4.6:** Demonstration of the double-threshold detector on 40 measurement samples from which the auxiliary sequence  $\{z(i)\}$  is plotted in the upper subplot. The sliding window of length  $m = 10$  is shown at sample position  $i = 12$  in which two samples exceed the threshold  $\zeta$ . In the middle subplot, the number of samples which exceed  $\zeta$  are plotted. At which samples the double-threshold detector detects EMG activity is shown in the third subplot. The output of the detector is shifted by  $r_0 - 1$  samples to the left, as this is the latest possible position of activity. In this example, samples at position eleven and twelve are above  $\zeta$ . When the window slides to position  $i = 12$ , the second threshold  $r_0 = 2$  is reached and an onset is detected at position eleven.

samples between two EMG bursts. An EMG burst is an interval in which the double-threshold detector detects EMG activity for all samples within. This approach has the disadvantage that the output of the detector is used for triggering a new noise variance estimation. If the noise variance is initialized wrongly, convergence to the real noise variance may never happen or is at least slow.

The onset detector by Xu et al. [155] build the auxiliary sequence from the squared sum of independent EMG trials. These trials were taken from several EMG recordings of different muscles from the leg during running on a treadmill. The recordings were synchronized using a food switch. By taking the sum of several squared and aligned EMG signals, the detector accuracy could be improved and a whitening filter could be omitted, but is only applicable on cyclic activities such as walking, running, or cycling. The auxiliary sequence with one degree of freedom corresponds to the auxiliary sequence  $\{z(i)\}$  used within this work.

In both studies (Bonato et al. [152] and Xu et al. [155]), the sampling frequency of the EMG signal was limited to 1 kHz by the experimental setup. A time resolution of the detector of 10 ms was demanded. This resulted in a specific window length  $m = 5$ . In Figure 4.7, the curves which are



**Figure 4.7:** Receiver operating characteristic (ROC) curve of the double-threshold detector presented by Bonato et al. [152] ( $m = 5$ ) in comparison to the ROC curve of the approach which is used within this thesis ( $m = 10$ ). The SNR was set to 3 dB.

spanned by  $P_{fa}$  and  $P_d$  are plotted for the approach by Bonato et al. [152] ( $m = 5$ ) and the approach which is used in this thesis ( $m = 10$ , one degree of freedom). The performance of the detector by Bonato et al. [152] is slightly worse for  $P_{fa} < 0.1$  than the onset detector presented in this work with  $m = 10$  and  $r_0 = 2$ . Thus, summing up two neighbored samples in order to increase the degree of freedom of the  $\chi^2$  distribution does not increase the performance for low  $P_{fa}$  values which are desired. Higher sampling frequencies than 1 kHz would allow larger sliding windows in term of samples without decreasing the time resolution of the detector.

The  $\chi^2$  distribution of the auxiliary sequence  $\{z(i)\}$  presented in this work has one degree of freedom. In order to still obtain acceptable values for  $P_{fa}$  and  $P_d$ , a higher sampling frequency for EMG will be applied that allows sliding windows of short time duration and large sample number  $m$ . In the following, an iterative method for tuning  $m$ ,  $r_0$  and  $\zeta$  of the double-threshold detector based on the statistical properties will be presented. Finally, results of benchmarks will be shown.

#### 4.3.4.2 Statistical Properties of the Double-Threshold Detector

In the following it is assumed that the preprocessed and whitened EMG signal  $x(i)$  at sample instance  $i$  can be modeled as a zero-mean Gaussian process.  $x(i)$  consists of an independent Gaussian noise

$n(i) \sim \mathcal{N}(0, \sigma_n^2)$ , an independent non-stationary Gaussian disturbance  $d(i) \sim \mathcal{N}(0, \sigma_d^2(i))$  which models components that are not related to EMG activity with the disturbance variance  $\sigma_d^2(i) = [0, \bar{\sigma}_d^2]$  for all  $i$  and a non-stationary Gaussian process  $s(i) \sim \mathcal{N}(0, \sigma_s^2(i))$  which correlates with the real EMG signal.  $s(i)$  reflects muscle activity with signal variance  $\sigma_s^2(i) = [0, \tilde{\sigma}_s^2(i)]$  for all  $i$ , where  $0 < \underline{\sigma}_s^2 \leq \tilde{\sigma}_s^2(i) \leq \bar{\sigma}_s^2$  and  $\underline{\sigma}_s^2$  is the smallest EMG activity that shall be detected.

The following assumptions are made: The signal  $x(i)$  contains at least one continuous sequence of at least  $N_{\tilde{\sigma}_n^2}$  samples with  $x(i) \sim \mathcal{N}(0, \sigma_n^2)$ , i.e. this sequence contains only noise without any disturbances ( $\sigma_d^2(i) = 0$ ) or EMG activity ( $\sigma_s^2(i) = 0$ ). Furthermore,  $x(i)$  contains at least  $th_{\bar{\sigma}_d} \cdot N$  samples (not required to be successive) with  $x(i) \sim \mathcal{N}(0, \sigma_n^2 + \bar{\sigma}_d^2)$ , where  $N$  is the number of samples of  $x(i)$  and  $0 < th_{\bar{\sigma}_d} < 1$ . The parameters  $N_{\tilde{\sigma}_n^2}$  and  $th_{\bar{\sigma}_d}$  are important for estimating  $\sigma_n^2$  and  $\bar{\sigma}_d^2$  and explained later in detail. The SNR from noise and disturbances to EMG activity is

$$SNR(\tilde{\sigma}_s^2) = 10 \log_{10} \left( \frac{\tilde{\sigma}_s^2(i)}{\bar{\sigma}_d^2 + \sigma_n^2} \right). \quad (4.8)$$

The worst-case SNR is  $SNR(\underline{\sigma}_s^2) = SNR_{min}$  for which  $SNR_{min} > 0$  must be hold in order to detect EMG activity sufficiently. The worst-case SNR  $SNR_{min}$  is set by observation for the specific application.

As  $n(i)$ ,  $d(i)$ , and  $s(i)$  are independent random variables that are normally distributed,  $x(i)$  is also normally distributed with  $x(i) \sim \mathcal{N}(0, \sigma_n^2 + \sigma_d^2(i) + \sigma_s^2(i)) = \mathcal{N}(0, \sigma_i^2)$ . For a more compact presentation,  $x_i = x(i)$  and  $z_i = z(i)$  will be used in the following. Given the assumed Gaussian distribution  $\mathcal{N}(0, \sigma_i^2)$  at sampling instance  $i$ , the non-stationary probability density function of  $x_i$  is defined as:

$$f(x_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{x_i^2}{2\sigma_i^2}}. \quad (4.9)$$

The probability density function  $\tilde{f}$  for the previously defined auxiliary time series  $\{z(i)\}$  (cf. Eq. (4.2)) is:

$$\tilde{f}(z_i, \sigma_i^2) = \frac{f_{\chi^2, \nu=1}(z_i/\sigma_i^2)}{\sigma_i^2} = \frac{\left(\frac{z_i}{\sigma_i^2}\right)^{1/2-1} e^{-\frac{z_i}{2\sigma_i^2}}}{\sigma_i^2 2^{1/2} \Gamma(1/2)}, \quad (4.10)$$

where  $f_{\chi^2, \nu=1}$  is the probability density function of a  $\chi^2$  distribution (based on Gaussian distribution  $\mathcal{N}(0, 1)$ ) with one degree of freedom ( $\nu = 1$ ) and  $\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx$  is the gamma function.  $\tilde{f}$  can now be used for calculating the probabilities that  $z_i$  exceeds threshold  $\zeta$ .

---

<sup>2</sup> $\mathcal{N}(\mu, \sigma^2)$  this is a short notation for a Gaussian distribution with mean value  $\mu$  and variance  $\sigma^2$ .

The probability  $P_{z_i > \zeta}$  that a specific signal sample exceeds the first threshold  $\zeta$  is calculated as follows:

$$P_{z_i > \zeta} = P[z_i > \zeta, x_i \sim \mathcal{N}(0, \sigma_i^2)], \quad (4.11)$$

$$P_{z_i > \zeta} = \int_{\zeta}^{\infty} \tilde{f}(z, \sigma_i^2) dz, \quad (4.12)$$

$$P_{z_i > \zeta} = 1 - \int_0^{\zeta} \tilde{f}(z, \sigma_i^2) dz. \quad (4.13)$$

$$(4.14)$$

Then, the probability density function (4.10) can be inserted into Equation (4.14) yielding

$$P_{z_i > \zeta} = 1 - \int_0^{\zeta} \frac{\left(\frac{z}{\sigma_i^2}\right)^{1/2-1} e^{-\frac{z}{2\sigma_i^2}}}{\sigma_i^2 2^{1/2} \Gamma(1/2)} dz. \quad (4.15)$$

The subsequent variable substitution  $y = \frac{z}{\sigma_i^2}$  results in

$$P_{z_i > \zeta} = 1 - \int_0^{\frac{\zeta}{\sigma_i^2}} \frac{y^{1/2-1} e^{-\frac{y}{2}}}{2^{1/2} \Gamma(1/2)} dy. \quad (4.16)$$

By using the lower incomplete gamma function with  $\gamma(a, x) = \int_0^x t^{a-1} e^{-t} dt$ , the Equation (4.16) can be rewritten as

$$P_{z_i > \zeta} = 1 - \frac{1}{\Gamma(1/2)} \gamma\left(1/2, \frac{\zeta}{2\sigma_i^2}\right). \quad (4.17)$$

The terms of (4.17) can be easily calculated using standard numerical software packages. In order to solve Equation (4.17) for  $\zeta$ , Equation (4.16) will be further reformulated. Using the  $\chi^2$  probability density function  $f_{\chi^2, \nu=1}(y) = \frac{y^{1/2-1} e^{-\frac{y}{2}}}{2^{1/2} \Gamma(1/2)}$ , we can write (4.16) as

$$P_{z_i > \zeta} = 1 - \underbrace{\int_0^{\frac{\zeta}{\sigma_i^2}} f_{\chi^2, \nu=1}(y) dy}_{F_{\chi^2, \nu=1}\left(\frac{\zeta}{\sigma_i^2}\right)}. \quad (4.18)$$

The integral in (4.18) represents the cumulative distribution function  $F_{\chi^2, \nu=1}$ .

In the next step, the second threshold  $r_0$  is included into the probability analysis.  $P_r$  is the probability that  $r$  samples out of  $m$  are above the first threshold  $\zeta$  and can be calculated by using repetitions of Bernoulli trials:

$$P_r = \binom{m}{r} P_{z_i > \zeta}^r (1 - P_{z_i > \zeta})^{m-r}. \quad (4.19)$$

The probability  $P_{r \geq r_0}$  that at least  $r_0$  samples out of  $m$  are above the threshold  $\zeta$  can be calculated by the following sum:

$$P_{r \geq r_0} = \sum_{k=r_0}^m \binom{m}{k} P_{z_i > \zeta}^k (1 - P_{z_i > \zeta})^{m-k}. \quad (4.20)$$

The parameters of the double-threshold detector are optimized for a worst-case scenario where disturbances are always present and where only EMG activity at the lowest intensity ( $\underline{\sigma}_s^2$ ) is present.

The false-alarm probability  $P_{fa}$  is defined as the probability that the detector wrongly classifies  $x(i)$  with noise and disturbances ( $\sigma_d^2(i) = \bar{\sigma}_d^2$ ) and without any EMG activity ( $\sigma_s^2(i) = 0$ ) as EMG activity:

$$P_{fa} = P_{r \geq r_0} \text{ for } x(j) \sim \mathcal{N}(0, \sigma_n^2 + \bar{\sigma}_d^2), \quad j = i - m + 1, \dots, i. \quad (4.21)$$

The probability  $P_\zeta$  defines the probability that a sample exceeds the threshold  $\zeta$  for the same distribution of  $x(i)$ :

$$P_\zeta = P_{z_j > \zeta} \text{ for } x(j) \sim \mathcal{N}(0, \sigma_n^2 + \bar{\sigma}_d^2). \quad (4.22)$$

The detection probability  $P_d$  describes the lower bound of the probability that the time series at instant  $i$  with  $\sigma_s^2(i) = \underline{\sigma}_s^2 > 0$ ,  $\sigma_d^2(i) = \bar{\sigma}_d^2$  and noise is correctly classified as EMG activity:

$$P_d = P_{r \geq r_0} \text{ for } x(j) \sim \mathcal{N}(0, \sigma_n^2 + \bar{\sigma}_d^2 + \underline{\sigma}_s^2), \quad j = i - m + 1, \dots, i. \quad (4.23)$$

As mentioned earlier,  $\underline{\sigma}_s^2$  is the lowest intensity of all EMG activities that shall be detected. For the same distribution of  $x(i)$ , the probability that a sample exceeds the threshold  $\zeta$  is named  $P_{dk}$ :

$$P_{dk} = P_{z_j > \zeta} \text{ for } x(j) \sim \mathcal{N}(0, \sigma_n^2 + \bar{\sigma}_d^2 + \underline{\sigma}_s^2). \quad (4.24)$$

It is assumed that  $\mathcal{N}(0, \sigma_i^2)$  is constant within the sliding window for the instances  $j = i - m + 1, \dots, i$  in order to be able to calculate the relation between  $P_{r \geq r_0}$  and  $P_{z_j > \zeta}$ .

For a given probability  $P_\zeta$ , the threshold  $\zeta$  can be calculated depending on the noise variance  $\sigma_n^2$  and the disturbance variance  $\bar{\sigma}_d^2$  by rewriting Equation (4.18):

$$\zeta = (\bar{\sigma}_d^2 + \sigma_n^2) F_{\chi^2, v=1}^{-1}(1 - P_\zeta), \quad (4.25)$$

where  $F_{\chi^2, v=1}^{-1}$  is the inverse  $\chi^2$  cumulative distribution function with one degree of freedom.

According to Equation (4.17), the integral for calculating  $P_{dk}$  can be written as

$$P_{dk} = 1 - \frac{1}{\Gamma(1/2)} \gamma \left( 1/2, \frac{(\bar{\sigma}_d^2 + \sigma_n^2) F_{\chi^2, v=1}^{-1}(1 - P_\zeta)}{2(\underline{\sigma}_s^2 + \bar{\sigma}_d^2 + \sigma_n^2)} \right), \quad (4.26)$$

$$= 1 - \frac{1}{\Gamma(1/2)} \gamma \left( 1/2, \frac{F_{\chi^2, v=1}^{-1}(1 - P_\zeta)}{2(1 + 10^{SNR(\underline{\sigma}_s^2)/10})} \right), \quad (4.27)$$

where  $SNR(\underline{\sigma}_s^2) \geq SNR_{min}$  is defined as the ratio of the minimal EMG signal variance to the noise and disturbance variances (cf. Eq. (4.8)).

#### 4.3.4.3 Method for Tuning the Double-Threshold Detector

The following procedure is applied to optimally tune the double-threshold detector:

1. For a predefined  $SNR_{min}$  value, a given false-alarm probability  $P_{fa}$  and a desired offset-detection latency  $tr_{max}$ , values for  $m$  and  $r_0$  are calculated such that the detection probability  $P_d$  at  $P_{fa}$  is maximally.
2. Measurement data is divided into windows of length  $N_{\hat{\sigma}_n^2}$  from which their specific variances are calculated. The smallest variance is taken as estimation of the noise variance  $\sigma_n$  and named  $\hat{\sigma}_n$ .
3. A value  $\hat{\zeta}$  for the first threshold of the detector that yields a desired  $P_{fa}$  also in the presence of disturbances is determined in the next step. Since the disturbance variance  $\bar{\sigma}_d^2$  is not known a-prioro, an iterative data-based procedure is performed. Starting from a calculated initial  $\hat{\zeta}_{\bar{\sigma}_d^2=0}$  (assuming no disturbance and taking into account the estimated noise variance  $\hat{\sigma}_n$ ), the threshold  $\hat{\zeta}$  will be stepwise increased until the desired  $P_{fa}$  is achieved. The  $P_{fa}$  is determined as follows: First, windows of step 2 that have a variance lower than the  $th_{\bar{\sigma}_d^2}$  percentile of all the variance values are selected. Then,  $N_{\hat{\zeta}}/m$  windows of length  $m$  are randomly taken from the selected windows and the data are shuffled. The double-threshold detector is applied. The windows which have more than  $r_0$  samples over  $\hat{\zeta}$  are counted and the false-alarm probability  $P_{fa}$  is estimated for the given  $\hat{\zeta}$ .
4. In a last step, the SNR values are determined and it is checked if the assumption of an minimal SNR value  $SNR_{min}$  was justified. The SNR is calculated based on the estimated noise and the estimated disturbance variance which is calculated from  $\hat{\zeta}$ . In case of a negative result, the complete dataset will be rejected.

#### 4.3.4.4 Optimization of the Parameters $r_0$ and $m$

For finding optimal values for parameter  $r_0$  and for  $m$ , a new approach is presented. An iterative procedure is used to find values for  $m$  and  $r_0$  such that  $P_d$  is maximized at a desired  $P_{fa}$  for a given worst case  $SNR_{min}$  and a desired maximal allowed offset-detection latency  $tr_{max}$  of the detector. A feasible offset-detection latency of the double-threshold detector for detecting EMG activity is 10 ms as proposed by Bonato et al. [152]. The detector output  $act(i)$  is shifted such that the onset of a continuous EMG activity is correctly estimated without latency. Thus, the onset-offset transition is detected in the worst case  $m - 2r_0 + 1$  time instances too late. The worst-case offset-detection latency  $tr$  of the detector is defined as the maximum time difference between the true onset-offset transition time and the detected onset-offset transition time for a period of continuous activity. The window length  $m$  and the threshold  $r_0$  must fulfill

$$m - 2r_0 + 1 \leq \frac{tr}{T_s^{EMG}} \leq \frac{tr_{max}}{T_s^{EMG}}, \quad tr \bmod T_s^{EMG} = 0, \quad (4.28)$$

where the largest possible window length  $m$  is preferred as this maximizes the detection probability  $P_d$ . Optimization of  $r_0$  and  $m$  includes the following steps:

#### 4. Automatic Segmentation and Classification of Swallowing

1. Use of an initial value of  $m = \frac{tr_{max}}{T_s^{EMG}} + 1$  and  $r_0 = 1$  which fulfills condition from Eq. (4.28).
2. Calculate  $P_d$  at the given  $P_{fa}$  for  $r_0 = 1, \dots, m$  using Equations (4.27) and (4.20). Select  $r_0$  that gives the largest detection probability  $P_d$ .
3. Use new  $m = \frac{tr_{max}}{T_s^{EMG}} + 2r_0 - 1$  (cf. Eq. (4.28)), and go to step 2 until  $r_0$  convergences.

Optimally results for  $m$  and  $r_0$  are obtained in a few iteration steps and can be stored for later use.

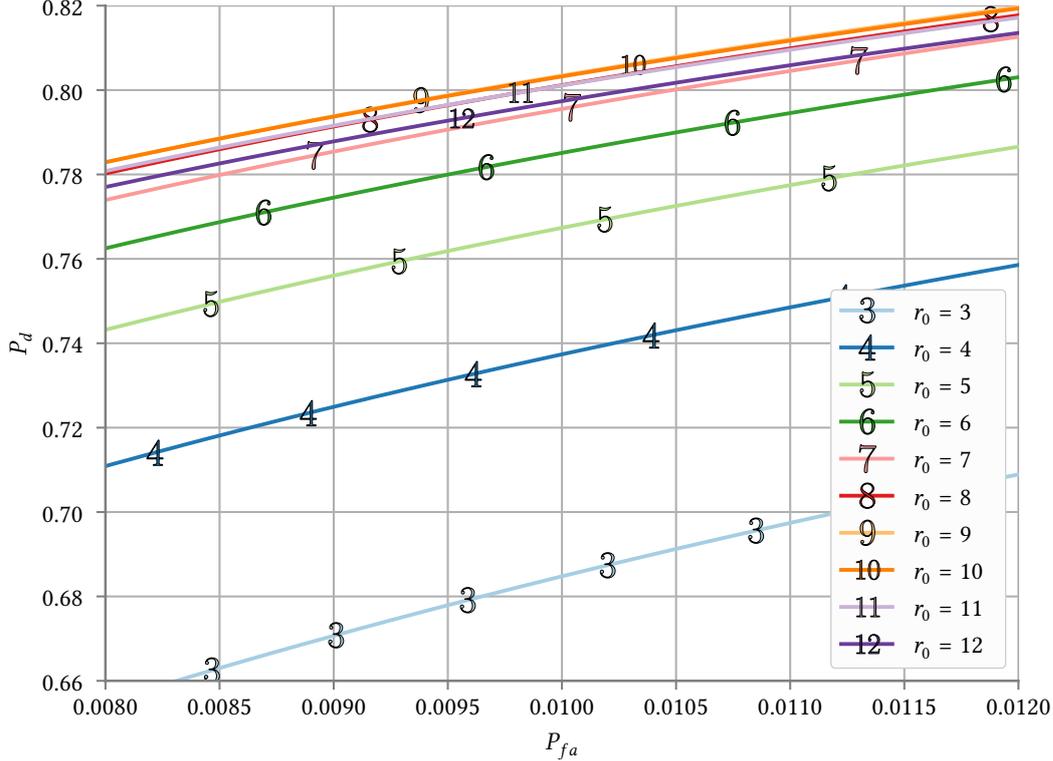


Figure 4.8: ROC of the double-threshold detector for different values of  $r_0$  and a window length of  $m = 60$  and an SNR of 3 dB. Optimization of  $r_0$  is performed by selecting the value which gives the highest detection probability for a given false-alarm probability  $P_{fa} = 0.01$ . The highest detection probability  $P_d$  is achieved for  $r_0 = 9$ .

Figure 4.8 shows exemplarily the selection of  $r_0$  for a fixed window length  $m = 60$  and  $SNR_{min} = 3$  dB. A value of  $r_0 = 9$  maximizes  $P_d$  at  $P_{fa} = 0.01$  and is therefore selected as second threshold  $r_0$ .

##### 4.3.4.5 Determination of an Optimal Threshold $\hat{\zeta}$

In order to calculate an optimal threshold  $\hat{\zeta}$ , the true noise variance  $\sigma_n^2$  is estimated by  $\hat{\sigma}_n^2$  from the measurement data. The noise variance is estimated by finding sufficient small signal parts of length  $N_{\hat{\sigma}_n^2}$  where the disturbance variances is zero ( $\bar{\sigma}_d^2 = 0$ ). In order to automatically estimate the noise variance  $\hat{\sigma}_n^2$ , the complete time series  $\{x(i)\}$  with length  $N$  is reshaped into a matrix

$$\mathbf{X}(i, j) = x((i-1)N_{\hat{\sigma}_n^2} + j), \quad j = 1, \dots, N_{\hat{\sigma}_n^2}, \quad i = 1, \dots, N/N_{\hat{\sigma}_n^2}, \quad N \bmod N_{\hat{\sigma}_n^2} = 0, \quad (4.29)$$

where the column size  $N_{\hat{\sigma}_n^2}$  should be large enough in order to allow a robust variance estimation and small enough in order to find signal parts without EMG activity and disturbances. The variance of each row  $i$  is calculated by:

$$\text{VAR}_x(i) = \frac{1}{N_{\hat{\sigma}_n^2} - 1} \sum_{k=1}^{N_{\hat{\sigma}_n^2}} \left( X(i, k) - \frac{1}{N_{\hat{\sigma}_n^2}} \sum_{l=1}^{N_{\hat{\sigma}_n^2}} X(i, l) \right)^2, \quad i = 1, \dots, N/N_{\hat{\sigma}_n^2}. \quad (4.30)$$

It is assumed that at least the variance of one row  $i$  is almost equal to the true noise variance. The variance  $\hat{\sigma}_n^2$  can now be estimated by taking the smallest variance value from all rows:

$$\hat{\sigma}_n^2 = \min \text{VAR}_x(i), \quad i = 1, \dots, N/N_{\hat{\sigma}_n^2}. \quad (4.31)$$

The disturbance variance  $\bar{\sigma}_d^2$  is not known a-priori. In order to estimate the threshold  $\zeta$  an iterative data-based procedure is performed by which a value  $\hat{\zeta}$  for the first threshold of the detector is determined such that the desired false-alarm probability  $P_{fa}$  can be observed in the presence of disturbances. Finally, a disturbance variance  $\hat{\sigma}_d^2$  which is the estimation of the true variance  $\bar{\sigma}_d^2$  can be calculated from  $\hat{\zeta}$ .

Firstly, the estimated noise variance  $\hat{\sigma}_n^2$  is used for calculating an initial value for the threshold

$$\hat{\zeta}_{\bar{\sigma}_d^2=0} = \hat{\sigma}_n^2 F_{\chi^2, \nu=1}^{-1} (1 - P_{\zeta}). \quad (4.32)$$

Then, the  $P_{fa}$  is determined by randomly taking parts of the measurement data and counting windows of length  $m$  which have more than  $r_0$  samples over  $\hat{\zeta}$ . Beginning from the initial value  $\hat{\zeta}_{\bar{\sigma}_d^2=0}$ ,  $\hat{\zeta}$  is increased until the measured  $P_{fa}$  is close to the given one.

In order to be able to select parts of the measurement data which contains only noise and disturbances, a ratio between  $\text{VAR}_x(i)$  and the estimated noise variance  $\hat{\sigma}_n^2$  is defined as

$$R_{\hat{\zeta}}(i) = 10 \log_{10} \left( \frac{\text{VAR}_x(i)}{\hat{\sigma}_n^2} \right). \quad (4.33)$$

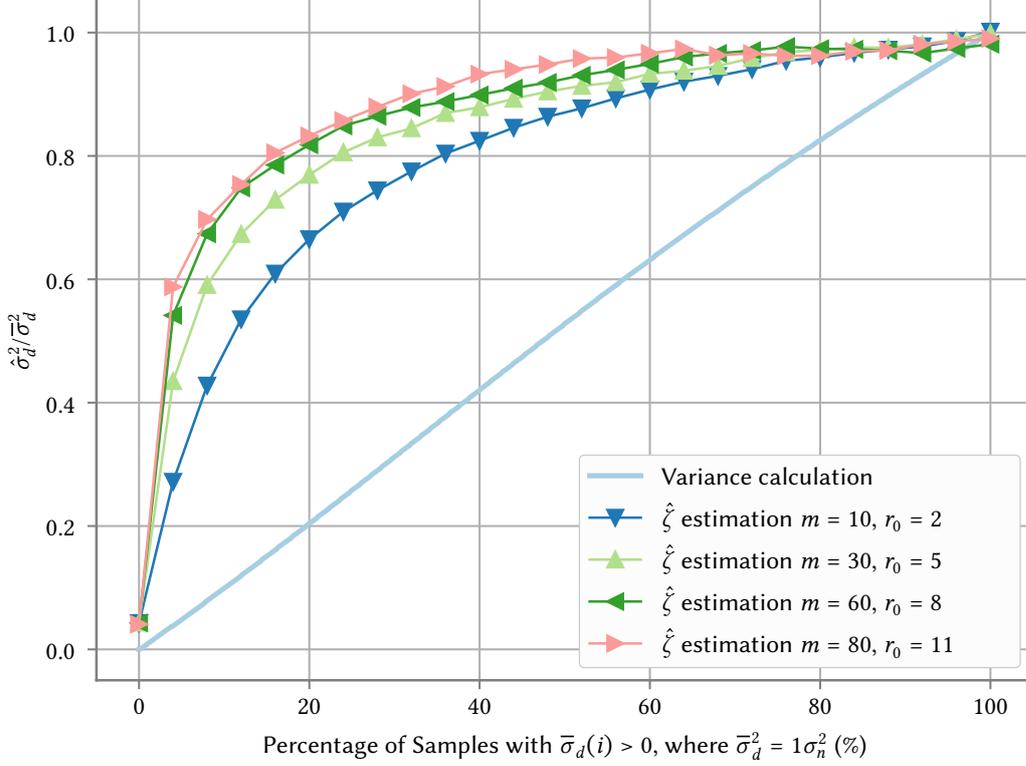
A percentage share  $th_{\bar{\sigma}_d}$  of signal parts with disturbances and without real EMG activity is defined a-priori based on the observed activity. Parts of the measurement data with a variance  $\text{VAR}_x(i)$  smaller than a threshold  $R_{\hat{\zeta}}^{max}$  which is derived from  $th_{\bar{\sigma}_d}$  are then assumed to contain not any real EMG activity. The threshold  $R_{\hat{\zeta}}^{max}$  is set to the corrected  $th_{\bar{\sigma}_d}$  percentile of all  $R_{\hat{\zeta}}$ :

$$R_{\hat{\zeta}}^{max} = \text{percentile} \left( R_{\hat{\zeta}}, th_{\bar{\sigma}_d} \right) + \text{ppt} \left( 0.999, 0, \frac{e^{-0.5}}{2\sqrt{2}} \right) - \text{ppt} \left( th_{\bar{\sigma}_d}, 0, \frac{e^{-0.5}}{2\sqrt{2}} \right), \quad (4.34)$$

where  $\text{percentile}(R_{\hat{\zeta}}, th_{\bar{\sigma}_d})$  gives percentile of vector  $R_{\hat{\zeta}}$  at  $th_{\bar{\sigma}_d}$  and  $\text{ppt}(p, \mu, \sigma)$  is the percent point function for a given probability  $p$  of a Gaussian distribution with mean  $\mu$  and standard deviation  $\sigma$ <sup>3</sup> [156].

<sup>3</sup>The standard deviation which is used in both percent point functions increases  $\text{percentile}(R_{\hat{\zeta}}, th_{\bar{\sigma}_d})$  in a way that almost all samples from the Gaussian process which describes noise are below  $R_{\hat{\zeta}}^{max}$  and are used for estimating the disturbance variance.

Randomly taken rows of  $\mathbf{X}$  that have a ratio  $R_{\hat{\zeta}}(i)$  smaller than  $R_{\hat{\zeta}}^{max}$  are merged into a new matrix  $\mathbf{X}_{\hat{\zeta}}$  until the number of elements is  $N_{\hat{\zeta}}$ . By taking not all possible rows, computation time of estimating  $\hat{\zeta}$  is significantly reduced.



**Figure 4.9:** On a time series containing 10000 Gaussian samples from which one part has a variance  $\sigma_n^2$  and the other part has a variance of  $\sigma_d^2 + \sigma_n^2$ , the disturbance variance  $\sigma_d^2$  should be estimated when  $\sigma_n^2$  is already known. It can be seen that estimation of  $\sigma_d^2$  depends on the window length  $m$  and the number of disturbance samples in the signal. The proposed  $\hat{\zeta}$  estimation leads to better results in comparison to direct variance calculation.

Beginning from the initial threshold  $\hat{\zeta}_{\bar{\sigma}_d^2=0}$ ,  $\hat{\zeta}$  is step-wised increased until the estimated false-alarm probability  $\hat{P}_{fa}$  is at least  $P_{fa}$  or a predefined threshold of  $\hat{\zeta}_{\bar{\sigma}_d^2=0}(1 + \kappa_{max})$  is reached:

$$\hat{\zeta} = \hat{\zeta}_{\bar{\sigma}_d^2=0}(1 + \kappa), \quad \kappa = 0, \kappa_{\delta}, 2\kappa_{\delta}, \dots, \kappa_{max}. \quad (4.35)$$

$\kappa_{max}$  and  $\kappa_{\delta}$  have to be chosen accordingly such that the desired  $P_{fa}$  can be achieved by applying  $\hat{\zeta}$  to the double-threshold detector. In practice, a value for  $\kappa_{max}$  of around 25 has proven to be sufficient.  $\kappa_{\delta}$  is set to a value which results in around 500 iterations which gives an accurate estimation of  $\hat{\zeta}$  in recorded EMG data during swallowing. For each iteration step, the double-threshold detector is applied on each row of  $\mathbf{X}_{\hat{\zeta}}$ . The number of rows where more than  $r_0$  samples are above  $\hat{\zeta}$  are saved

into  $onsets_{\hat{\zeta}}$ . The false-alarm probability  $\hat{P}_{fa}$  is calculated by:

$$\hat{P}_{fa} = \frac{onsets_{\hat{\zeta}} \cdot m}{N_{\hat{\zeta}}}, \quad \frac{m}{N_{\hat{\zeta}}} < P_{fa}. \quad (4.36)$$

If  $\hat{P}_{fa}$  is less or equal to the desired false-alarm probability  $P_{fa}$ , the current value of  $\hat{\zeta}$  is then taken as optimal value for the first threshold  $\zeta$ . Otherwise,  $\hat{\zeta}$  is increased by  $\kappa_{\delta}$  and the described steps are repeated.

Finally, the disturbance variance  $\hat{\sigma}_d^2$  can be calculated from the optimal threshold  $\hat{\zeta}$ :

$$\hat{\sigma}_d^2 = \frac{\hat{\zeta}}{F_{\chi^2, v=1}^{-1}(1 - P_{\zeta})} - \hat{\sigma}_n^2. \quad (4.37)$$

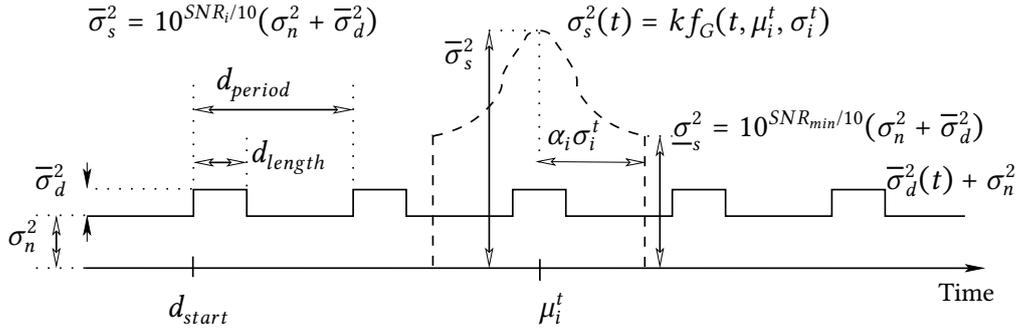
Figure 4.9 shows the performance of the presented  $\hat{\zeta}$ -estimation in comparison to variance calculation of data samples. A Gaussian disturbance  $\bar{\sigma}_d^2(i)$  with  $\bar{\sigma}_d^2 = \sigma_n^2$  is added to a noise vector, which contains 20000 samples with Gaussian noise (noise variance is  $\sigma_n^2$ ). The percentage of the 20000 data samples on which  $\bar{\sigma}_d^2(i) > 0$  is varied between 0 % to 100 %. The disturbance variance  $\hat{\sigma}_d^2$  is calculated from the estimated threshold  $\hat{\zeta}$  according to Equation (4.37). The threshold  $\hat{\zeta}$  is chosen such that the resulting  $\hat{P}_{fa}$  is at least  $P_{fa}$  which is set to 1%. The estimated  $\hat{\sigma}_d^2$  values are plotted against the percentage on which  $\bar{\sigma}_d^2(i) > 0$ .

The procedure is repeated for different window length from  $m = 10$  to  $m = 80$  for which the corresponding  $r_0$  is set to the particular optimal value. It can be seen that  $\hat{\sigma}_d^2$  increases nonlinear with a greater slope in the beginning. Thus,  $\hat{\sigma}_d^2$  can be robustly estimated although only parts of the signal are influenced by disturbances. It can be seen that the presented  $\hat{\zeta}$ -estimation is slightly dependent on the window length  $m$  and provides a better estimation for greater window length  $m$ . For e.g. when only 20 % of disturbances are present, a  $\hat{\sigma}_d^2$  value which correspond to  $0.8\bar{\sigma}_d^2$  is estimated for window lengths  $m \geq 60$ . The presented  $\hat{\zeta}$  estimation can estimate disturbances more accurately than a direct variance calculation, which estimates only a disturbance variance corresponding to  $0.2\bar{\sigma}_d^2$  when only 20 % of disturbances are present.

#### 4.3.4.6 Performance Benchmark Tests

The double-threshold detector benchmark from Severini et al. [154] is applied which consists of several trials with different noise, EMG, and disturbance signal realizations. The presented double-threshold detector is applied on each trial and its ability to estimate noise and disturbances is investigated.

Each trial of the benchmark consists of 60000 Gaussian noise samples modeling noise, disturbances and EMG activity. The length of each trial corresponds to 20 s as sampling time  $T_s^{EMG} = 1/4000$  s is used. It is shown in Figure 4.10 how the variances of noise, disturbances and EMG vary over time. The disturbance sequence with  $\sigma_d^2(t) > 0$  starts at time point  $d_{start}$  which is randomly chosen for each trial in a range from 0.5 s to 1 s. The time distance between two Gaussian disturbance periods is set to  $d_{period}$  such that 75 disturbance periods fit into one trial. The length  $d_{length}$  of a disturbance is set randomly for each trial to a value in a range from 27 ms to 60 ms. The



**Figure 4.10:** Exemplary variance profile of a trial. Each trail consists of Gaussian noise with variance  $\sigma_n^2$ , Gaussian disturbances  $\bar{\sigma}_d^2(t)$  and EMG bursts  $\sigma_s^2(t)$ . Each trial contains 75 Gaussian disturbances. The resulting variance for each disturbance is  $\sigma_n^2 + \bar{\sigma}_d^2$ . EMG activity is simulated by bursts of Gaussian shape. The minimal SNR of each bursts is  $SNR_{min}$ , whereas the maximal SNR varies randomly from 6 to 12 dB. The SNR values are calculated related to noise and disturbances.

total length of all disturbance periods corresponds to 10 % to 30 % of the complete trial duration. The disturbance variance  $\bar{\sigma}_d^2$  is set to  $1\sigma_n^2$ ,  $2\sigma_n^2$ , and  $3\sigma_n^2$ .

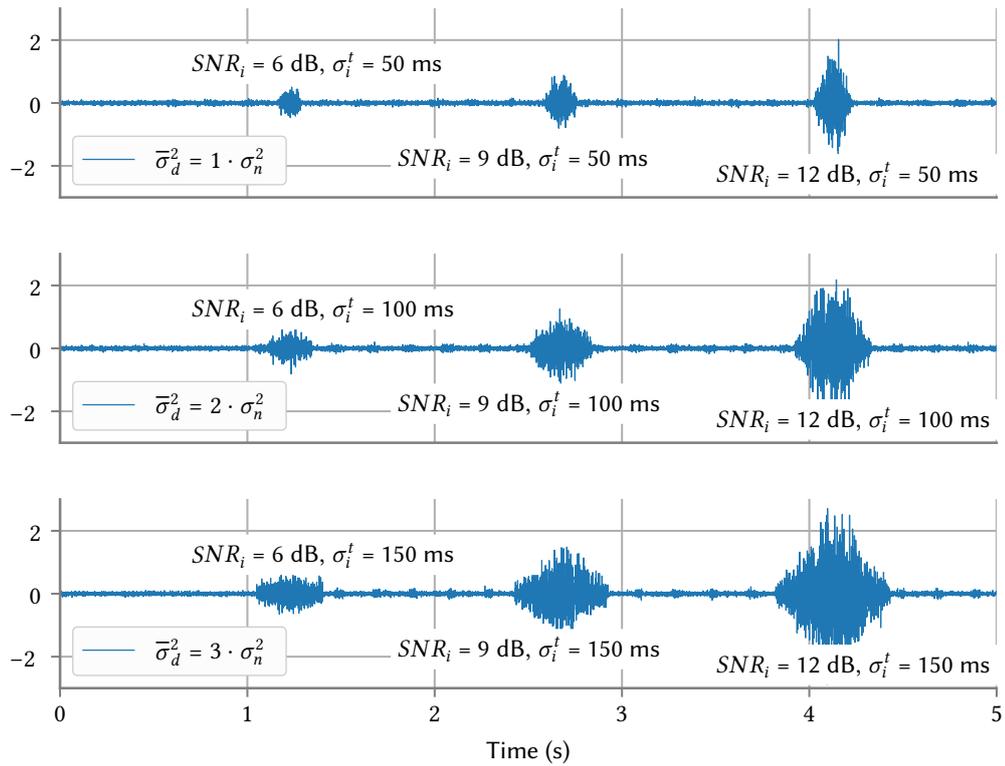
In each of the resulting 3000 trials, ten Gaussian shaped bursts of EMG activity with different SNR values have to be detected by the double-threshold detector. Each burst  $i$  has its peak at the time point  $\mu_i^t$  and a total duration of  $2\alpha_i\sigma_i^t$  seconds. The shape is modulated by a Gaussian function

$$f_G(t_i, \mu_i^t, \sigma_i^t) = \frac{e^{-\left(\frac{t_i - \mu_i^t}{2\sigma_i^t}\right)^2}}{\sigma_i^t \sqrt{2\pi}}, \quad t_i = \mu_i^t - \alpha_i\sigma_i^t, \dots, \mu_i^t + \alpha_i\sigma_i^t, \quad (4.38)$$

where the time support  $t_i$  has a sampling resolution of  $T_s^{EMG}$ . The center  $\mu_i^t$  is chosen such that ten bursts are produced within the trial length of 20 s. Each burst  $i$  is multiplied by a scaling factor  $k$  which is chosen such that  $\underline{\sigma}_s^2 = k f_G(\mu_i^t - \alpha_i\sigma_i^t, \mu_i^t, \sigma_i^t)$  corresponds to a SNR of  $SNR_{min}$ . The SNR is calculated in relation to the sum of the noise variance  $\sigma_n^2$  and the disturbance variance  $\bar{\sigma}_d^2$ . The amplitude of the peak at  $\mu_i^t$  is also influenced by  $\alpha_i$  which is set to a value such that the peak correspond to a SNR value  $SNR_i$ . The maximal SNR  $SNR_i$  is randomly set to 6 dB, 9 dB, or 12 dB. Additionally,  $\sigma_i^t$  is set randomly to 50 ms, 100 ms, or 150 ms. The resulting width of a burst lays in a range from 0.119 s to 0.611 s for all possible combination of  $SNR_i$  and  $\sigma_i^t$ .

It is assumed that noise, disturbances, and EMG bursts have their sources inside the tissue. According to Stulen et al. [151], the spectrum of a EMG signal will be shaped by recording it through surface electrodes. Shaping of the signal spectrum is simulated by filtering the complete sequence by the Stulen-De Luca filter (see Eq. (4.3)). Then, the double-threshold detector is applied to the signal after applying the whitening filter, estimating noise variance, and finally finding the best value for the threshold  $\hat{\zeta}$  (cf. Section 4.3.4.5). This procedure is repeated for each trial.

The results of the benchmark tests are shown in Table 4.1. For  $tr_{max} = 0.01$  s and  $P_{fa} = 0.01$ , the window length  $m = 66$  and threshold  $r_0 = 9$  are obtained by optimization (cf. Section 4.3.4.4) and used for all trials. The mean and the standard deviation of the difference between the detected EMG onset and the true EMG onset for different  $SNR_i$  levels are displayed. The same calculation is performed for the offset times. For each of the three disturbance levels, the ratio of estimated to true noise and the ratio of the sum of true noise and disturbance variance to the estimated ones are



**Figure 4.11:** Signals of the double-threshold detector benchmark. Exemplary, the beginning of three trials is shown in different sub-plots for three different disturbance variance levels. In each sub-plot, only the first three bursts are shown. Normally, the  $SNR_i$  value and the parameter  $\sigma_i^t$  are randomly selected. For better illustration, in this plot, the  $SNR_i$  value of the first burst is set to 6 dB, then to 9 dB and for the third burst to 12 dB. In the first sub-plot, the parameter  $\sigma_i^t$  is set to 50 ms, in the second sub-plot to 100 ms and in the third sub-plot to 150 ms. The signals are shown after applying the Stulen-DeLuca filter as well as the whitening filter.

displayed. Finally, the obtained probabilities  $P_d$  and  $P_{fa}$  are presented in the table for each of the three disturbance levels. For the false-alarm probability  $P_{fa}$ , a second estimated value  $P_{fa}^*$  is shown, which is the result of the a double-onset detector with exactly the same parameters applied on a trial with the same noise and disturbance sequence but without any EMG bursts.

It could be shown that a double-threshold detector with optimally chosen threshold  $\hat{\zeta}$  is able in detecting all bursts with high accuracy. Furthermore, disturbances are introduced into the benchmark which leads to adaption of the threshold  $\hat{\zeta}$  such that the false-alarm probability  $P_{fa}$  remains almost equal to the desired  $P_{fa}$  of 1%. The false-alarm probability  $P_{fa}^*$  which is recorded on only noise and disturbances with the same double-onset detector and the same threshold  $\hat{\zeta}$  indicates that the threshold  $\hat{\zeta}$  is optimally chosen. The resulting  $P_{fa}$  of all trials is higher than 1% as the Stulen-DelUCA filter and the whitening filter leads to time delays and broadening of the bursts. Furthermore, previous samples are influencing the output of the double-threshold detector and leading to an increased  $P_{fa}$  after each burst.

**Table 4.1:** Results of the double-threshold detector benchmark. The mean and the standard deviation of the difference between the detected onset / offset and the true onset / offset at different  $SNR_i$  levels and disturbances are shown. Each trial contains bursts of all  $SNR_i$  levels. The detection probability  $P_d$  and the false-alarm probability  $P_{fa}$  are calculated across all  $SNR_i$  levels. For comparison, the equivalent false-alarm probability  $P_{fa}^*$  is calculated on a trial with pure noise and disturbance without any EMG bursts. The quotient  $\frac{\hat{\sigma}_n^2}{\sigma_n^2}$  between true and estimated noise variance and the quotient  $\frac{\hat{\sigma}_n^2 + \hat{\sigma}_d^2}{\sigma_n^2 + \bar{\sigma}_d^2}$  between the true and estimated sum of noise and disturbances are displayed.  $\frac{\hat{\sigma}_d^2}{\hat{\sigma}_n^2}$  is the quotient between estimated disturbance variance and noise variance. The sampling rate is set to  $T_s^{EMG} = 1/4000$  s. The results in this table were obtained with the presented detector using  $tr_{max} = 10$  ms,  $SNR_{min} = 3$  dB,  $P_{fa} = 0.01$ ,  $N_{\hat{\zeta}} T_s^{EMG} = 8$  s,  $N_{\hat{\sigma}_n^2} T_s^{EMG} = 0.2$  s, and  $th_{\bar{\sigma}_d^2} = 0.3$ .

$\frac{\bar{\sigma}_d^2}{\sigma_n^2}$	SNR	Onset error	Offset error	$\frac{\hat{\sigma}_n^2}{\sigma_n^2}$	$\frac{\hat{\sigma}_d^2}{\hat{\sigma}_n^2}$	$\frac{\hat{\sigma}_n^2 + \hat{\sigma}_d^2}{\sigma_n^2 + \bar{\sigma}_d^2}$	$P_{fa} / P_{fa}^*$	$P_d$
1.0	12	2.85 ± 1.68 ms	9.14 ± 3.21 ms				1.65 %	99.03 %
1.0	9	3.29 ± 2.49 ms	8.15 ± 2.05 ms	0.98	1.12	1.04	1.03 %	
1.0	6	3.05 ± 1.91 ms	8.34 ± 1.95 ms					
2.0	12	3.16 ± 2.69 ms	7.99 ± 2.41 ms				1.78 %	98.95 %
2.0	9	3.49 ± 1.91 ms	9.12 ± 3.45 ms	1.05	1.81	0.98	1.29 %	
2.0	6	3.39 ± 2.60 ms	8.92 ± 3.25 ms					
3.0	12	3.20 ± 1.80 ms	8.39 ± 1.93 ms				1.71 %	98.98 %
3.0	9	3.38 ± 2.14 ms	8.24 ± 1.92 ms	0.98	3.14	1.01	1.09 %	
3.0	6	2.65 ± 1.35 ms	8.27 ± 2.67 ms					

This effect can also be seen on the mean onset and offset times. A mean onset error of around 3 to 4 ms is observed. As the previous samples influence the detection result, the mean offset error is around 10 ms. Both mean error differences are below the requested offset-detection latency  $tr_{max}$ .

Larger disturbances  $\bar{\sigma}_d^2$  lead to higher thresholds ( $\hat{\zeta}$ ) such that the EMG bursts are correctly identified and the onset and offset times are independent from the disturbance level. The detection probability  $P_d$  remains almost constant across the different disturbance levels.

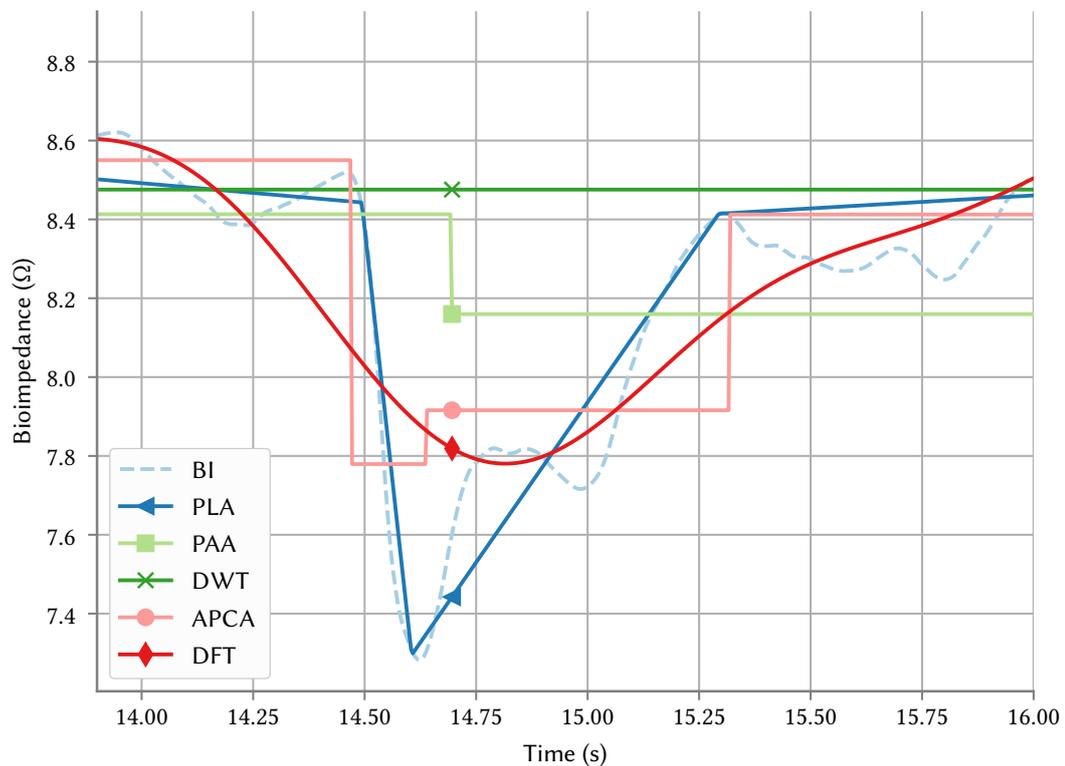
The estimated noise variance  $\hat{\sigma}_n^2$  is close to the true variance. The disturbance variance  $\hat{\sigma}_d^2$  could also be accurately estimated such that the quotient  $\frac{\hat{\sigma}_n^2 + \hat{\sigma}_d^2}{\sigma_n^2 + \bar{\sigma}_d^2}$  is close to one. The quotient of the sum of estimated noise and disturbance variance to the true sum of noise and disturbance variance is also almost one.

#### 4.3.5 Piece-wise Linear Approximation of BI

Due to the observation that the swallowing process causes EMG activity and a valley shaped deflection in the BI signal, a valley search algorithm seems a meaningful approach for finding all swallowing events in the measurement data. In order to perform a heuristic valley search, the complexity of the BI time series has to be reduced. Otherwise, each small deflection in the BI signal could be marked as valley.

Several time series representations have been introduced in the last years, including discrete Fourier transform (DFT) [157], discrete wavelet transform (DWT) [158], piece-wise linear approxi-

mation (PLA) [159], piece-wise aggregate approximation (PAA) [160], adaptive piece-wise constant approximation (APCA) [161], and singular value decomposition (SVD) [162]. Figure 4.12 shows a comparison of these approximation methods. The SVD approximation is not shown, as this method can not be applied to a one-dimensional time series. The complete time series has a length of 28 seconds, from which two seconds around a swallow are displayed. The compression ratio<sup>4</sup> of all considered approximation methods was set to 350. Only the PLA method can sufficiently approximate the swallow in such a way that begin, minimum, and end position are preserved in time and value.



**Figure 4.12:** Comparison of different approximation methods with the same compression ratio of 350. The root mean square error of the complete sequence (duration is 28 s) is comparable for all approximation methods (PLA: 271.2, PAA: 321.4, DWT: 413.2, APCA: 295.4, DFT: 215.36).

Motivated by this pre-analysis, the PLA method which is described by Keogh et al. [159] is applied to the pre-filtered (de-noised and down-sampled) BI measurement signal. Using this method, the time series is approximated by several lines of different lengths which are always an integer of the sampling time  $T_s^{BI}$ . The parameters of a specific line can be either determined by linear interpolation or linear regression. If linear interpolation is used, start and end points of all segments are sub-sets of the time series and each line is therefore defined by two sampling instances. Increasing the length of a segment is computational inexpensive as the end point is set to the next sampling instance

<sup>4</sup>The compression ratio is the ratio of the original time series length to the number of data points after applying the compression.

of the time series. If a line would be determined by linear regression instead, its start and end points are optimized by minimizing the sum of the squared error between the line and the time series. Generally, linear regression leads to a lower approximation error in comparison to linear interpolation but has higher computational cost. Therefore, linear interpolation is used within this thesis due to its lower computational demands.

For a time series with  $N^{BI}$  elements:  $\{BI\} = \{BI(1), BI(2), \dots, BI(N^{BI})\}$ , the line segments  $P_l$ ,  $l \in \{1, \dots, N^P\}$  line up an approximated time series  $\{BI^P\}$ . A line segment  $P_l$  is defined by its first sampling instance  $a_l$  and its last sampling instance  $b_l$ . Each segment can be written as:

$$P_l = [b_l - a_l, BI(a_l), BI(b_l)], \quad l \in \{1, \dots, N^P\}, \quad a_l \in \mathbb{N}_{>0}, \quad b_l \in \mathbb{N}_{>0},$$

$$a_l = \begin{cases} \sum_{k=1}^{l-1} P_k(1) + 1 & \text{if } l > 1 \\ 1 & \text{if } l = 1. \end{cases}, \quad b_l = \sum_{k=1}^l P_k(1), \quad (4.39)$$

where  $P_l(1) = b_l - a_l$  defines the length of the segment in samples,  $P_l(2)$  is the first value and  $P_l(3)$  is the last value of the line segment. The time distance between the end and the start of two adjacent lines is the sampling time ( $a_{l+1} = b_l + 1$ ,  $l \in \{2, \dots, N^P\}$ ). A line segment  $P_l$  approximates the subset  $\{BI[a_l, b_l]\}$  which contains the time series from  $BI(a_l)$  to  $BI(b_l)$ .  $\{BI^{P_l}[a_l, b_l]\}$  is reconstructed in between  $a_l$  and  $b_l$  as follows:

$$BI^{P_l}(k) = BI(a_l) + \frac{k - a_l}{b_l - a_l} (BI(b_l) - BI(a_l)), \quad k = a_l, \dots, b_l. \quad (4.40)$$

To determine the segments a bottom-up algorithm is used as suggested by Keogh et al. [159]. It will be explained in detail after this paragraph. Beside the bottom-up algorithm, other algorithms as the top-down algorithm or the sliding window algorithm can be applied in order to calculate the approximating line segments. The top-down algorithm starts with one line segment for approximating the complete time series. Then, the algorithm recursively splits the segments until all segments have an approximation error below a defined threshold  $max\_error$ . This algorithm has a quadratic complexity which is higher than the logarithmic complexity of the bottom-up and sliding window approach [159]. The sliding window approach does not optimize the segments globally and can therefore be better used for online PLA. In the sliding window approach, each segment is consecutively obtained such that the length of the current segment is repeatedly extended by one sampling instance until the sum of the squared error between this line and the time series exceeds a pre-defined threshold  $max\_error$ . Then, the current segment is stored and a new segment is initialized. Keogh et al. [159] have shown that the bottom-up algorithm has the best performance in comparison to the sliding window and the top-down algorithm.

The implementation of the bottom-up algorithm is shown in Figure 4.13. The algorithm starts with the maximal number of segments (each segment contains only two sampling instances). For all pairs of adjacent lines the merge costs are calculated and saved into a vector. The merge costs are defined as the squared error sum between the measurement data and the merged line segment from the beginning of the first line to the end of the second line. Thus, the merge cost vector contains the information which line merging leads to the lowest increase in approximation error. The squared

```

1: procedure PLA_BOTTOM_UP( {BI})
2:    $j \leftarrow 1$ 
3:   for  $i \leftarrow 1 : 2 : N^{BI} - 1$  do ▷ Time series {BI} has length  $N^{BI}$ 
4:      $P_j \leftarrow [2, BI(i), BI(i + 1)]$  ▷ Create initial fine approximation.
5:      $j \leftarrow j + 1$ 
6:   end for
7: ▷  $N^P$  is now  $j - 1$ 
8:   for  $i \leftarrow 1$  to  $N^P - 1$  do ▷ Find the cost of merging...
9:      $merge\_cost(i) \leftarrow \text{CALCULATE\_ERROR}(\text{MERGE}(P_i, P_{i+1}))$  ▷ For each pair of segments.
10:  end for
11:  while  $\min(merge\_cost) < max\_error$  AND  $N^P > 1$  do
12:     $i \leftarrow \arg \min(merge\_cost)$  ▷ Find line pair with the lowest merge cost.
13:     $P_i \leftarrow \text{MERGE}(P_i, P_{i+1})$ 
14:    remove element  $P_{i+1}$  ▷  $N^P$  decreases by one.
15:    remove element of  $merge\_cost$  at  $i$ 
16:    if  $N^P \geq i + 1$  then
17:       $merge\_cost(i) \leftarrow \text{CALCULATE\_ERROR}(\text{MERGE}(P_i, P_{i+1}))$ 
18:    end if
19:    if  $i > 1$  then
20:       $merge\_cost(i - 1) \leftarrow \text{CALCULATE\_ERROR}(\text{MERGE}(P_{i-1}, P_i))$ 
21:    end if
22:  end while
23:  return  $P_l$  ▷  $l = 1, \dots, N^P$ 
24: end procedure

```

**Figure 4.13:** Implementation of the bottom-up algorithm. The bottom-up algorithm returns the piecewise linear approximation  $P_l$  for the time series  $\{BI\}$ . The parameter  $max\_error$  is the maximal allowed squared error.

error sum  $e_i$  for a line segment  $i$  is calculated by the following equation

$$e_i = \text{CALCULATE\_ERROR}(P_i) = \sum_{k=a_i}^{b_i} (BI(k) - BI^{P_i}(k))^2, \quad (4.41)$$

where  $a_i$  is the first sampling instance of segment  $P_i$  and  $b_i$  is the last sampling instance of this segment. Merging of adjacent segments means that a new segment  $P_i$  is created by the combination of the segment  $P_i$  and the adjacent segment  $P_{i+1}$ . As linear interpolation is used, merging does not cost much computational time:

$$P_i = \text{MERGE}(P_i, P_{i+1}) = [P_i(1) + P_{i+1}(1), P_i(2), P_{i+1}(3)]. \quad (4.42)$$

After initializing the  $merge\_cost$  vector, the pair of lines with the lowest merging cost are merged and the number of lines and the elements within  $merge\_cost$  are reduced by one. The merging costs of the newly merged line with respect to its preceding line and its successive line have to be recalculated. The merging of line segments with partly recalculation of  $merge\_cost$  is repeated as long as the lowest cost in  $merge\_cost$  does not exceed the maximal allowed squared error  $max\_error$ .

The valley detection involves a searching of local minima within the BI time series. The time distance between the line end point  $P_l(3)$  and the start point  $P_{l+1}(2)$  of the subsequent line is only one sampling instance. The values between the last sampling point of a line and starting point of the subsequent line are usually very similar. Therefore, a further approximation is applied where the time instance of the ending point of a line segments is set to the starting point of the subsequent line. The new point is named connection point. The time instances of all connection points are stored separately in a vector which is defined as

$$P_l(k) = \begin{cases} 1 & \text{if } k = 1 \\ 1 + b_{k-1} = 1 + \sum_{l=1}^{k-1} P_l(1) & \text{if } k = 2, \dots, N^P, \quad P_l \in \mathbb{N}^{N^P+1}. \\ b_{k-1} = \sum_{l=1}^{k-1} P_l(1) & \text{if } k = N^P + 1 \end{cases} \quad (4.43)$$

The values at the newly defined connection points are set to the average of the original end and start points. The vector which contains all connection point values is then defined as

$$P_C(k) = \begin{cases} P_1(2) & \text{if } k = 1 \\ (P_{k-1}(3) + P_k(2))/2 & \text{if } k = 2, \dots, N^P, \quad P_C \in \mathbb{R}^{N^P+1}. \\ P_{N^P}(3) & \text{if } k = N^P + 1 \end{cases} \quad (4.44)$$

Finally,  $P_S \in \mathbb{R}^{N^P}$  contains the slope of each segment:

$$P_S(l) = \frac{P_C(l+1) - P_C(l)}{P_l(l+1) - P_l(l)}, \quad l = 1 \dots N^P. \quad (4.45)$$

The connection point vector  $P_C$  can then be used to easily find all local minimum values.

Muscle activity is a physiological requirement for swallowing. Thus, the presence of EMG activity will be an indicator for a potential swallow and only valleys which coincide with EMG activity should be segmented<sup>5</sup>. To assign EMG activity to the BI line segments, a vector  $P_{EMG}$  with  $N^P$  entries is created.

Around the starting point of each line segment  $l$ , the number of samples with detected EMG activity (output *act* of the double-threshold detector) will be counted in a time interval  $[P_l(l)T_s^{BI} - VS_{emg}, P_l(l)T_s^{BI} + VS_{emg}]$ . The parameter  $VS_{emg}$ <sup>6</sup> describes the width of the interval. If  $VS_{onset}$  percentage of samples in this interval indicate EMG activity, the vector entry  $P_{EMG}(l)$  is set to one:

$$P_{EMG}(l) = \begin{cases} 1 & \text{if } \left( \sum_{k=\frac{P_l(l)T_s^{BI}-VS_{emg}}{T_s^{EMG}}}^{\frac{P_l(l)T_s^{BI}+VS_{emg}}{T_s^{EMG}}} act(k) \right) \geq \frac{2VS_{emg}VS_{onset}}{T_s^{EMG}}, \quad l = 2 \dots N^P. \\ 0 & \text{otherwise} \end{cases} \quad (4.46)$$

The parameters  $VS_{emg}$  and  $VS_{onset}$  are optimally chosen when almost all swallowing related valleys are found and most of the non swallowing related valley are rejected.

<sup>5</sup>Marking all potential swallows in the measurement data is called segmentation in this context.

<sup>6</sup>VS stands for valley search.

### 4.3.6 Segmentation of Swallowing by a Heuristic Valley Search

It is assumed that swallowing leads to a valley shaped deflection in the BI measurement signal while EMG activity is present. It is furthermore assumed that the PLA preserves the valley in the BI signal during a swallow, whereas the starting point of the first line segment marks always the beginning of a swallow induced valley and the ending point of a subsequent line segment marks the ending of this valley.

A valley in the approximated BI signal consists of a starting point  $i$  with  $0 < i < N^P - 1$  of a line segment and an ending point  $j$  with  $i + 1 \leq j \leq N^P + 1$  which is related to a subsequent line segment. The line segments from  $i$  to  $j$  form a valley when the following conditions are fulfilled:

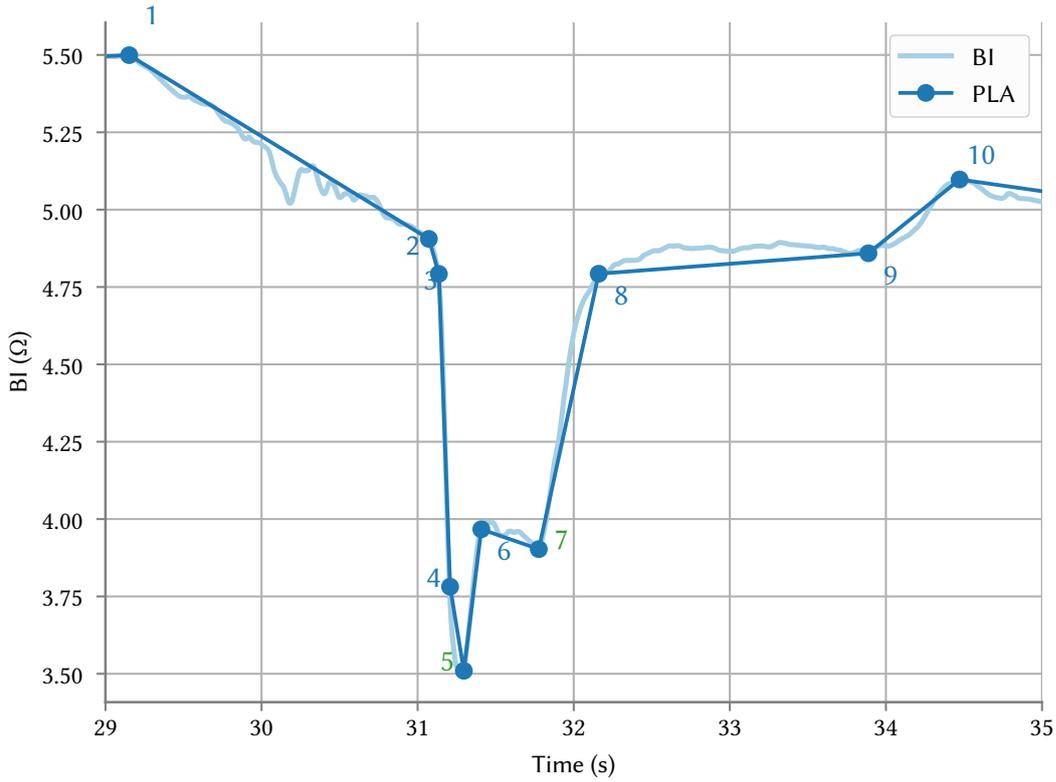
1. the first line has a negative and the last line has a positive slope,
2. the valley contains only one minimum,
3. the time duration of the valley is within a given bound  $VS_{bound} = [\underline{vs}, \overline{vs}]$  which consists of a lower bound  $\underline{vs}$  and an upper bound  $\overline{vs}$ ,
4. the first line of the valley shows sufficient EMG activity at the beginning (indicated by a one in  $P_{EMG}(i)$ ),
5. a connection from the beginning to the end of the valley does not cross any other line segment.

Within the approximated BI signal, all valleys which fulfill these conditions are searched within the vectors  $P_S$ ,  $P_I$ ,  $P_{EMG}$ , and  $P_C$ . All valley combinations are stored into an upper triangle matrix  $T \in \mathbb{N}^{N^P \times N^P}$ . If the conditions for a valley are fulfilled between  $P_I(i)$  and  $P_I(j)$ , the corresponding segment connection number  $m$  of the corresponding minimum is written into  $T(i, j)$ ; otherwise,  $T(i, j)$  is set to 0. Thus, for  $T(i, j) = m > 0$  with row index  $i \in \{1 \dots N^P\}$  and column index  $j \in \{i \dots N^P\}$ , a valley is defined by the following segment interconnection points: start point  $V_{start} = i$ , minimum point  $V_{min} = m$  and end point  $V_{end} = j$ . All minimum points which belong to a valley are additionally stored into the vector  $M \in \mathbb{N}_{>0}$ .

**Table 4.2:** Matrix  $T$  and vector  $M$  from the example in Fig. 4.14.

$T$	$j=1$	2	3	4	5	6	7	8	9	10	$M$
$i=1$	0	0	0	0	0	0	0	0	0	0	5
2		0	0	0	0	0	0	0	0	0	7
3			0	0	0	5	0	0	0	0	
4				0	0	5	0	0	0	0	
5					0	0	0	0	0	0	
6						0	0	7	0	0	
7							0	0	0	0	
8								0	0	0	
9									0	0	
10										0	

The definition for a valley may lead to different possible start and end points which have the same local minimum. This can easily be seen from the example (Fig. 4.14 and Table 4.2). The sequence notation is  $(V_{start}-V_{min}-V_{end})$ . For the second minimum ( $m = 7$ ), only one sequence of segments



**Figure 4.14:** Numbered line segments of a BI measurement around a swallow. Within this time interval, two minima are found at the connection time points 5 and 7. Possible valleys are (3–5–6, 4–5–6) for the minimum 5 and (6–7–8) for the minimum 7. Possible valley combination are written into the matrix  $T$  and shown in Table 4.2.

can be found (6–7–8); whereas for the first minimum ( $m = 5$ ), two possible sequences of segments exist (3–5–6 and 4–5–6).

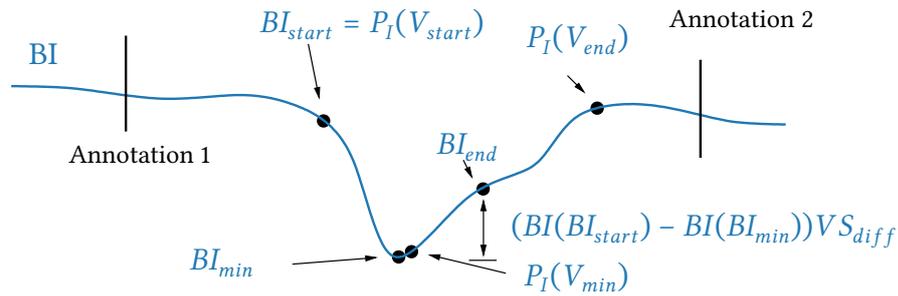
Subsequently, the best start and end point is selected by maximizing an utility function with respect to all detected valleys for the corresponding minimum  $m$  within  $M$ . The utility function is chosen heuristically in such a way that valleys with large area and short distance between the start and end point are preferred. The utility function  $f(i, j)$  is maximized in order to find the best valley to minimum point  $m$

$$\begin{aligned} \arg \max_{i,j} f(i, j), \quad f(i, j) &= \frac{\text{area}(i, j)}{\text{length}(i, j)}, \\ \text{subject to } T(i, j) &= m, \\ i &= 1, \dots, N^P, \quad j = i, \dots, N^P, \quad m \in M, \end{aligned} \quad (4.47)$$

where  $\text{area}(i, j)$  is the area enclosed by the approximated BI segments from  $P_I(i)$  to  $P_I(j)$  and the closing line of the valley which has the length

$$\text{length}(i, j) = \sqrt{((P_I(j) - P_I(i))T_s^{BI})^2 + (P_C(j) - P_C(i))^2}. \quad (4.48)$$

All other entries in  $T$  which correspond to the minimum point  $m$  but are not the optimal points  $i$  and  $j$  are overwritten with 0.



**Figure 4.15:** Adjustment of minimum and end point of a swallow. Each swallow was manually marked by a hand switch and labeled by the investigator. Based on these information, a start and stop annotation was set for all recorded swallows. The manually added annotations mark the begin and end of a swallow related valley and are used for training and testing of a classifier. The figure is based on Fig. 4 from [132].

After finding one optimal valley for each minimum point  $m$  in  $M$ , the previously found start, minimum, and end points ( $V_{start}$ ,  $V_{min}$ , and  $V_{end}$ ) of each valley are improved in a post-processing step. The post-processing starts with the first valley within the BI time series. The revised points of the valley are the sampling points  $BI_{start}$ ,  $BI_{min}$ , and  $BI_{end}$  (cf. Fig. 4.15). For the recalculation of these points, the original BI signal is analyzed between the sampling instances  $P_I(V_{start})$  and  $P_I(V_{end})$ . The start point  $BI_{start}$  is set to  $P_I(V_{start})$ .  $BI_{min}$  is set to the local minimum of the non-approximated BI signal between  $P_I(V_{start})$  and  $P_I(V_{end})$ :

$$BI_{min} = \arg \min_k BI(k), P_I(V_{start}) < k < P_I(V_{end}). \quad (4.49)$$

The end point  $BI_{end}$  is located after  $BI_{min}$  and is defined as the sampling instance where the BI recovers to  $0 < VS_{diff} < 1$  of the swallow-related drop:

$$BI_{end} = \arg \min_k \left( (BI_{start} - BI_{min})VS_{diff} - (BI(k) - BI_{min}) \right), \quad k > BI_{min}. \quad (4.50)$$

The parameter  $VS_{diff}$  determines the position of  $BI_{end}$  and should be set to at least 0.5 as otherwise the end point will be located near to the minimum point  $BI_{min}$ .

The newly found valley with the sampling instances  $BI_{start}$ ,  $BI_{min}$ , and  $BI_{end}$  is stored. All remaining valleys in  $T$  are checked whether they coincide with the interval starting from  $BI_{start}$  to  $BI_{end}$ . All valleys which begin or end in this interval are removed from the matrix  $T$  by overwriting the corresponding entry with zero. All minimum points in  $M$  which do not have any corresponding entry within  $T$  are removed from  $M$ .

Post-processing is proceeded with the next remaining entry in  $M$  which belongs to the next nearest valley until all entries have been processed.

#### 4.3.7 Feature Extraction

The preprocessed but not whitened EMG signal is high-pass filtered by a non-causal fourth-order high-pass filter with a cut-off frequency of 40 Hz. Then the signal is rectified and filtered by a non-

causal low-pass filter of 4th order with a cut-off frequency of 10 Hz. The filter parameters were chosen such that the filter output approximates the envelope of the rectified EMG signal. Finally, the rectified and filtered EMG sequence is down-sampled to 250 Hz. This signal is used for feature extraction and denoted as  $sEMG$ .

For each segmented valley  $i$  in the BI signal, the sampling instances  $BI_{start,i}$ ,  $BI_{min,i}$ , and  $BI_{end,i}$  are given. The corresponding EMG activity is defined as the interval from  $sEMG_{start,i}$  to  $sEMG_{end,i}$  of the signal  $sEMG$ . The EMG intervals are extracted from the double-threshold detector output  $act$ .  $sEMG_{start,i}$  and  $sEMG_{end,i}$  are set to that EMG interval which is nearest to the starting point  $BI_{start,i}$ .  $sEMG_{max,i}$  is set to the sampling instance of the maximum value of  $sEMG$  in this interval.

$\mathbf{BI}_{SM}$  is defined as the BI measurement vector of length  $N_{SM}$  from  $BI_{start,i}$  to  $BI_{min,i}$ .  $\mathbf{BI}_{ME}$  is defined as the BI measurement vector of length  $N_{ME}$  containing all measurement samples from  $BI_{min,i}$  to  $BI_{end,i}$ . In the following, the upper index  $V$  refers to the value and the upper index  $T$  to time of a BI or EMG point of a valley (cf. Fig. 4.16). Using this convention, the following features which are related to time and amplitude can be calculated:

- $t_{min,i} = BI_{min,i}^T - BI_{start,i}^T$ ,  $t_{end,i} = BI_{end,i}^T - BI_{start,i}^T$
- $t_{emgdiff,i} = sEMG_{start,i}^T - BI_{start,i}^T$
- $t_{emgmax,i} = sEMG_{max,i}^T - sEMG_{start,i}^T$
- $t_{emgend,i} = sEMG_{end,i}^T - sEMG_{start,i}^T$
- $\Delta_{min,i} = BI_{min,i}^V - BI_{start,i}^V$
- $sEMG_{max,i}^V$

The slope  $S_{1,i}$  is defined as the steepest slope in the range from  $BI_{start,i}$  to  $BI_{min,i}$  and the slope  $S_{2,i}$  is set to the steepest slope from  $BI_{min,i}$  to  $BI_{end,i}$ .  $S_{1,i}$  and  $S_{2,i}$  are included as features. The slope  $S_{1,i}$  is calculated by

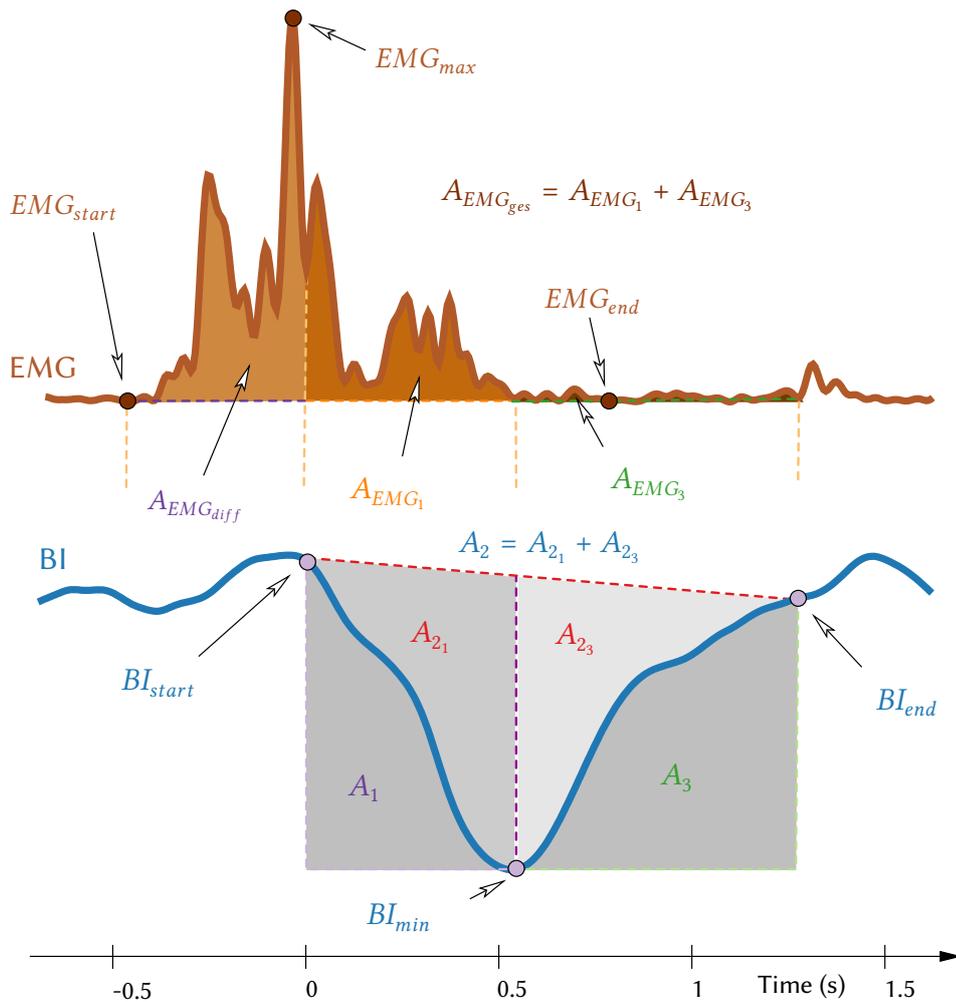
$$S_{1,i} = \min_k \frac{\mathbf{BI}_{SM}^V((k+1)\frac{N_{SM}}{N_{slope}}) - \mathbf{BI}_{SM}^V(k\frac{N_{SM}}{N_{slope}})}{\mathbf{BI}_{SM}^T((k+1)\frac{N_{SM}}{N_{slope}}) - \mathbf{BI}_{SM}^T(k\frac{N_{SM}}{N_{slope}})}, \quad 0 < k < N_{slope} - 1, \quad (4.51)$$

where  $N_{slope}$  defines in how many parts the vector  $\mathbf{BI}_{SM}$  is fragmented. For each part, the slope is calculated and finally the steepest is taken for  $S_{1,i}$ . The second slope can be found as follows:

$$S_{2,i} = \max_k \frac{\mathbf{BI}_{ME}^V((k+1)\frac{N_{ME}}{N_{slope}}) - \mathbf{BI}_{ME}^V(k\frac{N_{ME}}{N_{slope}})}{\mathbf{BI}_{ME}^T((k+1)\frac{N_{ME}}{N_{slope}}) - \mathbf{BI}_{ME}^T(k\frac{N_{ME}}{N_{slope}})}, \quad 0 < k < N_{slope} - 1. \quad (4.52)$$

For determining the slopes during swallowing, a value  $N_{slope} = 25$  has been found to give good results. Areas ( $A$ ) which are determined by the BI and  $sEMG$  signals are also used as features. Figure 4.16 shows the areas that are used and here defined.

The areas under the  $sEMG$  curve are calculated as sum of the measured samples multiplied by the sampling frequency. The area  $A_{sEMG_{diff,i}}$  is the area from  $sEMG_{start,i}$  to  $BI_{start,i}$ . The next area



**Figure 4.16:** Definition of area ( $A$ ) based features. The elevation of the larynx takes place between  $BI_{start,i}$  and  $BI_{min,i}$ . The figure is based on Fig. 5 from [132].

is area  $A_{sEMG_1,i}$  which goes from  $BI_{start,i}$  to  $BI_{min,i}$ . And finally, the area  $A_{sEMG_3,i}$  is calculated from  $BI_{min,i}$  to  $BI_{end,i}$ .

In addition, the symbolic aggregate approximation (SAX) [163] of the BI and the EMG curves is used as features. SAX is a technique for efficient reduction of time series in order to preserve the essential characteristics of the signal. The method is widely used in time-series pattern analysis. Wang et al. [164] used SAX for feature extraction in combination with an SVM for online word segmentation and recognition. SAX words were used as features for classifying different drilling operation by Esmael et al. [165]. Another field of application is pattern matching by which computational complexity can be reduced by the SAX method. This allows indexing and searching of large datasets [166] and can be used for finding repeated patterns [167].

For each valley  $i$ , the pre-processed BI signal (cf. Section 4.3.3) is analyzed within the interval from  $BI_{start,i}^T - 0.4$  s with a length  $T_{SAX}^{BI} = 2$  s. The interval was chosen such that swallow induced changes are included. The BI sequence is normalized within these ranges (zero mean value, standard deviation of one). The resulting sequence  $BI_{norm,i}$  is down-sampled by piece-wise aggregate approximation (PAA) by which the sequence is subdivided in equally spaced segments

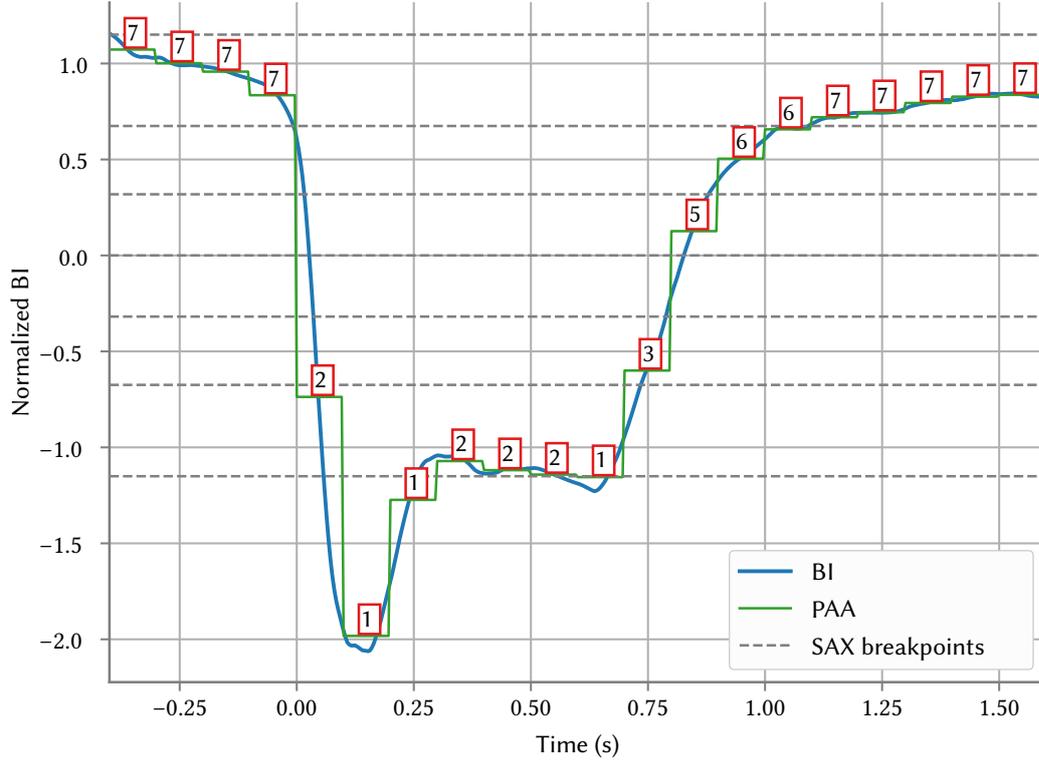


Figure 4.17: Features generated by the SAX method from the normalized BI sequence of a recorded swallow. As parameters  $N_{sym}^{BI} = 8$  and  $N_{PAA}^{BI} = 20$  were used. The figure is based on Fig. 6 from [132].

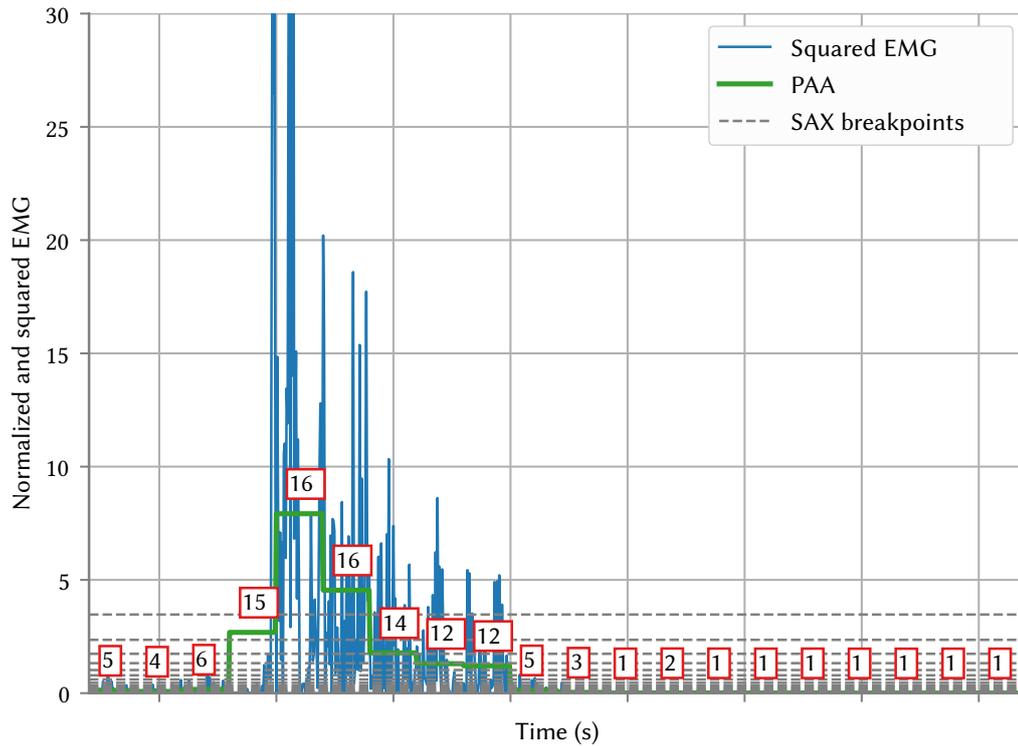
along the time axis and each segment is approximated by its mean value. The normalized BI sequence  $BI_{norm,i}$  with  $T_{SAX}^{BI}/T_s^{BI}$  samples is reduced to the number of PAA coefficients  $N_{PAA}^{BI}$ :

$$BI_{PAA,i}(k) = \frac{1}{w_{PAA}} \sum_{j=w_{PAA}(k-1)+1}^{w_{PAA}k} BI_{norm,i}(j), \quad (4.53)$$

$$w_{PAA} = \frac{T_{SAX}^{BI}}{T_s^{BI} N_{PAA}^{BI}} \in \mathbb{N}, \quad k = 1, \dots, N_{PAA}^{BI}.$$

The pre-processed (cf. Section 4.3.3) EMG signal is filtered by a zero-phase low-pass filter of fourth order with a cut-off frequency of 250 Hz. The filtered EMG signal is down-sampled to 500 Hz in order to reduce the necessary storage which is needed to store all EMG sequences. The EMG sequence for valley  $i$  starts from  $BI_{start,i}^T - 0.8$  s and has a length of  $T_{SAX}^{EMG} = 3.6$  s. The interval was chosen such that EMG activity are included.

Then, the EMG sequence is normalized (zero mean value, standard deviation of one). Additionally, each normalized EMG sequence  $EMG_{norm,i}$  is squared, as otherwise an approximation with the PAA would not give any useful results. The normalized and squared EMG sequence  $EMG_{norm,i}^2$  with



**Figure 4.18:** Features generated by the SAX method from a normalized and squared EMG sequence of a recorded swallow ( $N_{sym}^{EMG} = 16$  and  $N_{PAA}^{EMG} = 20$ ).

$T_{SAX}^{EMG}/T_s^{EMG}$  samples is approximated by

$$EMG_{PAA,i}(k) = \frac{1}{w_{PAA}} \sum_{j=w_{PAA}(k-1)+1}^{w_{PAA}k} EMG_{norm,i}^2(j), \quad (4.54)$$

$$w_{PAA} = \frac{T_{SAX}^{EMG}}{T_s^{EMG} N_{PAA}^{EMG}} \in \mathbb{N}, \quad k = 1, \dots, N_{PAA}^{EMG},$$

where  $N_{PAA}^{EMG}$  is the number of PAA coefficients.

In a next step, the SAX method maps the PAA coefficients to SAX symbols. This step reduces the resolution across the BI/EMG-coordinate. All PAA coefficients that belong to the same interval on the BI/EMG-axis are mapped to the same SAX symbol.

Breakpoints  $\beta_l^{BI}$  define to which symbol a certain BI sample is mapped. The alphabet size for the BI sequence is called  $N_{sym}^{BI}$ .  $N_{sym}^{BI}$  different breakpoints  $\beta_l^{BI}$  are defined as a sorted list such that an area under an  $\mathcal{N}(0, 1)$  Gaussian curve from  $\beta_l^{BI}$  to  $\beta_{l+1}^{BI}$  is  $1/N_{sym}^{BI}$ . The first and the last breakpoints are defined as  $\beta_0^{BI} = -\infty$  and  $\beta_{N_{sym}^{BI}}^{BI} = \infty$ . By using the inverse cumulative distribution function of the normal distribution  $F_{norm}^{-1}$ , which is defined over the cumulative distribution function with

$$x = F_{norm}^{-1}(p, \mu, \sigma) = \{x : F_{norm}(X \leq x, \mu, \sigma) = p\} \quad (4.55)$$

with mean  $\mu = 0$  and standard deviation  $\sigma = 1$ , the breakpoints can be numerically estimated:

$$\beta_l^{BI} = F_{norm}^{-1}(l/N_{sym}^{BI}, 0, 1), \quad l = 2, \dots, N_{sym}^{BI} - 1. \quad (4.56)$$

Using these breakpoints, mapping from the PAA approximation  $BI_{PAA,i}$  to  $BI_{SAX,i}$  is performed for each valley  $i$  according to Lin et al. [163] by:

$$BI_{SAX,i}(k) = \alpha_l^{BI} \text{ if } \beta_{l-1}^{BI} \leq BI_{PAA,i}(k) < \beta_l^{BI}, \quad k = 1, \dots, N_{PAA}^{BI}, \quad (4.57)$$

where the set  $\alpha^{BI}$  contains  $N_{sym}^{BI}$  symbols with a defined ranking. The SAX symbols are defined as follows:

$$\alpha_l^{BI} = l, \quad l = 1, 2, \dots, N_{sym}^{BI}. \quad (4.58)$$

Figure 4.17 illustratively shows the features generated by the SAX method from the normalized BI sequence using  $N_{sym}^{BI} = 8$  and  $N_{PAA}^{BI} = 20$ .

The squared EMG signal has a  $\chi^2$  distribution. In order to produce symbols which are equally distributed for a  $\chi^2$  distributed input, the breakpoints  $\beta_l^{EMG}$  are chosen such that the area under a  $\chi^2$  cumulative distribution function with one degree of freedom from  $\beta_l$  to  $\beta_{l+1}$  is  $1/N_{sym}^{EMG}$ , where  $N_{sym}^{EMG}$  is the alphabet size. The first and the last breakpoints are defined as  $\beta_0^{EMG} = 0$  and  $\beta_{N_{sym}^{EMG}}^{EMG} = \infty$ . By using the inverse cumulative distribution function of the  $\chi^2$  distribution, which is defined over the cumulative distribution function with

$$x = F_{\chi^2}^{-1}(p, \nu) = \{x : F_{\chi^2}(X \leq x, \nu) = p\} \quad (4.59)$$

with degree of freedom  $\nu = 1$ , the breakpoints can be numerically estimated:

$$\beta_l = F_{\chi^2}^{-1}(l/N_{sym}^{EMG}, 1), \quad l = 2, \dots, N_{sym}^{EMG} - 1. \quad (4.60)$$

Using these breakpoints, mapping the approximated EMG sequence  $EMG_{PAA,i}$  to  $EMG_{SAX,i}$  for each valley  $i$  is performed by:

$$EMG_{SAX,i}(k) = \alpha_l^{EMG} \text{ if } \beta_{l-1}^{EMG} \leq EMG_{PAA,i}(k) < \beta_l, \quad k = 1, \dots, N_{PAA}^{EMG}, \quad (4.61)$$

where the set  $\alpha^{EMG}$  contains  $N_{sym}^{EMG}$  symbols and is defined by

$$\alpha_l^{EMG} = l, \quad l = 1, 2, \dots, N_{sym}^{EMG}. \quad (4.62)$$

An exemplary EMG sequence with corresponding SAX symbols is shown in Figure 4.18.

For each segmented valley  $i$ , a SAX sequence is generated for the related BI and EMG time series. In total, the feature vector has a length of  $17 + N_{PAA}^{BI} + N_{PAA}^{EMG}$ . The first 17 features are related to times, distances and areas.

### 4.3.8 Support Vector Machine

In the previous sections, it is described how valleys are segmented in the BI and EMG measurement data. A C-support vector machine (C-SVM) is used for classifying these valleys in swallows and non-swallowing related valleys. As the C-SVM outperformed all other kinds of classifier, it was decided to focus on the C-SVM. A C-SVM was firstly described by Cortes et al. [168] and within this thesis the C-SVM realization by Chang et al. [169] is used.

For each valley  $i$ , a feature vector<sup>7</sup>  $\mathbf{x}_i$  and the corresponding class membership  $y_i$  are defined by

$$\mathbf{x}_i = [\mathbf{EMG}_{SAX,i}, \mathbf{BI}_{SAX,i}, \mathbf{t}_i, \mathbf{A}_i, \mathbf{\Delta}_i, emg_{max,i}], \quad (4.63)$$

$$y_i = \begin{cases} 1 & \text{if valley } v \text{ is marked as swallow} \\ -1 & \text{otherwise} \end{cases}, \quad (4.64)$$

where  $i$  is the index of the valley from the feature vector and class membership is calculated.

The vector  $\mathbf{t}_i$  contains all time related features for valley  $i$ ,  $\mathbf{A}_i$  consists of all area based features of valley  $i$  and  $\mathbf{\Delta}_i$  contains all value difference features of valley  $i$ .  $\mathbf{EMG}_{SAX,i}$  and  $\mathbf{BI}_{SAX,i}$  contain the SAX words of the EMG and BI time series around the valley  $i$ . The corresponding class membership  $y_i$  is found by checking each valley weather it coincide with the manually set swallow annotations. The swallow annotations are set based on the state of a hand switch which is manually pressed during all recorded swallows.

Within a C-SVM, the feature vector  $\mathbf{x}_i$  is transformed to a higher-dimensional feature space by the feature map  $\phi(\mathbf{x}_i)$  which is derived from a kernel function  $K$  which holds  $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$ . As implementation of the C-SVM, LIBSVM [169] with a radial basis function (RBF) kernel function is used. A RBF kernel function is defined as:

$$K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2} = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j), \quad \gamma > 0, \quad (4.65)$$

where  $\gamma$  is the kernel width of the RBF kernel function and  $\phi$  is the mapping vector function which is found by taylor series expansion.

The higher-dimensional feature space is split by a hyperplane which is defined by a normal vector  $\mathbf{w}$  and an offset  $b$ . The membership of valley  $i$  is determined by calculating on which site of the hyperplane the transfered feature vector  $\phi(\mathbf{x}_i)$  lays:

$$y_i^{pred} = \text{sgn}(\mathbf{w}^T \phi(\mathbf{x}_i) + b), \quad (4.66)$$

where  $\text{sgn}$  is the sign function. The hyperplane is optimally chosen when the predicted class membership  $y_i^{pred}$  is equal to the true membership  $y_i$  for almost all valleys.

The actual class membership  $y_i$  is compared with the predicted membership  $y_i^{pred}$  for all  $i$  and written into a 2x2 confusion matrix. The trained C-SVM is evaluated by the sensitivity, specificity, precision and accuracy which are obtained by the confusion matrix according to Table 4.3.

<sup>7</sup>In order to avoid numerical problems, the feature vector is normalized to a range of  $[-1, 1]$ . The normalization parameter are calculated from the training data and later applied to the test data.

#### 4. Automatic Segmentation and Classification of Swallowing

**Table 4.3:** Accuracy, sensitivity, and specificity can be calculated from the confusion matrix. TP - The number of swallows which were correctly identified as swallow. FP - The number of non swallow related events which were incorrectly identified as swallow. TN - The number of non swallow related events which were correctly identified as non-swallow. FN - The number of swallows which were incorrectly identified as non swallow related events (non-swallows).

		Predicted	
		Swallows	Non-swallows
Actual	Swallows	TP	FN
	Non-swallows	FP	TN
Sensitivity:		$TP/(TP + FN)$	
Specificity:		$TN/(TN + FP)$	
Precision:		$TP/(TP + FP)$	
Accuracy:		$\frac{(TP+TN)}{(TP+FP+FN+TN)}$	

Given  $l$  valleys together with the training set  $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_l, y_l)\}$  containing feature vector  $\mathbf{x}_i$  and the corresponding membership  $y_i$  for  $i = 1, \dots, l$ , the C-SVM can be trained by solving the following optimization problem<sup>8</sup>:

$$\begin{aligned} & \underset{\mathbf{w}, b, \xi}{\text{minimize}} && \frac{1}{2} \mathbf{w}^T \mathbf{w} + W^- C \sum_{k=1}^l \begin{cases} \xi_k, & \text{if } y_k = -1 \\ 0, & \text{if } y_k = 1 \end{cases} + W^+ C \sum_{k=1}^l \begin{cases} 0, & \text{if } y_k = -1 \\ \xi_k, & \text{if } y_k = 1 \end{cases} \\ & \text{subject to} && y_i \cdot (\mathbf{w}^T \phi(\mathbf{x}_i) + b) \geq 1 - \xi_i, \\ & && \xi_i \geq 0, \quad i = 1, \dots, l. \end{aligned}$$

Minimizing the first term  $\frac{1}{2} \mathbf{w}^T \mathbf{w}$  corresponds to maximizing the margin  $\frac{2}{\|\mathbf{w}\|}$  which is defined as the distance between the hyperplane and the nearest correctly classified feature  $\phi(\mathbf{x})$ . As it is not always possible to find  $\mathbf{w}$  and  $b$  that satisfy  $y_i \cdot (\mathbf{w}^T \phi(\mathbf{x}_i) + b) \geq 1$ , an error term  $\xi_i$  which is needed for solving the optimization problem together with  $\mathbf{w}$  and  $b$  is introduced. The number of all  $\xi_i$  which are zero correspond to the number of correctly classified training data. The learning parameter  $C$  has to be chosen by the user. Increasing  $C$  means that penalizing misclassified training data become more important than maximizing the margin between the hyperplane and the nearest feature vector  $\phi(\mathbf{x})$ .

For some classification problems the number of both classes are unbalanced, the classification penalty parameter  $C$  can be differently weighted by both classes by  $W^-$  and  $W^+$ . Normally  $W^-$  and  $W^+$  is equal to one. Depending on the selected segmentation parameters, both used classes may become unbalanced as the number of non-swallows may be greater than the number of swallows. The in-homogeneous number of data in both classes can be compensated by using class dependent weighting parameters  $W^+$  and  $W^-$ . Each weighting parameter is set to the inverse frequency of the

<sup>8</sup>In order to solve the optimization problem efficiently, an equivalent lower dimensional optimization problem is solved which leads to a solution of the original optimization problem [169].

corresponding class such that the sum of  $W^+$  and  $W^-$  is always two:

$$W^+ = \frac{2}{l} \sum_{k=1}^l \begin{cases} 1, & \text{if } y_k = -1 \\ 0, & \text{if } y_k = 1 \end{cases}, \quad (4.67)$$

$$W^- = \frac{2}{l} \sum_{k=1}^l \begin{cases} 0, & \text{if } y_k = -1 \\ 1, & \text{if } y_k = 1 \end{cases}, \quad (4.68)$$

where  $l$  is the number of valleys in the training data.

The training parameters  $C$  and  $\gamma$  have to be optimized for each specific classification problem by a grid search. The parameter  $C = 2^n$ , which is related to the cost for false classification, and the parameter  $\gamma = 2^p$ , which is the kernel width, should be varied such that  $n$  and  $p$  are linearly spaced. The parameter  $n$  is normally searched in a range from -5 to 15, whereas the parameter  $p$  is searched in a range from -15 to 5. The parameter combination which leads to the the best cross-validation result according to a k-fold cross validation should then lead to the best classification results for the specific problem.

#### 4.3.9 Validation of the Support Vector Machine

The C-SVM was trained and tested by the leave-one-subject-out approach. The complete data for a subject are removed from the training set and the trained classifier is then tested on the data from this subject. For each valley  $i$  the corresponding subject  $subject_i \in \{1, \dots, N_{subject}\}$  is denoted, where  $N_{subject}$  is the number participating subjects. The training parameters  $C$  and  $\gamma$  are kept constant, in order to have equal condition for each subject  $s$ .

Thus, the training set which excludes all data from subject  $s$  becomes

$$\{(\mathbf{x}_i, y_i, subject_i \neq s)\}, \quad i = 1, \dots, l. \quad (4.69)$$

The training set contains  $l$  valleys. The remaining valleys are in the testing set

$$\{(\mathbf{x}_j, y_j, subject_j = s)\}, \quad j = 1, \dots, m. \quad (4.70)$$

The testing set consists of  $m$  valleys.

The prediction  $y_j^{pred}$ , which is obtained by applying the trained C-SVM on the test set, is compared with the true class membership  $y_j$  for all valleys which are recorded from subject  $s$ . A confusion matrix

$$cm_s = \begin{pmatrix} TP & FN \\ FP & TN \end{pmatrix} \quad (4.71)$$

according to Table 4.3 is calculated.

Training and testing is performed for all subjects by the leave-one-subject-out approach and the resulting confusion matrix  $cm$  is the sum of all obtained confusion matrices from all subjects:

$$cm = \sum_{s=1}^{N_{subject}} cm_s. \quad (4.72)$$

Table 4.3 shows how sensitivity, specificity, and accuracy can be calculated from the finally obtained confusion matrix *cm*.

## 4.4 Results

### 4.4.1 Data Collection

Healthy subjects without any swallowing disorder and patients with oropharyngeal swallowing disorder were examined at the Technische Universität Berlin and the Unfallkrankenhaus Berlin in collaboration with Rainer O. Seidl and Corinna Schultheiss from the Unfallkrankenhaus Berlin<sup>9</sup>. The exclusion criteria were: pregnancy, implanted cardiac pacemaker or defibrillator, metallic implants, or central venous catheters. All subjects had given informed consent and were informed about the used methods. This study was approved by the ethic board at the Charité Berlin (EA1/019/10 and EA1/161/09).

**Table 4.4:** Number of recorded swallows, movements, recordings, and complete recording duration of the datasets I, II, III, and IV.

Dataset	Subjects	Mean age	Swallows	Movements	Duration
I	20 (12 ♂, 8 ♀)	30.5 ± 7.7	862	331	3.70 h
II	15 (11 ♂, 4 ♀)	29 ± 4.5	1912	0	7.13 h
III	9 (2 ♂, 7 ♀)	38.6 ± 9.4	139	0	0.21 h
IV	41 (26 ♂, 15 ♀)	63.4 ± 13.8	748	0	2.51 h

20 healthy subjects participated in the first study yielding dataset I, in which the effects of head and tongue movements, electric conductivity of liquid, and different bolus volumes on BI measurements have been investigated. One subgroup of the participating subjects was asked to swallow 20 ml water while their head was oriented in a specific position, move their head without swallowing, speak, and apply pressure with their tongue and jaw. A second subgroup was asked to swallow saliva and 200 ml of liquid in 20 ml portions. A third subgroup swallowed saliva, eat yogurt and bread in 5 g portions. A fourth subgroup of subjects should swallow liquid in 5 ml, 10 ml, 20 ml, and 30 ml portions. Finally, a fifth subgroup swallowed water with different conductivity.

For the second dataset II, 15 subjects should drink 200 ml of water in their own speed. Ten subjects of dataset II underwent four measurements within one day while using newly placed electrodes for each measurement. Four measurements were conducted on four successive days on all 15 subjects of dataset II.

<sup>9</sup>The collected measurement data are also described in the doctoral thesis of Corinna Schultheiss [135] and a related paper [147]. For the analysis in [135] and [147], the collected measurements were segmented by the heuristic valley search algorithm presented in this thesis (cf. Section 4.3.6), and swallowing related valleys were selected based on the manually added annotations. Time, value and area based features which have been described in Section 4.3.7 have been statistically analyzed, to evaluate the reproducibility of the measurements, differences between swallowing and head movements, and the influence of various factors such as gender, bolus amount, consistency, conductivity, repeatability, and intra-rater reliability.

Nine healthy subjects were included in the third dataset III which was recorded by four different investigators on every subject in order to compare the influence of different investigators. The participants swallowed water with a bolus size of 20 ml.

In datasets I to III, each swallow or movement was manually marked by a hand switch and labeled by the investigator. The mean age and the gender of the participating subjects can be seen in Table 4.4. In total, 31 healthy subjects<sup>10</sup> (15 female, 16 male) were examined in the datasets I, II, and III. The age of the healthy subjects ranged from 24 to 51 years while the mean age was  $32.5 \pm 7.8$ . 2913 swallows and 331 non-swallowing related movements were recorded and labeled resulting in a complete recording time of 11 hours and 4 minutes.

41 patients<sup>11</sup> (15 female and 26 male) suffering from swallowing disorder were recruited for the dataset IV. The age was in a range from 24 to 93 years (mean age  $63.4 \pm 13.8$ ). 24 patients suffered from neurological diseases while 17 had ear, nose, and throat (ENT) specific disorders. Depending on the ability of the patient, swallows of saliva, small portions of dyed water, green jelly, or bread were recorded. Patients were evaluated by FEES in an upright position in combination with BI and EMG measurements. A physician, proficient in FEES, evaluated all endoscopic examinations, marked all swallows in the BI and EMG measurements and mapped the swallows into swallows with and without penetration<sup>12</sup>. In total, 748 swallows were recorded which consists of 163 times saliva, 239 times a small portion of colored water, 242 times a semisolid bolus, and 99 times a solid bolus. Additionally, five clearing post-swallows were marked. From these 748 swallows, 85 were marked by the investigator as swallows with penetration.

#### 4.4.2 Detection of EMG Activity Periods

The double-threshold detector (cf. Section 4.3.4) was used for automatic EMG activity detection in all datasets. The desired offset-detection latency was set to  $tr_{max} = 10$  ms and a  $P_{fa}$  of 0.01 was requested, assuming a minimal SNR  $SNR_{min}$ <sup>13</sup> of 3 dB.

In order to achieve a high accuracy, noise, and disturbance variances have been estimated individually for each recorded EMG measurement. In Section 4.3.4.6, a synthetic benchmark was used for evaluating the performance of the detector. Based on this analysis, the parameters of the noise and disturbance estimation were set to  $N_{\hat{\sigma}_n^2} T_s^{EMG} = 0.2$  s and  $N_{\hat{\sigma}_d^2} T_s^{EMG} = 8$  s. In swallowing activity recordings a value of  $th_{\hat{\sigma}_d} = 0.25$ <sup>14</sup> produced good results and is therefore used for all datasets.

According to Section 4.3.4.4, optimization of the double-threshold detector yields  $m = 56$  and  $r_0 = 8$ .

<sup>10</sup>The 31 healthy subjects are described in the thesis of C. Schultheiss [135] in table B.4.

<sup>11</sup>The patients are described in the thesis of C. Schultheiss [135] in table B.5 and are named by the following codes: 113, 114, 115, 116, 117, 118, 131, 132, 133, 138, 139, 141, 142, 143, 144, 147, 148, 150, 152, 153, 154, 155, 158, 159, 165, 166, 167, 176, 194, 198, 199, 200, 201, 203, 205, 206, 208, 209, 210, 211, and 212.

<sup>12</sup>Penetration describes the observation that the bolus enters the glottis up to the vestibule above the true vocal folds. The act of a bolus passing the true vocal folds is described as aspiration. Aspiration could not be observed in any patient and is not considered.

<sup>13</sup>The worst-case SNR is  $SNR_{min} = 10 \log_{10} \left( \frac{\sigma_s^2}{\sigma_d^2 + \sigma_n^2} \right)$  for which  $SNR_{min} > 0$  must be hold in order to be able to detect EMG activity.

<sup>14</sup>The percentage share  $th_{\hat{\sigma}_d}$  of signal parts with disturbances and without real EMG activity is defined a-priori based on the observed activity.

#### 4.4.3 Piece-wise Linear Approximation of BI

**Table 4.5:** Datasets I, II, III, and IV were used for comparing the influence of  $max\_error$  on the PLA of the BI signal. The root mean square error is calculated between the original and the approximated BI time series between all manually marked swallows. The mean compression  $\frac{N^{BI}}{2N^P}$  show the data point reduction by the PLA.

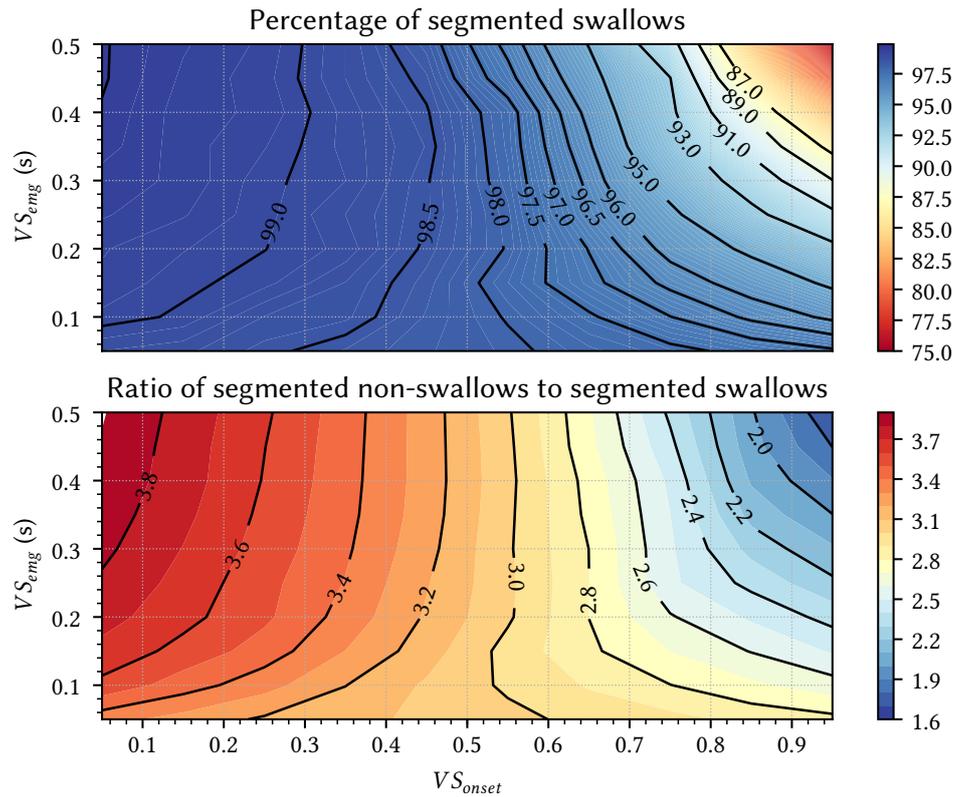
$max\_error$	$\frac{N^{BI}}{2N^P}$	Maximum root mean square error
0.1	$34.76 \pm 12.38$	$0.06 \pm 0.02 \Omega$
0.5	$54.02 \pm 20.28$	$0.11 \pm 0.03 \Omega$
1.0	$66.15 \pm 26.04$	$0.14 \pm 0.04 \Omega$
2.5	$87.42 \pm 36.12$	$0.18 \pm 0.06 \Omega$
5.0	$109.00 \pm 46.37$	$0.23 \pm 0.08 \Omega$
8.5	$130.28 \pm 55.99$	$0.27 \pm 0.09 \Omega$
10.0	$137.65 \pm 59.16$	$0.28 \pm 0.10 \Omega$
15.0	$159.06 \pm 69.05$	$0.32 \pm 0.11 \Omega$
25.0	$191.96 \pm 85.14$	$0.36 \pm 0.14 \Omega$
50.0	$251.31 \pm 113.90$	$0.43 \pm 0.17 \Omega$

The parameter  $max\_error$  influences the number of the segments which were obtained by PLA on the filtered and de-noised BI measurement. The number of obtained line segments influences the number of found valleys. More segments means that it is more likely that more valleys are found by the subsequent heuristic valley search. All valleys that were found by the valley search must then be classified into swallows and non-swallowing events. A higher number of valleys, whereas the number of swallow related valley remains constant, increases the demands on the later applied classifier. Thus, it is important to keep the number of found valleys as small as possible without excluding valleys which are caused by swallowing.

It is known from previous investigations that a BI valley has a difference of  $1.52 \pm 0.56 \Omega$  between the start point and the minimum point [147]. Thus, the maximal error between the original BI time series and the approximated time series must be smaller than  $1 \Omega$ . In Table 4.5, different  $max\_error$  values have been investigated on the datasets I, II, III, and IV. For each parameter value, the compression and the maximum root mean square error between the original and the approximated time series for all manually marked swallows are shown. A good compromise for a compression rate above 100 and a root mean square error below  $0.25 \Omega$  is achieved for  $max\_error = 5$ .

#### 4.4.4 Segmentation of BI Data by a Heuristic Valley Search

After finding and segmenting all valleys in the BI time series, it is not known which of these valleys are related to swallowing. A subsequently applied classification has to reject all non-swallowing related valleys in order to be able to automatically detect swallows with high specificity. The classification accuracy can be improved a-priori by reducing the number of segmented non-swallowing events. The number of segmented swallows and non-swallowing events depends



**Figure 4.19:** Influence of  $VS_{onset}$  and  $VS_{emg}$  on segmentation is exemplarily shown on measurement data from datasets I, II, and III. On the first sub-plot, the percentage of correctly segmented swallows is plotted. The ratio of segmented non-swallows to segmented swallows is shown on the second sub-plot.

strongly on  $VS_{onset}$ <sup>15</sup> and slightly on  $VS_{emg}$ <sup>16</sup>. In Figure 4.19, the percentage of segmented swallows and the ratio of non-swallowing events to swallows are plotted depending on  $VS_{onset}$  and  $VS_{emg}$ . Datasets I, II, and III were used for this comparison. The same procedure was repeated on the dataset IV and is shown in Figure 4.20.

It can be seen that a higher  $VS_{onset}$  almost linearly reduces the ratio of non-swallowing valleys to swallowing valleys. On the other hand, a higher  $VS_{onset}$  value leads to a lower number of successfully segmented swallows. A good trade-off between a low number of non-swallow events and a low number of missed swallows is achieved for  $VS_{onset} = 0.3$  and  $VS_{emg} = 0.35$  for all datasets I, II, III, and IV. This combination leads to approximately 1% of missed swallows and reduces the ratio of non-swallow events to swallow to 3.5 for healthy subjects. For patients (dataset IV), 5% of missed swallows and a ratio of non-swallow events to swallow of 4.05 is obtained by the selected parameters.

The parameters  $VS_{diff}$  and  $VS_{bound}$  were not further optimized. The parameter  $VS_{diff}$  is heuristically set to 0.5, which leads successfully segmentation of all swallowing related valleys. It is not likely that varying this parameters lead to a different number of segmented swallows or reduces

<sup>15</sup>The parameter  $VS_{onset}$  defines the minimal EMG activity percentage in the beginning of a valley.

<sup>16</sup>The parameter  $VS_{emg}$  describes the width of the interval around the beginning of a valley in which EMG activity must be present.

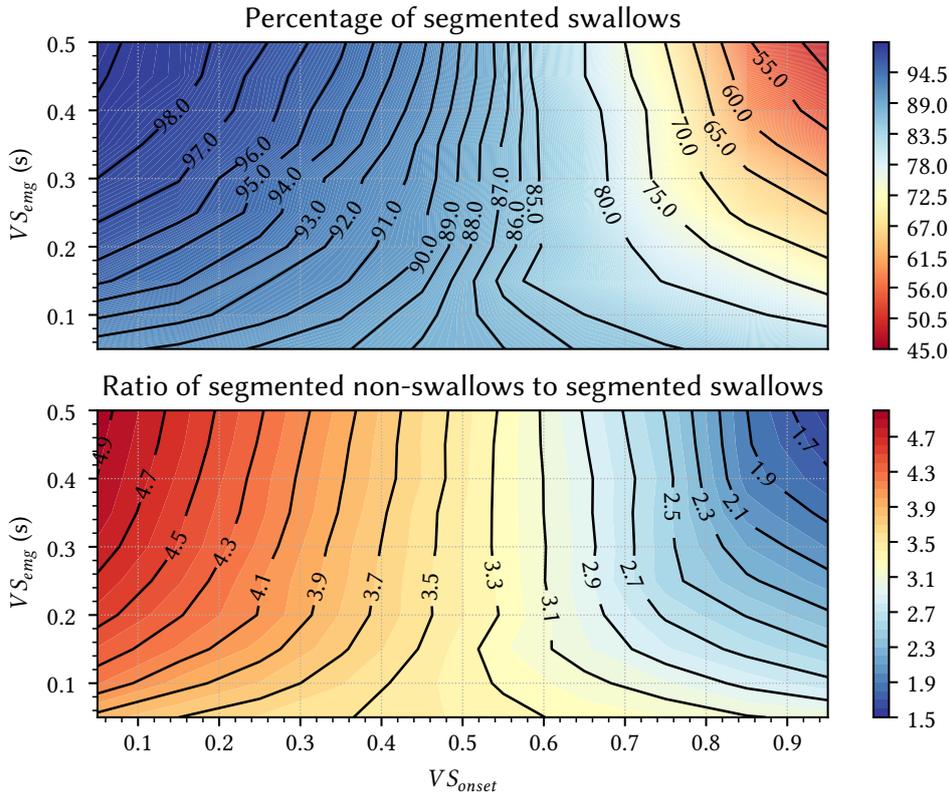


Figure 4.20: Influence of  $VS_{onset}$  and  $VS_{emg}$  on segmentation is exemplarily shown on measurement data from dataset IV. On the first sub-plot, the percentage of segmented swallows is plotted. The ratio of the number of segmented non-swallows to segmented swallows is shown on the second sub-plot.

the number of non-swallowing events. This can also be assumed for the lower and upper bound parameter  $VS_{bound} = [\underline{vs}, \overline{vs}]$ <sup>17</sup> which is set to  $\underline{vs} = 0.2$  s for the lower bound and to  $\overline{vs} = 3.6$  s for the higher bound. This chosen parameter values of  $VS_{bound}$  leads to inclusion of all recorded swallows.

#### 4.4.5 Swallowing Segmentation

The segmentation algorithm with the parameter set listed in Table 4.6 was applied to all datasets. 2879 (98.8 %) from 2913 marked swallows could be successfully segmented from the datasets I, II, and III which contains the measurement data from healthy subjects. Simultaneously, 9762 non-swallowing related valleys are included as segmented valleys. Within the dataset IV (containing data from patients), 705 (94.3 %) from 748 marked swallows could be segmented, while 2894 non-swallowing valleys were also segmented by the algorithm.

The numbers show that almost all swallows are included by the segmentation. The results show also that a swallow segmentation based on EMG and BI measurement is not sufficient in separating swallows from non-swallowing events with EMG activity. Thus, a subsequent classification in which the segmented valleys are classified into swallows and non-swallows is necessary.

<sup>17</sup>The time duration of a swallowing valley is within a given bound  $VS_{bound} = [\underline{vs}, \overline{vs}]$  which consists of a lower bound  $\underline{vs}$  and an upper bound  $\overline{vs}$ .

**Table 4.6:** List of the parameters for the swallowing segmentation of datasets I, II, III, and IV.

Name	Default	Reason for parameter choice
$tr_{max}$	10 ms	desired offset-detection latency of the double-threshold detector
$P_{fa}$	0.01	results in a sufficiently high detection probability $P_d$
$SNR_{min}$	3 dB	minimal EMG that shall be detected
$N_{\hat{\sigma}_n^2} T_s^{EMG}$	0.2 s	sufficient window length for estimating $\hat{\sigma}_n^2$
$N_{\hat{\zeta}} T_s^{EMG}$	8 s	sufficient measurement data length for estimating $\hat{\zeta}$
$th_{\hat{\sigma}_d^2}$	0.25	expected percentage of periods in the EMG signal without activity
$max\_error$	5	good ratio between approximation error and compression
$VS_{bound}$	[0.2, 3.6] s	decreases the number of non-swallow valleys
$VS_{diff}$	0.5	leads to a reasonable valley end point
$VS_{emg}$	0.35 s	results in a high percentage of segmented swallows
$VS_{onset}$	0.3	results in a high percentage of successfully segmented swallows

#### 4.4.6 SVM and Feature Parameter Optimization

A C-SVM with RBF kernel was trained on manually categorized valleys. The hand switch for the datasets I, II, and III and the FEES for the dataset IV were evaluated and a group vector was created for all valleys by which all valleys have been classified into swallows and non-swallowing events.

Remaining parameters which have to be optimized and are not already defined in Table 4.6 are the parameters of the SAX features<sup>18</sup> and the learning parameter<sup>19</sup> of the C-SVM. As this represents a nonlinear multi-dimensional optimization problem which is not solvable within an acceptable time frame by a complete grid search approach, the following strategy was applied to find an approximate solution.

At first, the best parameter set (word length and alphabet size) for the SAX based feature generation were estimated by a grid search by using default learning parameters for the C-SVM. Due to the unbalanced data and as the learning parameters are not yet optimized, the area under the ROC curve (AUC) metric should be used for evaluating the classifier performance, as proposed in [170].

Then the obtained SAX parameter set was used to generate the feature vectors for all valleys and the learning parameters were optimized by sensitivity in a second grid search.

The ROC curve is a two-dimensional measure of the classifier performance by which the true positive rate is plotted against the false positive rate. Each point in the ROC curve represents one realization of the trained classifier by which the hyperplane of the C-SVM is shifted by varying its offset. The most commonly used measure for characterizing the ROC curve is area under the ROC curve (AUC). A maximum value of one can be achieved by the AUC measure which means that the

<sup>18</sup>From each segmented valley, the BI time series are transformed into a SAX word with the specific SAX word length  $N_{PAA}^{BI}$  and alphabet size  $N_{sym}^{BI}$ . Within the same time interval, the EMG time series is also transformed into a SAX word with length  $N_{PAA}^{EMG}$  and alphabet size  $N_{sym}^{EMG}$ . These parameters directly influence the feature vector which is generated for each valley.

<sup>19</sup>The learning parameters  $C$  and  $\gamma$  are influencing the training process of the classifier and have to be optimized individually for each problem.

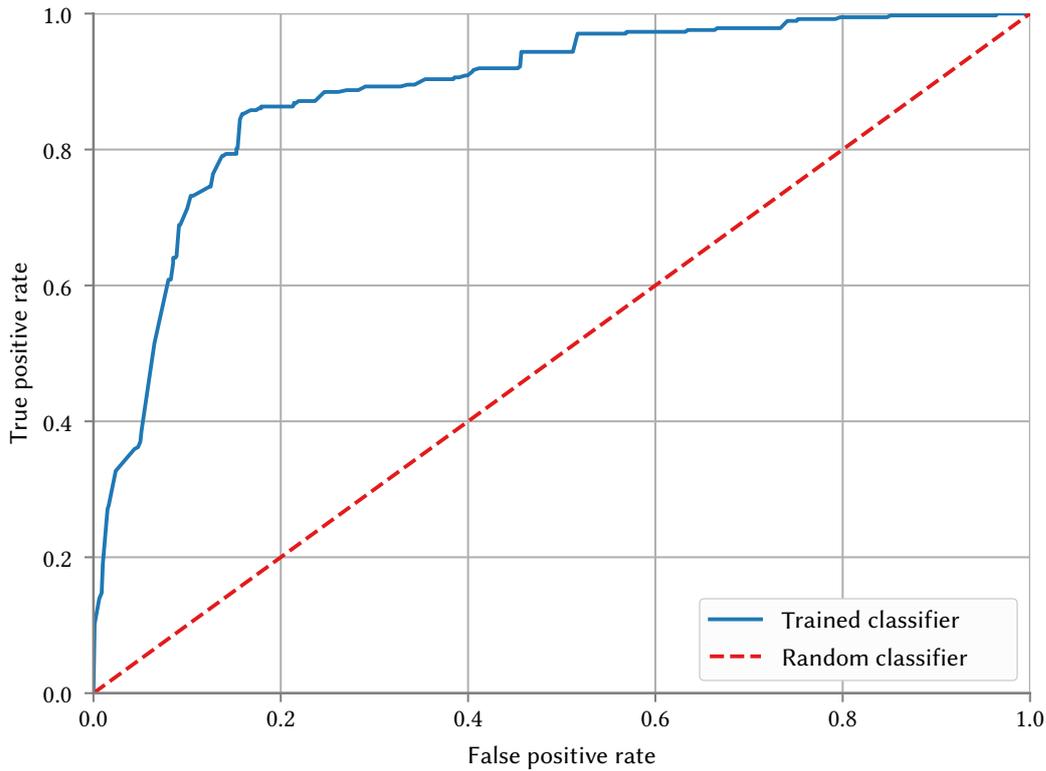


Figure 4.21: Shown is an exemplary ROC curve of a trained classifier. The straight ROC curve of a random classifier is shown as comparison.

classification error is zero. The worst performance is achieved by a random classifier which leads to a AUC of 0.5. An exemplary ROC curve is shown in Figure 4.21.

For the first grid search, default values  $C = 1$ ,  $\gamma = 1/\text{num\_features}$  for training the C-SVM have been used, where  $\text{num\_features}$  is the dimension of the feature vector. As the swallowing and the non-swallowing class were unbalanced, the weighting parameters within the C-SVM were set to the inverse frequency of the corresponding class according to the Equations (4.67) and (4.68). A five-fold cross-validation<sup>20</sup> was carried out for  $N_{PAA}$  in a range from 5 to 85 and for  $N_{sym}$  in a range from 4 to 64. The word length and the alphabet size were independently set for the SAX words and all possible combinations were tested. The AUC was then calculated from the mean ROC curve which has been obtained for each specific SAX parameter configuration. The SAX parameter configuration with the highest AUC is then selected. The highest AUC was 0.95728 and was obtained with the SAX parameters  $N_{PAA}^{BI} = 40$ ,  $N_{PAA}^{EMG} = 40$ ,  $N_{sym}^{BI} = 8$ , and  $N_{sym}^{EMG} = 16$ . The ten best SAX parameter combinations are shown in Table 4.7. The resulting SAX words of all valleys from the healthy subjects (datasets I, II, and III) are shown in Figure 4.23. The same SAX parameters were applied to the patient dataset IV, which is shown in Figure 4.24.

<sup>20</sup>The dataset is randomly divided into five equal parts with the same number of swallows in it. Then, a C-SVM is evaluated five times on different combinations using four parts for training and the remaining data part for testing. For each trained C-SVM, a ROC curve is calculated. Finally, a mean ROC is calculated from all five testing results.

**Table 4.7:** Cross-validation AUC for the ten best SAX parameter combinations. Data from datasets I, II, III, and IV were used.

$N_{PAA}^{BI}$	$N_{sym}^{BI}$	$N_{PAA}^{EMG}$	$N_{sym}^{EMG}$	AUC
40	8	40	16	0.95728
40	8	25	8	0.95724
40	8	55	8	0.95720
40	16	25	8	0.95719
25	16	25	8	0.95717
25	8	25	8	0.95716
25	64	25	8	0.95715
40	64	25	8	0.95708
40	32	25	8	0.95708
25	32	25	8	0.95708

After obtaining optimal SAX parameters, it was evaluated if the complete feature vector leads to the best results. In addition to both SAX words of the BI and EMG sequence, 17 further features which are based on certain areas, timing intervals, slopes and amplitudes of the extracted valley and the corresponding EMG activity were used (cf. Section 4.3.7). In Table 4.8, the contribution of these additional used features is reviewed by a five-fold cross-validation on the datasets I, II, III, and IV. The AUC was again used as performance indicator. The best results were obtained by using the complete feature vector which has a length of 97 entries.

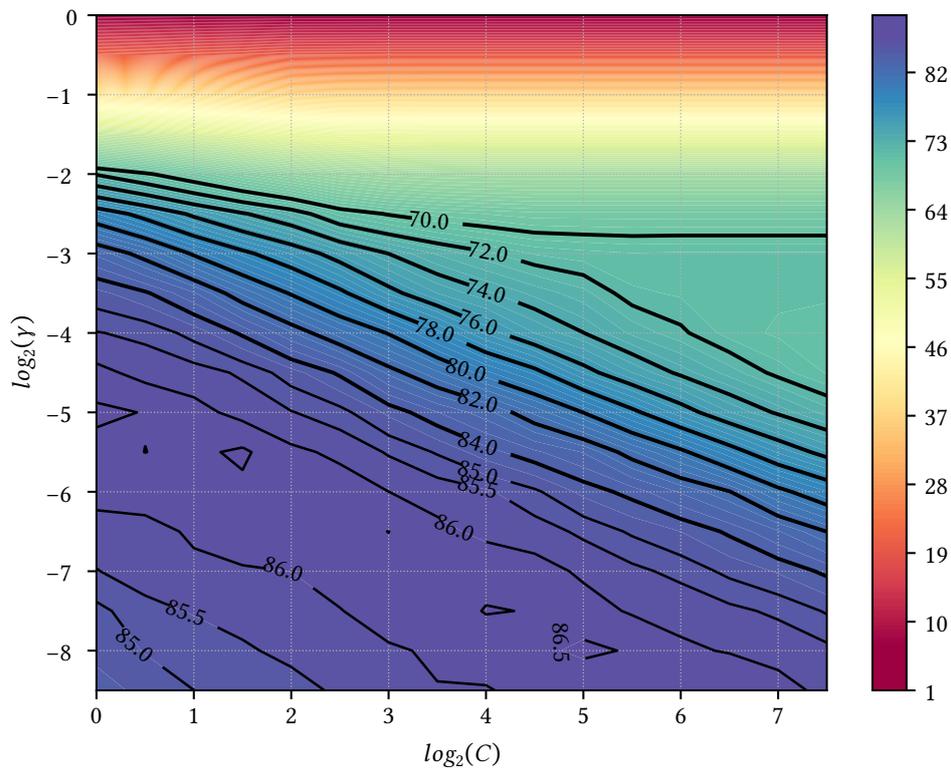
**Table 4.8:** Cross-validation AUC for different feature vectors. Data from datasets I, II, III, and IV were used. SAX parameters were set to  $N_{PAA}^{BI} = N_{PAA}^{EMG} = 40$ ,  $N_{sym}^{BI} = 8$ , and  $N_{sym}^{EMG} = 16$ . The learning parameter were set to default values  $C = 1$  and  $\gamma = 1/num\_features$ . Included features are marked by an “x”.

SAX - BI	SAX - EMG	Valley features	ROC AUC
x	x	x	0.95728
x	x		0.95046
	x	x	0.94062
x		x	0.92638
	x		0.91718
x			0.91325
		x	0.87763

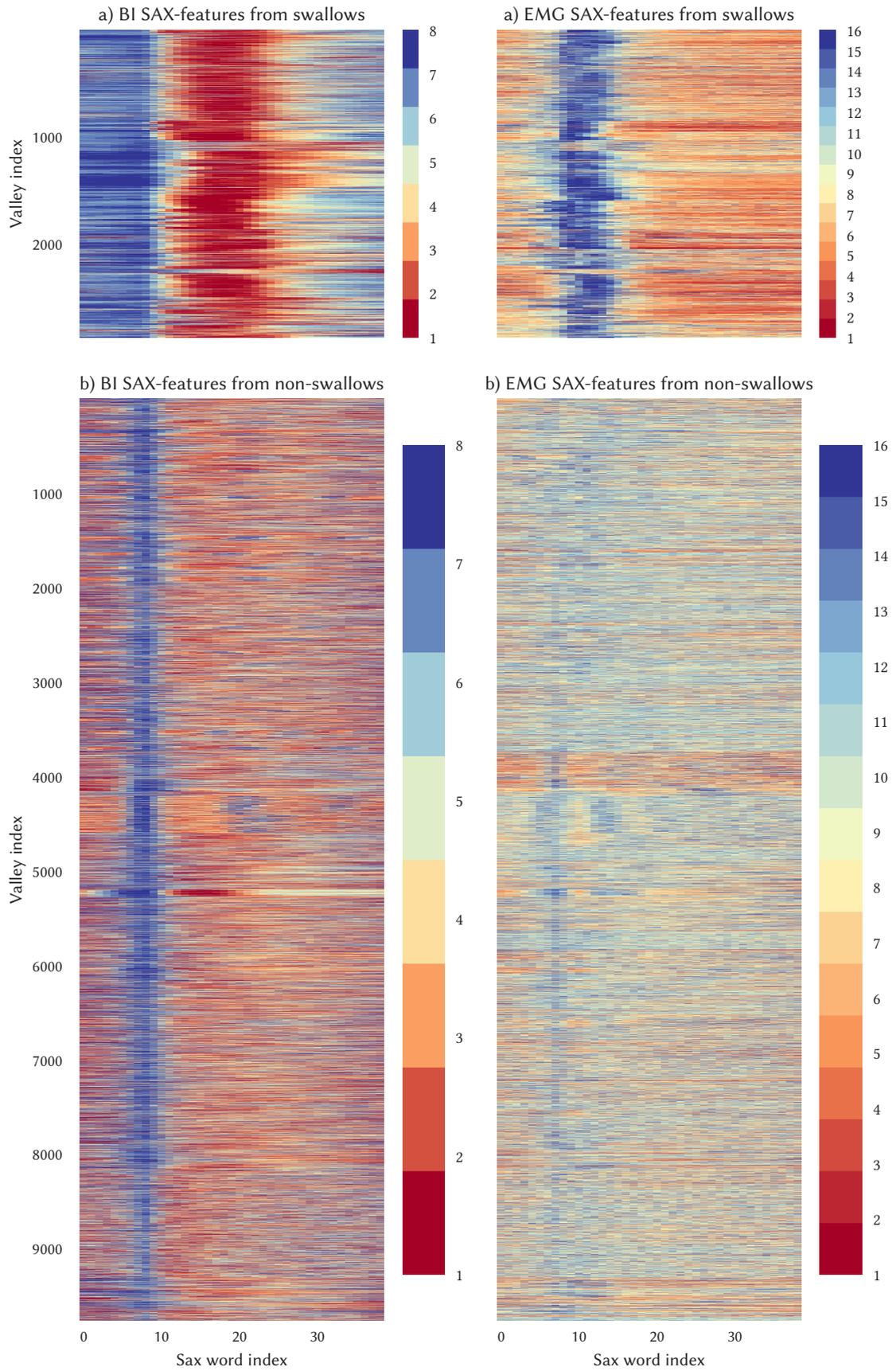
The second grid search was performed on the learning parameters  $C$  and  $\gamma$  of the C-SVM with RBF kernel. As the learning parameters are optimized in this grid search, sensitivity was used as a performance indicator for the classifier. A five-fold cross-validation was carried out for each parameter pair and the mean sensitivity was calculated. Finally, the parameter set was selected which leads to the highest sensitivity value. Based on some exemplary performed training and testing sessions, it was decided to search for  $C$  in a range from  $2^0$  to  $2^{12}$  and for  $\gamma$  in a range from  $2^{-12}$  to  $2^0$ . The best performance was achieved for  $C = 2^{1.5}$  and  $\gamma = 2^{-5.5}$ . The results of the grid search are shown in Figure 4.22.

**Table 4.9:** SAX and training parameters used for swallowing classification of datasets I, II, III, and IV. The weighting parameters  $W^+$  and  $W^-$  are set to the inverse frequency of the corresponding class according to the Equations (4.67) and (4.68).

SAX parameters			C-SVM parameters	
	BI	EMG		
$N_{PAA}$	40	40	$C$	$2^{1.5}$
$N_{sym}$	8	16	$\gamma$	$2^{-5.5}$
			$W^+$	1.559
			$W^-$	0.441

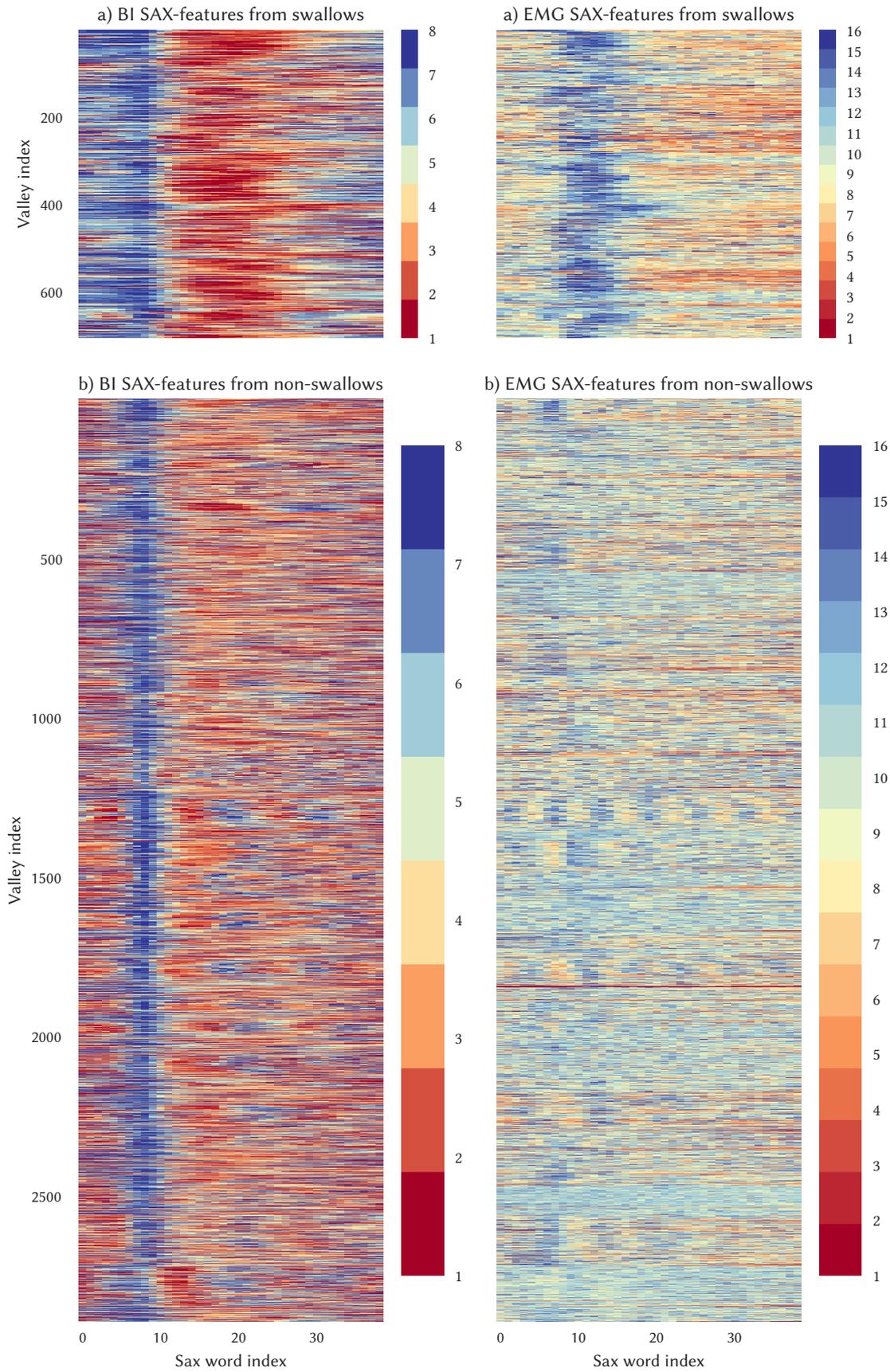


**Figure 4.22:** Grid search for finding the best learning parameters for datasets I, II, III, and IV. The mean sensitivity of a five-fold cross-validation is shown for different learning parameters  $C$  and  $\gamma$ .



**Figure 4.23:** SAX-features for EMG and BI of 2879 swallows (a) and 9762 non-swallowing events from datasets I, II, and III ( $N_{PAA}^{BI} = 40$ ,  $N_{PAA}^{EMG} = 40$ ,  $N_{sym}^{BI} = 8$ , and  $N_{sym}^{EMG} = 16$ ).

#### 4. Automatic Segmentation and Classification of Swallowing



**Figure 4.24:** SAX-features for EMG and BI of 705 swallows (a) and 2894 non-swallowing events (b) from dataset IV. ( $N_{PAA}^{BI} = 40$ ,  $N_{PAA}^{EMG} = 40$ ,  $N_{sym}^{BI} = 8$ , and  $N_{sym}^{EMG} = 16$ ).

#### 4.4.7 Classification Performance

As described in Section 4.3.6, valleys were extracted from the recorded EMG and BI time series by the presented heuristic valley search algorithm. Although almost all swallow related valleys could be segmented, several valleys with EMG activity which are not related to swallowing were also segmented. This classification problem is solved by applying the trained C-SVM to the feature vectors of all valleys which are classified into swallowing related valleys and non-swallowing related valleys. The used parameters for feature calculation and training the C-SVM are summarized in the Tables 4.6 and 4.9.

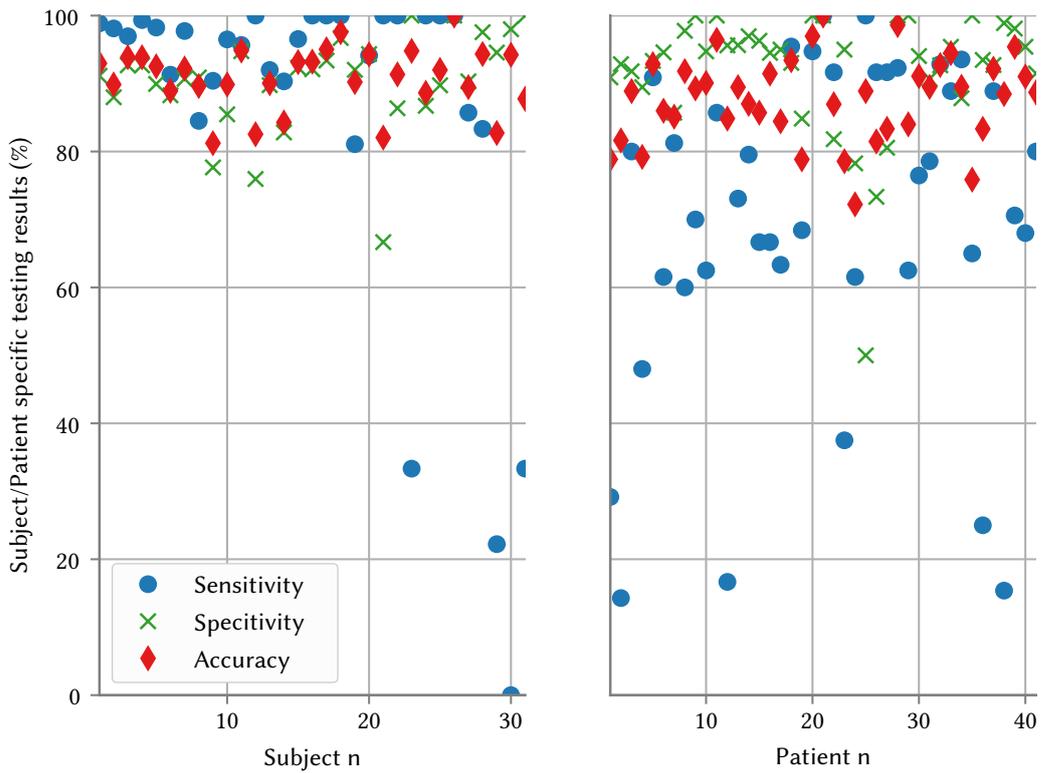
The performance of the classifier was evaluated by the leave-one-subject-out approach. The leave-one-subject-out approach has the advantage that all patients/subjects are equally used to evaluate the performance of a classifier for handling unknown measurement data. In this approach, all valleys of one patient/subject are excluded from training and exclusively used for testing. This was repeated for all patients/subjects (datasets I, II, III, and IV) and the test results of all the individually trained C-SVM were summed up in a confusion matrix<sup>21</sup>.

The classification results are shown as confusion matrix in Table 4.10. The results of all patients/subjects were obtained by applying the leave-one-subject-out approach on the datasets I, II, III, and IV. The testing results for all datasets as well as parts of the testing results containing only healthy subjects or only patients are presented as confusion matrices. The leave-one-subject-out approach was also applied exclusively on the dataset IV in order to evaluate the performance of the classifier and its features when only patient data are used for training. The classification results are also plotted individually for each patient/subject in Figure 4.25.

**Table 4.10:** Performance of the classifier on the datasets I, II, III, and IV. The results are obtained by the leave-one-subject-out approach by which every time the data from all datasets except one patients/subjects are used for training. The confusion matrices are then the sum of all testing results. The confusion matrix containing the number of correctly classified swallows (Sw.) and non-swallows (Non-Sw.) is shown for the complete data, only for the data of the healthy subjects (datasets I - III) and twice for the data of the patients (dataset IV). The classifier was trained by the leave-one-subject-out approach on all datasets (I - IV) and only on the dataset IV. The results are obtained with the parameters defined in the Tables 4.6 and 4.9.

Trained on:		I, II, III, and IV	I, II, III, and IV	I, II, III, and IV	IV				
Tested on:		I, II, III, and IV	I, II, and III	IV	IV				
		Predicted							
		Sw.	Non-Sw.	Sw.	Non-Sw.	Sw.	Non-Sw.	Sw.	Non-Sw.
Actual	Sw.	3219	365	2742	137	477	228	577	128
	Non-Sw.	1082	11574	909	8853	173	2721	379	2515
Sensitivity:		89.9 %		95.2 %		67.7 %		81.8 %	
Specificity:		91.5 %		90.7 %		94.0 %		86.9 %	
Precision:		74.8 %		75.1 %		73.4 %		60.4 %	
Accuracy:		91.1 %		91.7 %		88.9 %		85.9 %	

<sup>21</sup>The number of correctly classified swallows (TP), correctly classified non-swallows (TN), wrongly as swallow classified non-swallows (FP), and wrongly as non-swallow classified swallows (FN) are calculated for each tested patient/subject by comparing predicted and actual swallows and non-swallows. The confusion matrix is then obtained by counting TP, TN, FP, and FN across all tested patients/subjects.



**Figure 4.25:** The sensitivity, specificity, and accuracy are plotted for each subject/patient. The results are obtained by testing the data of one subject/patient on the C-SVM which is trained on the datasets I, II, III, and IV (see Table 4.10). The testing results for datasets I, II, and III are shown in the left subplot. The results for testing the dataset IV by the leave-one-subject-out approach are shown in the right subplot.

The testing results for the different types of recorded swallows and movements for the healthy subjects of datasets I, II, and III are summarized in Table 4.11. Beside the number of recorded and segmented events, the testing output of the trained classifier is shown by the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) counts.

The number of correctly segmented and classified swallows for dataset IV trained on datasets I, II, III, and IV is shown for all different kind of swallowing consistencies in Table 4.12. The results for training exclusively on the dataset IV are shown in Table 4.13. The different swallowing consistencies are divided into swallows in which an expert determined a penetration and into swallows without penetration. Beside the number of recorded and segmented events, the testing output of the trained C-SVM is shown by the true positive (TP) and false negative (FN) counts.

**Table 4.11:** Performance of the classifier on the datasets I, II, and III (healthy subjects). The results were obtained by the same parameter set and classifier as used in Table 4.10. Beside the number of recorded and segmented events, the testing output of the trained classifier is shown by the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) counts.

Type	Recorded	Segmented	C-SVM trained on datasets I - IV			
			TP	TN	FN	FP
Saliva	928	925	856	0	69	0
Clearing post-swallow	152	146	135	0	11	0
5 ml water	28	28	28	0	0	0
10 ml water	95	95	94	0	1	0
20 ml water	922	898	867	0	31	0
30 ml water	29	29	29	0	0	0
Yogurt	397	397	387	0	10	0
Bread	71	71	66	0	5	0
20 ml water at different conductibility	131	131	126	0	5	0
Swallowing in different head positions	120	120	118	0	2	0
Mendelsohn maneuver	40	39	36	0	3	0
Sum swallowing	2913	2879	2742	0	137	0
Speaking	28	28	0	28	0	0
Head movements	219	63	0	51	0	12
Tongue / jaw movements	84	27	0	21	0	6
Sum movements	331	118	0	100	0	18

#### 4. Automatic Segmentation and Classification of Swallowing

**Table 4.12:** Achieved performance of the trained classifier for different bolus types of dataset IV. The results were obtained by the same parameter set and the classifier which was trained on the datasets I, II, III and IV as used in Table 4.10. The different consistencies are grouped into swallows where a penetration occurred and into swallows where no penetration was observed by the expert. The testing output of the classifiers is shown by the true positive (TP) and the false negative (FN) counts.

Type	Recorded	Segmented	C-SVM trained on datasets I - IV	
			TP	FN
Clearing post-swallow	5	5	2	3
Saliva	130	115	83	32
Water	226	211	143	68
Semisolid	204	197	135	62
Bread	98	97	64	33
Sum of swallows without penetration	663	625	427	198
Saliva with penetration	33	31	13	18
Water with penetration	13	13	12	1
Semisolid with penetration	38	35	24	11
Bread with penetration	1	1	1	0
Sum swallows with penetration	85	80	50	30

**Table 4.13:** Performance for different bolus types of the classifier trained and tested on the dataset IV by the leave-one-subject-out approach. The results were obtained by the same parameter set and the classifier which was trained on the dataset IV as used in Table 4.10. The different consistencies are grouped into swallows where a penetration occurred and into swallows where no penetration was observed by the expert. The testing output of the classifiers are shown by the true positive (TP) and the false negative (FN) counts.

Type	Recorded	Segmented	C-SVM trained on dataset IV	
			TP	FN
Clearing post-swallow	5	5	3	2
Saliva	130	115	99	16
Water	226	211	172	39
Semisolid	204	197	159	38
Bread	98	97	81	16
Sum of swallows without penetration	661	625	514	111
Saliva with penetration	33	31	21	10
Water with penetration	13	13	13	0
Semisolid with penetration	38	35	28	7
Bread with penetration	1	1	1	0
Sum swallows with penetration	85	80	63	17

## 4.5 Conclusions

It is shown that BI and EMG measurement data which are recorded at the neck can be used for automatic detection of swallowing events. Recorded data from healthy subjects and patients who suffered from oropharyngeal dysphagia are used for evaluating the developed algorithms. In total, several hours of recorded EMG and BI time series from more than 70 different subjects could be successfully processed by the presented algorithms such that swallows could be found in the time series with high sensitivity and specificity.

Due to the observation that swallowing is initiated by EMG activity and followed by a deflection in the BI signal, the problem of finding all swallows in the measurement data is solved as follows: Firstly, EMG activity is detected. Then, all valleys in the BI measurement signal which coincide with EMG activity at the beginning are segmented. Finally, a C-SVM is used to distinguish valleys with EMG activity which are caused by swallowing from all other valleys.

EMG activity is detected by an optimally tuned double-threshold detector. The window length  $m$  and the second threshold  $r_0$  are estimated based on the desired false-alarm probability  $P_{fa}$ , a worst-case SNR  $SNR_{min}$ , and the desired offset-detection latency  $tr_{max}$ . On each recorded EMG measurement, noise, and disturbance variance are automatically estimated and the optimal threshold  $\hat{\zeta}$  is calculated. The feasibility of this approach was successfully demonstrated on a performance benchmark test.

In order to sufficiently and robustly segment all valleys in the BI signal, a PLA algorithm is applied on the BI signal. It is shown that this approximation method can sufficiently approximate the BI signal such that the dimensionality is reduced and the maximal error is lower than the expected mean drop ( $1.52 \pm 0.56 \Omega$ ) in the BI signal during a swallow. By setting the  $max\_error = 5$  in the PLA algorithm, the root mean square error is below  $0.25 \Omega$  which ensures that all swallows will be approximation by at least two segments.

A heuristic valley search algorithm is introduced which segments all valleys in the approximated BI time series. A valley consists of a starting line segment with negative slope and an ending line segment with a positive slope. It is also demanded that enough EMG activity around the starting point exists and that only one minimum exists in the approximated BI time series. Almost all valleys which are caused by swallowing were found. Besides, several valleys which do not belong to any swallow were also segmented.

It is shown that the ratio between segmented swallows and non-swallows is strongly influenced by the parameter  $VS_{onset}$  which defines how much EMG activity must be present in the beginning of a valley. By increasing this parameter, the ratio of the expected occurrence of non-swallowing related valleys to swallows could be reduced to 3.40 for healthy subjects and 4.11 for patients. The parameter  $VS_{onset}$  influences also the detection rate. For healthy subjects, the ratio of segmented swallows to all recorded swallows is only slightly reduced to 98.9%. For the patient dataset IV, the resulting ratio is 94.3%.

Finally, all segmented valleys are classified into swallows and non-swallows by a trained classifier. In order to achieve good classification performance, meaningful features are selected. Features are derived from the valley itself such as value and time differences, areas and slopes. Furthermore, features that are related to EMG activity as the area under the envelope and time

differences to the BI valley are used. It has been found that symbolic aggregate approximation (SAX) words from the BI and EMG time series around the corresponding valley are additional meaningful features. Each time series part is normalized and translated into a word with fixed word and alphabet size.

In order to calculate SAX words for the EMG time series, the method from Li et al. [163] was extended for signals with a  $\chi^2$  distribution like the squared and preprocessed EMG time series. Optimal alphabet sizes and word lengths have been determined by a grid search. From all used features, both SAX words are the most important features.

Segmented valleys which belong to swallowing are separated from non-swallowing related valleys by applying a trained classifier on the corresponding feature vectors. In this work, only a C-SVM with RBF basis function is evaluated for classifying swallow related valleys. A grid search for finding optimally learning parameter values have been performed. Both considered classes, swallows and non-swallows, are not well balanced. Weighting depending on the class frequency is therefore applied.

It was observed that the swallowing patterns are strongly correlated within the same subject/patient. For that reason, approaches which just divide the dataset into equal parts for training and testing would benefit from this intra-subject correlation. The same would be held for a simple 80 % / 20 % split where the unsolvable question remains which part of the dataset should be used for testing. In order to evaluate the performance of the C-SVM in a more fair way, the leave-one-out approach is modified to the leave-one-subject-out approach. An own classifier for each subject/patient without using the specific data from this patient/subject is trained and only tested on the data from this patient/subject. By repeating this for each patient/subject and summing up all results, it is evaluated how good the classifier can correctly classify swallow related valleys for an unknown subject/patient.

A C-SVM with RBF kernel and optimally chosen learning parameters and feature weights can classify swallows and non-swallowing related valleys with good sensitivity and specificity. By evaluating the datasets I, II, III, and IV (healthy subjects and patients) with the leave-one-subject-out, method a high sensitivity (88.9 %) and specificity (91.5 %) could be obtained. This balanced sensitivity and specificity could be achieved by weighting the feature vectors depending on the number of swallows and non-swallows. Due to the high number of non-swallowing related valleys (12656) in comparison to segmented swallows (3584), the obtained precision (74.8 %) is slightly lower than the sensitivity and the specificity. When the number of not successfully segmented swallows is taken into account (77 of 3661 swallows could not be successfully segmented and are not included in the classification), the sensitivity is slightly reduced from 89.9 % to 87.9 %.

The individual performance for each subject and patient is shown in Figure 4.25. The sensitivity for four healthy subjects and seven patients is below 50 %. For all other 27 subjects and 34 patients a better swallowing detection performance could be achieved. The achieved accuracy is at least 80 % for all subjects and 70 % for all patients.

The obtained classification results have been analyzed separately for the healthy subjects, for which a sensitivity of 95.2 % and a specificity of 90.7 % could be achieved. The sensitivity which is obtained when the missing swallows are taken into account (34 from 2913 could not be segmented)

remains high with 94.1%. In the following the sensitivity<sup>22</sup> for different kinds of swallowing is reported: saliva swallows (92.2 %), clearing post-swallows (88.8 %), 5 ml water swallows (100 %), 10 ml water swallows (98.9 %), 20 ml water swallows (94.0 %), 30 ml water swallows (100 %), yogurt (97.4 %), bread (93.0 %), swallows of water with different conductibilities (96.2 %), swallowing in different head positions (98.3 %), and Mendelsohn maneuvers (90.0 %). The sensitivity of all different types of swallows is at least 88 %, where clearing post-swallows and bread swallows show the least best performance. Due to the low number of recorded swallows for some types, 100 % could be reached.

A clearing post-swallow, which is observed in some subjects as an immediately executed second swallow, is not an entirely voluntary swallow. Some features depend on the EMG and the BI signal in the period before the beginning of a swallow. Thus, a clearing post-swallow may contain deviated features which differ from all other kinds of swallows and may therefore not be sufficiently detected by the classifier.

A Mendelsohn maneuver differs from all other kinds of swallows as the total swallowing time is significantly larger than normal swallowing time. Several features are affected by the length of a swallow. Thus, the sensitivity is expected to be lower than for detecting Mendelsohn maneuvers. Nevertheless, the observed performance of the C-SVM is surprisingly good in detecting Mendelsohn maneuvers.

Movements can be distinguished from swallows with a specificity of 94.6 %. All marked movements except twenty tongue and jaw moments could be correctly classified. Tongue and jaw movements lead to a deflection in the BI signal, but contrary to swallowing such movements lead to temporary rising within the BI signal. Between two peaks which are caused by tongue or jaw movements, a valley with EMG activity may segment in some cases. Thus, the specificity is slightly reduced for tongue and jaw movements.

For the dataset IV (patients) a sensitivity of 67.7 %, specificity of 94.0 %, and precision of 73.4 % could be achieved. Whereas the achieved precision and specificity are good, the sensitivity is low. The sensitivity drops down to 63.8 % when the results are compared to the total number of swallows and not only on the number of segmented swallows. The following sensitivity values are defined with respect to the total number of swallows. The C-SVM can classify swallows with penetration (sensitivity of 58.8 %) almost as good as swallows without penetration (sensitivity of 64.4 %). The different bolus consistencies are almost equally classified: Saliva (58.9 %), water (64.9 %), semisolid (65.7 %) and bread (65.7 %).

For the previously presented results, a C-SVM was trained on the complete datasets I, II, III, and IV except for the tested subject/patient. When only the dataset IV was used for training and testing, a sensitivity of 81.8 %, a specificity of 86.9 % and a precision of 60.4 % were achieved. Thus, the sensitivity was improved by using only data from dataset IV. When the number of swallows which could not be segmented is included, the sensitivity becomes 77.3 %. In the following, the sensitivity is defined with respect to the total number of swallows. Swallows with penetration and without penetration are again almost equally classified with a sensitivity of 77.6 % for swallows without penetration and 74.1 % for swallows with penetration. Saliva could be classified with a sensitivity of 73.6 %, water swallows with 77.4 %, semisolid swallows with 77.3 %, and bread swallows with 82.8 %.

<sup>22</sup>The sensitivity is calculated by comparing the correctly classified swallows to the total number of recorded swallows.

The presented methods for finding the positions of recorded swallows could be used for enhancing FEES examination. When EMG and BI is additionally recorded at the neck, the examiner would have access to all recorded swallowing positions. This could significantly improve the reliability of FEES for an inexperienced examiner. Swallowing can then be analyzed in the pharyngeal phase where the video of the FEES shows mainly a “white-out”.

Swallowing therapy could be improved by continuously recording all swallows and giving feedback of the last recorded swallows back to the patients and therapist. Specific characteristics which are calculated from the BI and EMG time series during swallowing could be represented in reference of recorded mean values. Spontaneous swallowing frequency of the patient could be measured which could be an important feature in indicating dysphagia [171].

# 5

## Assessment of Pharyngeal Swallowing by Bioimpedance and EMG

### 5.1 Summary

**Aim:** The correlation between bioimpedance (BI) / electromyography (EMG) measurements at the neck and the displacements of the hyoid bone and the laryngeal cartilage during swallowing is evaluated.

**Methods:** BI and EMG measurements at the neck were performed simultaneously to a videofluoroscopy (VF) of the swallowing process. The displacements of the hyoid bone and the laryngeal cartilage in superior and anterior direction were extracted from the recorded VF frames. The displacement directions were aligned to the coordinate system which is spanned by the vertebrae from C2 to C4. Additionally, start and stop times of bolus flow through the upper esophagus sphincter were obtained by VF. Recorded swallows were automatically segmented in the EMG and BI time series, and segmented swallows which coincide with bolus transit were selected for analysis. Pearson's correlation coefficients were calculated for maximal displacement amplitudes and the resulting change in BI related to the swallowing start. In a further step, Pearson's correlation coefficients were calculated for the displacement trajectories and the BI time series during all selected swallows. The inter-patient variability was assessed by calculating Pearson's correlation coefficients for each patient individually.

**Results:** In total, 92 swallows with bolus passage were segmented from the recorded measurement of all 17 patients. All patients had a disturbed pharyngeal swallowing phase. The Pearson's correlation coefficient between the maximum amplitude change in BI and the maximum displacement of the larynx in anterior direction was  $-0.55 \pm 0.17$ . A correlation of  $-0.7 \pm 0.14$  was found for the superior direction. For the maximum displacement of the hyoid, a Pearson's correlation coefficient of  $-0.52 \pm 0.16$  in the anterior direction and  $-0.64 \pm 0.15$  in the superior direction was obtained.

The Pearson's correlation coefficients for the complete hyoid trajectory and the BI time series was  $-0.65 \pm 0.05$  in anterior direction and  $-0.65 \pm 0.04$  in superior direction, respectively. For the larynx, the obtained correlation was  $-0.6 \pm 0.05$  in anterior direction, whereas in superior direction a value of  $0.71 \pm 0.03$  was observed. The mean value of the correlation coefficients, which were individually calculated for each patient, between BI and the hyoid displacement was  $-0.72 \pm 0.21$  in anterior direction and  $-0.46 \pm 0.45$  in superior direction. Mean Pearson's correlation coefficients of  $-0.6 \pm 0.36$  and  $-0.7 \pm 0.23$  were obtained for the laryngeal displacement in anterior and superior direction, respectively.

**Conclusion:** A deeper understanding of the underlying effects, which cause a deflection within the BI time series during swallowing, was developed. The results imply that BI measurements at the neck reflect the pharyngeal phase of swallowing. As BI is mostly correlated to the superior displacement of hyoid bone and larynx, it may be a meaningful measurement method for therapy and diagnosis of dysphagia.

**Contribution:** The author contributed all the presented methods for extracting the position of hyoid bone and the laryngeal cartilage in the recorded videofluoroscopic frames and performed the statistical data analysis. The recorded videofluoroscopy and the recorded EMG and BI measurements were analyzed independently using different methods in [172], [135], and [147].

## 5.2 Motivation

As described in Section 2.1.6.7, a four-electrode transcutaneous BI measurement at the neck was firstly used by Kusuhara et al., Nakamura et al., and Yamamoto et al. [22], [96], [97] in order to assess swallowing induced movements. They observed that swallowing causes a repeatable deflection in the BI signal. The measured BI signal was interpreted by Kusuhara et al. [22] as a reflection of the entire swallowing process (oral, pharyngeal, esophageal phase) caused by movement of the larynx, pharynx, throat, and esophagus. The end of the oral phase is indicated by the begin of the drop-off in bioimpedance, followed by the pharyngeal phase which ends at the lowest measured BI value. Finally, the bioimpedance is going back to its baseline during the esophageal phase. However, they did not verify their assumption on a gold standard as VF.

Decreased hyoid and laryngeal movements within the pharyngeal swallowing phase may lead to disturbances in oropharyngeal filling, bolus transport, airway closure, and upper esophagus sphincter (UES) opening [173]. Thus, insufficient hyoid and laryngeal displacement may lead to penetration or aspiration. Steele et al. [174] scaled the hyoid and laryngeal displacement of 28 patients with respect to their specific C2-C4 distance. The anterior and superior displacement was measured in percentage of the C2-C4 distance and correlated to the presence of penetration/aspiration and post-swallow pharyngeal residuals. A significant higher occurrence of penetration/aspiration was found in patients which had an anterior movement of the hyoid lower than the 25 % percentile. Post-swallow pharyngeal residuals were found to be associated with patients with an anterior laryngeal displacement which is lower than the 25 % percentile of all measured anterior laryngeal displacements.

The work of Steele et al. [174] shows that a significant correlation between reduced hyoid and laryngeal displacement and penetration/aspiration exists. Thus, other measurement methods which sensor output correlates to the hyoid and laryngeal displacement may be used to estimate the risk for penetration/aspiration and post-swallow pharyngeal residuals.

## 5.3 Methods

### 5.3.1 Subjects

Patients were recruited who required a videofluoroscopic swallowing study (VFSS) for examination of their swallowing ability. All patients were given an explanation and signed a consent. The exclusion criteria were pregnancy, implanted cardiac pacemaker or defibrillator, stent, or central venous catheter. All subjects were awake and could sit upright during the examination. This study was approved by the ethic board at the Charité Berlin (EA1/160/09).

### 5.3.2 Measurement Setup

Each subject was sitting while a lateral VF (Pulsera, Philips, 15 frames/seconds) of their neck area was performed. BI and EMG were simultaneously measured at the neck. The EMG and the BI measurement signals were continuously recorded by a personal computer (PC). The analog video stream was converted to a digital video stream by a universal serial bus (USB) video grabber and finally saved on the same PC together with a time stamp for each received frame for synchronization.

The patients were asked to take a contrast agent (Gastrografin®) of 10 ml into their mouth and swallow it on command after the recording was started. The recording was stopped when the patient successfully swallowed the liquid or the maximal recording time of 10 seconds was reached. In average three recordings were performed by each patient.

For the measurement of the trans-pharyngeal BI at the neck, the four-electrode method at a frequency of 50 kHz was used. Additionally, EMG was measured across both voltages measuring electrodes. For this, Blue Sensor N ECG electrodes (Ambu A/S, Denmark) were used. The electrode positions which are used for all measurements are defined as follows (see also Fig. 4.1):

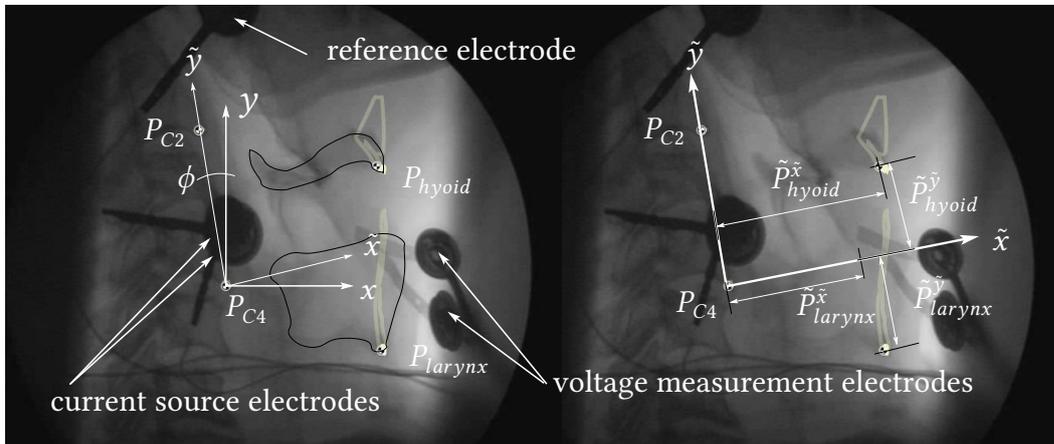
- The current electrodes are placed bilaterally on the upper onset of the sternocleidomastoid muscle below the mastoid process.
- The voltage measurement electrodes are placed laterally on the gap between hyoid bone and the thyroid cartilage symmetrically on both sides.

### 5.3.3 Extraction of Hyoid and Laryngeal Displacement

Structures which should be tracked across all frames are hyoid bone, larynx, C2 vertebra, and C4 vertebra (see Fig. 5.1). The point  $P_{hyoid}$  is set to the most anterior-inferior point of the hyoid bone. The point  $P_{larynx}$  is set to the most anterior-inferior point of the laryngeal cartilage. The reference points  $P_{C2}$  and  $P_{C4}$  are the most anterior-inferior points of both vertebral bodies.

The video clips were processed by the software tool Kinovea<sup>1</sup> for structural movement tracking. The software contains a semi-automatic tracking function for manually marked pixels. Beginning from the manually set start position, each marker is automatically transferred to the next frame to the position which has the same brightness distribution as in the frame before. The automatically proposed position is then manually reviewed and corrected if necessary. The position data of all tracked structures are exported with sub-pixel accuracy. A conversion of pixel coordinates into metric length measurements was carried out individually for each recording based on the known diameter (12 mm) of the used round electrodes, which are visible in the VF. Additionally, frames were marked in which the UES opened and trans-UES flow was visible.

The orthogonal coordinate system with  $C_4$  as origin and the  $C_2$ - $C_4$  line as one axis is used for measuring hyoid and laryngeal displacements [175]. The points  $P_{hyoid}$  and  $P_{larynx}$  are transformed from the camera aligned  $x$ - $y$ -coordinate system (with origin  $P_{C_4}$ ) into the  $\tilde{x}$ - $\tilde{y}$ -coordinate system which is rotated such that the  $\tilde{y}$  axis goes through the point  $P_{C_2}$ . The notation in the  $C_2$ - $C_4$  coordinate system is  $\tilde{P}_{hyoid}$  and  $\tilde{P}_{larynx}$ .



**Figure 5.1:** Definition of the tracking points and their coordination system.  $\tilde{P}_{hyoid}$  and  $\tilde{P}_{larynx}$  are finally extracted from the VF with respect to the coordinate system defined by  $P_{C_2}$  and  $P_{C_4}$ . The recorded VF image is also shown in [172] and used in Fig. 3.4 from [147].

All four points were tracked by two experts ( $e = \{1, 2\}$ ) across all available frames  $k = 1, \dots, N_k$ , where  $N_k$  is the number of recorded frames. The obtained time series are denoted as  $\{P_{C_2,e}(k)\}$ ,  $\{P_{hyoid,e}(k)\}$ , and  $\{P_{larynx,e}(k)\}$ <sup>2</sup>. In the next step, the angle  $\phi_e(k)$  between the  $C_2$ - $C_4$  oriented coordinate system and the video frame related coordinate system is calculated. The angle  $\phi_e(k)$  is calculated for each frame  $k$  and the calculation is repeated for each expert  $e$ :

$$\phi_e(k) = \text{atan2}(P_{C_2,e}^x(k), P_{C_2,e}^y(k)), \quad k = 1, \dots, N_k. \quad (5.1)$$

<sup>1</sup><http://www.kinovea.org>

<sup>2</sup>Each obtained point consists of a  $x$  and  $y$  component. The point  $\{P_{C_4}\}$  is the origin of the coordinate system and therefore always zero.

The coordinate system is changed in each frame by rotating all points towards the C2-C4 coordinate system using the rotation matrix:

$$rot_e(k) = \begin{bmatrix} \cos(\phi_e(k)) & -\sin(\phi_e(k)) \\ \sin(\phi_e(k)) & \cos(\phi_e(k)) \end{bmatrix}, \quad k = 1, \dots, N_k. \quad (5.2)$$

The rotation matrix is multiplied with the tracked points of larynx and hyoid in order to transform all points into the new coordination system.

$$\begin{bmatrix} \tilde{P}_{hyoid,e}^{\tilde{x}}(k) \\ \tilde{P}_{hyoid,e}^{\tilde{y}}(k) \end{bmatrix} = rot_e(k) \begin{bmatrix} P_{hyoid,e}^x(k) \\ P_{hyoid,e}^y(k) \end{bmatrix}, \quad (5.3)$$

$$\begin{bmatrix} \tilde{P}_{larynx,e}^{\tilde{x}}(k) \\ \tilde{P}_{larynx,e}^{\tilde{y}}(k) \end{bmatrix} = rot_e(k) \begin{bmatrix} P_{larynx,e}^x(k) \\ P_{larynx,e}^y(k) \end{bmatrix}, \quad k = 1, \dots, N_k. \quad (5.4)$$

Then, both independently obtained trajectories are combined by taken the mean value:

$$\tilde{P}_{hyoid} = ((\tilde{P}_{hyoid,e=1}(k) + \tilde{P}_{hyoid,e=2}(k)) / 2, \quad (5.5)$$

$$\tilde{P}_{larynx} = ((\tilde{P}_{larynx,e=1}(k) + \tilde{P}_{larynx,e=2}(k)) / 2, \quad k = 1, \dots, N_k. \quad (5.6)$$

$\{\tilde{P}_{hyoid}\}$  consists of  $\{\tilde{P}_{hyoid}^{\tilde{x}}\}$ , which is the anterior movement of the hyoid bone and  $\{\tilde{P}_{hyoid}^{\tilde{y}}\}$ , which is the superior movement of the hyoid bone.  $\{\tilde{P}_{larynx}\}$  consists of  $\{\tilde{P}_{larynx}^{\tilde{x}}\}$  and  $\{\tilde{P}_{larynx}^{\tilde{y}}\}$ , respectively. The movement definitions of  $\tilde{x}$  and  $\tilde{y}$  for the larynx are the same as for the hyoid bone.

### 5.3.4 Signal Pre-processing of EMG and BI

In order to apply the swallow segmentation algorithm to the measurement data, the pre-processing from Section 4.3.3 is applied to the recorded EMG and BI.

Additionally, the BI signal is filtered by a non-causal low pass of fourth order with a cut-off frequency of 7 Hz and resampled to the time stamps of the corresponding video frames from the VF. Therefore, one BI measurement sample at each recorded video time stamp is saved together with the three positions ( $P_{hyoid}$ ,  $P_{larynx}$ , and  $P_{C2}$ ) of the corresponding video frame. Recorded and preprocessed data are exemplarily shown in Figure 5.2.

### 5.3.5 Swallow Segmentation

Possible swallow positions in the measurement data are selected by the swallow segmentation algorithm (see Section 4.3.6). At first, all EMG activity periods are extracted by the double-threshold onset detector. Used parameters for the detector were  $tr_{max} = 0.01$  s,  $SNR_{min} = 3$  dB, and  $P_{fa} = 0.01$ . The noise variance  $N_{\hat{\sigma}_n^2}$  and the first threshold  $\hat{\zeta}$  are estimated for each recording as described in Section 4.3.4.4 using the parameters  $N_{\hat{\sigma}_n^2} T_s^{EMG} = 0.2$  s and  $N_{\hat{\zeta}} T_s^{EMG} = 8$  s.

The piece-wise linear approximation (PLA) is performed on the BI data using the bottom-up algorithm with  $max\_error = 5$ . Valleys containing EMG activity are searched and their start and end times are saved. The following parameters were used for the valley search algorithm:  $VS_{bound} = [0.2, 3.6]$ ,  $VS_{onset} = 0.3$ ,  $VS_{emg} = 0.25$  s, and  $VS_{diff} = 0.5$ . Finally, the values  $BI_{start,i}$ ,  $BI_{min,i}$ ,

$BI_{end,i}$ ,  $EMG_{start,i}$ , and  $EMG_{end,i}$  are extracted for swallow  $i$  which describes the corresponding EMG activity and BI valley (cf. Fig. 4.16).

### 5.3.6 Data Processing

All valleys are neglected which are shorter than the duration of three frames or do not contain any trans-UES bolus flow at all. The remaining valleys are considered as swallows. From each swallow sequences the first value is subtracted in order to remove the offset of the measurement which should not be considered in the analysis. The time series for changes in BI measurement during swallow  $i$  are written as

$$\{\Delta BI_i\} = BI_i(1) - BI_i(4), \dots, BI_i(N_i) - BI_i(4), \quad i = 1, \dots, S, \quad (5.7)$$

where  $S$  is the number of segmented swallows,  $BI_i(4)$  is defined as  $BI_{start}$ , and  $BI_i(N_i)$  is defined as  $BI_{end}$  of swallow  $i$ . Thus, three frames before  $BI_{start}$  are considered, as the movement trajectories sometimes start a little bit earlier than  $BI_{start}$ . The displacement trajectories of swallow  $i$  are build from the recorded points  $\{\tilde{P}_{hyoid}^{\tilde{x}}\}$  and  $\{\tilde{P}_{hyoid}^{\tilde{y}}\}$  and stored into  $\{H\tilde{x}_i\}$  and  $\{H\tilde{y}_i\}$  beginning from three frames before  $BI_{start}$  to the valley endpoint  $BI_{end}$ . The same was done for  $\{\tilde{P}_{larynx}^{\tilde{x}}\}$  and  $\{\tilde{P}_{larynx}^{\tilde{y}}\}$  and the resulting displacement trajectories for swallow  $i$  are denoted as  $\{L\tilde{x}_i\}$  and  $\{L\tilde{y}_i\}$ .

$$\{H\tilde{x}_i\} = \tilde{P}_{hyoid,i}^{\tilde{x}}(1) - \tilde{P}_{hyoid,i}^{\tilde{x}}(4), \dots, \tilde{P}_{hyoid,i}^{\tilde{x}}(N_i) - \tilde{P}_{hyoid,i}^{\tilde{x}}(4), \quad (5.8)$$

$$\{H\tilde{y}_i\} = \tilde{P}_{hyoid,i}^{\tilde{y}}(1) - \tilde{P}_{hyoid,i}^{\tilde{y}}(4), \dots, \tilde{P}_{hyoid,i}^{\tilde{y}}(N_i) - \tilde{P}_{hyoid,i}^{\tilde{y}}(4), \quad (5.9)$$

$$\{L\tilde{x}_i\} = \tilde{P}_{larynx,i}^{\tilde{x}}(1) - \tilde{P}_{larynx,i}^{\tilde{x}}(4), \dots, \tilde{P}_{larynx,i}^{\tilde{x}}(N_i) - \tilde{P}_{larynx,i}^{\tilde{x}}(4), \quad (5.10)$$

$$\{L\tilde{y}_i\} = \tilde{P}_{larynx,i}^{\tilde{y}}(1) - \tilde{P}_{larynx,i}^{\tilde{y}}(4), \dots, \tilde{P}_{larynx,i}^{\tilde{y}}(N_i) - \tilde{P}_{larynx,i}^{\tilde{y}}(4), \quad (5.11)$$

where  $i = 1, \dots, S$ .

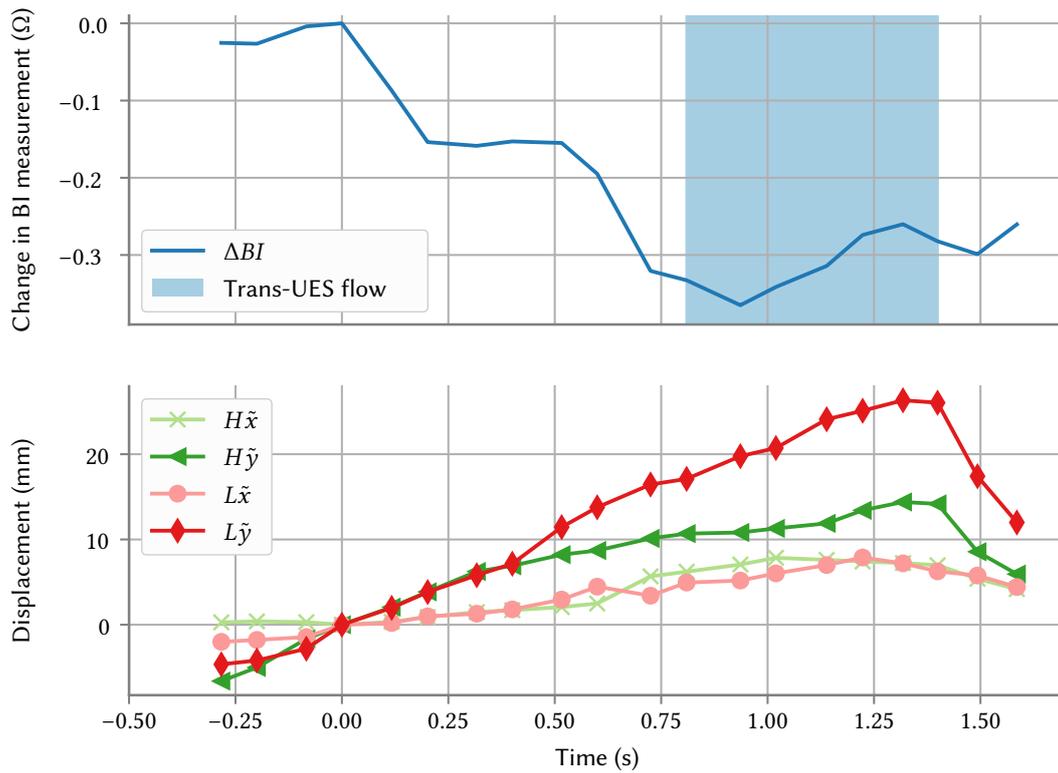
The linear relationship between the extracted displacement trajectories and the corresponding changes in BI measurements is analyzed using the Pearson correlation coefficient  $r$ . The swallow sequences are stringed together in order to be able to calculate a correlation. For BI, a time series is build by

$$\{\Delta BI\} = \{\Delta BI_1\}, \dots, \{\Delta BI_S\}. \quad (5.12)$$

The same procedure is repeated for all four displacement trajectories which results in the time series  $\{H\tilde{x}\}$ ,  $\{H\tilde{y}\}$ ,  $\{L\tilde{x}\}$ , and  $\{L\tilde{y}\}$ .

The minimum of  $\{\Delta BI_i\}$  (maximal BI drop) is correlated to the maximum of the hyoid and laryngeal displacement in anterior and superior direction for each swallow  $i$ . The Pearson correlation coefficient  $r$  is then calculated for the extracted minimum and maximum values across all swallows. The largest drop in BI during swallow  $i$  is found by

$$BI_i^{drop} = \arg \min_k \Delta BI_i(k), \quad 1 < k < N_i. \quad (5.13)$$



**Figure 5.2:** Pre-processed BI measurement data during a swallow together with the extracted displacement trajectories of hyoid and larynx. The timing of the observed trans-UES bolus flow is shown by the blue area. The time at  $BI_{start}$  is defined as zero.

For all of the four movement trajectories, the corresponding maximum displacement during a swallow is extracted ( $H\tilde{x}_i^{max}$ ,  $H\tilde{y}_i^{max}$ ,  $L\tilde{x}_i^{max}$ , and  $L\tilde{y}_i^{max}$ ):

$$T_i^{max} = \arg \max_k T_i(k), \quad 1 < k < N_i, \quad (5.14)$$

where  $T_i^{max}$  is standing for the four displacements  $H\tilde{x}_i^{max}$ ,  $H\tilde{y}_i^{max}$ ,  $L\tilde{x}_i^{max}$ , and  $L\tilde{y}_i^{max}$ . The values of the maximum BI amplitude drop and displacement amplitudes during each swallow are merged into a single set which results in  $\{BI^{drop}\}$ ,  $\{H\tilde{x}^{max}\}$ ,  $\{H\tilde{y}^{max}\}$ ,  $\{L\tilde{x}^{max}\}$ , and  $\{L\tilde{y}^{max}\}$ .

### 5.3.7 Statistical Analysis across all Patients

The linear relation between the displacement of hyoid and larynx and the measured BI values is expressed by the Pearson's  $r$ . Additionally, slope and intercept values are calculated. As the true distribution is unknown, the 95% confidence interval of  $r$  is estimated by bootstrapping [176]. All statistical values are calculated<sup>3</sup> with 10000 bootstrapping iterations. Considered are the complete time series  $\{\Delta BI\}$ ,  $\{H\tilde{x}\}$ ,  $\{H\tilde{y}\}$ ,  $\{L\tilde{x}\}$ , and  $\{L\tilde{y}\}$  as well as the BI drop  $\{BI^{drop}\}$  and maximal displacement amplitudes  $\{H\tilde{x}^{max}\}$ ,  $\{H\tilde{y}^{max}\}$ ,  $\{L\tilde{x}^{max}\}$ , and  $\{L\tilde{y}^{max}\}$ .

<sup>3</sup>The R project (<http://r-project.org>) with the boot package (<https://cran.r-project.org/web/packages/boot/>) are used for statistically calculations.

### 5.3.8 Individual Statistical Analysis for each Patient

The time series of all swallows belonging to patient  $p$  are grouped. The resulting time series are  $\{\Delta BI^p\}$ ,  $\{H\tilde{x}^p\}$ ,  $\{H\tilde{y}^p\}$ ,  $\{L\tilde{x}^p\}$ , and  $\{L\tilde{y}^p\}$ . For each patient  $p$ , Pearson's correlation coefficient  $r_p$  is calculated from the time series  $\{\Delta BI^p\}$  to all displacement time series.

The method from meta-analysis according to [177] can be used to calculate weighted mean coefficients across all patients. The weighted mean correlation  $\bar{r}$  is calculated by:

$$\bar{r} = \frac{\sum_{p=1}^k N_p r_p}{\sum_{p=1}^k N_p}, \quad (5.15)$$

where  $N_p$  is the number of data points of patient  $p$ ,  $k$  is the number of patients, and  $r_p$  is the correlation coefficient of the patient  $p$ . The variance of sample effect sizes can be calculated by:

$$\sigma_r^2 = \frac{\sum_{p=1}^k N_p (r_p - \bar{r})^2}{\sum_{p=1}^k N_p}. \quad (5.16)$$

The sampling error variance is calculated by

$$\sigma_e^2 = \frac{(1 - \bar{r}^2)^2}{N - 1}. \quad (5.17)$$

The variance in population effect size  $\hat{\sigma}_p^2 = \sigma_r^2 - \sigma_e^2$  can be used to calculate the 95 % confidence interval:

$$\text{Confidence Interval} = \bar{r} \pm 1.96 \sqrt{\hat{\sigma}_p^2}. \quad (5.18)$$

## 5.4 Results

### 5.4.1 Test Subjects

19 patients participated in this study. Due to synchronization problems between BI/EMG measurement and VF recording in two patients, 17 patients (mean age =  $67.4 \pm 10.5$ , 15 male, 2 female) were finally included in the analysis<sup>4</sup>. All patients showed abnormalities in the pharyngeal swallowing phase.

### 5.4.2 Correlation and Regression Analysis

The complete data set contained 2828 samples composed of VF, BI, and EMG measurements. The automatic segmentation algorithm found 238 valleys. In 92 from 103 manually marked trans-UES bolus flow sequences, an appropriate BI valley with EMG activity which contained at least three VF

---

<sup>4</sup>The patients are listed with more details in the thesis from C. Schultheiss [135] in table B.5. The corresponding codes of the patients are 130(1), 136(2), 137(3), 145(4), 146(5), 160(6), 162(7), 163(8), 164(9), 174(10), 175(11), 183(12), 207(13), 213(14), 214(15), 215(16), and 218(17). The recorded data from the patients with the codes 135 and 161 are not included in the analysis performed within this thesis as the recorded BI signals from both patients are not in phase with the extracted displacements of hyoid and larynx. It was concluded that during the recording synchronization problems must have occurred. The inclusion of the data from both excluded patients would have lead to outliers.

frames was found. The final data set with all 92 swallow sequences contained 1333 data samples. All swallow sequences started three VF frames before  $BI_{start}$  and ended at  $BI_{end}$ <sup>5</sup>.

**Table 5.1:** Time difference between the swallow related events  $BI_{start}^T$ ,  $BI_{min}^T$ ,  $BI_{end}^T$ ,  $EMG_{start}^T$ ,  $EMG_{end}^T$ , and the onset  $t_{UES_{start}}$  of the trans-UES flow. In the last row, the mean values across all swallows are displayed. All times are denoted in milliseconds.

p	$EMG_{start}^T$ $-t_{UES_{start}}$	$BI_{start}^T$ $-t_{UES_{start}}$	$EMG_{max}^T$ $-t_{UES_{start}}$	$BI_{min}^T$ $-t_{UES_{start}}$	$t_{UES_{end}}$ $-t_{UES_{start}}$	$BI_{end}^T$ $-t_{UES_{start}}$	$EMG_{end}^T$ $-t_{UES_{start}}$
1	-898	-514	18	10	359	950	351
2	-931	-731	-192	29	543	852	695
3	-1783	-486	171	102	335	725	1805
4	-605	-544	-197	-115	449	333	515
5	-628	-328	165	231	414	554	49
6	-526	-126	156	154	368	551	487
7	-2348	-504	272	118	336	546	568
8	-784	-244	68	130	350	440	821
9	-1351	-292	93	307	246	785	877
10	-461	-167	180	379	336	708	15
11	-497	-394	71	25	365	489	397
12	-879	-431	-224	237	371	660	430
13	-1802	-361	-146	49	290	373	1238
14	-244	-103	469	317	313	534	693
15	-333	-52	169	413	452	928	525
16	-883	-295	31	52	400	196	712
17	-461	-417	67	41	433	141	-109
$\emptyset$	-811	-315	62	130	335	514	530

In the following, the upper index  $V$  refers to the value and the upper index  $T$  to the time of a specific point.  $EMG_{start}$  and  $EMG_{end}$  are set to the associated EMG interval;  $EMG_{max}$  is set to the maximum in this interval. For each patient, the mean times of the BI valleys, EMG activity and offset of trans-UES bolus flow are extracted. The time differences of the related events with respect to the onset  $t_{UES_{start}}$  of the trans-UES bolus flow are shown in Table 5.1.

The bolus needs around 0.34 s for passing the laryngeal cartilage. The begin of a swallow based on the valley begin in the BI time series was found to be 0.32 s before any bolus was visible. The position of the minimum in the BI valley is around 0.13 s after a bolus flow was visible and 0.21 s before the bolus disappeared. The end point  $BI_{end}^T$  lays around 0.51 s after  $t_{UES_{start}}$  and 0.18 s after  $t_{UES_{end}}$ . The mean BI valley has a duration of 0.83 s. An EMG activity could be observed around 0.81 s before  $t_{UES_{start}}$  and thus 0.5 s before  $BI_{start}^T$ . The maximal EMG activity is located 0.1 s after  $t_{UES_{start}}$ . The mean EMG activity duration is 1.3 s.

Pearson's  $r$  were calculated between the largest drop in BI and the maximum displacement of hyoid bone and larynx in  $\tilde{x}$  and  $\tilde{y}$  direction. The results are summed up in Table 5.2. The correlation is also visualized as scatter plot (see Fig. 5.3) in which all data pairs are plotted together with the fitted linear slope. The regression parameters of these slopes are shown in Table 5.2. The best correlation

<sup>5</sup>Each data sample corresponded to a single VF frame.  $BI_{start}$  and  $BI_{end}$  were estimated for each swallow based on the swallow segmentation algorithm which is presented in Chapter 4 and associated to a VF frame.

## 5. Assessment of Pharyngeal Swallowing by Bioimpedance and EMG

**Table 5.2:** Pearson correlation coefficients of the correlation between the maximum drop in BI ( $\{BI^{drop}\}$ ) and the maximal displacement of the hyoid bone ( $\{H\tilde{x}^{max}\}$ ,  $\{H\tilde{y}^{max}\}$ ) and larynx ( $\{L\tilde{x}^{max}\}$ ,  $\{L\tilde{y}^{max}\}$ ) during all 92 swallows (N = 92) from 17 patients ( $\tilde{x}$  - anterior direction and  $\tilde{y}$  - superior direction).

	$\{H\tilde{x}^{max}\}$	$\{H\tilde{y}^{max}\}$	$\{L\tilde{x}^{max}\}$	$\{L\tilde{y}^{max}\}$
Pearson's $r$	$-0.55 \pm 0.17$	$-0.7 \pm 0.14$	$-0.52 \pm 0.16$	$-0.64 \pm 0.15$
Slope [mm/ $\Omega$ ]	-3	-7.9	-2	-10
Intercept [mm]	2.7	1.2	1.4	3.5

**Table 5.3:** Pearson correlation coefficients of the correlation between the BI time series  $\{\Delta BI\}$  and the displacement trajectory of the hyoid bone ( $\{H\tilde{x}\}$ ,  $\{H\tilde{y}\}$ ) and the larynx ( $\{L\tilde{x}\}$ ,  $\{L\tilde{y}\}$ ) during all 92 swallows (N = 1333) ( $\tilde{x}$  - anterior direction and  $\tilde{y}$  - superior direction).

	$\{H\tilde{x}\}$	$\{H\tilde{y}\}$	$\{L\tilde{x}\}$	$\{L\tilde{y}\}$
Pearson's $r$	$-0.65 \pm 0.048$	$-0.65 \pm 0.044$	$-0.6 \pm 0.046$	$-0.71 \pm 0.034$
Slope [mm/ $\Omega$ ]	-4.2	-8.4	-2.5	-13
Intercept [mm]	0.63	-0.88	0.27	0.52

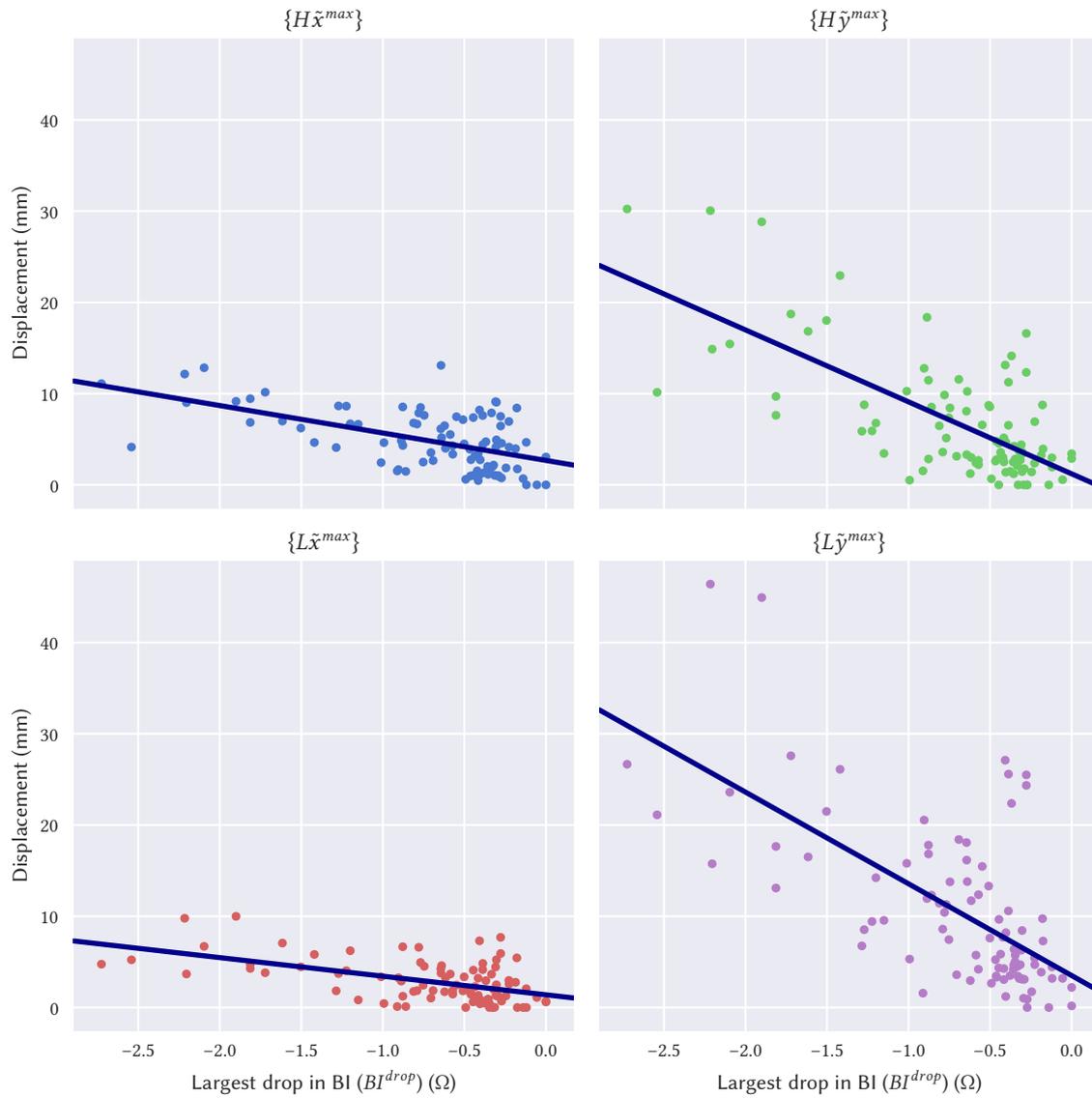
was found between the drop in BI during a swallow and the maximal displacement of the larynx in  $\tilde{y}$  direction. The Pearson's  $r$  is -0.69 with a bootstrap confidence interval of  $\pm 0.11$ .

In Table 5.3, a regression analysis was performed on all data samples of all patients. For  $\{H\tilde{x}\}$ ,  $\{H\tilde{y}\}$ ,  $\{L\tilde{x}\}$ , and  $\{L\tilde{y}\}$ , similar regression parameters were found as on the maximal value correlation (see Table 5.2).

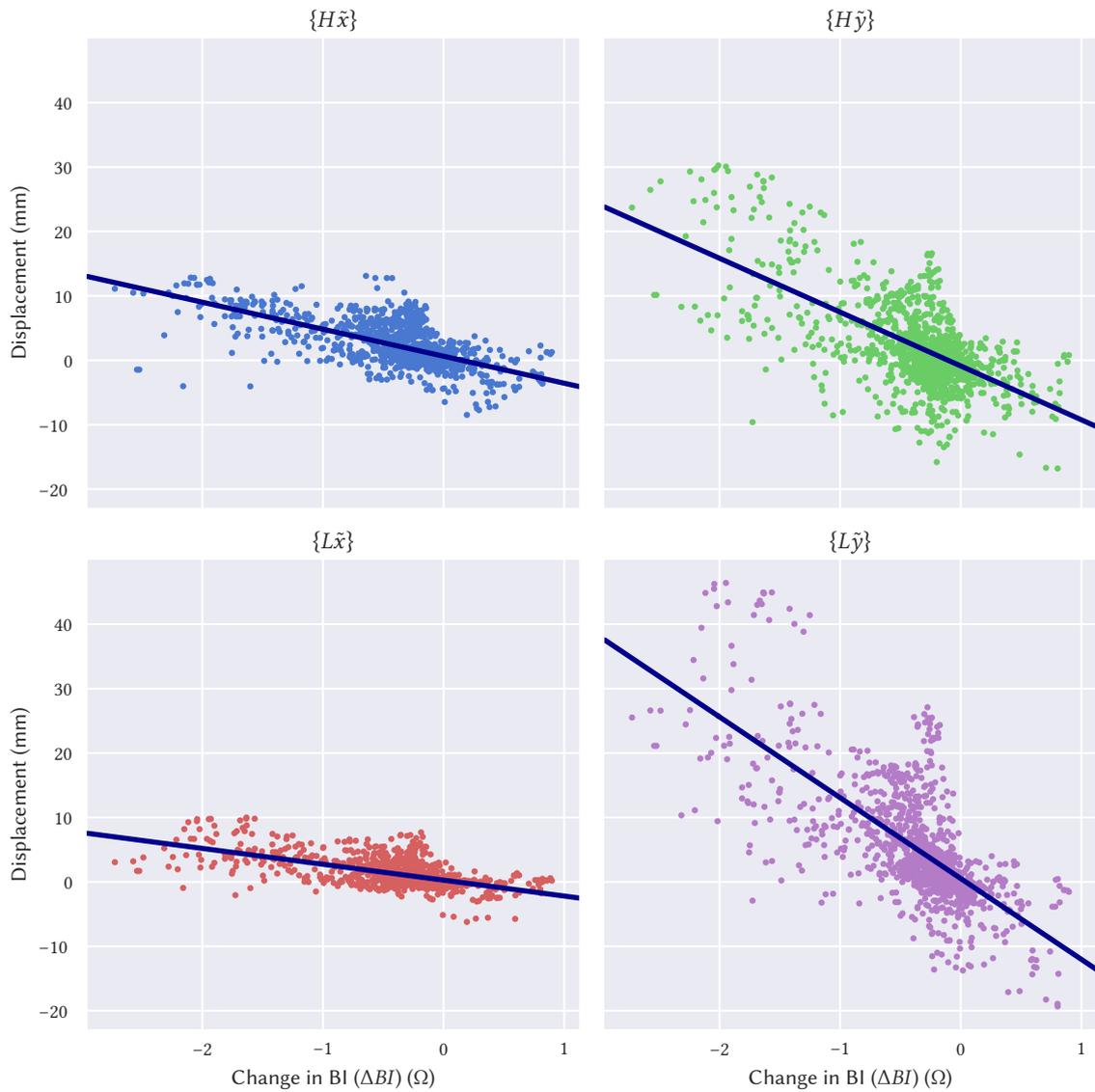
In Table 5.4, a linear regression analysis was performed individually for each patient. The regression coefficients between the change in BI  $\{\Delta BI^p\}$  to the displacement trajectories  $\{H\tilde{x}^p\}$ ,  $\{H\tilde{y}^p\}$ ,  $\{L\tilde{x}^p\}$ , and  $\{L\tilde{y}^p\}$  are displayed in this table. In the last row, the weighted mean correlation coefficients with 95 % confidence interval of all patients are given.

**Table 5.4:** Correlation coefficients of the correlation between the change of BI ( $\{\Delta BI^p\}$ ) and displacement trajectory of hyoid bone ( $H$ ) and larynx ( $L$ ) for each patient  $p$  ( $\tilde{x}$  - anterior direction and  $\tilde{y}$  - superior direction).  $S_p$  is the number of recorded swallows and  $N_p$  the number of stored data points for patient  $p$ . In the last row, the weighted mean correlation coefficients with 95 % confidence interval of all patients are given.

$p$	$S_p$	$N_p$	$\{H\tilde{x}^p\}$ $r_p$	$\{H\tilde{y}^p\}$ $r_p$	$\{L\tilde{x}^p\}$ $r_p$	$\{L\tilde{y}^p\}$ $r_p$
1	4	58	-0.501	-0.757	-0.507	-0.761
2	3	60	-0.806	-0.845	-0.872	-0.869
3	6	116	-0.869	-0.825	-0.754	-0.809
4	3	29	-0.681	-0.754	-0.708	-0.779
5	8	109	-0.605	-0.315	-0.578	-0.748
6	5	58	-0.824	-0.248	0.099	-0.733
7	2	36	-0.789	-0.0577	-0.575	-0.564
8	22	271	-0.74	-0.189	-0.579	-0.637
9	8	154	-0.607	-0.37	-0.428	-0.608
10	3	42	-0.632	-0.188	-0.563	-0.346
11	6	81	-0.757	-0.642	-0.743	-0.771
12	4	83	-0.817	-0.575	-0.759	-0.893
13	5	60	-0.371	-0.621	-0.554	-0.752
14	4	55	-0.774	-0.341	-0.774	-0.528
15	3	55	-0.85	-0.734	-0.863	-0.857
16	3	30	-0.783	-0.375	-0.405	-0.386
17	3	36	-0.853	-0.83	-0.801	-0.857
$\bar{r}$			$-0.72 \pm 0.21$	$-0.46 \pm 0.45$	$-0.6 \pm 0.36$	$-0.7 \pm 0.23$



**Figure 5.3:** The scatter plot shows the relationship of the largest drop in BI ( $\{BI^{drop}\}$ ) to the maximal displacement of the hyoid bone ( $\{H\tilde{x}^{max}\}, \{H\tilde{y}^{max}\}$ ) and larynx ( $\{L\tilde{x}^{max}\}, \{L\tilde{y}^{max}\}$ ) for all 92 swallows. The slope and interception of the regression line can be found in Table 5.2.



**Figure 5.4:** The scatter plot shows the relationship of the change in BI to the displacement trajectory of the hyoid bone ( $\{H\tilde{x}\}$ ,  $\{H\tilde{y}\}$ ) and larynx ( $\{L\tilde{x}\}$ ,  $\{L\tilde{y}\}$ ) during all 92 swallows (N=1333) ( $\tilde{x}$  - anterior direction and  $\tilde{y}$  - superior direction). The slope and interception of the regression lines can be found in Table 5.3.

### 5.5 Conclusions

In a VF swallowing examination, all movements during the swallowing process can be observed. The movements of the tongue and mouth floor muscles can be estimated particularly by means of the movement of the bony and cartilaginous structures which are visible in an X-ray examination. The swallowing process itself can be estimated by the coordination of these movements and by the flow of the bolus which is entirely visible due to contrast agents. Thus, VFSS is one gold standard in the diagnosis of swallowing disorders.

It was evaluated in which way the typically observed drop-off in BI during swallowing is linked to the displacement of hyoid bone and larynx. Therefore, displacements of hyoid bone and larynx were extracting by performing a VF while BI and EMG measurements were simultaneously recorded.

With the combination of EMG and BI measurement, it is possible to measure the muscular activation which induces a swallow and the hyolaryngeal movements during a swallow. The muscle activity which is necessary to swallow can be measured using EMG. Changes in the tissue which are caused by the swallow effect itself (lifting of the larynx and constriction) can be recorded using BI.

The connection between the displacement of hyoid bone and larynx and the measured BI was investigated using correlation analysis. The drop in BI during a swallow correlates moderately to strongly negative to the maximum displacement of hyoid and larynx during a swallow. The best correlation coefficients could be found between the drop in BI and the hyoid displacement in superior direction with 0.7. The second best correlation is observed between the drop in BI and the laryngeal displacement in superior direction with 0.64.

A correlation analysis between all measurement samples across all patients between the change in BI and the displacement trajectories of hyoid bone and larynx shows similar results. The best correlation is observed between the change in BI and the larynx displacement in superior direction with 0.71. The second best correlation is achieved between the change in BI and the hyoid displacement in anterior and superior direction each with a value of 0.65.

Better correlation results could be obtained for a correlation analysis individually for each patient. For some patients, a strong correlation between BI and one movement trajectory could be observed. The best mean correlation across all patients was achieved at the hyoid displacement in anterior direction with 0.72 and at the laryngeal displacement in superior direction with 0.7.

The presented results show that BI measurements are at least moderately correlated to the displacement of hyoid bone in anterior and superior direction and larynx cartilage in anterior and superior direction. Although the behavior of the bolus cannot be measured by BI and it is not possible to map the trajectories of BI measurements back to movement trajectories of hyoid and larynx, BI measurement can be used to describe the pharyngeal phase of swallowing. A steep gradient of the BI valley indicates a fast movement of the larynx and the hyoid bone. A valley with a large drop in BI suggests a great movement amplitude of hyoid bone and larynx.

The drop in BI during swallowing correlates to the movements of hyoid and larynx, which defines the beginning of the pharyngeal swallowing phase. Strong and sufficient fast elevation of hyoid and larynx are a condition for lower airway protection and safe swallowing. The begin of the esophageal phase is set to the point when the bolus is pulled behind the UES. Extracted timing

information of all swallows indicates that the end of the extracted BI valley coincides with the end of trans-UES bolus flow. Thus, BI and EMG measurement characterize mainly the pharyngeal phase of the swallowing process. Therefore, it may be used for diagnostic and therapy in pharyngeal swallowing disorders.

Swallowing accelerometry and swallowing sounds are somehow correlated to the movement of the larynx. In these measurement methods, the time series curve can not be easily interpreted (it shows either the amount of acceleration or the loudness during a swallow). The BI time series is directly correlated to the movement of hyoid and larynx and for this reason to the complex closing mechanism of the pharynx during swallowing. This allows the calculation and usage of deduced variables as timing information, displacement speed and the maximal deflection in swallowing diagnostic. Furthermore, EMG can be measured on the voltage measurement electrodes without further effort.

Melfenter et al. [178] showed the large variability of hyoid and laryngeal displacement assessment using VF. VF is a gold standard in evaluating swallowing, but using it for measuring the displacement of structures as hyoid and larynx is time-consuming and error-prone. Furthermore, the method is costly and the patient is contaminated by radiation during the examination. BI measurement at the neck in optimal conditions (exactly defined electrode position and the patient does not do any movement except swallowing during the examination) could be a measurement method for the structural movement during swallowing. It is difficult to obtain precise measurements as correlation is only moderate, but changes in a patient over therapy might be good monitored.

Fiber-optic endoscopic evaluation of swallowing (FEES) examinations have the disadvantage that the pharyngeal phase is crossfaded by the “white-out” which is caused by constricting of the pharynx. BI/EMG measurement could provide missing information about the pharyngeal phase. In combination with swallowing segmentation algorithms, it could improve the average inter-rater reliability of the FEES examination.

Bio-feedback based swallowing therapy could use BI/EMG measurement for giving real-time information about the last swallow to the patient. Speed and amount of drop in BI could be used to inform the patient about his/her current swallowing intensity and could encourage the patient to improve his/her swallowing performance.

BI/EMG measurement may also be used in a controlled swallowing prosthesis, in which a electrical stimulation is triggered by EMG activity detection and the beginning of a drop in BI. The stimulation intensity may be controlled by the drop in BI which is correlated to the movement of hyoid and larynx and the speed of BI changes which is correlated to the movement speed of hyoid and larynx.

In further studies, the reliability and validity of this method have to be investigated. It should be examined how dysphagia and aspiration can be assessed using BI and EMG.



# 6

## BI- and EMG-Based Control of FES For Supporting Swallowing

### 6.1 Summary

**Aim:** It is investigated if a combined bioimpedance (BI) and electromyography (EMG) measurement can control a supportive functional electrical stimulation (FES) at the neck in order to improve and secure swallowing.

**Methods:** BI and EMG are measured at the neck. Algorithms are presented which robustly detect EMG activity and changes in the BI signal to enable a detection of swallow onsets as early as possible. A hand switch is used for activation of the swallow onset detection. Parameters for swallow detection need to be tuned for every patient based on one baseline swallow. Whenever a swallow onset is detected by the algorithms and the hand switch is pressed, FES of the submental muscles is activated in order to support swallowing.

**Results:** It was demonstrated in a case study with one chronic stroke patient that the amount and speed of larynx and hyoid elevation can be improved by controlling FES based on EMG and BI measurements. The evaluation of the impaired elevation of larynx and hyoid was assessed by BI measurements. In comparison to healthy subjects, the drop in BI during swallowing could be improved from 37 % ( $-0.56 \Omega$ ) to 74 % ( $-1.12 \Omega$ ). The change in BI at the beginning of a swallow increased from 29 % ( $-1.87 \Omega/s$ ) to 47 % ( $-3.03 \Omega/s$ ).

**Conclusion:** The presented algorithms are able to detect a swallowing event early enough such that an FES can be applied to the submental muscles in order to give a biofeedback to the patient and to increase the lower airway protection by raising hyoid and larynx faster and higher. It could be shown that FES of the submental muscles was synchronized to voluntarily induced swallowing.

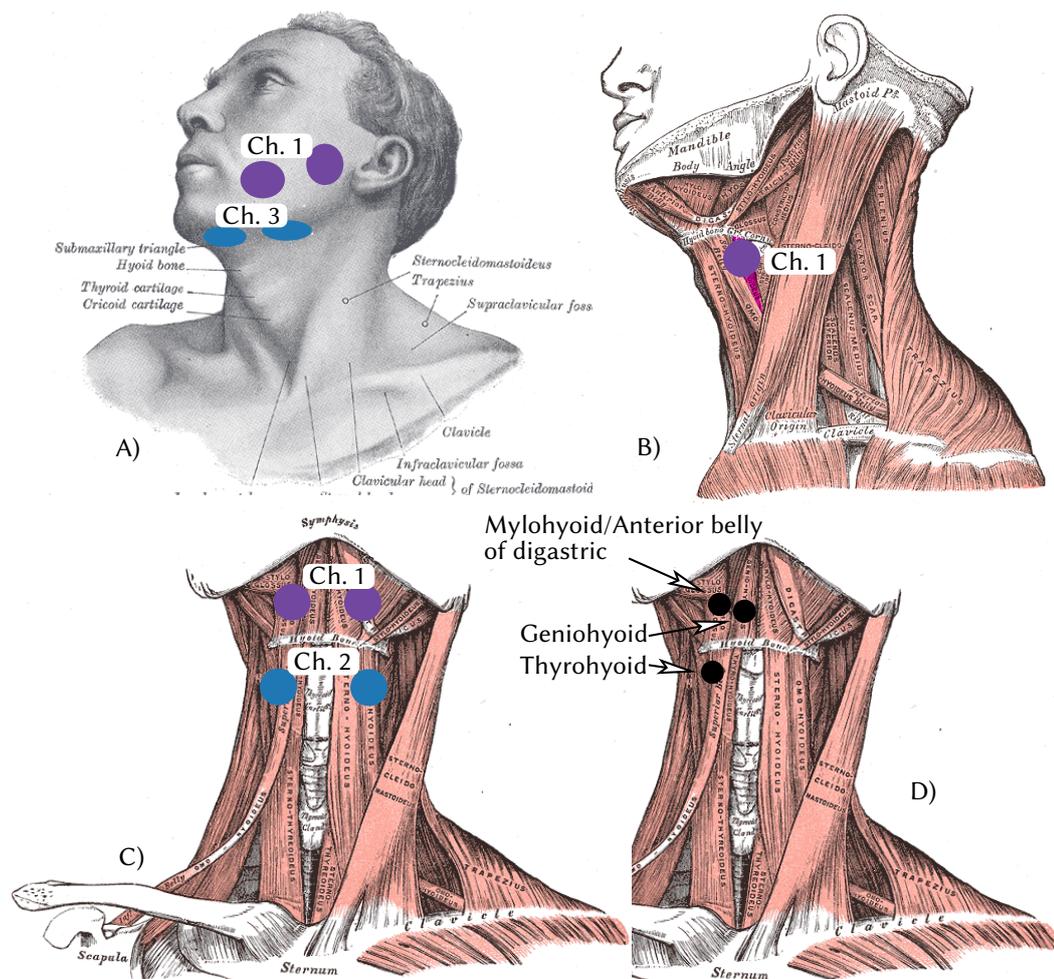
**Contribution:** All presented algorithms were developed by the author. Parts of the presented algorithm and the results were presented in [179] and [180]. The case study was performed in cooperation with the Unfallkrankenhaus Berlin (ukb).

### 6.2 Motivation

In the majority of studies, which were dealing with neuromuscular electrical stimulation (NMES) for the purpose of treating swallowing disorders, surface electrical stimulation was applied for several minutes in order to strengthen the muscles which are involved in swallowing. The NMES is thereby not functional as the stimulation is not applied timely correlated to swallowing itself. NMES of swallowing involved muscles was introduced by Freed et al. [123]. In this study, a fixed swallowing-independent electrical stimulation protocol was applied to patients 60 minutes per day. This stimulator was later made available under the name VitalStim®. This protocol was used in several successive studies with contradicting results [181]–[183]. It was shown by Ludlow et al. [33] and Humbert et al. [32] that NMES with the laryngeal electrode positions suggested in [123] can increase the risk of aspiration in some cases. Only the submental electrode positions cause no negative effect and lead to slight elevation of the laryngeal structures. Kagaya et al. [31] determined the positions of the motor points for the mylohyoid muscle / anterior belly of the digastric, geniohyoid muscle, and thyrohyoid muscle. Stimulation of the motor points for the mylohyoid muscle / anterior belly of the digastric and geniohyoid muscle leads to a superior and anterior movement of the hyoid and larynx. They also showed that surface electrical stimulation of the thyrohyoid motor points, which correspond to the laryngeal electrode position proposed by [123], leads to an inferior movement of the hyoid bone and the larynx. This electrode position should, therefore, be avoided as aspiration may be induced by it. The position of the stimulation electrodes used by the VitalStim® device and the motor points are shown in Figure 6.1.

Burnett et al. [36] showed that healthy subjects are able to trigger FES which was applied intramuscularly synchronized to their swallows by pressing a switch. However, it is questionable if patients are able to support all of their swallows by manually triggering FES. Leelaminit et al. [37] designed a study in which patients with reduced laryngeal elevation were stimulated at the thyrohyoid muscle. The stimulation was synchronized to EMG activity from the posterior tongue. This submental EMG, which is only an unspecific indicator for the pharyngeal phase, was used to trigger the stimulation. The study showed promising results, however, the patients had to be treated for 4 hours per day. The used stimulation electrode position may also lead to an inferior movement of the hyoid and larynx which increases the risk for aspiration as shown in [31]–[33]. The electrode position corresponds to the laryngeal position proposed in [123] and is shown in Figure 6.1.

In a study by Soon et al. [129], patients should swallow synchronized to a fixed stimulation pattern of the masseter muscle followed by a stimulation of the digastric muscle in the submental region. The electrode positions are also shown in Figure 6.1. The authors showed that the time, which the patients needed for drinking a certain amount of water, reduced in all patients. Thus, significant improvement in swallowing function can be gained over time when swallowing is performed synchronized to a fixed stimulation pattern. However, to swallow at fixed time points is demanding for patients and eventually even not possible for patients with cognitive difficulties.



**Figure 6.1:** A) Electrode setup by Soon et al. [129]. Channel 1 and 2 (not shown) stimulate the masseter muscle on both sides. Channel 3 and 4 (not shown) stimulate the digastric muscle. B) Electrode setup by Leelaminit et al. [37] for stimulating the thyrohyoid muscle. Only one electrode of the pair of electrodes is shown. C) VitalStim® stimulation setup which is used in [32], [33], [123] and several other studies. Channel 1 stimulates muscles in the submental region and channel 2 stimulates in the laryngeal region. Channel 2 correspond to setup B). D) Motor points investigated by Kagaya et al. [31]. Anatomical illustrations are from [2].

Besides swallowing therapy, immediate improvement in swallowing function may be achieved by NMES synchronized to the elevation of hyoid bone and larynx. As such a swallowing event-triggered stimulation recovers swallowing functions, it can be called FES. Applying FES in the very beginning of a swallow could improve lower airway protection by increasing the elevation of the hyoid bone and the laryngeal cartilage. Without appropriate elevation, aspiration may take place in which parts of the bolus may enter the airway. Recent studies [184], [185] showed a positive effect on motor relearning when FES is applied synchronously to voluntarily induced movements. In order to perform further studies in which FES is applied on patients with dysphagia, methods have to be developed that can be used to detect the beginning of a swallow. The detection has to be performed in real-time and should trigger FES such that the force output of the stimulated muscles is within the pharyngeal elevation period. The delay between stimulation onset and the force output of the stimulated muscle lays in a range of a few hundred milliseconds. As swallowing induced EMG

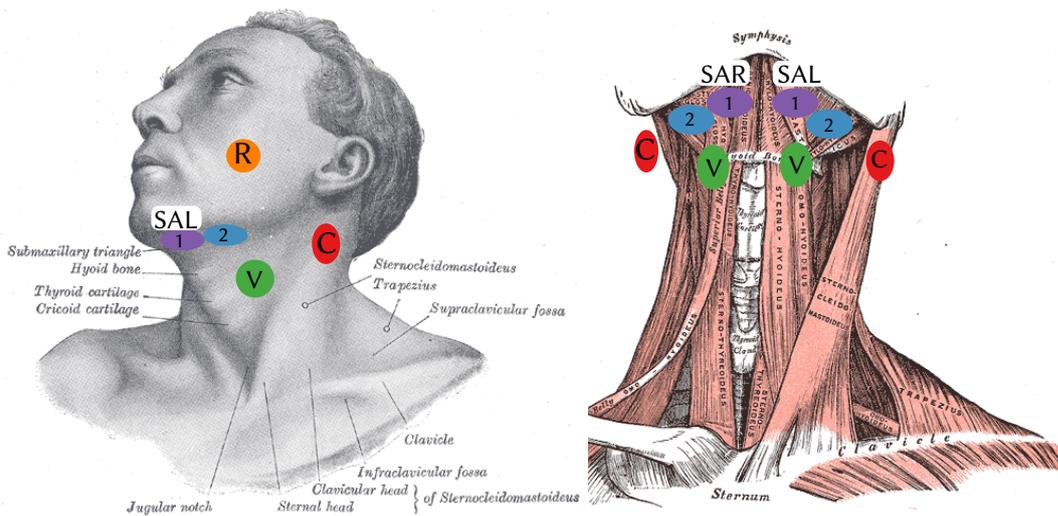
activity can be measured before the swallow starts, it is a good trigger for FES and was already used by Leelaminit et al. [37]. The disadvantages of using submental EMG as a trigger are that the EMG threshold changes the stimulation timing and that FES would be also triggered by nearby muscle activity (e.g. tongue-, masseter-, or head movements).

In Chapter 4, a swallowing detection has been presented which utilizes BI and EMG measurements at the neck. The presented algorithms for EMG activity detection and line segment approximation of the BI signal were optimized for recorded data. In this chapter, these algorithms are adapted such that EMG activity and changes in the BI signal are immediately detected. Finally, algorithms are presented which are able to trigger FES in the very beginning of a swallow, such that swallowing can be timely supported by FES.

### 6.3 Methods

#### 6.3.1 Experimental Setup

The measurement system PHYSIOSENSE, which has been described in detail in Chapter 3, was used for measuring BI and EMG at the neck. An electrical alternating current (AC) current at a frequency of 50 kHz was applied via two silver/silver-chloride (Ag/AgCl) electrodes and the resulting voltage was measured via an additional pair of Ag/AgCl electrodes. The EMG and BI measurement signal is then extracted from the measured voltage signal and send to the personal computer (PC). Blue Sensor N ECG electrodes (Ambu A/S, Denmark) were used. This setup and the electrode position have been described in detail in Section 4.3.2.



**Figure 6.2:** Electrode positions for electrical stimulation and BI/EMG measurements (C - current electrodes, V - voltage measurement electrodes, R - reference electrode, SAL / SAR - stimulation electrodes for the left/right anterior part of the submental muscles). The SAL and the SAR stimulation area consists of two (illustrated by 1 and 2) stimulation channels, which are activated always simultaneously. Anatomical illustrations are from [2].

For applying FES, one pair of surface stimulation electrodes (SAL/SAR) was attached to the submental region near the suprahyoid muscles. The electrodes were placed symmetrically around the median plane below the mandible and above the hyoid bone (cf. Fig. 6.2). Round hydro-gel

stimulation electrodes ( $\varnothing$  32 mm, KRAUTH+TIMMERMANN GmbH, Germany) were used. The positions of the stimulation electrodes and the BI/EMG measurement electrodes are shown in Figure 6.2.

A RehaStim device (HASOMED GmbH, Germany), in which the maximal available current amplitude of the stimulation pulses was limited to 25 mA, was used for applying electrical pulses to the suprahyoid muscles. The stimulator delivers bi-phasic pulses with a pulse width up to 500  $\mu$ s and is able to generate doublets, that consists of two subsequent bi-phasic stimulation pulses with a very short inter-pulse interval of 5 ms. The stimulator can send a transistor-transistor logic (TTL) trigger signal that is synchronized to the start and end of each stimulation pulse/doublet. This TTL signal was fed to the Galvanically isolated synchronization input of the measuring system PHYSIOSENSE and was transmitted to the PC simultaneously with all other measurement signals. This signal was used for blanking stimulation artifacts in the EMG and BI measurement signal.

The state of a manual hand switch was also recorded with the PHYSIOSENSE device. The manual switch was used for enabling the control of the FES.

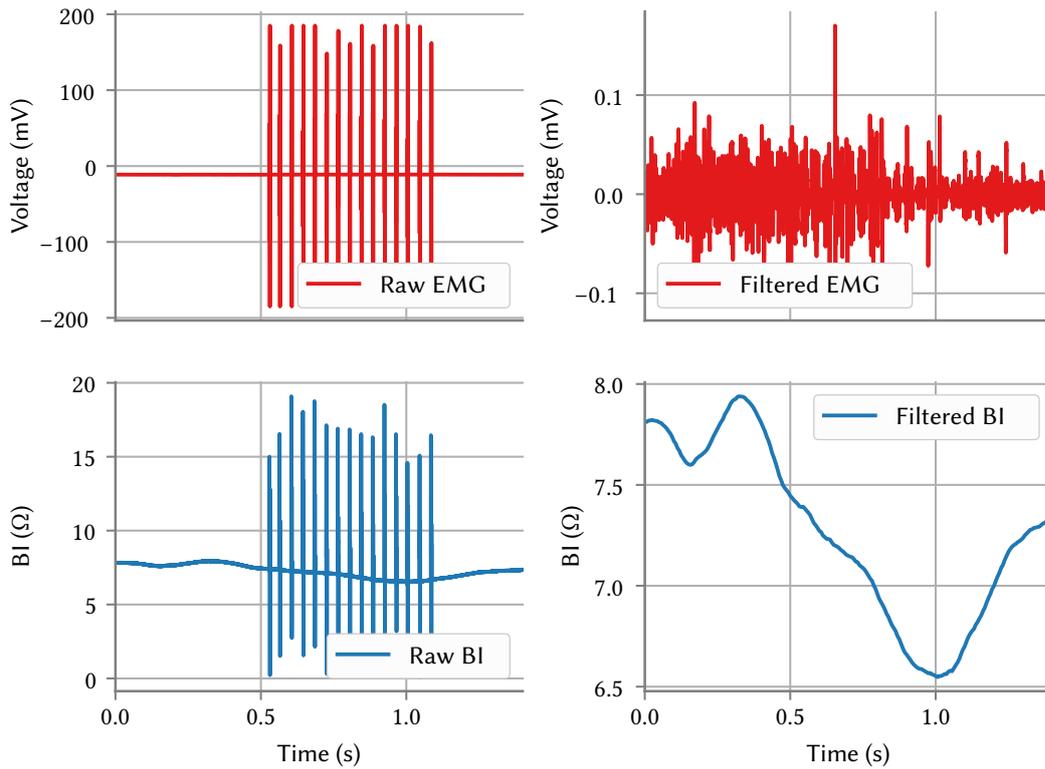
### 6.3.2 Stimulation Artifact Removing and Preprocessing of EMG and BI

During active FES the recorded EMG and BI signals are disturbed by the high voltage stimulation pulses. As described in Chapter 3, the measurement device PHYSIOSENSE was designed to enable the measurement of EMG and BI also in between stimulation pulses. After a stimulus, the measurement amplifier goes shortly into saturation and is able to measure again after 1 ms to 2 ms. The corresponding FES-evoked EMG activity is named M-wave. It starts directly after the stimulus and lasts for about 10 ms in the current setup. The amplitude of the M-wave depends on the stimulation intensity (defined by the product of current amplitude and pulse-width) and the muscle state [186]. The remaining EMG signal (after the M-wave has declined and until the next pulse/doublet will arrive) mainly contains normal volitional EMG activity which we are interested in.

Whenever a stimulation pulse is applied, the TTL-signal from the stimulator changes its level. A stimulation frequency of 20 Hz was used. The trigger signal arrives around 2 ms before any stimulation artifacts are visible in the signals and is registered by the PHYSIOSENSE device. The last BI- and EMG-values are held for 2 ms to 12.5 ms after the first trigger arrives. Thus, spikes induced by FES are suppressed from the BI and EMG signal by blanking.

The spike-free EMG signal is filtered by an elliptic infinite impulse response (IIR) high-pass filter of third order with a cut-off frequency of 80 Hz. This filter removes low frequent movement artifacts and supply voltage noise around 50 Hz. The selected cut-off frequency of 80 Hz minimizes the filter transients caused by the stimulation pulses and signal muting, as the step response time is sufficiently small. The normal frequency range of EMG lays in a range from 5 Hz to 400 Hz [187]. It is therefore accepted that the high-pass filter reduces the EMG signal power by approximately two [187], as a sufficient amount of EMG signal power remains.

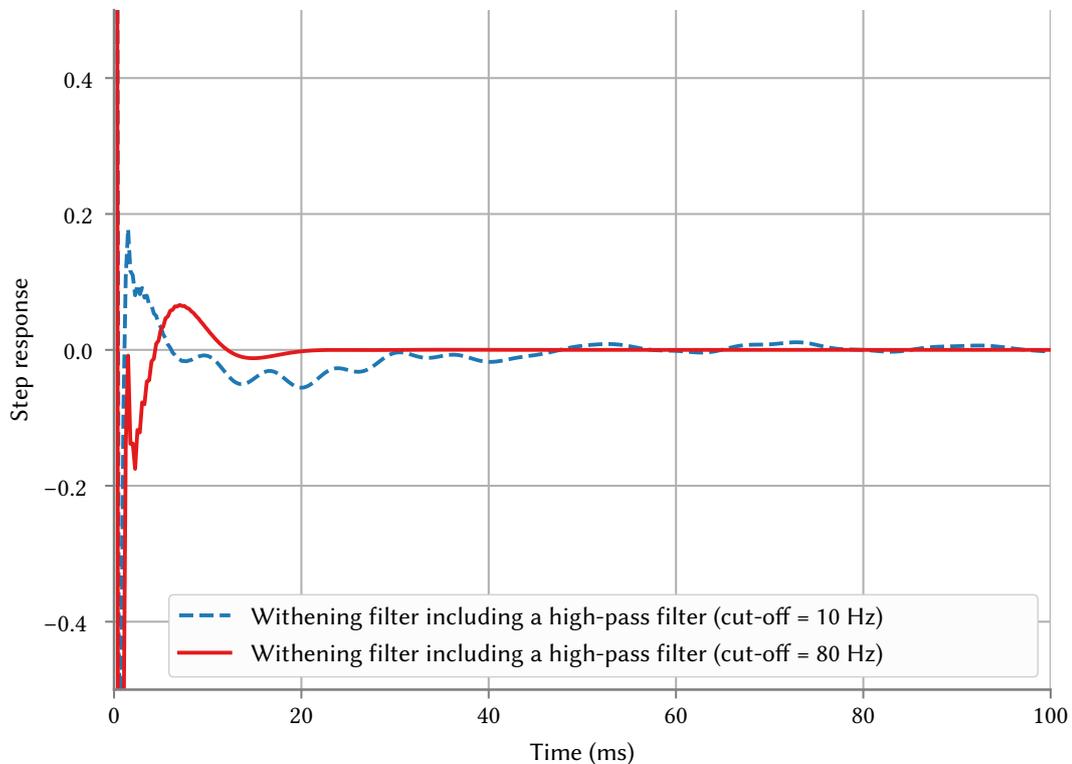
In order to increase the signal to noise ratio and to reduce the sample rate, a PAA is applied to the BI signal. The PAA divides the signal into equally spaced segments of 25 samples and uses the mean value as new sample value for each segment. Applying the PAA method to the BI signal



**Figure 6.3:** Signal processing of BI/EMG signals during active FES. In the upper left subplot, an EMG recording is shown in which stimulation artifacts are visible. By transmitting the TTL output of the simulator synchronized to the BI/EMG signal to the PC, the time positions of the stimulation artifacts are known and the stimulation artifacts can be muted. In the upper left subplot, a EMG recording with visible stimulation artifacts is displayed. In the upper right subplot, the result after muting and high-pass filtering with a cut-off frequency of 80 Hz is shown. In the lower left subplot, the BI recording with visible stimulation artifacts is displayed. The result after muting and approximation with piece-wise aggregate approximation (PAA) is shown in the lower right subplot. By applying PAA, the resulting sampling frequency is reduced from 4 kHz to 160 Hz.

reduces the sampling rate from 4 kHz to 160 Hz. The signal processing of BI/EMG signals during electrical stimulation is exemplarily shown in Figure 6.3.

### 6.3.3 Online Detection of EMG Activity Periods



**Figure 6.4:** Step response of the modified online EMG filter including whitening and high-pass filter (cut-off frequency of 80 Hz). For comparison, the step response of the complete offline EMG high-pass filter with an additional whitening is shown.

Muscle activity periods are detected by using a modification of the double-threshold detector that is described in Section 4.3.4. The modified version is able to detect EMG activity on streamed EMG measurement data in real-time.

As described in Section 4.3.4, a whitening filter (cf. Eq. (4.3)) is applied to the pre-processed EMG signal in order to give the measured EMG signal Gaussian characteristics. Figure 6.4 shows a comparison of two EMG filter step responses. One is the offline version from Chapter 4.3.4 including the filters from pre-processing together with the whitening filter. The other is the modified online version including also the filters from pre-processing together with the whitening filter which is introduced in this chapter. By increasing the cut-off frequency of the high-pass and removing both stop-band filters, the transit time of the step response could be successfully decreased.

Any EMG activity at the current measurement sample  $x$  is detected and detector output  $act$  is set to one if in the previous  $m$  samples  $r_0$  (called second threshold) samples of the auxiliary sequence  $z(i) = x^2(i)$  are above the first threshold  $\hat{\zeta}$ . The current detector output  $act$  is set to zero when less than  $r_0$  samples from the previous  $m$  samples are above  $\hat{\zeta}$ .

Similar to Section 4.3.4.3, the following procedure is applied to find the optimal parameters  $m$ ,  $r_0$  and  $\hat{\zeta}$  for the online double-threshold detector:

1. The parameters  $m$  and  $r_0$  of the double-threshold detector are optimized once before the detector is applied to the measurement data. The optimization is performed as described in Section 4.3.4.4 for a desired false-alarm probability  $P_{fa}$ , a given worst case  $SNR_{min}$  and a pre-defined offset-detection latency  $tr_{max}$  of the detector.
2. The measurement data are stacked up in windows of length  $N_{\hat{\sigma}_n^2}$  (cf. Eq. (4.29)). Whenever a new window is filled up with measurement data, the window index  $k$  is increased by one and its variance  $\text{VAR}_x(k)$  is calculated according to Equation (4.30):

$$\text{VAR}_x(k) = \frac{1}{N_{\hat{\sigma}_n^2} - 1} \sum_{i=n-N_{\hat{\sigma}_n^2}}^n \left( x(i) - \frac{1}{N_{\hat{\sigma}_n^2}} \sum_{l=n-N_{\hat{\sigma}_n^2}}^n x(l) \right)^2, \quad (6.1)$$

where  $x(n)$  is the newest sample with  $n > N_{\hat{\sigma}_n^2}$ . The minimum of all variances of windows within the last  $T_{var}$  seconds is taken as estimate for  $\hat{\sigma}_n$ .

3. At the time point  $N_{\hat{\sigma}_n^2} T_s^{EMG}$ , the first variance estimation took place and an initial threshold  $\hat{\zeta}_{\hat{\sigma}_d^2=0}$  (assuming no disturbance and taking into account the estimated noise variance) is calculated according to Equation (4.32):

$$\hat{\zeta}_{\hat{\sigma}_d^2=0} = \hat{\sigma}_n^2 F_{\chi^2, \nu=1}^{-1} (1 - P_\zeta). \quad (6.2)$$

$T_s^{EMG}$  is the sampling time of the EMG signal. Thus, the online double-threshold detector can be applied to the measurement data after  $N_{\hat{\sigma}_n^2} T_s^{EMG}$  seconds.

4. Disturbance estimation for improving the threshold  $\hat{\zeta}$  is started after  $N_\zeta > N_{\hat{\sigma}_n^2}$  measurement samples have been collected. As described in Section 4.3.4.5, the threshold  $\hat{\zeta}$  will be stepwise increased until the desired  $P_{fa}$  is achieved. The optimal  $\hat{\zeta}$  is determined as follows:

The ratio  $R_\zeta(k)$  of the previously calculated variances  $\text{VAR}_x(k)$  to  $\hat{\sigma}_n^2$  as shown in Equation (4.33) is calculated. The index  $k$  goes from the oldest window within the last  $T_{var}$  seconds to the newest one. All windows with  $R_\zeta(k)$  lower than threshold  $R_\zeta^{max}$  which is calculated by using  $th_{\hat{\sigma}_d}$  (cf. Eq. (4.34)) are then selected. Then,  $N_\zeta$  samples are randomly taken from the selected data. Beginning from  $\hat{\zeta}_{\hat{\sigma}_d^2=0}$ ,  $\hat{\zeta}$  is increased and applied to the randomly selected data set. The onsets are counted which have more than  $r_0$  samples over  $\hat{\zeta}$ . The false-alarm probability  $\hat{P}_{fa}$  is estimated according to Equation (4.36).  $\hat{\zeta}$  is increased by  $\kappa_\delta$  (cf. Eq. (4.35)) until the estimated false-alarm probability  $\hat{P}_{fa}$  reaches the desired false-alarm probability  $P_{fa}$  or  $\kappa_{max}$  is reached. Optimization of  $\hat{\zeta}$  is repeated whenever  $N_{\hat{\sigma}_n^2}$  samples arrive.

### 6.3.4 Online Segmentation of BI Measurements

In order to detect changes in the BI signal, which are caused by swallowing, a piece-wise linear approximation (PLA)<sup>1</sup> is applied to the BI time series. The bottom-up algorithm which was used in

<sup>1</sup>In Chapter 4, the bottom-up algorithm was used for approximating the BI time series by line segments. These line segments were used in the presented valley detection algorithm, which was able to find all swallowing induced signal changes.

Chapter 4 cannot be used without modification, as this algorithm needs the complete data sequence for signal approximation. Thus, an online capable version is required in which the BI signal is successively approximated.

The Sliding Window And Bottom-up (SWAB) segmentation algorithm, which is described by Keogh et al. [159], has a better performance than a sliding window approach. It combines the sliding window approach with the bottom-up algorithm. Similar to the sliding window approach, the segments are consecutively created such that the squared error sum between each line to the corresponding time series is equal or is smaller than a predefined value.

In the beginning of each swallow, the BI signal starts to drop due to the movement of hyoid bone and larynx. In order to guarantee that a drop in the BI signal is detected as fast as possible, a modification of the SWAB algorithm is proposed. This modified PLA-method allows to detect changes within the BI signal with only a small delay.

The modified SWAB algorithm approximates gradually the time series with the segments  $P_l, l = 1, 2, \dots, N^P$  by iteratively applying the following steps while the number of line segments  $N^P$  increases with time:

1. Newly arrived measurement data are buffered until the squared error sum between a temporary line to the corresponding time series exceeds just a predefined error *max\_error*. Additionally, the maximum length of a segment is bounded by *max\_length* and the value difference of the same segment is bounded by *max\_diff*. The algorithm that checks these conditions is called the BEST\_LINE function (cf. Fig. 6.5). Thus, a new temporary line is created when the BEST\_LINE function returns.

A buffer  $w$  with varying length is filled up with data samples from the original pre-processed time series whenever such a new temporary line is created. The number of data samples which are added to  $w$  is defined by the length of the temporary line.

2. The bottom-up algorithm (cf. Fig. 4.13) is applied to the data within the buffer  $w$  and creates  $N^{P^w}$  new line segments  $P_j^w, j = 1, 2, \dots, N^{P^w}$  approximating the data in the buffer  $w$ . The slope of the most recent line segment is named *act\_slope* and is used later as a feature for the swallow onset detection.

The number of lines is reduced to *max\_seg* and only the newest *max\_seg* lines are remaining in the buffer. The modified SWAB function continuously gives back all line segments within the buffer which exceed the number of lines *max\_seg*. The oldest  $N^{P^w} - \text{max\_seg}$  segments are appended to the segment list  $P_l, l = 1, 2, \dots, N^P$ . The data samples in  $w$  which correspond to these first  $N^{P^w} - \text{max\_seg}$  lines are removed from the buffer  $w$ .

If *max\_seg* is set to zero, the modified SWAB algorithm converts to the sliding window algorithm. If *max\_seg* is set to infinity, the results of the SWAB-algorithm are equal to the results of the bottom-up algorithm. The implementation of the modified SWAB algorithm is shown in Figure 6.6.

The BEST\_LINE function uses the CALCULATE\_ERROR function for calculating the squared error sum  $e_i$  of the samples  $BI^{P_i}(k)$  of line segment  $P_i$  and the BI signal  $BI(k)$  between the starting

```

1: procedure BEST_LINE({BI}, i)                                ▷ streaming time series {BI}, time series index i
2:   j = i + 2
3:   P = [j - i, BI(i), BI(j)]
4:   while CALCULATE_ERROR(P) <= max_error AND data at input do
5:     j ← j + 1
6:     P = [j - i, BI(i), BI(j)]
7:   end while
8:   return P ▷ line segment P for which the squared error sum from i to i + P(1) of {BI} just
   exceeds max_error.
9: end procedure

```

**Figure 6.5:** Implementation of the BEST\_LINE function. The CALCULATE\_ERROR function is defined in Equation (6.3). The parameter *max\_error* is the maximal allowed squared error.

point  $a_i$  of the line segment and the ending point  $b_i$ :

$$e_i = \text{CALCULATE\_ERROR}(P_i) = \begin{cases} \max\_error + \epsilon & \text{if } (P_i(3) - P_i(2)) \geq \max\_diff \\ \max\_error + \epsilon & \text{else if } P_i(1) \geq \max\_length \\ \sum_{k=a_i}^{b_i} (BI(k) - BI^{P_i}(k))^2 & \text{otherwise} \end{cases} \quad (6.3)$$

The squared error sum  $e_i$  is set to  $\max\_error + \epsilon^2$  for a segment  $i$  with length equal to or greater than  $\max\_length$  or for a value difference between the starting point and the ending point equal to or greater than  $\max\_diff$ , where  $\epsilon > 0$  is a small number. Both measures prevent lines with a length greater than  $\max\_length$  and a value difference greater than  $\max\_diff$ . The maximum size of the buffer  $w$  is bounded by the product  $\max\_length \cdot \max\_seg$  which bounds also the computational complexity.

When  $\max\_diff$  is set smaller than the BI valley height during swallowing, the modified SWAB algorithm will create several line segments in the beginning of a swallow when the BI starts to drop. Without defining a difference value boundary  $\max\_diff$ , the algorithm may return only one line segment from the starting point to the minimum point. As only at the end of a line segment its slope is known, FES can only be activated at the end of a line segment. When line segments end at the minimum point of a swallow, it would be too late for supporting the swallow with FES. Thus, a bounded line difference by using  $\max\_diff$  leads to more lines in the beginning of a swallow and the detection algorithm is able to start the swallowing support by FES in time.

In Figure 6.7, the influence of the  $\max\_diff$  parameter on the online slope detection is shown. The threshold  $\max\_error$  and the number of segments  $\max\_seg$  is set to one. The  $\max\_diff$  parameter is set to  $0.25\Omega$  and to infinity. It is shown at which time points slope informations (*act\_slope*) are available, beginning from the swallow start. Whenever the BEST\_LINE function gives back that a new segment can be created from the most recent measurement data, the bottom-up algorithm is called and the data within the buffer is approximated by segments. The slope of the most recent segment in the buffer is called *act\_slope*. In the second and third subplot, the *act\_slope* is plotted whenever the bottom-up algorithm function was called and more than one segment was

---

<sup>2</sup>An error greater than  $\max\_error$  stops the approximation process immediately.

```

1: procedure SWAB( $\{BI\}$ ) ▷ streaming time series  $\{BI\}$ 
2:    $w_{end} \leftarrow 0$  ▷ buffer  $w$  contains data from 1 to  $w_{end}$ 
3:   repeat
4:      $P^{SW} \leftarrow \text{BEST\_LINE}(\{BI\}, w_{end} + 1)$ 
5:      $w_{end} \leftarrow w_{end} + P^{SW}(1)$ 
6:      $P_j^w \leftarrow \text{PLA\_BOTTOM\_UP}(BI[1, w_{end}])$  ▷  $j = 1, \dots, N^{P^w}$ 
7:   until  $N^{P^w} < \text{max\_seg}$  AND  $w_{end} < N^{BI}$  ▷ The buffer  $w$  contains at least  $\text{max\_seg}$  line segments within  $P_j^w$ .
8:    $w_{start} \leftarrow 1$  ▷ buffer  $w$  contains data from  $w_{start}$  to  $w_{end}$ 
9:   repeat
10:    while  $N^{P^w} > \text{max\_seg}$  do
11:       $P_l \leftarrow \text{CONCAT}(P_l, P_1^w);$  ▷ Add segment  $P_1^w$  to  $P_l$ ,  $l = 1, \dots, N^P$ .
12:       $w_{start} \leftarrow w_{start} + P_1^w(1)$  ▷ Remove entries from  $w$ .
13:      remove first segment  $P_1^w$  from  $P^w$ 
14:    end while
15:     $P^{SW} \leftarrow \text{BEST\_LINE}(\{BI\}, w_{end} + 1)$  ▷ Waits for enough data samples.
16:     $w_{end} \leftarrow w_{end} + P^{SW}(1)$ 
17:     $P_j^w \leftarrow \text{PLA\_BOTTOM\_UP}(BI[w_{start}, w_{end}])$  ▷  $j = 1, \dots, N^{P^w}$ 
18:     $\text{act\_slope} \leftarrow$  slope of the last segment of  $P^w$ 
19:  until data at input  $\{BI\}$ 
20:   $P_l \leftarrow \text{CONCAT}(P_l, P_j^w)$  ▷ add remaining segments to  $P_l$ ,  $l = 1, \dots, N^P$ 
21:  return  $P_l$  ▷  $l = 1, \dots, N^P$ 
22: end procedure

```

**Figure 6.6:** Implementation of the modified SWAB algorithm. The algorithm returns the piece-wise linear approximation of  $\{BI\}$ . While the algorithm is running, the segments  $P_l$ ,  $l = 1, \dots, N^P$  and the most recent line segment  $\text{act\_slope}$  are used as features for the swallow onset detection. The CONCAT function is used for merging segments to  $P_l$  by appending them. The number of segments  $N^P$  is then increased by the number of appended segments. The BEST\_LINE function is described in Figure 6.5. The parameter  $\text{max\_error}$  is the maximal allowed squared error and  $\text{max\_seg}$  is the maximal allowed number of segments in the buffer  $w$ . The PLA\_BOTTOM\_UP function is shown in Figure 4.13.

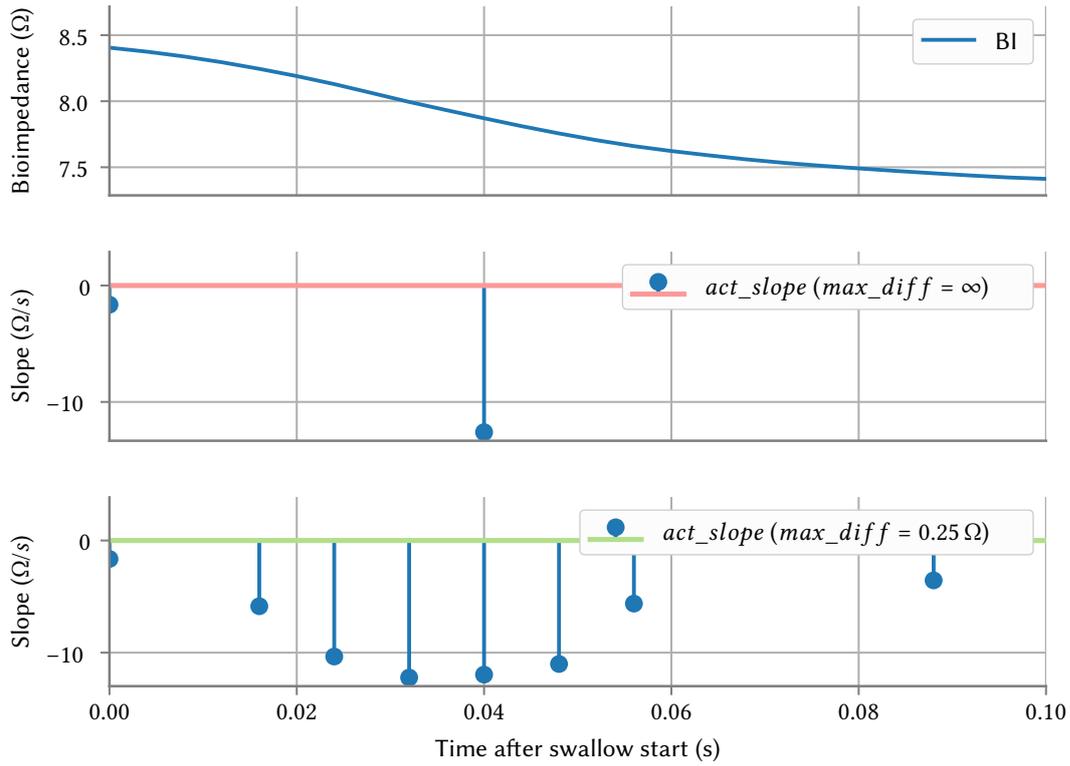
calculated. It can be seen that the  $\text{max\_diff}$  parameter decreases the time delay from 40 ms to 16 ms as the  $\text{max\_diff}$  parameter forces the BEST\_LINE function and the bottom-up algorithm to create more segments when the signal changes.

### 6.3.5 Control of FES

The algorithm for controlling FES for supporting swallowing should detect the voluntary intention of swallowing far before the minimum in the BI signal is reached. The algorithm should robustly ignore all other kinds of artifacts such as jaw, tongue, or head movements.

Whenever EMG activity has been detected by the double-threshold detector and a sufficient drop in the BI signal is sensed, a stimulation of the submental muscles may be activated until the BI signal increases again. This allows supporting the movement of hyoid bone and larynx by means of FES within the pharyngeal swallowing phase.

The line segments  $P_j^w$ ,  $j = 1, \dots, N^{P^w}$  within the buffer  $w$  and the line segments  $P_l$ ,  $l = 1, \dots, N^P$  outside the buffer can be used for feature calculation. Whenever the BEST\_LINE function creates



**Figure 6.7:** The real-time performance of the modified SWAB algorithm for detection of slope changes is shown on exemplary data. Besides the number of segments ( $max\_seg = 1$ ) in the buffer  $w$  and the maximal approximation error ( $max\_error = 1$ ), the modified SWAB algorithm has an additionally parameter  $max\_diff$ . The influence of this parameter is compared by setting it to  $0.25\Omega$  and to  $\infty$ . This parameter ensures that a new segment is created whenever the difference of the start point to the end point exceeds  $max\_diff$ . The BI signal is shown in the first subplot. In the second and the third subplot, the slope of the most recent segment in the buffer, which is called  $act\_slope$ , is plotted whenever more than  $max\_seg$  segments were calculated within the buffer. Two segments were firstly calculated after 40 ms after the begin of the swallow by setting the parameter  $max\_diff$  to infinity, whereas the algorithm with  $max\_diff = 0.25\Omega$  could give out a slope after 16 ms after the swallow start.

a new temporary line by which the buffer  $w$  is increased and the line segments  $P_j^w$  and  $P_l$  changes, FES is activated when following conditions are fulfilled:

- EMG activity must be present. This is the case if the output  $act$  from the online double-threshold detector is continuously one for at least  $emg_{activity}^{MIN}$  seconds within the newest line segment  $P_{N^{pw}}^w$  in the buffer  $w$ .
- The slope  $act\_slope$  of the newest line segment  $P_{N^{pw}}^w$  fulfills  $th_{lower}^{act\_slope} < act\_slope < th_{upper}^{act\_slope}$  and
- the slope value  $p\_slope$  of the newest line segments  $P_{N^p}$  fulfills  $th_{lower}^{p\_slope} < p\_slope < th_{upper}^{p\_slope}$ .

The threshold values  $th_{upper}^{act\_slope}$ ,  $th_{lower}^{act\_slope}$ ,  $th_{upper}^{p\_slope}$ , and  $th_{lower}^{p\_slope}$  are individually adjusted for each patient by recording one swallow and fitting all four thresholds such that an FES would have

been successfully activated in the beginning of this swallow. The algorithms from Section 4 are used to detect the exact start point of the recorded swallow. Then, the line approximation is applied to the BI measurement signal and the slopes of the lines before and after the swallow onset are calculated. The threshold values  $th_{upper}^{act\_slope}$  and  $th_{lower}^{act\_slope}$  are fitted to the slope of the first line slope after the swallow onset such that the distance between both thresholds and the slope itself is sufficient. The thresholds  $th_{upper}^{p\_slope}$  and  $th_{lower}^{p\_slope}$  are fitted to the slope of the first line segment before the swallow in an analogous manner.

When the conditions for activating FES are fulfilled and the manual switch is pressed by the therapist, FES is activated. The FES is set inactive again at the first positive value of  $act\_slope$ . It is assumed that a change in the slope indicates that hyoid bone and larynx have reached their maximum elevation and moving back to their rest position. Therefore, additionally stimulation support is normally not useful. For some patient, a longer stimulation could be helpful when e.g. the esophagus sphincter opens delayed. In these cases, the FES is extended by a fixed time period beginning from the first positive value of  $act\_slope$ .

## 6.4 Results

One post stroke patient (62 years, male), who is still suffering on the locked in syndrome and had difficulties in swallowing such that his trachea had to be blocked by a tracheal cannula participated in this study. The trial was performed at the Unfallkrankenhaus Berlin in collaboration with Rainer O. Seidl.

The swallowing performance of the patient with and without FES support was investigated. The patient was given an explanation and he signed a consent. The study was approved by the ethic board at the Charité Berlin (EA1/160/09).

The stimulation electrode setup with two stimulation channels as shown in Figure 6.2 was applied to the patient. The biphasic stimulation pattern with doublets (inter-pulse-interval of 5 ms) was set to a pulse width of 200  $\mu$ s and a stimulation frequency of 20 Hz. These simulation parameters were fixed during the whole experiment for both stimulation channels. In the beginning, the maximally tolerated stimulation current was determined. The applied current amplitudes were carefully chosen for each channel such that the patient felt comfortable and a clearly visible movement of the stimulated muscles could be observed. Both stimulation channels were activated and deactivated simultaneously.

The switch was operated by a therapist. Whenever the patient had prepared the bolus, the therapist pressed the switch and asked the patient to swallow. The control of FES was activated when the switch was pressed. The switch did not control the FES itself. This was necessary, as patients may have a long preparation phase in which head, jaw, or tongue movements could lead to a similar drop in BI with EMG activity as swallowing.

FES was then automatically activated as soon as EMG activity was present and the BI began to decrease until the BI was rising again. In Figure 6.8, an FES-supported swallow is compared to swallow without stimulation.

In total, the BI drop and the slope during the swallow onset were calculated from 24 recorded swallows from which 19 were supported by controlled FES based on EMG and BI measurements.

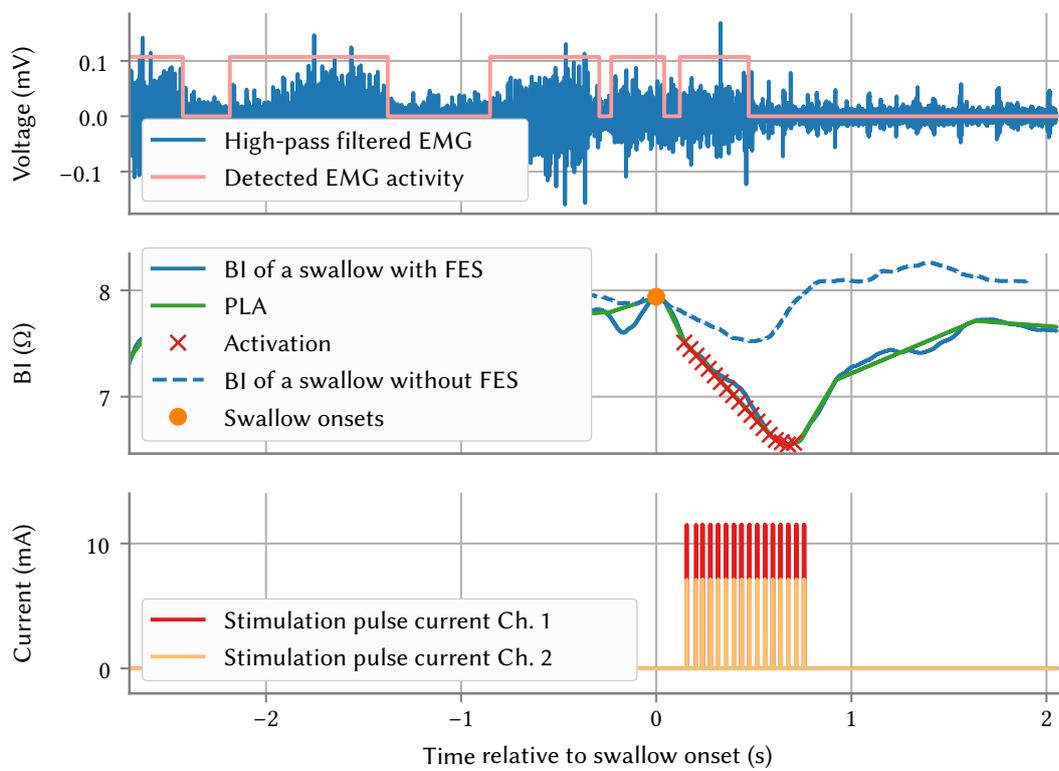


Figure 6.8: Comparison of a swallow with and without stimulation support.

**Table 6.1:** List of parameters for the online swallowing segmentation.

Name	Default	Reason for parameter choice
$tr_{max}$	11 ms	good offset-detection latency of the detector
$P_{fa}$	0.01	results in a sufficiently high detection probability $P_d$
$SNR_{min}$	3 dB	minimal EMG that shall be detected
$N_{\hat{\sigma}_n^2} T_s^{EMG}$	0.25 s	good window length for estimating $\hat{\sigma}_n^2$
$N_{\hat{\zeta}} T_s^{EMG}$	2 s	good window length for estimating $\hat{\zeta}$
$T_{var}$	10 s	results in a good adaption of $\hat{\zeta}$ and $\hat{\sigma}_n^2$
$th_{\bar{\sigma}_d}$	0.3	results in a good adaption of $\hat{\zeta}$
$emg_{activity}^{MIN}$	30 ms	smallest possible muscle activation period
$\kappa_{max}$	2.2	limits computational time sufficiently
$\kappa_{delta}$	0.1	leads to sufficient $\hat{\zeta}$ estimation
$max\_error$	0.02	leads to good BI approximation
$max\_diff$	0.02	reduces the time delay for slope estimation
$max\_length T_s^{EMG}$	1 s	bounding the maximum line segment length
$max\_seg$	1	decreases time delay much as possible
$th_{upper}^{act\_slope}$	-0.5	leads to sufficient swallow onset detection
$th_{lower}^{act\_slope}$	-10	prevents stimulation on data outliers
$th_{upper}^{p\_slope}$	-0.59	leads to sufficient stimulation in the onset of a swallow
$th_{lower}^{p\_slope}$	-10	leads to sufficient swallow onset detection

**Table 6.2:** Effect of different stimulation currents for the first stimulation channel 1 and the second stimulation channel 2 during swallowing stimulation patterns with doublets (inter-pulse-interval of 5 ms), a pulse width of 200  $\mu$ s and a frequency of 20 Hz. The electrode setup is shown in Figure 6.2. Stimulation is only activated during the elevation of the larynx. The normative values of healthy people are provided as a comparison. The given percentages show the relation between the patient data and the normative values of healthy people. \*\* - significant level  $p = 0.01$ . The results were also presented in [179] and [180].

	healthy	patient			
Pulse current Ch. 1	without	without	7.1 mA	9.4 mA	11.8 mA
Pulse current Ch. 2			0 mA	7.1 mA	7.1 mA
Number of swallows	328	5	6	9	4
Drop in BI ( $\Omega$ )	$-1.52 \pm 0.56$	-0.56 (37 %)	-0.55 (36 %)	-0.78 (51 %)	-1.12** (74 %)
Rate of change ( $\Omega/s$ )	$-6.43 \pm 2.57$	-1.87 (29 %)	-2.56 (40 %)	-2.82** (44 %)	-3.03** (47 %)

The patient was fed with thickened liquid. Three different stimulation intensities were applied to the patient. The lowest applied stimulation current did not lead to an increase in the BI drop and the slope increased only slightly; the highest applied stimulation intensity caused larger and faster BI drops during swallowing. The measured drop and slope of each swallow related BI valley

are displayed in Table 6.2. Additionally, the normative values of healthy people are provided as a comparison [147]. For the highest applied stimulation current, 74 % of the BI drop of healthy people and 47 % of the rate of change from healthy people could be achieved.

### 6.5 Conclusions

The previously developed algorithms for the swallowing detection from Chapter 4 could be adapted for automatically activating FES in the very beginning of a swallow in real-time. Signal processing for the EMG signal had to be reduced to a high-pass filter. During active stimulation, artifacts could be muted as the stimulation periods were transmitted simultaneously to the PC. The BI is processed by a PAA reduction, which is used as a simple low-pass filter. The double-threshold detector could be applied, but the estimation of the noise variance and the parameter  $\hat{\zeta}$  had to be adapted. The valley detection algorithm which was successfully used in the swallowing segmentation could not be utilized as FES should be triggered at the beginning of the valley. A modified Sliding Window And Bottom-up (SWAB) algorithm was proposed and was able to estimate the slope of the BI signal such that directly after the onset of a swallow an FES could be activated.

On one patient it has been shown, that the presented system is able to detect the beginning of a voluntarily induced swallow. The time point in which FES was activated was set early enough in respect to the swallow onset in order to stimulate the submental musculature for achieving a positive effect regarding elevation and acceleration of hyoid and larynx. The patient easily adapted to the stimulation system.

It can be seen that due to the stimulation a significant faster and deeper decrease in the BI measurement signal could be produced. As shown in Chapter 5, BI correlates to the movement of hyoid and larynx. A greater and a faster decrease in the BI signal can, therefore, be interpreted as a wider and faster movement of hyoid and larynx in anterior and superior direction.

As the system was only tested on one patient, a study with a larger number of patients should be performed in the future. The stimulation parameters and placement of the stimulation electrodes have to be optimized. The methods from Chapter 4 could be used for segment already recorded swallows. Thus, the classification thresholds for FES control could be adapted from swallow to swallow.

In the presented approach, control of the stimulation was only activated after the patient had prepared the bolus and was ready to swallow. A more robust control which does not react to chewing and head movements could greatly improve the application. This could be accomplished by adding additional sensors.

It has been shown that increasing the stimulation intensity has also a positive effect on the BI drop during swallowing and the elevation speed. Thus, a feedback system could be implemented by which the stimulation intensity is automatically controlled to maintain the desired elevation of larynx and hyoid.

# 7

## Conclusion and Future Work

A measurement device has been developed which can be used for recording electromyography (EMG) and bioimpedance (BI) at the neck for the assessment of the pharyngeal swallowing phase in patients with swallowing disorders (dysphagia). A heuristic BI valley search algorithm was presented in order to find all swallows in the measurement data. The algorithm for segmenting all valleys in the BI data is based on an efficient piece-wise linear approximation (PLA) which is an approximation of the time series into linked line segments of different length. The obtained line segments were combined with the extracted EMG activities which are detected by a robust and adaptive double-threshold detector. All line segments which form a valley and coincide with EMG activity are marked as possible swallows. From these valleys, a feature set is extracted and a trained support vector machine (SVM) could separate non-swallowing related artifacts from swallows. The offline approach was tested on the data from healthy subjects which results in a very good accuracy regarding sensitivity and specificity. The swallowing detection algorithm was also successfully tested on data from patients.

By applying the valley search algorithm on measurement data which were recorded parallel to a videofluoroscopy (VF), the correlation between BI and the movement of hyoid bone and larynx was investigated. It was shown that the movement of both structures correlates to BI during a swallow.

By rendering the previously developed offline methods real-time capable, control of functional electrical stimulation (FES) based on BI and EMG measurements could be realized. It was possible to activate FES correctly timed to voluntary evoked swallows. An electrode position for effectively stimulating the pharyngeal muscles was found. EMG and BI could be measured also during active stimulation with the presented measurement system as all stimulation artifacts in the measurement data could be muted. In a pilot study, FES of the pharyngeal muscles controlled by the developed algorithm could successfully improve swallowing on a single post-stroke patient.

### Future Work

It could be shown that a combined BI/EMG measurement at the neck can be used for automatic segmentation of swallows. In combination with a video-based diagnostic system, such as fiber-

optic endoscopic evaluation of swallowing (FEES), a more powerful system can be created to support the examination of swallowing disorders. Periods in the recorded video in which the patient had swallowed could be automatically selected and presented to the examiner. It would be possible to automatically collect all passages which an examiner could use for diagnostic purposes. Due to the automatic segmentation of swallows, database-based machine learning approaches for evaluating the video images would become feasible.

Swallowing parameters could be extracted from recorded BI/EMG measurements promptly after the swallow by using the presented segmentation algorithms. Thus, biofeedback of the swallowing performance could be provided to the patient for training purposes. In combination with a deep image camera, posture correction together with swallowing segmentation could help patients in training swallowing in different head positions and swallowing maneuvers. Due to the easily attachable sensor concept, the outcome of therapeutic measures could be regularly monitored. Systems based on BI/EMG measurements could be utilized also outside of a clinical environment.

As it could be shown that the measurement setup is feasible for controlling FES accurately timed for swallowing support, a combined measurement/stimulation system could be used for swallowing therapy and long-term swallowing support.

## References

- [1] C. Ertekin and I. Aydogdu, “Neurophysiology of swallowing”, *Clinical Neurophysiology*, vol. 114, no. 12, pp. 2226–2244, 2003, ISSN: 1388-2457. DOI: 10.1016/S1388-2457(03)00237-2.
- [2] H. Gray and W. H. Lewis, *Anatomy of the human body*, 20th ed. Philadelphia: Lea & Febiger and Bartleby.com, 1918, vol. 36, ISBN: 9781587341021. DOI: 10.5962/bhl.title.20311.
- [3] D. W. Buchholz, J. F. Bosma, and M. W. Donner, “Adaptation, compensation, and decompensation of the pharyngeal swallow”, *Gastrointestinal radiology*, vol. 10, no. 3, pp. 235–239, 1985, ISSN: 0364-2356. DOI: 10.1007/BF01893106.
- [4] R. Shaker, J. Ren, E. Bardan, C. Easterling, K. Dua, P. Xie, and M. Kern, “Pharyngoglottal closure reflex: characterization in healthy young, elderly and dysphagic patients with predeglutitive aspiration”, *Gerontology*, vol. 49, no. 1, pp. 12–20, 2003, ISSN: 0304-324X. DOI: 10.1159/000066504.
- [5] I. J. Cook, W. J. Dodds, R. O. Dantas, B. Massey, M. K. Kern, I. M. Lang, J. G. Bresseur, and W. J. Hogan, “Opening mechanisms of the human upper esophageal sphincter”, *The American journal of physiology*, vol. 257, no. 5 Pt 1, G748–G759, 1989, ISSN: 0002-9513.
- [6] O. Ekberg, *Dysphagia: diagnosis and treatment*, ser. Medical radiology. Berlin and New York: Springer, 2012, ISBN: 978-3-642-17887-0.
- [7] J. A. Y. Cichero and B. E. Murdoch, *Dysphagia: foundation, theory and practice*. Chichester, England and New York: Wiley, 2006, ISBN: 9781861565051.
- [8] R. Martino, N. Foley, S. Bhogal, N. Diamant, M. Speechley, and R. Teasell, “Dysphagia after stroke: incidence, diagnosis, and pulmonary complications”, *Stroke*, vol. 36, no. 12, pp. 2756–2763, 2005. DOI: 10.1161/01.STR.0000190056.76543.eb.
- [9] C. D. van der Maarel-Wierink, J. N. O. Vanobbergen, E. M. Bronkhorst, J. M. G. A. Schols, and C. de Baat, “Meta-analysis of dysphagia and aspiration pneumonia in frail elders”, *Journal of dental research*, vol. 90, no. 12, pp. 1398–1404, 2011, ISSN: 1544-0591. DOI: 10.1177/0022034511422909.
- [10] G. Mann, G. J. Hankey, and D. Cameron, “Swallowing function after stroke: prognosis and prognostic factors at 6 months”, *Stroke; a journal of cerebral circulation*, vol. 30, no. 4, pp. 744–748, 1999, ISSN: 1524-4628. DOI: 10.1161/01.STR.30.4.744.

## REFERENCES

---

- [11] U. Frank, M. Mäder, and H. Sticher, “Dysphagic patients with tracheotomies: a multidisciplinary approach to treatment and decannulation management”, *Dysphagia*, vol. 22, no. 1, pp. 20–29, 2007, ISSN: 0179-051X. DOI: 10.1007/s00455-006-9036-5.
- [12] B. R. Garon, M. Engle, and C. Ormiston, “Silent Aspiration: Results of 1,000 videofluoroscopic swallow evaluations”, *Neurorehabilitation and Neural Repair*, vol. 10, no. 2, pp. 121–126, 1996. DOI: 10.1177/154596839601000207.
- [13] A. M. Kelly, P. Leslie, T. Beale, C. Payten, and M. J. Drinnan, “Fibreoptic endoscopic evaluation of swallowing and videofluoroscopy: does examination type influence perception of pharyngeal residue severity?”, *Clinical Otolaryngology*, vol. 31, no. 5, pp. 425–432, 2006, ISSN: 1749-4486. DOI: 10.1111/j.1749-4486.2006.01292.x.
- [14] H. Tohara, A. Nakane, S. Murata, S. Mikushi, Y. Ouchi, Y. Wakasugi, M. Takashima, Y. Chiba, and H. Uematsu, “Inter- and intra-rater reliability in fibroptic endoscopic evaluation of swallowing”, *Journal of oral rehabilitation*, vol. 37, no. 12, pp. 884–891, 2010, ISSN: 1365-2842. DOI: 10.1111/j.1365-2842.2010.02116.x.
- [15] K. R. Mills, “The basics of electromyography”, *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 76, no. suppl 2, pp. ii32–ii35, 2005, ISSN: 1468-330X. DOI: 10.1136/jnnp.2005.069211.
- [16] A. Rainoldi, J. Bullock-Saxton, F. Cavarretta, and N. Hogan, “Repeatability of maximal voluntary force and of surface EMG variables during voluntary isometric contraction of quadriceps muscles in healthy subjects”, *Journal of Electromyography and Kinesiology*, vol. 11, no. 6, pp. 425–438, 2001, ISSN: 10506411. DOI: 10.1016/S1050-6411(01)00022-0.
- [17] J. R. Daube and D. I. Rubin, “Needle electromyography”, *Muscle & Nerve*, vol. 39, 2009. DOI: 10.1002/mus.21180.
- [18] M. Vaiman and E. Eviatar, “Surface electromyography as a screening method for evaluation of dysphagia and odynophagia”, *Head & face medicine*, vol. 5, no. 1, p. 9, 2009, ISSN: 1746-160X. DOI: 10.1186/1746-160X-5-9.
- [19] S. Grimnes and Ø. G. Martinsen, *Bioimpedance and bioelectricity Basics: Second Edition*. Academic Press, 2008, ISBN: 9780123740045. DOI: 10.1016/B978-0-12-374004-5.00007-6.
- [20] U. G. Kyle, I. Bosaeus, A. D. de Lorenzo, P. Deurenberg, M. Elia, J. M. Gomez, B. L. Heitmann, L. Kent-Smith, J.-C. Melchior, M. Pirlich, H. Scharfetter, A. M. W. J. Schols, and C. Pichard, “Bioelectrical impedance analysis—part I: review of principles and methods”, *Clinical nutrition (Edinburgh, Scotland)*, vol. 23, no. 5, pp. 1226–1243, 2004, ISSN: 0261-5614. DOI: 10.1016/j.clnu.2004.06.004.
- [21] M. J. Osypka, D. P. Bernstein, M. J. Osypka, and D. P. Bernstein, “Electrophysiologic principles and theory of stroke volume determination by thoracic electrical bioimpedance”, *AACN Clinical Issues: Advanced Practice in Acute and Critical Care*, vol. 10, no. 3, pp. 385–399, 1999, ISSN: 1079-0713. DOI: 10.1097/00044067-199908000-00008.

- [22] T. Kusuhara, T. Nakamura, Y. Shirakawa, K. Mori, Y. Naomoto, and Y. Yamamoto, "Impedance pharyngography to assess swallowing function", *Journal of International Medical Research*, vol. 32, no. 6, pp. 608–616, 2004, ISSN: 0300-0605. DOI: 10.1177/147323000403200605.
- [23] D. N. Rushton, "Functional electrical stimulation", *Physiological Measurement*, vol. 18, no. 4, pp. 241–275, 1997, ISSN: 0967-3334. DOI: 10.1088/0967-3334/18/4/001.
- [24] E. Henneman, G. Somjen, and D. O. Carpenter, "Excitability and inhibibility of motoneurons of different sizes", *Journal of neurophysiology*, vol. 28, no. 3, pp. 599–620, 1965, ISSN: 0022-3077.
- [25] H. S. Milner-Brown, R. B. Stein, and R. Yemm, "The orderly recruitment of human motor units during voluntary isometric contractions", *The Journal of Physiology*, vol. 230, no. 2, pp. 359–370, 1973.
- [26] P. H. Gorman and J. T. Mortimer, "The effect of stimulus parameters on the recruitment characteristics of direct nerve stimulation", *Biomedical Engineering, IEEE Transactions on*, vol. BME-30, no. 7, pp. 407–414, 1983, ISSN: 0018-9294. DOI: 10.1109/TBME.1983.325041.
- [27] M. Levy, J. Mizrahi, and Z. Susak, "Recruitment, force and fatigue characteristics of quadriceps muscles of paraplegics isometrically activated by surface functional electrical stimulation", *Journal of Biomedical Engineering*, vol. 12, no. 2, pp. 150–156, 1990, ISSN: 01415425. DOI: 10.1016/0141-5425(90)90136-B.
- [28] P. E. Crago, P. H. Peckham, and G. B. Thrope, "Modulation of muscle force by recruitment during intramuscular stimulation", *IEEE Transactions on Biomedical Engineering*, vol. BME-27, no. 12, pp. 679–684, 1980, ISSN: 0018-9294. DOI: 10.1109/TBME.1980.326592.
- [29] M. Bülow, R. Speyer, L. Baijens, V. Woisard, and O. Ekberg, "Neuromuscular electrical stimulation (NMES) in stroke patients with oral and pharyngeal dysfunction", *Dysphagia*, vol. 23, no. 3, pp. 302–309, 2008, ISSN: 0179-051X. DOI: 10.1007/s00455-007-9145-9.
- [30] T. A. Burnett, E. A. Mann, S. A. Cornell, and C. L. Ludlow, "Laryngeal elevation achieved by neuromuscular stimulation at rest", *Journal of applied physiology (Bethesda, Md. : 1985)*, vol. 94, no. 1, pp. 128–134, 2003, ISSN: 8750-7587. DOI: 10.1152/japplphysiol.00406.2002.
- [31] H. Kagaya, M. Baba, E. Saitoh, S. Okada, M. Yokoyama, and Y. Muraoka, "Hyoid bone and larynx movements during electrical stimulation of motor points in laryngeal elevation muscles: a preliminary study", *Neuromodulation : journal of the International Neuromodulation Society*, vol. 14, no. 3, 278–83, discussion 283, 2011, ISSN: 1525-1403. DOI: 10.1111/j.1525-1403.2011.00331.x.
- [32] I. A. Humbert, C. J. Poletto, K. G. Saxon, P. R. Kearney, L. Crujido, W. Wright-Harp, J. Payne, N. Jeffries, B. C. Sonies, and C. L. Ludlow, "The effect of surface electrical stimulation on hyolaryngeal movement in normal individuals at rest and during swallowing", *Journal of applied physiology (Bethesda, Md. : 1985)*, vol. 101, no. 6, pp. 1657–1663, 2006, ISSN: 8750-7587. DOI: 10.1152/japplphysiol.00348.2006.

## REFERENCES

---

- [33] C. L. Ludlow, I. Humbert, K. Saxon, C. Poletto, B. Sonies, and L. Crujido, "Effects of surface electrical stimulation both at rest and during swallowing in chronic pharyngeal Dysphagia", *Dysphagia*, vol. 22, no. 1, pp. 1–10, 2007, ISSN: 0179-051X. DOI: 10.1007/s00455-006-9029-4.
- [34] A. J. Hadley, I. Kolb, and D. J. Tyler, "Laryngeal elevation by selective stimulation of the hypoglossal nerve", *Journal of neural engineering*, vol. 10, no. 4, p. 46 013, 2013, ISSN: 1741-2552. DOI: 10.1088/1741-2560/10/4/046013.
- [35] M. Broniatowski, S. Grundfest-Broniatowski, D. J. Tyler, P. Scolieri, F. Abbass, H. M. Tucker, and S. Brodsky, "Dynamic laryngotracheal closure for aspiration: a preliminary report", *The Laryngoscope*, vol. 111, no. 11, pp. 2032–2040, 2001, ISSN: 0023-852X. DOI: 10.1097/00005537-200111000-00031.
- [36] T. A. Burnett, E. A. Mann, J. B. Stoklosa, and C. L. Ludlow, "Self-triggered functional electrical stimulation during swallowing", *Journal of neurophysiology*, vol. 94, no. 6, pp. 4011–4018, 2005, ISSN: 0022-3077. DOI: 10.1152/jn.00025.2005.
- [37] V. Leelamanit, C. Limsakul, and A. Geater, "Synchronized electrical stimulation in treating pharyngeal dysphagia", *The Laryngoscope*, vol. 112, no. 12, pp. 2204–2210, 2002, ISSN: 0023-852X. DOI: 10.1097/00005537-200212000-00015.
- [38] B. a. Mathers-Schmidt and M. Kurlinski, "Dysphagia evaluation practices: inconsistencies in clinical assessment and instrumental examination decision-making", *Dysphagia*, vol. 18, no. 2, pp. 114–125, 2003, ISSN: 0179-051X. DOI: 10.1007/s00455-002-0094-z.
- [39] B. Martin-Harris and B. Jones, "The videofluorographic swallowing study", *Physical medicine and rehabilitation clinics of North America*, vol. 19, no. 4, pp. 769–85, viii, 2008, ISSN: 1047-9651. DOI: 10.1016/j.pmr.2008.06.004.
- [40] G. J. Bours, R. Speyer, J. Lemmens, M. Limburg, and R. de Wit, "Bedside screening tests vs. videofluoroscopy or fiberoptic endoscopic evaluation of swallowing to detect dysphagia in patients with neurological disorders: systematic review", *Journal of Advanced Nursing*, vol. 65, no. 3, pp. 477–493, 2009, ISSN: 1365-2648. DOI: 10.1111/j.1365-2648.2008.04915.x.
- [41] D. J. C. Ramsey, D. G. Smithard, and L. Kalra, "Can pulse oximetry or a bedside swallowing assessment be used to detect aspiration after stroke?", *Stroke*, vol. 37, no. 12, pp. 2984–2988, 2006. DOI: 10.1161/01.STR.0000248758.32627.3b.
- [42] S. K. Daniels, L. A. Ballo, M. C. Mahoney, and A. L. Foundas, "Clinical predictors of dysphagia and aspiration risk: outcome measures in acute stroke patients", *Archives of Physical Medicine and Rehabilitation*, vol. 81, no. 8, pp. 1030–1033, 2000, ISSN: 0003-9993.
- [43] K. L. DePippo, M. A. Holas, and M. J. Reding, "The Burke dysphagia screening test: validation of its use in patients with stroke", *Archives of Physical Medicine and Rehabilitation*, vol. 75, no. 12, pp. 1284–1286, 1994, ISSN: 0003-9993.

- [44] R. Martino, F. Silver, R. Teasell, M. Bayley, G. Nicholson, D. L. Streiner, and N. E. Diamant, “The Toronto bedside swallowing screening test (TOR-BSST): development and validation of a dysphagia screening tool for patients with stroke”, *Stroke*, vol. 40, no. 2, pp. 555–561, 2009. DOI: 10.1161/STROKEAHA.107.510370.
- [45] M. Trapl, P. Enderle, M. Nowotny, Y. Teuschl, K. Matz, A. Dachenhausen, and M. Brainin, “Dysphagia bedside screening for acute-stroke patients: the Gugging swallowing screen”, *Stroke; a journal of cerebral circulation*, vol. 38, no. 11, pp. 2948–2952, 2007, ISSN: 1524-4628. DOI: 10.1161/STROKEAHA.107.483933.
- [46] C. Schultheiss, R. Nusser-Müller-Busch, and R. O. Seidl, “The semisolid bolus swallow test for clinical diagnosis of oropharyngeal dysphagia: a prospective randomised study”, *Eur Arch Otorhinolaryngol (European Archives of Oto-Rhino-Laryngology)*, vol. 268, no. 12, pp. 1837–1844, 2011. DOI: 10.1007/s00405-011-1628-5.
- [47] J. C. Rosenbek, J. A. Robbins, E. B. Roecker, J. L. Coyle, and J. L. Wood, “A penetration-aspiration scale”, *Dysphagia*, vol. 11, no. 2, pp. 93–98, 1996, ISSN: 0179-051X. DOI: 10.1007/BF00417897.
- [48] D. H. Kim, K. H. Choi, H. M. Kim, J. H. Koo, B. R. Kim, T. W. Kim, J. S. Ryu, S. Im, I. S. Choi, S. B. Pyun, J. W. Park, J. Y. Kang, and H. S. Yang, “Inter-rater reliability of videofluoroscopic dysphagia scale”, *Annals of rehabilitation medicine*, vol. 36, no. 6, pp. 791–796, 2012, ISSN: 2234-0645. DOI: 10.5535/arm.2012.36.6.791.
- [49] S. O’Donoghue and A. Bagnall, “Videofluoroscopic evaluation in the assessment of swallowing disorders in paediatric and adult populations”, *Folia Phoniatrica et Logopaedica*, vol. 51, no. 4-5, pp. 158–171, 1999, ISSN: 1421-9972. DOI: 10.1159/000021494.
- [50] J. E. Aviv, S. T. Kaplan, J. E. Thomson, J. Spitzer, B. Diamond, and L. G. Close, “The safety of flexible endoscopic evaluation of swallowing with sensory testing (FEESST): an analysis of 500 consecutive evaluations”, *Dysphagia*, vol. 15, no. 1, pp. 39–44, 2000, ISSN: 0179-051X.
- [51] S. G. Hiss and G. N. Postma, “Fiberoptic endoscopic evaluation of swallowing”, *Laryngoscope*, vol. 113, pp. 1386–1393, 2003.
- [52] P. M. Zenner, D. S. Losinski, and R. H. Mills, “Using cervical auscultation in the clinical dysphagia examination in long-term care”, *Dysphagia*, vol. 10, no. 1, pp. 27–31, 1995, ISSN: 0179-051X. DOI: 10.1007/BF00261276.
- [53] C. Borr, M. Hielscher-Fastabend, and A. Lücking, “Reliability and validity of cervical auscultation”, *Dysphagia*, vol. 22, no. 3, pp. 225–234, 2007, ISSN: 0179-051X. DOI: 10.1007/s00455-007-9078-3.
- [54] P. Leslie, M. J. Drinnan, P. Finn, G. A. Ford, and J. A. Wilson, “Reliability and validity of cervical auscultation: a controlled comparison using videofluoroscopy”, *Dysphagia*, vol. 19, no. 4, pp. 231–240, 2004, ISSN: 0179-051X. DOI: 10.1007/s00455-004-0007-4.

- [55] C.-L. Peng, P.-G. Jost-Brinkmann, N. Yoshida, H.-H. Chou, and C.-T. Lin, "Comparison of tongue functions between mature and tongue-thrust swallowing—an ultrasound investigation", *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 125, no. 5, pp. 562–570, 2004, ISSN: 0889-5406. DOI: 10.1016/j.ajodo.2003.06.003.
- [56] K. Yabunaka, H. Sanada, S. Sanada, H. Konishi, T. Hashimoto, H. Yatake, K. Yamamoto, T. Katsuda, and M. Ohue, "Sonographic assessment of hyoid bone movement during swallowing: a study of normal adults with advancing age", *Radiological physics and technology*, vol. 4, no. 1, pp. 73–77, 2011, ISSN: 1865-0341. DOI: 10.1007/s12194-010-0107-9.
- [57] D. R. Scarborough, S. Waizenhofer, L. Siekemeyer, and M. Hughes, "Sonographically measured hyoid bone displacement during swallow in preschool children: a preliminary study", *Journal of clinical ultrasound : JCU*, vol. 38, no. 8, pp. 430–434, 2010, ISSN: 1097-0096. DOI: 10.1002/jcu.20733.
- [58] M. Komori, M. Hyodo, and K. Gyo, "A swallowing evaluation with simultaneous videoesndoscopy, ultrasonography and videofluorography in healthy controls", *ORL; journal for otorhino-laryngology and its related specialties*, vol. 70, no. 6, pp. 393–398, 2008, ISSN: 1423-0275. DOI: 10.1159/000163036.
- [59] V. Kuhl, B. M. Eicke, M. Dieterich, and P. P. Urban, "Sonographic analysis of laryngeal elevation during swallowing", *Journal of neurology*, vol. 250, no. 3, pp. 333–337, 2003, ISSN: 0340-5354. DOI: 10.1007/s00415-003-1007-2.
- [60] C. Ertekin, "Oropharyngeal swallowing in craniocervical dystonia", *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 73, no. 4, pp. 406–411, 2002, ISSN: 00223050. DOI: 10.1136/jnnp.73.4.406.
- [61] E. S. Sazonov, O. Makeyev, S. Schuckers, P. Lopez-Meyer, E. L. Melanson, and M. R. Neuman, "Automatic detection of swallowing events by acoustical means for applications of monitoring of ingestive behavior", *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 3, pp. 626–633, 2010, ISSN: 0018-9294. DOI: 10.1109/TBME.2009.2033037.
- [62] O. Amft and G. Tröster, "Methods for detection and classification of normal swallowing from muscle activation and sound", in *Pervasive Health Conference and Workshops, 2006*, 2006, pp. 1–10. DOI: 10.1109/PCTHEALTH.2006.361624.
- [63] —, "Recognition of dietary activity events using on-body sensors", *Artificial intelligence in medicine*, vol. 42, no. 2, pp. 121–136, 2008, ISSN: 0933-3657. DOI: 10.1016/j.artmed.2007.11.007.
- [64] M. Aboofazeli and Z. K. Moussavi, "Comparison of recurrence plot features of swallowing and breath sounds", *Chaos, Solitons & Fractals*, vol. 37, no. 2, pp. 454–464, 2008, ISSN: 09600779. DOI: 10.1016/j.chaos.2006.09.026.
- [65] M. Aboofazeli and Z. Moussavi, "Swallowing sound detection using hidden markov modeling of recurrence plot features", *Chaos, Solitons & Fractals*, vol. 39, no. 2, pp. 778–783, 2009, ISSN: 09600779. DOI: 10.1016/j.chaos.2007.01.071.

- [66] L. J. Lazareck and Z. M. K. Moussavi, "Classification of normal and dysphagic swallows by acoustical means", *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 12, pp. 2103–2112, 2004, ISSN: 0018-9294. DOI: 10.1109/TBME.2004.836504.
- [67] S. Sarraf Shirazi and Z. Moussavi, "Silent aspiration detection by breath and swallowing sound analysis", *Conf Proc IEEE Eng Med Biol Soc (Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society)*, vol. 2012, pp. 2599–2602, 2012. DOI: 10.1109/EMBC.2012.6346496.
- [68] S. Sarraf Shirazi, A. H. Birjandi, and Z. Moussavi, "Noninvasive and automatic diagnosis of patients at high risk of swallowing aspiration", *Medical & biological engineering & computing*, vol. 52, no. 5, pp. 459–465, 2014, ISSN: 0140-0118. DOI: 10.1007/s11517-014-1151-0.
- [69] K. Yatani and K. N. Truong, "BodyScope: a wearable acoustic sensor for activity recognition", in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, A. K. Dey, Ed., New York, NY: ACM, 2012, pp. 341–350, ISBN: 978-1-4503-1224-0. DOI: 10.1145/2370216.2370269.
- [70] T. Olubanjo and M. Ghovanloo, "Real-time swallowing detection based on tracheal acoustics", in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, IEEE, 2014, pp. 4384–4388.
- [71] K. Takahashi, M. E. Groher, and K.-i. Michi, "Methodology for detecting swallowing sounds", *Dysphagia*, vol. 9, no. 1, pp. 54–62, 1994, ISSN: 0179-051X. DOI: 10.1007/BF00262760.
- [72] J. A. Cichero and B. E. Murdoch, "Detection of swallowing sounds: methodology revisited", *Dysphagia*, vol. 17, no. 1, pp. 40–49, 2002, ISSN: 0179-051X. DOI: 10.1007/s00455-001-0100-x.
- [73] N. P. Reddy, A. Katakam, V. Gupta, R. Unnikrishnan, J. Narayanan, and E. P. Canilang, "Measurements of acceleration during videofluorographic evaluation of dysphagic patients", *Medical engineering & physics*, vol. 22, no. 6, pp. 405–412, 2000, ISSN: 1873-4030.
- [74] J. Lee, S. Blain, M. Casas, D. Kenny, G. Berall, and T. Chau, "A radial basis classifier for the automatic detection of aspiration in children with dysphagia", *Journal of neuroengineering and rehabilitation*, vol. 3, no. 14, p. 14, 2006, ISSN: 1743-0003. DOI: 10.1186/1743-0003-3-14.
- [75] E. Sejdić, C. M. Steele, and T. Chau, "Segmentation of dual-axis swallowing accelerometry signals in healthy subjects with analysis of anthropometric effects on duration of swallowing activities", *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 1090–1097, 2009, ISSN: 0018-9294. DOI: 10.1109/TBME.2008.2010504.
- [76] E. Sejdić, T. H. Falk, C. M. Steele, and T. Chau, "Vocalization removal for improved automatic segmentation of dual-axis swallowing accelerometry signals", *Medical engineering & physics*, vol. 32, no. 6, pp. 668–672, 2010, ISSN: 1873-4030. DOI: 10.1016/j.medengphy.2010.04.008.
- [77] J. Lee, C. M. Steele, and T. Chau, "Swallow segmentation with artificial neural networks and multi-sensor fusion", *Medical engineering & physics*, vol. 31, no. 9, pp. 1049–1055, 2009, ISSN: 1873-4030. DOI: 10.1016/j.medengphy.2009.07.001.

- [78] S. Damouras, E. Sejdic, C. M. Steele, and T. Chau, "An online swallow detection algorithm based on the quadratic variation of dual-axis accelerometry", *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3352–3359, 2010, ISSN: 1053-587X. DOI: 10.1109/TSP.2010.2043972.
- [79] J. Lee, C. M. Steele, and T. Chau, "Classification of healthy and abnormal swallows based on accelerometry and nasal airflow signals", *Artificial intelligence in medicine*, vol. 52, no. 1, pp. 17–25, 2011, ISSN: 0933-3657. DOI: 10.1016/j.artmed.2011.03.002.
- [80] C. Merey, A. Kushki, E. Sejdić, G. Berall, and T. Chau, "Quantitative classification of pediatric swallowing through accelerometry", *Journal of neuroengineering and rehabilitation*, vol. 9, no. 1, p. 34, 2012, ISSN: 1743-0003. DOI: 10.1186/1743-0003-9-34.
- [81] C. M. Steele, E. Sejdić, and T. Chau, "Noninvasive detection of thin-liquid aspiration using dual-axis swallowing accelerometry", *Dysphagia*, vol. 28, no. 1, pp. 105–112, 2012, ISSN: 0179-051X. DOI: 10.1007/s00455-012-9418-9.
- [82] E. Sejdić, C. M. Steele, and T. Chau, "Classification of penetration–aspiration versus healthy swallows using dual-axis swallowing accelerometry signals in dysphagic subjects", *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 7, pp. 1859–1866, 2013, ISSN: 0018-9294. DOI: 10.1109/TBME.2013.2243730.
- [83] V. Gupta, N. P. Reddy, and E. P. Canilang, "Surface EMG measurements at the throat during dry and wet swallowing", *Dysphagia*, vol. 11, no. 3, pp. 173–179, 1996, ISSN: 0179-051X.
- [84] A. L. Perlman, P. M. Palmer, T. M. McCulloch, and D. J. Vandaele, "Electromyographic activity from human laryngeal, pharyngeal, and submental muscles during swallowing", *J Appl Physiol (Journal of Applied Physiology)*, vol. 86, pp. 1663–1669, 1999.
- [85] M. Vaiman, G. Shoval, and H. Gavriel, "The electrodiagnostic examination of psychogenic swallowing disorders", *Eur Arch Otorhinolaryngol (European Archives of Oto-Rhino-Laryngology)*, vol. 265, no. 6, pp. 663–668, 2008. DOI: 10.1007/s00405-007-0519-2.
- [86] A. J. Fourcin and E. Abberton, "First applications of a new laryngograph", *Medical & biological illustration*, vol. 21, no. 3, pp. 172–182, 1971, ISSN: 0025-6978.
- [87] A. Nieto, I. Cobeta, and P. Kitzing, "Electroglottography in the research and physiology of larynx", *Acta Otorrinolaringol Esp*, vol. 44, pp. 257–263, 1993.
- [88] R. Sorin, M. D. McClean, F. Ezerzer, and B. Meissner-Fishbein, "Electroglottographic evaluation of the swallow", *Archives of physical medicine and rehabilitation*, vol. 68, no. 4, pp. 232–235, 1987, ISSN: 0003-9993.
- [89] A. L. Perlman and H. X. Liang, "Frequency response of the Fourcin electroglottograph and measurement of temporal aspects of laryngeal movement during swallowing", *J Speech Lang Hear Res (Journal of speech, language, and hearing research : JSLHR)*, vol. 34, no. 4, pp. 791–795, 1991.
- [90] A. L. Perlman and J. P. Grayhack, "Use of the electroglottograph for measurement of temporal aspects of the swallow: preliminary observations", *Dysphagia*, vol. 6, pp. 88–93, 1991, ISSN: 0179-051X.

- [91] J. L. Schultz, A. L. Perlman, and D. J. Vandaele, "Laryngeal movement, oropharyngeal pressure, and submental muscle contraction during swallowing", *Archives of physical medicine and rehabilitation*, vol. 75, pp. 183–188, 1994, ISSN: 0003-9993.
- [92] R. Ding, C. R. Larson, J. A. Logemann, and A. W. Rademaker, "Surface electromyographic and electroglottographic studies in normal subjects under two swallow conditions: normal and during the Mendelsohn maneuver", *Dysphagia*, vol. 17, no. 1, pp. 1–12, 2002, ISSN: 0179-051X.
- [93] M. Kob and T. Frauenrath, "A system for parallel measurement of glottis opening and larynx position", *Biomedical Signal Processing and Control*, vol. 4, no. 3, pp. 221–228, 2009. DOI: 10.1016/j.bspc.2009.03.004.
- [94] T. A. Hughes, P. Liu, H. Griffiths, B. W. Lawrie, and C. M. Wiles, "An analysis of studies comparing electrical impedance tomography with x-ray videofluoroscopy in the assessment of swallowing", *Physiological Measurement*, vol. 15 Suppl 2a, A199–209, 1994, ISSN: 0967-3334.
- [95] —, "Simultaneous electrical impedance tomography and videofluoroscopy in the assessment of swallowing", *Physiological Measurement*, vol. 17, no. 2, pp. 109–119, 1996, ISSN: 0967-3334.
- [96] T. Nakamura, Y. Yamamoto, and H. Tsugawa, "Measurement system for swallowing based on impedance pharyngography and swallowing sound", in *Instrumentation and Measurement Technology Conference, 2000. IMTC 2000. Proceedings of the 17th IEEE*, vol. 1, IEEE, 2000, pp. 191–194.
- [97] Y. Yamamoto, T. Nakamura, T. Kusuhara, and Adli, "Consideration of conditions required for multi-channel simultaneous bioimpedance measurement", *Instrumentation and Measurement Technology Conference, 1998. IMTC/98. Conference Proceedings. IEEE*, vol. 1, pp. 231–234, 1998. DOI: 10.1109/IMTC.1998.679763.
- [98] Y. Yamamoto, T. Nakamura, Y. Seki, K. Utsuyama, K. Akashi, and K. Jikuya, "Neck electrical impedance for measurement of swallowing", *Electrical Engineering in Japan*, vol. 130, no. 4, pp. 35–44, 2000, ISSN: 0424-7760. DOI: 10.1002/(SICI)1520-6416(200003)130:4<35::AID-EEJ5>3.0.CO;2-W.
- [99] N. Morimoto, T. Yamashita, K. Sato, T. Kurata, Y. Ikeda, T. Kusuhara, N. Murata, and K. Abe, "Assessment of swallowing in motor neuron disease and Asidan/SCA36 patients with new methods", *Journal of the neurological sciences*, vol. 324, no. 1-2, pp. 149–155, 2013, ISSN: 1878-5883. DOI: 10.1016/j.jns.2012.10.025.
- [100] L. C. Ward, W. A. McCullagh, and J. Cichero, "The use of bioimpedance analysis for the study of dysphagia", in *13th International Conference on Electrical Bioimpedance and the 8th Conference on Electrical Impedance Tomography*, ser. IFMBE Proceedings, vol. 17, Springer Berlin Heidelberg, 2007, pp. 683–686, ISBN: 978-3-540-73841-1. DOI: 10.1007/978-3-540-73841-1\_176.
- [101] C. J. Chester, P. T. Gaynor, R. D. Jones, and M.-L. Huckabee, "Electrical bioimpedance measurement as a tool for dysphagia visualisation", *Healthcare technology letters*, vol. 1, no. 3, pp. 115–118, 2014, ISSN: 2053-3713. DOI: 10.1049/htl.2014.0067.

## REFERENCES

---

- [102] K. V. Kuhlemeier, J. B. Palmer, and D. Rosenberg, "Effect of liquid bolus consistency and delivery method on aspiration and pharyngeal retention in dysphagia patients", *Dysphagia*, vol. 16, no. 2, pp. 119–122, 2001, ISSN: 0179-051X. DOI: 10.1007/s004550011003.
- [103] A. Rasley, J. A. Logemann, P. J. Kahrilas, A. W. Rademaker, B. R. Pauloski, and W. J. Dodds, "Prevention of barium aspiration during videofluoroscopic swallowing studies: value of change in posture", *AJR. American journal of roentgenology*, vol. 160, no. 5, pp. 1005–1009, 1993, ISSN: 0361-803X. DOI: 10.2214/ajr.160.5.8470567.
- [104] J. Robbins, S. G. Butler, S. K. Daniels, R. Diez Gross, S. Langmore, C. L. Lazarus, B. Martin-Harris, D. McCabe, N. Musson, and J. Rosenbek, "Swallowing and dysphagia rehabilitation: translating principles of neural plasticity into clinically oriented evidence", *J Speech Lang Hear Res (Journal of speech, language, and hearing research : JSLHR)*, vol. 51, no. 1, pp. 276–300, 2008. DOI: 10.1044/1092-4388(2008/021).
- [105] I. A. Humbert and S. Joel, "Tactile, gustatory, and visual biofeedback stimuli modulate neural substrates of deglutition", *NeuroImage*, vol. 59, no. 2, pp. 1485–1490, 2012, ISSN: 1095-9572. DOI: 10.1016/j.neuroimage.2011.08.022.
- [106] R. Shaker, C. Easterling, M. Kern, T. Nitschke, B. Massey, S. Daniels, B. Grande, M. Kazandjian, and K. Dikeman, "Rehabilitation of swallowing by exercise in tube-fed patients with pharyngeal dysphagia secondary to abnormal UES opening", *Gastroenterology*, vol. 122, no. 5, pp. 1314–1321, 2002, ISSN: 0016-5085. DOI: 10.1053/gast.2002.32999.
- [107] J.-H. Kang, R.-Y. Park, S.-J. Lee, J.-Y. Kim, S.-R. Yoon, and K.-I. Jung, "The effect of bedside exercise program on stroke patients with Dysphagia", *Annals of rehabilitation medicine*, vol. 36, no. 4, pp. 512–520, 2012, ISSN: 2234-0653. DOI: 10.5535/arm.2012.36.4.512.
- [108] K. M. Wheeler-Hegland, J. C. Rosenbek, and C. M. Sapienza, "Submental sEMG and hyoid movement during Mendelsohn maneuver, effortful swallow, and expiratory muscle strength training", *J Speech Lang Hear Res (Journal of speech, language, and hearing research : JSLHR)*, vol. 51, no. 5, pp. 1072–1087, 2008. DOI: 10.1044/1092-4388(2008/07-0016).
- [109] G. H. McCullough, E. Kamarunas, G. C. Mann, J. W. Schmidley, J. A. Robbins, and M. A. Crary, "Effects of Mendelsohn maneuver on measures of swallowing duration post stroke", *Topics in stroke rehabilitation*, vol. 19, no. 3, pp. 234–243, 2012, ISSN: 1074-9357. DOI: 10.1310/tsr1903-234.
- [110] M. Bülow, R. Olsson, and O. Ekberg, "Supraglottic swallow, effortful swallow, and chin tuck did not alter hypopharyngeal intrabolus pressure in patients with pharyngeal dysfunction", *Dysphagia*, vol. 17, no. 3, pp. 197–201, 2002, ISSN: 0179-051X. DOI: 10.1007/s00455-002-0050-y.
- [111] C. M. Steele and A. J. Miller, "Sensory input pathways and mechanisms in swallowing: a review", *Dysphagia*, vol. 25, no. 4, pp. 323–333, 2010, ISSN: 0179-051X. DOI: 10.1007/s00455-010-9301-5.
- [112] R. O. Seidl, R. Nusser-Müller-Busch, W. Hollweg, M. Westhofen, and A. Ernst, "Pilot study of a neurophysiological dysphagia therapy for neurological patients", *Clinical rehabilitation*, vol. 21, no. 8, pp. 686–697, 2007, ISSN: 0269-2155. DOI: 10.1177/0269215507076393.

- [113] C. L. Park, P. A. O'Neill, and D. F. Martin, "A pilot exploratory study of oral electrical stimulation on swallow function following stroke: an innovative technique", *Dysphagia*, vol. 12, no. 3, pp. 161–166, 1997, ISSN: 0179-051X.
- [114] S. Hamdy, Q. Aziz, J. C. Rothwell, M. Power, K. D. Singh, D. A. Nicholson, R. C. Tallis, and D. G. Thompson, "Recovery of swallowing after dysphagic stroke relates to functional reorganization in the intact motor cortex", *Gastroenterology*, vol. 115, no. 5, pp. 1104–1112, 1998, ISSN: 0016-5085.
- [115] S. Hamdy, J. C. Rothwell, Q. Aziz, K. D. Singh, and D. G. Thompson, "Long-term reorganization of human motor cortex driven by short-term sensory stimulation", *Nature neuroscience*, vol. 1, no. 1, pp. 64–68, 1998, ISSN: 1097-6256. DOI: 10.1038/264.
- [116] C. Fraser, M. Power, S. Hamdy, J. Rothwell, D. Hobday, I. Hollander, P. Tyrell, A. Hobson, S. Williams, and D. Thompson, "Driving plasticity in human adult motor cortex is associated with improved motor function after brain injury", *Neuron*, vol. 34, no. 5, pp. 831–840, 2002, ISSN: 0896-6273.
- [117] V. Jayasekeran, S. Singh, P. Tyrrell, E. Michou, S. Jefferson, S. Mistry, E. Gamble, J. Rothwell, D. Thompson, and S. Hamdy, "Adjunctive functional pharyngeal electrical stimulation reverses swallowing disability after brain lesions", *Gastroenterology*, vol. 138, no. 5, pp. 1737–1746, 2010, ISSN: 0016-5085. DOI: 10.1053/j.gastro.2010.01.052.
- [118] M. Power, C. Fraser, A. Hobson, J. C. Rothwell, S. Mistry, D. A. Nicholson, D. G. Thompson, and S. Hamdy, "Changes in pharyngeal corticobulbar excitability and swallowing behavior after oral stimulation", *American journal of physiology. Gastrointestinal and liver physiology*, vol. 286, no. 1, G45–50, 2004, ISSN: 0193-1857. DOI: 10.1152/ajpgi.00114.2003.
- [119] M. L. Power, C. H. Fraser, A. Hobson, S. Singh, P. Tyrrell, D. A. Nicholson, I. Turnbull, D. G. Thompson, and S. Hamdy, "Evaluating oral stimulation as a treatment for dysphagia after stroke", *Dysphagia*, vol. 21, no. 1, pp. 49–55, 2006, ISSN: 0179-051X. DOI: 10.1007/s00455-005-9009-0.
- [120] A. J. Miller, "Characteristics of the swallowing reflex induced by peripheral nerve and brain stem stimulation", *Experimental Neurology*, vol. 34, no. 2, pp. 210–222, 1972, ISSN: 0014-4886. DOI: 10.1016/0014-4886(72)90168-9.
- [121] G. Chi-Fishman, N. F. Capra, and G. N. McCall, "Thermomechanical facilitation of swallowing evoked by electrical nerve stimulation in cats", *Dysphagia*, vol. 9, no. 3, pp. 149–155, 1994, ISSN: 0179-051X.
- [122] M. Broniatowski, N. Z. Moore, S. Grundfest-Broniatowski, H. M. Tucker, E. Lancaster, K. Krival, A. J. Hadley, and D. J. Tyler, "Paced glottic closure for controlling aspiration pneumonia in patients with neurologic deficits of various causes", *The Annals of otology, rhinology, and laryngology*, vol. 119, no. 3, pp. 141–149, 2010, ISSN: 0003-4894. DOI: 10.1016/j.yoto.2011.03.052.
- [123] M. L. Freed, L. Freed, R. L. Chatburn, and M. Christian, "Electrical stimulation for swallowing disorders caused by stroke", *Respiratory care*, vol. 46, no. 5, pp. 466–474, 2001, ISSN: 0020-1324.

## REFERENCES

---

- [124] L. Blumenfeld, Y. Hahn, A. Lepage, R. Leonard, and P. C. Belafsky, "Transcutaneous electrical stimulation versus traditional dysphagia therapy: a nonconcurrent cohort study", *Otolaryngology-Head and Neck Surgery*, vol. 135, no. 5, pp. 754–757, 2006. DOI: 10.1016/j.otohns.2006.04.016.
- [125] B.-M. Oh, D.-Y. Kim, and N.-J. Paik, "Recovery of swallowing function is accompanied by the expansion of the cortical map", *The International journal of neuroscience*, vol. 117, no. 9, pp. 1215–1227, 2007, ISSN: 0020-7454. DOI: 10.1080/00207450600936254.
- [126] K.-B. Lim, H.-J. Lee, S.-S. Lim, and Y.-I. Choi, "Neuromuscular electrical and thermal-tactile stimulation for dysphagia caused by stroke: a randomized controlled trial", *Journal of rehabilitation medicine*, vol. 41, no. 3, pp. 174–178, 2009, ISSN: 1651-2081. DOI: 10.2340/16501977-0317.
- [127] D. M. Suiter, S. B. Leder, and J. L. Ruark, "Effects of neuromuscular electrical stimulation on submental muscle activity", *Dysphagia*, vol. 21, no. 1, pp. 56–60, 2006, ISSN: 0179-051X. DOI: 10.1007/s00455-005-9010-7.
- [128] L. W. J. Baijens, R. Speyer, V. L. Passos, W. Pilz, N. Roodenburg, and P. Clavé, "The effect of surface electrical stimulation on swallowing in dysphagic Parkinson patients", *Dysphagia*, vol. 27, no. 4, pp. 528–537, 2012, ISSN: 0179-051X. DOI: 10.1007/s00455-011-9387-4.
- [129] K.-S. Soon, M.-Y. Lee, W.-W. Tsai, C.-F. Lin, and C.-H. Chen, "Development of a swallowing electrical stimulation system for treatment of dysphagia in stroke patients", *Journal of Medical and Biological Engineering*, vol. 33, no. 5, pp. 497–503, 2013, ISSN: 1609-0985. DOI: 10.5405/jmbe.1117.
- [130] I. A. Humbert, C. J. Poletto, K. G. Saxon, P. R. Kearney, and C. L. Ludlow, "The effect of surface electrical stimulation on vocal fold position", *The Laryngoscope*, vol. 118, no. 1, pp. 14–19, 2008, ISSN: 0023-852X. DOI: 10.1097/MLG.0b013e318155a47d.
- [131] H. Nahrstaedt, T. Schauer, and R. O. Seidl, "Bioimpedance based measurement system for a controlled swallowing neuro-prosthesis", in *Proc. of 15th Annual International FES Society Conference and 10th Vienna Int. Workshop on FES*, 2010, pp. 49–51.
- [132] H. Nahrstaedt, C. Schultheiss, R. O. Seidl, and T. Schauer, "Swallow detection algorithm based on bioimpedance and EMG measurements", in *8th IFAC Symposium on Biological and Medical Systems*, Budapest, Hungary, 2012, pp. 91–96. DOI: 10.3182/20120829-3-HU-2029.00067.
- [133] Z.-y. Li, C.-s. Ren, S. Zhao, H. Sha, and J. Deng, "Gastric motility functional study based on electrical bioimpedance measurements and simultaneous electrogastrography", *Journal of Zhejiang University - Science B*, vol. 12, no. 12, pp. 983–989, 2011, ISSN: 1673-1581. DOI: 10.1631/jzus.B1000436.
- [134] T. Vuorela, J. Vanhala, V.-P. Seppa, and J. Hyttinen, "Two portable long-term measurement devices for ECG and bioimpedance", in *Pervasive Computing Technologies for Healthcare, 2008. PervasiveHealth 2008. Second International Conference on*, IEEE, Ed., 2008, pp. 169–172. DOI: 10.1109/PCTHEALTH.2008.4571061.

- [135] C. Schultheiss, "Die Bewertung der pharyngalen Schluckphase mittels Bioimpedanz: Evaluation eines Mess- und Diagnostikverfahrens", Dissertation, Universität Potsdam, Potsdam, 2013-11-07.
- [136] R. O. Seidl, T. Schauer, H. Nahrstaedt, and C. Schultheiss, *Evaluation eines Bioimpedanzmesssystems zur automatisierten Schluckerkennung*, Freiburg, 2011. DOI: 10.3205/11hnod062. [Online]. Available: <http://bigdyspro.de/de/wp-content/uploads/2011/02/Vortrag-HNO-Freiburg-2011.pdf>.
- [137] S. Grimnes and Ø. G. Martinsen, "Sources of error in tetrapolar impedance measurements on biomaterials and other ionic conductors", *Journal of Physics D: Applied Physics*, vol. 40, no. 1, pp. 9–14, 2007, ISSN: 0022-3727. DOI: 10.1088/0022-3727/40/1/S02.
- [138] D. Holder, *Electrical impedance tomography: Methods, history, and applications*, ser. Series in medical physics and biomedical engineering. Bristol and Philadelphia: Institute of Physics Pub, 2005, ISBN: 978-0-7503-0952-3.
- [139] G. S. Pomeroy, "Differential current source with active common mode reduction", US6501255B2, 2002. [Online]. Available: <https://patents.google.com/patent/US6501255B2>.
- [140] P. Bertemes-Filho, B. H. Brown, and A. J. Wilson, "A comparison of modified Howland circuits as current generators with current mirror type circuits", *Physiological Measurement*, vol. 21, no. 1, pp. 1–6, 2000, ISSN: 0967-3334. DOI: 10.1088/0967-3334/21/1/301.
- [141] H. Kalvøy, C. Tronstad, B. Nordbotten, S. Grimnes, and Ø. G. Martinsen, "Electrical impedance of stainless steel needle electrodes", *Annals of biomedical engineering*, vol. 38, no. 7, pp. 2371–2382, 2010, ISSN: 1573-9686. DOI: 10.1007/s10439-010-9989-2.
- [142] D. P. Dobrev, T. Neycheva, and N. Mudrov, "Bootstrapped two-electrode biosignal amplifier", *Medical & biological engineering & computing*, vol. 46, no. 6, pp. 613–619, 2008, ISSN: 0140-0118. DOI: 10.1007/s11517-008-0312-4.
- [143] R. Merletti and P. A. Parker, *Electromyography*. John Wiley & Sons, Inc, 2004.
- [144] L. K. Maheshwari and M. M. S. Anand, *Analog electronics*. New Delhi: PHI Learning and Prentice-Hallof India Pvt. Ltd, 2005, ISBN: 9788120327221.
- [145] B. Kemp and J. Olivan, "European data format 'plus' (EDF+), an EDF alike standard format for the exchange of physiological data", *Clinical Neurophysiology*, vol. 114, no. 9, pp. 1755–1761, 2003, ISSN: 1388-2457. DOI: 10.1016/S1388-2457(03)00123-8.
- [146] H. Nahrstaedt and T. Schauer, "A bioimpedance measurement device for sensing force and position in neuroprosthetic systems", in *4th European Conference of the International Federation for Medical and Biological Engineering*, R. Magjarevic, J. H. Nagel, J. Vander Sloten, P. Verdonck, M. Nyssen, and J. Haueisen, Eds., vol. 22, Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 1642–1645. DOI: 10.1007/978-3-540-89208-3\_390.

## REFERENCES

---

- [147] C. Schultheiss, T. Schauer, H. Nahrstaedt, and R. O. Seidl, "Evaluation of an EMG bioimpedance measurement system for recording and analysing the pharyngeal phase of swallowing", *European Archives of Oto-Rhino-Laryngology*, vol. 270, no. 7, pp. 2149–2156, 2013, ISSN: 0937-4477, 1434-4726. DOI: 10.1007/s00405-013-2406-3.
- [148] —, "Automated detection and evaluation of swallowing using a combined EMG/bioimpedance Measurement System", *The Scientific World Journal*, vol. 2014, p. 405 471, 2014. DOI: 10.1155/2014/405471.
- [149] D. L. DONOHO and J. M. JOHNSTONE, "Ideal spatial adaptation by wavelet shrinkage", *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994. DOI: 10.1093/biomet/81.3.425.
- [150] S. G. Mallat, *A wavelet tour of signal processing: The Sparse way*, 3rd ed. Amsterdam and Boston: Elsevier /Academic Press, 2009, ISBN: 978-0-12-374370-1.
- [151] F. B. Stulen and C. J. de Luca, "Frequency parameters of the myoelectric signal as a measure of muscle conduction velocity", in *1981 Index IEEE Transactions on Biomedical Engineering Vol. BME-28*, vol. 28, IEEE, 1981, pp. 515–523. DOI: 10.1109/TBME.1981.324738.
- [152] P. Bonato, T. D'Alessio, and M. Knaflitz, "A statistical method for the measurement of muscle activation intervals from surface myoelectric signal during gait", *Biomedical Engineering, IEEE Transactions on*, vol. 45, no. 3, pp. 287–299, 1998, ISSN: 0018-9294. DOI: 10.1109/10.661154.
- [153] G. Staude, C. Flachenecker, M. Daumer, and W. Wolf, "Onset detection in surface electromyographic signals: a systematic comparison of methods", *EURASIP Journal on Advances in Signal Processing*, vol. 2001, no. 2, pp. 67–81, 2001, ISSN: 1687-6172. DOI: 10.1155/S1110865701000191.
- [154] G. Severini, S. Conforto, M. Schmid, and T. D'Alessio, "Novel formulation of a double threshold algorithm for the estimation of muscle activation intervals designed for variable SNR environments", *Journal of Electromyography and Kinesiology*, vol. 22, no. 6, pp. 878–885, 2012, ISSN: 10506411. DOI: 10.1016/j.jelekin.2012.04.010.
- [155] L. Xu and A. Adler, "An improved method for muscle activation detection during gait", in *Canadian Conference on Electrical and Computer Engineering 2004 (IEEE Cat. No.04CH37513)*, vol. 1, IEEE, 2004, pp. 357–360, ISBN: 0-7803-8253-6. DOI: 10.1109/CCECE.2004.1345029.
- [156] W. Gilchrist, *Statistical modelling with quantile functions*. Boca Raton: Chapman & Hall/CRC, 2000, ISBN: 9781584881742.
- [157] R. Agrawal, C. Faloutsos, and A. Swami, "Efficient similarity search in sequence databases", in *Foundations of Data Organization and Algorithms*, ser. Lecture Notes in Computer Science, D. Lomet, Ed., vol. 730, Springer Berlin Heidelberg, 1993, pp. 69–84, ISBN: 978-3-540-57301-2. DOI: 10.1007/3-540-57301-1\_5.
- [158] K.-P. Chan and A. W.-C. Fu, "Efficient time series matching by wavelets", *Data Engineering, 1999. Proceedings., 15th International Conference on*, pp. 126–133, 1999, ISSN: 1063-6382. DOI: 10.1109/ICDE.1999.754915.

- [159] E. Keogh, S. Chu, D. Hart, and M. Pazzani, “An online algorithm for segmenting time series”, in *Proceedings 2001 IEEE International Conference on Data Mining*, IEEE, 2001, pp. 289–296, ISBN: 0-7695-1119-8. DOI: 10.1109/ICDM.2001.989531.
- [160] E. Keogh, K. Chakrabarti, M. Pazzani, and S. Mehrotra, “Dimensionality reduction for fast similarity search in large time series databases”, *Knowledge and Information Systems*, vol. 3, no. 3, pp. 263–286, 2001, ISSN: 0219-1377. DOI: 10.1007/PL00011669.
- [161] K. Chakrabarti, E. Keogh, S. Mehrotra, and M. Pazzani, “Locally adaptive dimensionality reduction for indexing large time series databases”, *ACM Transactions on Database Systems*, vol. 27, no. 2, pp. 188–228, 2002, ISSN: 03625915. DOI: 10.1145/568518.568520.
- [162] F. Korn, H. V. Jagadish, and C. Faloutsos, “Efficiently supporting ad hoc queries in large datasets of time sequences”, in *the 1997 ACM SIGMOD international conference*, J. M. Peckman, S. Ram, and M. Franklin, Eds., 1997, pp. 289–300. DOI: 10.1145/253260.253332.
- [163] J. Lin, E. Keogh, S. Lonardi, and B. Chiu, “A symbolic representation of time series, with implications for streaming algorithms”, in *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*, ACM, 2003, p. 2. DOI: 10.1145/882082.882086.
- [164] J. Wang, A. Balasubramanian, Mojica de la Vega, Luis, J. R. Green, A. Samal, and B. Prabhakaran, “Word recognition from continuous articulatory movement time-series data using symbolic representations”, in *Proceedings of SLPAT-2013*, Stroudsburg, PA: Association for Computational Linguistics, 2013, pp. 119–127, ISBN: 978-1-937284-93-0.
- [165] B. Esmael, A. Arnaout, R. K. Fruhwirth, and G. Thonhauser, “Multivariate time series classification by combining trend-based and value-based approximations”, in *Computational Science and Its Applications – ICCSA 2012*, vol. 7336, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 392–403. DOI: 10.1007/978-3-642-31128-4\_29.
- [166] J. Shieh and E. Keogh, “i SAX: indexing and mining terabyte sized time series”, in *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, 2008, pp. 623–631. DOI: 10.1145/1401890.1401966.
- [167] D. Yankov, E. Keogh, J. Medina, B. Chiu, and V. Zordan, “Detecting time series motifs under uniform scaling”, in *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, 2007, p. 844. DOI: 10.1145/1281192.1281282.
- [168] C. Cortes and V. Vapnik, “Support-vector networks”, *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995. DOI: 10.1023/A:1022627411411.
- [169] C.-C. Chang and C.-J. Lin, “LIBSVM: a library for support vector machines”, *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 1–27, 2011, ISSN: 21576904. DOI: 10.1145/1961189.1961199.
- [170] L. A. Jeni, J. F. Cohn, and F. de La Torre, “Facing Imbalanced Data Recommendations for the Use of Performance Metrics”, in *Int Conf Affect Comput Intell Interact Workshops*, vol. 2013, pp. 245–251. DOI: 10.1109/ACII.2013.47.

## REFERENCES

---

- [171] M. A. Crary, G. D. Carnaby, I. Sia, A. Khanna, and M. F. Waters, "Spontaneous swallowing frequency has potential to identify dysphagia in acute stroke", *Stroke; a journal of cerebral circulation*, vol. 44, no. 12, pp. 3452–3457, 2013, ISSN: 0039-2499. DOI: 10.1161/STROKEAHA.113.003048.
- [172] S. Wolter, H. Nahrstaedt, T. Schauer, C. Schultheiss, and R. O. Seidl, *Überprüfung der Reliabilität eines Bioimpedanzmesssignals für die Regelung*, Mainz, 2012. DOI: 10.3205/12hnod045. [Online]. Available: <http://bigdyspro.de/de/wp-content/uploads/2012/05/Poster-VFC-und-BI-HNO-Mainz-2012.pdf>.
- [173] W. J. Dodds, J. A. Logemann, and E. T. Stewart, "Radiologic assessment of abnormal oral and pharyngeal phases of swallowing", *AJR. American journal of roentgenology*, vol. 154, no. 5, pp. 965–974, 1990, ISSN: 0361-803X.
- [174] C. M. Steele, G. L. Bailey, T. Chau, S. M. Molfenter, M. Oshalla, A. A. Waito, and D. C. B. H. Zoratto, "The relationship between hyoid and laryngeal displacement and swallowing impairment", *Clinical Otolaryngology*, vol. 36, no. 1, pp. 30–36, 2011, ISSN: 1749-4486. DOI: 10.1111/j.1749-4486.2010.02219.x.
- [175] I. Sia, P. Carvajal, G. D. Carnaby-Mann, and M. A. Crary, "Measurement of hyoid and laryngeal displacement in video fluoroscopic swallowing studies: variability, reliability, and measurement error", *Dysphagia*, vol. 27, no. 2, pp. 192–197, 2012, ISSN: 0179-051X. DOI: 10.1007/s00455-011-9352-2.
- [176] A. C. Davison and D. V. Hinkley, *Bootstrap methods and their application*, ser. Cambridge series in statistical and probabilistic mathematics. Cambridge and New York, NY, USA: Cambridge University Press, 1997, ISBN: 9780521574716.
- [177] J. E. Hunter and F. L. Schmidt, *Methods of meta-analysis: correcting error and bias in research findings*, 2nd ed. Thousand Oaks, Calif.: Sage, 2004, ISBN: 9781412904797.
- [178] S. M. Molfenter and C. M. Steele, "Physiological variability in the deglutition literature: hyoid and laryngeal kinematics", *Dysphagia*, vol. 26, no. 1, pp. 67–74, 2011, ISSN: 0179-051X. DOI: 10.1007/s00455-010-9309-x.
- [179] H. Nahrstaedt, C. Schultheiss, T. Schauer, and R. O. Seidl, "Bioimpedance- and EMG-triggered FES for improved protection of the airway during swallowing", *Biomedical Engineering / Biomedizinische Technik*, 2013, ISSN: 1862-278X, 0013-5585. DOI: 10.1515/bmt-2013-4025.
- [180] R. O. Seidl, H. Nahrstaedt, C. Schultheiss, and T. Schauer, *Entwicklung einer Neuroprothese zur Therapie von Schluckstörungen - Erste Ergebnisse*, Nürnberg, 2013. DOI: 10.3205/13hnod031. [Online]. Available: <http://bigdyspro.de/de/wp-content/uploads/2011/02/Vortrag-HNO-Nürnberg-Stimulation-2013.pdf>.
- [181] C. M. Steele, A. T. Thrasher, and M. R. Popovic, "Electric stimulation approaches to the restoration and rehabilitation of swallowing: a review", *Neurological research*, vol. 29, no. 1, pp. 9–15, 2007, ISSN: 0161-6412. DOI: 10.1179/016164107X171555.

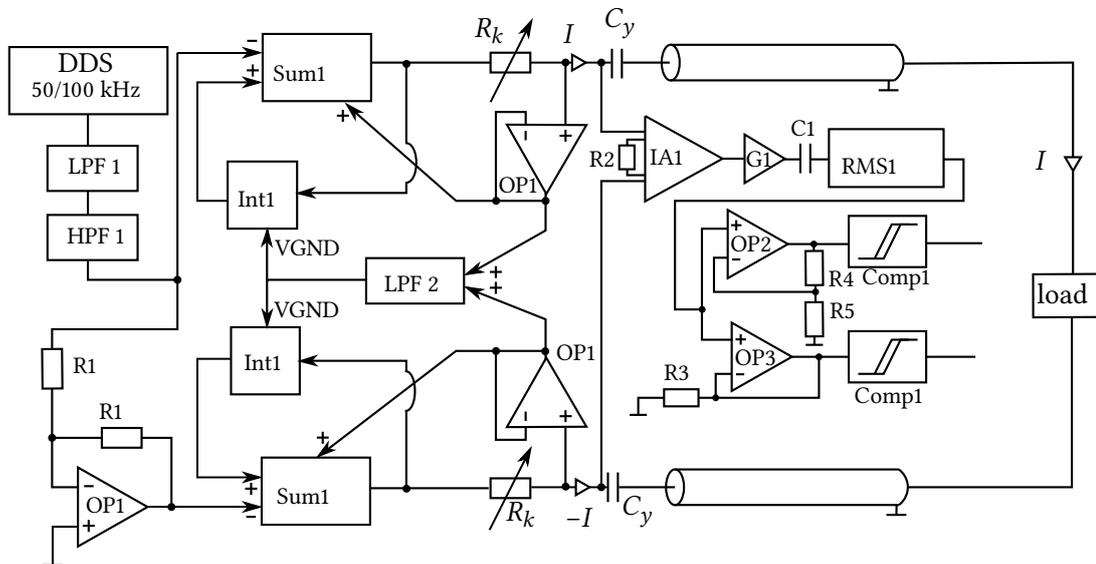
- 
- [182] K. J. Ayala and K. J. Cruz, "Transcutaneous electrical stimulation in dysphagia treatment: is there sufficient evidence?", *Journal of Audiology and Speech-Language Pathology*, vol. TEJAS Volume XXXI, pp. 37–54, 2008.
- [183] Y.-W. Chen, K.-H. Chang, H.-C. Chen, W.-M. Liang, Y.-H. Wang, and Y.-N. Lin, "The effects of surface neuromuscular electrical stimulation on post-stroke dysphagia: a systemic review and meta-analysis", *Clinical Rehabilitation*, vol. 30, no. 1, pp. 24–35, 2016, ISSN: 0269-2155. DOI: 10.1177/0269215515571681.
- [184] M. Lotze, C. Braun, N. Birbaumer, S. Anders, and L. G. Cohen, "Motor learning elicited by voluntary drive", *Brain : a journal of neurology*, vol. 126, no. Pt 4, pp. 866–872, 2003, ISSN: 0006-8950. DOI: 10.1093/brain/awg079.
- [185] M. S. Christensen and M. J. Grey, "Modulation of proprioceptive feedback during functional electrical stimulation: an fMRI study", *The European journal of neuroscience*, vol. 37, no. 11, pp. 1766–1778, 2013, ISSN: 0953-816X. DOI: 10.1111/ejn.12178.
- [186] R. Merletti, M. Knaflitz, and C. J. DeLuca, "Electrically evoked myoelectric signals", *Critical reviews in biomedical engineering*, vol. 19, no. 4, pp. 293–340, 1992, ISSN: 0278-940X.
- [187] A. Boxtel, "Optimal signal bandwidth for the recording of surface EMG activity of facial, jaw, oral, and neck muscles", *Psychophysiology*, vol. 38, no. 1, pp. 22–34, 2001, ISSN: 1469-8986. DOI: 10.1111/1469-8986.3810022.





# Component Lists for PHYSIOSENSE

## A.1 Current Source for BI Measurement

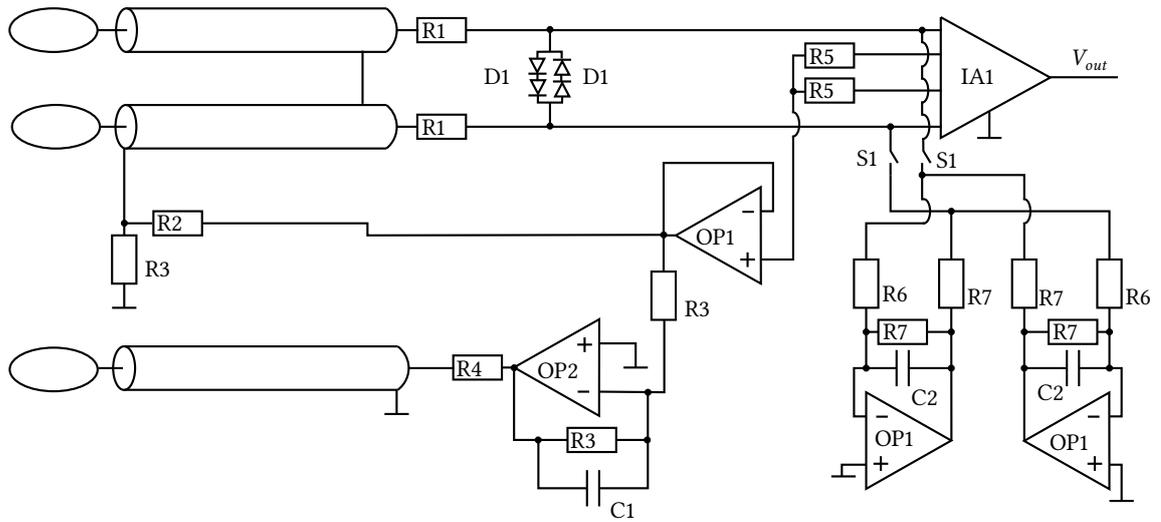


**Figure A.1:** Circuit layout with components of the floating current source from Figure 3.4. The output of the low pass LPF 2 is used as virtual ground (VGND) inside both integrators. The operating voltages are  $-5\text{ V}$  and  $5\text{ V}$ .

Table A.1: Component list for the floating current source.

Name	Description	Parameters	Type / Manufacturer
R1	resistor	10 k $\Omega$ , 1 %	
R2	resistor	open	
R3	resistor	10 k $\Omega$ , 1 %	
R4	resistor	100 k $\Omega$ , 1 %	
R5	resistor	10 k $\Omega$ , 1 %	
C1	capacitor	0.1 $\mu$ F	
DDS	programmable waveform generator	set to sinusoidal oscillation with a frequency of either 50 kHz or 100 kHz	AD9833BRMZ, Analog Devices
$R_k$	analog multiplexer	switching to either 2.7 k $\Omega$ , 3.3 k $\Omega$ , 4.7 k $\Omega$ , 6.8 k $\Omega$ , 8.2 k $\Omega$ , 27 k $\Omega$ , 56 k $\Omega$ , or 110 k $\Omega$ .	LTC1391CGN#PBF, Linear Technology
OP1	operational amplifier		LM6144BIM, National Semiconductor
$C_y$	Y1 safety capacitor		440LD47, Vishay
IA1	instrumentation amplifier		INA128UAE4, Texas Instruments
RMS1	RMS-to-DC converter	operating voltages are -2.5 V and 2.5 V	LTC1968, Linear Technology
OP2	operational amplifier		LT6221CS8#PBF, Linear Technology
OP3	operational amplifier		LT6220CS5#PBF, Linear Technology
Comp1	comparator		LTC1440CS8#PBF, Linear Technology
LPF 1	low-pass filter	3. order, $f_{\text{cut-off}} = 1$ MHz, Bessel, MFB single ended, gain = 10	LT1800, Linear Technology
LPF 2	low-pass filter	1. order, $f_{\text{cut-off}} = 160$ Hz, gain = 1	LM6144, National Semiconductor
HPF 1	high-pass filter	2. order, $f_{\text{cut-off}} = 20$ Hz, Butterworth, Sallen-Key, gain = 1,	LT6221CS8#PBF, Linear Technology
Int1	integrator	$\tau = 0.1$ ms	LT6221CS8#PBF, Linear Technology
Sum1	differential amplifier	gain = 1	LM6144, National Semiconductor
G1	amplifier	gain = 0.56	LT6221CS8#PBF, Linear Technology

## A.2 Voltage Measurement



**Figure A.2:** The circuit design of the voltage measurement input stage with the active shielding AS, the reference driver RD and the common mode rejection circuit CMRC in detail from Figure 3.6. The operating voltages are  $-5\text{ V}$  and  $5\text{ V}$ .

Table A.2: Component list for the voltage measurement input stage.

Name	Description	Parameters	Type / Manufacturer
R1	resistor	100 k $\Omega$ , 0.1 %	
R2	resistor	100 $\Omega$ , 1 %	
R3	resistor	10 k $\Omega$ , 1 %	
R4	resistor	400 k $\Omega$ , 1 %	
R5	resistor	2.7 k $\Omega$ , 1 %	
R6	resistor	300 k $\Omega$ , 1 %	
R7	resistor	620 k $\Omega$ , 1 %	
C1	capacitor	0.1 $\mu$ F	
C2	capacitor	6.8 nF	
D1	low-leakage double diode		BAV199, Infineon Technologies AG
S1	Photomos solid state relay		AQW210EHA, Panasonic EW
OP1	operational amplifier		OPA2131UJ, Texas Instruments
OP2	operational amplifier		MAX492CSA, Maxim integrated
IA1	instrumentation amplifier		INA128UAE4, Texas Instruments

### A.3 Analog Signal Processing

#### A.3.1 EMG Filtering

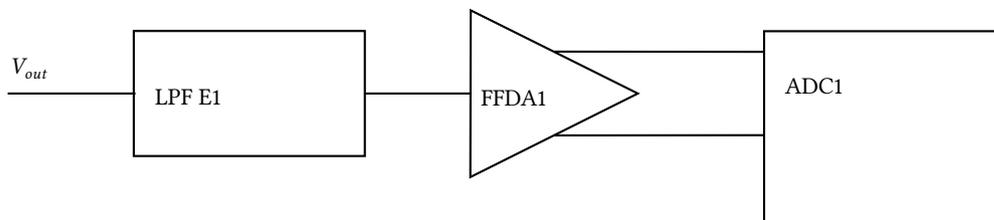


Figure A.3: The filter and amplifier block FiltEMG in detail from Figure 3.8. The analog to digital converter ADC1 also used within Figures A.4 and A.5. The operating voltages are  $-5$  V and  $5$  V.

Table A.3: Component list for the filter and amplifier block FiltEMG.

Name	Description	Parameters	Type / Manufacturer
LPF E1	low-pass filter	4. order, $f_{\text{cut-off}} = 13 \text{ kHz}$ , Butterworth, MFB single ended, gain = 1	OPA2131UJ, Texas Instruments
FFDA1	feedback fully differential low-pass filter	1. order, $f_{\text{cut-off}} = 13 \text{ Hz}$ , Bessel, MFB fully differential, gain = 0.5	THS4521, Texas Instruments
ADC1	analog digital converter	24-bit, sampling frequency 12 kHz	ADS1278, Texas Instruments

A.3.2 BI Filtering for Measurement at 50 kHz

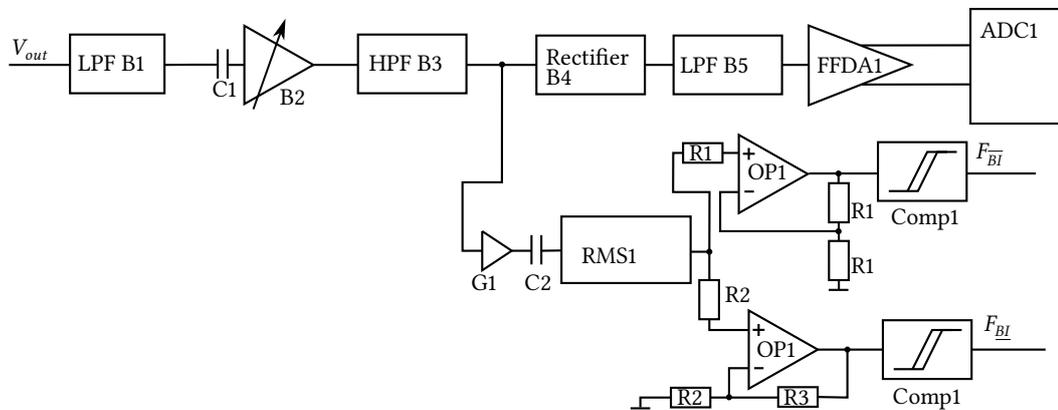
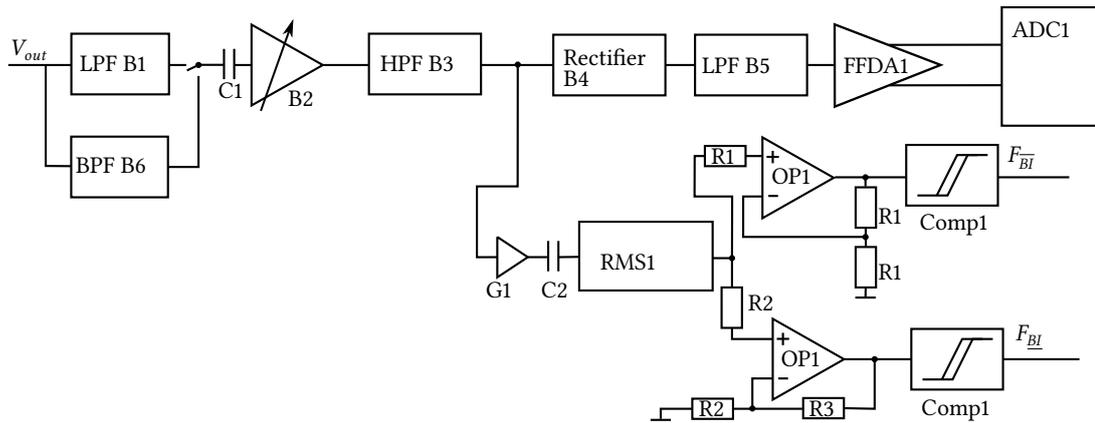


Figure A.4: The filter and amplifier block FiltBI and the fault detection block FaultBI in detail from Figure 3.10. The analog to digital converter ADC1 is also used within Figures A.3 and A.5. The operating voltages are  $-5 \text{ V}$  and  $5 \text{ V}$ .

**Table A.4:** Component list for the filter and amplifier block FiltBI1 and the fault detection block FaultBI1.

Name	Description	Parameters	Type / Manufacturer
R1	resistor	10 k $\Omega$ , 1 %	
R2	resistor	1 k $\Omega$ , 1 %	
R3	resistor	240 k $\Omega$ , 1 %	
C1	capacitor	0.1 $\mu$ F	
C2	capacitor	0.1 $\mu$ F	
OP1	operational amplifier		OPA2131UJ, Texas Instruments
RMS1	RMS-to-DC converter	operating voltages are -2.5 V and 2.5 V	LTC1968, Linear Technology
Comp1	comparator		LTC1442, Linear Technology
LPF B1	low-pass filter	4. order, $f_{\text{cut-off}} = 55$ kHz, Chebychev with ripple of 0.5 dB, MFB single ended, gain = 1	LT1801, Linear Technology
B2	programmable amplifier	gain = 1x, 5x, 20x, or 100x	LTC6910, Linear Technology
LPF B3	high-pass filter	2. order, $f_{\text{cut-off}} = 26$ kHz, Butterworth, Sallen-Key, gain = 1	LT1679, Linear Technology
B4	rectifier	described in Section A.3.4	LT1679, Linear Technology
LPF B5	low-pass filter	4. order, $f_{\text{cut-off}} = 13$ kHz, Butterworth, MFB single ended, gain = 1	OPA2131UJ, Texas Instruments
FFDA1	feedback fully differential low-pass filter	1. order, $f_{\text{cut-off}} = 13$ Hz, Bessel, MFB fully differential, gain = 0.5	THS4521, Texas Instruments
ADC1	analog digital converter	24-bit, sampling frequency 12 kHz	ADS1278, Texas Instruments
G1	amplifier	gain = 0.42	LT6221CS8#PBF, Linear Technology

## A.3.3 BI Filtering for Measurement at 100 kHz

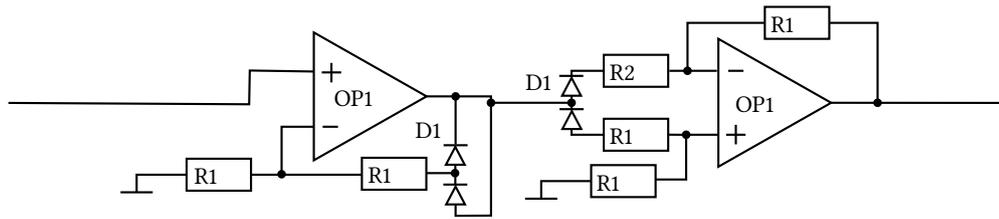


**Figure A.5:** The filter and amplifier block FiltBI2 and the fault detection block FaultBI2 in detail from Figure 3.11. The analog to digital converter ADC1 is also used within Figures A.3 and A.4. The operating voltages are  $-5\text{ V}$  and  $5\text{ V}$ .

**Table A.5:** Component list for the filter and amplifier block FiltBI2 and the fault detection block FaultBI2.

Name	Description	Parameters	Type / Manufacturer
R1	resistor	10 k $\Omega$ , 1 %	
R2	resistor	1 k $\Omega$ , 1 %	
R3	resistor	240 k $\Omega$ , 1 %	
C1	capacitor	0.1 $\mu$ F	
C2	capacitor	0.1 $\mu$ F	
OP1	operational amplifier		OPA2131UJ, Texas Instruments
RMS1	RMS-to-DC converter	operating voltages are -2.5 V and 2.5 V	LTC1968, Linear Technology
Comp1	comparator		LTC1442, Linear Technology
LPF B1	low-pass filter	4. order, $f_{\text{cut-off}} = 55$ kHz, Chebychev with ripple of 0.5 dB, MFB single ended, gain = 1	LT1802, Linear Technology
B2	programmable amplifier	gain = 1x, 5x, 20x, or 100x	LTC6910, Linear Technology
LPF B3	high-pass filter	2. order, $f_{\text{cut-off}} = 26$ kHz, Butterworth, Sallen-Key, gain = 1	LT1679, Linear Technology
B4	high-precision full-wave rectifier	described in Section A.3.4	LT1679, Linear Technology
LPF B5	low-pass filter	4. order, $f_{\text{cut-off}} = 13$ kHz, Butterworth, MFB single ended, gain = 1	OPA2131UJ, Texas Instruments
HPF B6	band-pass filter	2. order, $f_{\text{center}} = 100$ kHz, $Q = 16$ , gain = 2.3	LT1802, Linear Technology
FFDA1	feedback fully differential low-pass filter	1. order, $f_{\text{cut-off}} = 13$ Hz, Bessel, MFB fully differential, gain = 0.5	THS4521, Texas Instruments
ADC1	analog digital converter	24-bit, sampling frequency 12 kHz	ADS1278, Texas Instruments
G1	amplifier	gain = 0.42	LT6221CS8#PBF, Linear Technology

## A.3.4 Amplitude Demodulation



**Figure A.6:** Circuit layout with components of the high-precision full-wave rectifier B4 shown in Figure 3.12. The operating voltages are  $-5\text{ V}$  and  $5\text{ V}$ .

**Table A.6:** Component list for the high precision full-wave rectifier.

Name	Description	Parameters	Type / Manufacturer
R1	resistor	$10\text{ k}\Omega$ , 1%	
R2	resistor	$20\text{ k}\Omega$ , 1%	
D1	low-leakage double diode		BAV199, Infineon Technologies AG
OP1	operational amplifier		LT1679, Linear Technology