Adaptive Feedback Cancellation in Hearing Aids

Vom Fachbereich 18
Elektrotechnik und Informationstechnik
der Technischen Universität Darmstadt
zur Erlangung der Würde eines
Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Dissertation

von
Dipl.-Ing. Falco Strasser
geboren am 17.11.1985 in Hanau (Deutschland)

Referent:	Prof. Dr.-Ing. Henning Puder
Korreferent:	Prof. Dr.-Ing. Abdelhak M. Zoubir
Tag der Einreichung:	05.09.2016
Tag der mündlichen Prüfung:	30.01.2017

D 17
Darmstadt, 2017

To my Family and Friends.

Acknowledgments

I would like to thank all the people who helped and supported me during my doctoral study.

I especially thank Prof. Dr.-Ing. Henning Puder for his supervision. I can not think of a better guidance as Prof. Puder showed so much enthusiasm and interest in my work. He always provided me with very helpful scientific comments and feedback.

I wish to thank Prof. Dr.-Ing. Abdelhak Zoubir for his co-supervision during my doctoral study and for his advise during my diploma studies.
I also thank Prof. Dr.-Ing. Helmut Schlaak, Prof. Dr.-Ing. Christian Damm, and Prof. Dr.-Ing. Gerd Griepentrog for their work and time as Board of Examiners.

I am grateful to Dr.-Ing. Michael Muma for his guidance during my diploma thesis. I learned a lot about scientific work and research in that time.

Special thanks go to ALL my former colleagues at the Graduate School of Computational Engineering at TU Darmstadt. Especially, I want to mention Tobias Ritter, Stephan Krämer-Eis, Felix Loosmann, Crispin Deul, Dr.-Ing. Nicklas Linder, Dr.-Ing. Martin Lilienthal, Dr.-Ing. Fritz Kretzschmar, Zeger Bontnick, Jacopo Corno, Vera Bommer, Dr.rer.nat. Stefanie Bott, Christopher Müller, Christopher Spannring, Daniel Rettenmaier, Dan Koschier, Thorben Casper, Patrick Charrier, Dr.-Ing. Stefan Kneil, Anastasia Kondratyuk, Dr.-Ing. Laura Lukassen, and Dr.-Ing. Henning Spiegelberg. I had a really great and unforgettable time at the GSC because of you.
I am also thankful to all staff members at the GSC, namely Dr. Markus Lazanowski, Dr.-Ing. Melanie Gattermayer, Christian Schmitt, and Carina Schuster. You all helped and supported me during my time at GSC.

I would like to thank ALL current and former members of the Signal Processing Group at TU Darmstadt for their help.
Especially, I am grateful to Renate Koschella, the heart and the soul of the SPG, and Prof. em. Dr.-Ing. Eberhard Hänsler for his support during my doctoral study and encouraging words before my PhD presentation.

I want to thank my former roommates and great friends Dr.-Ing. Frederik Deckwerth, for the countless inspiring discussions, and Mousie Fasil, who offered much effort and time for comments and feedback on this thesis and my PhD presentation. I had a lot of fun living with both of you.

I wish to express my gratitude to my parents Wilma and Reinhard Strasser for their support throughout my life. You helped me to become the person I am today.

Finally, the most thanks go to Edwina for her love, support, understanding, and joy. Your love helped me to finish this thesis.

Berlin, 28.02.2017

Kurzfassung

Bei Hörgeräten, aber auch bei Redner- und Konzertanlagen, spielt die akustische Rückkopplung eine wichtige Rolle. Sie führt unter bestimmten Umständen zu einem lauten Pfeifen (engl. *Howling*), welches für den Hörgerätenutzer sehr unangenehm ist. Um das Pfeifen zu verhindern, muss die maximale Verstärkung des Hörgerätes reduziert werden, wodurch natürlich auch die Leistungsfähigkeit des Hörgerätes verringert wird. Die am häufigsten eingesetzte Methode zur Verringerung des Pfeifens ist die adaptive Rückkopplungsunterdrückung, welche in der Theorie die Rückkopplung vollständig beseitigen kann. Allerdings adaptiert sie zu einer Lösung mit Bias, wenn das Eingangssignal korreliert ist. Dadurch entstehen Verzerrungsartefakte (engl. *Entrainment*), die wiederum die Tonqualität drastisch schmälern können.
Mehrere wissenschaftliche Arbeiten befassen sich mit der Minimierung des Bias, um die Verzerrungsartefakte zu verhindern. Die meisten dieser Methoden funktionieren aber nur bei Sprache und nicht bei Musik, da die Tonalität und die Korrelation bei Musik viel größer ist. Dies hat einen größeren Bias und damit stärkere Verzerrungsartefakte als bei Sprache zur Folge. Andere Methoden, bei denen selbst bei Musik nur selten deutliche Artefakte auftreten, benutzen eine sehr langsame Adaptionsgeschwindigkeit. Dadurch wird aber die Fähigkeit, auf Änderungen des akustischen Rückkopplungspfades zu reagieren, deutlich reduziert. Infolgedessen tritt das Pfeifen bei Änderung des Rückkopplungspfades wesentlich länger auf.

In dieser Dissertation wird ein neues, adaptives Rückkopplungssystem für Hörgeräteanwendungen prsentiert. Um eine stabile Adaption ohne Pfeifen und Verzerrungsartefakte zu gewährleisten, werden Dekorrelationsmethoden mit einer neuen Realisierung einer nicht-parametrischen, variablen Schrittweite kombiniert.
Die Adaption findet in Teilbändern statt, wodurch der Rechenaufwand reduziert und gleichzeitig die Leistungsfähigkeit gesteigert wird.
Als Dekorrelationsmethoden werden Prädiktionsfehlerfilter und eine Frequenzverschiebung benutzt. Beides sind weit verbreitete Methoden, um den Bias zu reduzieren. Aber eine Kombination aus beiden wird das erste Mal in dieser Arbeit vorgeschlagen.
Zur Berechnung der empfohlenen Schrittweite wird die Signalleistung des gewünschten Eingangssignals, d.h. ohne Rückkopplungskomponenten, welches auch Quellsignal genannt wird, benötigt. Da das Quellsignal in der Realität nicht zugänglich ist, wird ein Weg über dessen Leistungsschätzung gewählt. Eine Methode zur Schätzung, welche in Kombination mit den Dekorrelationsmethoden zuverlässig ist, wird in dieser Arbeit theoretisch hergeleitet.
Außerdem werden drei zusätzliche Methoden zur Adaptionskontrolle vorgestellt,

mit denen die Leistungsfähigkeit weiter gesteigert werden kann. Mit der ersten Methode, der Impulsdetektion, werden Breitbandimpulse detektiert. Dadurch werden Verzerrungsartefakte, die infolge der Impulse auftreten können, verringert. Als zweites wird eine modifizierte Schätzung des Quellsignals vorgestellt. Diese modifizierte Schätzung stabilisiert das System, wenn schrille, schreiende Stimmen o.ä. auftreten. Die letzte Methode ist eine Korrelationsdetektion, die feststellt, ob das Quellsignal stark korreliert ist. Damit kann der Kompromiss zwischen Adaptionsstabilität und Adaptionsgeschwindigkeit verbessert werden.

Das gesamte System wird abschließend für verschiedene Sprach- und Musiksignale sowie für verschiedene Rückkopplungssituationen optimiert und evaluiert. Dazu werden unter realistischen Bedingungen gemessene Rückkopplungspfade und Methoden zum Hörverlustausgleich, wie sie in echten Hörgeräten benutzt werden, eingesetzt. Außerdem wird das System auch mit einem Echtzeitsystem getestet.
Zur objektiven Bestimmung der Leistungsfähigkeit des Systems werden zwei Maße definiert: Zum einen die Fähigkeit, Verzerrungsartefakte zu verhindern (Adaptionsstabilität) und zum anderen die Fähigkeit, auf die Änderung des Rückkopplungspfades zu reagieren (Adaptionsgeschwindigkeit).
Das gesamte adaptive Rückkopplungsunterdrückungssystem zeigt eine exzellente Leistungsfähigkeit. Dabei benötigt das System nur wenige Parameter und zeigt einen geringen Rechenaufwand. Diese Eigenschaften machen das System sehr relevant für die Praxis.

Abstract

Acoustic feedback is a well-known phenomenon in hearing aids and public address systems. Under certain conditions it causes the so-called howling effect, which is highly annoying for the hearing aid user and limits the maximum amplification of the hearing aid. The most common choice to prevent howling is the adaptive feedback cancellation algorithm, which is able to completely eliminate the feedback signal. However, standard adaptive feedback cancellation algorithms suffer from a biased adaptation if the input signal is spectrally colored, as it is for speech and music signals. Due to this bias distortion artifacts (entrainment) are generated and consequently, the sound quality is significantly reduced.
Most of the known methods to reduce the bias have focused on speech signals. However, those methods do not cope with music, since the tonality and correlation are much stronger for such signals. This leads to a higher bias and consequently, to stronger entrainment for music than for speech. Other methods, which deal with music signals, work only satisfactorily when using a very slow adaptation speed. This reduces the ability to react fast to feedback path changes. Hence, howling occurs for a longer time when the feedback path is changing.

In this thesis, a new sub-band adaptive feedback cancellation system for hearing aid applications is proposed. It combines decorrelation methods with a new realization of a non-parametric variable step size.
The adaptation is realized in sub-bands which decreases the computational complexity and increases the adaptation performance of the system simultaneously.
The applied decorrelation methods, prediction error filter and frequency shift, are well known approaches to reduce the bias. However, the combination of both is first proposed in this thesis.
To apply the proposed step size in the context of adaptive feedback cancellation, a method to estimate the signal power of the desired input signal, i.e., without feedback, also referred to as source signal power is necessary. This estimate is theoretically derived and it is demonstrated that it is a reliabe estimate if the decorrelation methods are additionally applied.
In order to further improve the performance of the system three additional control methods are derived: The first one is an impulse detection to detect wideband impulses, which could lead to misadaptation. Secondly, a modified estimate of the source signal power to stabilize the system in case of jarring voices is proposed. Lastly, a correlation detection, which is applied to improve the trade-off between adaptation

stability and tracking behavior, is developed.

The complete system is optimized and evaluated for several speech and music signals as well as for different feedback scenarios in simulations with feedback paths measured under realistic situations. Additionally, the system is tested by real-time simulations with hearing aid dummies and a torso and head simulator. For both simulation setups hearing loss compensation methods as applied in realistic hearing aids are used. The performance is measured in terms of being able to prevent entrainment (adaptation stability) and reacting to feedback path changes (tracking behavior).
The complete adaptive feedback cancellation system shows an excellent performance. Furthermore, the system relies only on few parameters, shows a low computational complexity, and therefore has a strong practical relevance.

Contents

1 Introduction **1**
1.1 Introduction and Motivation 1
1.1.1 Aims 2
1.1.2 Publications 2
1.2 Contributions of this Doctoral Project 3
1.3 Definitions and Notations 4

2 Overview of Acoustic Feedback Cancellation **7**
2.1 The Acoustic Feedback Problem 7
2.2 Acoustic Feedback Control 9
2.3 Adaptive Feedback Cancellation 12
2.4 Evaluation Procedures and Quality Measures 14

3 Adaptive Feedback Cancellation: Proposed System **21**
3.1 Basic System 21
3.1.1 Polyphase Filter Bank 23
3.1.2 Evaluation of the Basic System 26
3.1.3 Parameter Setting 28
3.2 Decorrelation Methods 28
3.2.1 Prediction Error Filter 29
3.2.2 Frequency Shift 31
3.2.3 Effect on the Bias and the Adaptation 33
3.3 Variable Step Sizes 41
3.3.1 Overview 42
3.3.2 Optimal Step Size 47
3.3.3 Non-Parametric Variable Step Size 49
3.3.4 Evaluation of Variable Step Sizes 54
3.3.5 Conclusion: Proposed Step Size Control 60
3.4 Additional Control Methods 61
3.4.1 Impulse Detection 61
3.4.2 Modified Estimate of Source Signal Power 65
3.4.3 Correlation Detection 67
3.5 Conclusion 78

4 Adaptive Feedback Cancellation: Real-time Implementation **79**
4.1 Experimental Setup 79
4.2 Graphical User Interface 82

5 Parameter Optimization and Evaluation **85**
5.1 Overview: Overall System and Parameters 86
5.1.1 Simulation Settings 86
5.1.2 Parameters 88
5.2 Parameter Optimization 90
5.2.1 Parameters for Variable Step Size 90
5.2.2 Parameter for Modified Estimate 98
5.2.3 Parameter for Correlation Detection 100
5.2.4 Conclusion 102
5.3 Performance Analysis of Decorrelation Methods 102
5.3.1 Evaluation based on Estimation Error of $\hat{\sigma}_x^2$ 103
5.3.2 Evaluation based on Effective Closed Loop Gain 106
5.4 Performance Analysis of Additional Control Methods 108
5.4.1 Impulse Detection 108
5.4.2 Modified Estimate of Source Signal Power 110
5.4.3 Correlation Detection 112
5.5 Evaluation using Real-time System 115
5.5.1 Analyses of Specific Situations 115
5.5.2 Evaluation based on Estimation Error of $\hat{\sigma}_x^2$ 138
5.6 Conclusion 139

6 Summary, Conclusions and Future Work **141**
6.1 Summary and Conclusion 141
6.2 Future Directions 144

List of Acronyms **145**

List of Symbols **147**

Bibliography **151**

Curriculum Vitae **157**

Chapter 1

Introduction

1.1 Introduction and Motivation

Whenever a sound signal is recorded by a microphone, processed (e.g. amplified) and then immediately played back by a loudspeaker nearby the microphone, it is unavoidable that the loudspeaker signal is fed back to the microphone. The described phenomenon is called acoustic feedback and it appears in case of, e.g., public address systems or hearing aids. Acoustic feedback creates a closed signal loop, which, under certain conditions, causes the so-called howling effect. These conditions depend on the magnitude and the phase of the feedback loop, which consists of the acoustic feedback path and the amplification of the system. In the case of hearing aids, the howling effect is highly annoying for the hearing aid user and his environment. Moreover, the acoustic feedback limits the maximum amplification that can be used in hearing aids. Despite much research in the last fifty to sixty years [1] the reduction of the acoustic feedback is still a very important and interesting task in the context of hearing aids.
The most promising and common approach to avoid acoustic feedback is the adaptive feedback cancellation (AFC) algorithm [1–4] as it provides the largest benefit of all known methods. However, due to the correlation between the source signal and the loudspeaker (or receiver) signal of the hearing aid a biased adaptation occurs, c.f. [1, 5, 6]. If the source signal is spectrally colored, which is the case for all kinds of real world signals, but especially tonal signals, this bias becomes very large. Hence, the bias can generate distortion artifacts. These artifacts are referred to as entrainment in this thesis.
Most of the known methods to reduce the bias have focused on speech signals [7–10]. However, those methods do not cope with music, since the tonality and correlation are much stronger for such signals leading to stronger entrainment than for speech. Other methods, which deal with music signals, work only satisfactorily when using a very slow adaptation speed [11]. This reduces the ability to react fast to feedback path changes. Hence, howling occurs for a longer time when the feedback path is changing. In this thesis a system is proposed which performs exceptionally well for all kinds of real world audio signals, even music signals. It shows very high adaptation stability and tracking ability at the same time.

1.1.1 Aims

The primary aim of this doctoral project is to develop a complete AFC system, which is embedded in a realistic hearing aid setup.
This system should have the abilities to prevent howling respectively entrainment and to track feedback path changes. Additionally, it has to work for all kinds of real world signals and under realistic conditions.
To evaluate the performance of the system an objective measure should be used. In order to obtain meaningful results, multiple kinds of audio signals and feedback paths need to be considered for the evaluations. Furthermore, the system needs to be implemented on a real-time system. This allows to test the performance of the system for even more realistic and time-varying scenarios.
Since the AFC system is implemented for hearing aids the computational costs of the system is limited. Hence, the proposed methods should not be to much computational demanding.

1.1.2 Publications

The following publications have been produced during this doctoral project.

- F. Strasser and H. Puder, "Correlation detection for adaptive feedback cancellation in hearing aids", *IEEE Signal Processing Letters*, Vol 23, No 7, pp. 979–983, July 2016.
- F. Strasser and H. Puder, "Adaptive feedback cancellation for realistic hearing aid applications", *IEEE/ACM Trans. on Audio, Speech, and Language Processing*, Vol 23, No 12, pp. 2322–2333, December 2015.
- F. Strasser and H. Puder, "Sub-band feedback cancellation with variable step sizes for music signals in hearing aids", In the Proc. of the *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2014*, in Florence, Italy, pp. 8207–8211, May 2014.

1.2 Contributions of this Doctoral Project

The main contribution is a complete AFC system which shows an impressive performance. The system is not only implemented offline but also on a real-time setup. As a consequence, it is evaluated by offline simulations as well as by real-time simulations.

Other important contributions:

- **Simulation procedure:** For the simulations an objective and realistic evaluation procedure is implemented. Therefore multiple representative audio signals and feedback paths along with a meaningful performance measure are used.
- **Combination of PEF and FS:** The combination of the prediction error filter (PEF) and the frequency shift (FS) is proposed the first time in literature and the benefit is clearly illustrated through simulations.
- **Variable step size investigations:** Several variable step sizes to control the adaptation are examined. Leading to the proposal to use the non-parametric variable step size (NPVSS) by Benesty et al. [12]. Furthermore, it is shown that, in theory, the AFC system adapts to the right solution with the NPVSS as no biased adaptation occurs. The fact that an optimal step size control can to lead to a unbiased adaptation in case of feedback cancellation is a new discovery.
- **Estimation of source signal power:** In order to implement the NPVSS in a concrete application a novel estimate for the source signal is theoretical derived and effectively implemented in the AFC system.
- **Correlation detection:** Another new contribution is the correlation detection. With this method one is able to detect if the source signal is spectrally colored, which leads to the biased adaptation. The correlation detection is used to additionally control the adaptation. Therewith, the trade-off between adaption stability and speed is improved.
- **Additional control methods:** An impulse detection to deal with wideband impulses and a modified source signal estimation to deal with jarring voices are developed and implemented in the system.
- **Implementation on real-time system:** The complete developed AFC system with all proposed methods in this thesis is implemented on the real-time system Speedgoat.

1.3 Definitions and Notations

This section gives an overview about the definitions and notations used in this thesis. In general, scalar values are denoted by lower case letters, vectors by bold lower case letters and matrices by bold upper case letters.
First the representations of signals are described:

- A scalar sampled signal is represented for example by $x(n)$ with n denoting the sampling time index, which is defined by the discrete sampling time nT_s of the continuous time signal $x(t)$. T_s is the sampling time and is given by $T_s = 1/f_s$, where f_s is the sampling frequency.

- A vector of a sampled signal has the length N and contains the current and the $N-1$ previous samples of the signal. Thus, the vector of $x(n)$ is given by $\mathbf{x}(n) = [x(n), x(n-1), \ldots, x(n-N+1)]^{\mathrm{T}}$.

- The spectrum of a signal $x(n)$ is given by $X(\Omega) = \mathcal{F}\{x(n)\}$, where $\mathcal{F}$ denotes the Fourier transform. Note that Ω denotes the normalized frequency. The spectrum can also depend on time by taking the short-term discrete time Fourier transform of a signal vector of time instance n, resulting in $X(\Omega, n)$.

- The power of a signal $x(n)$ with zero mean is defined as $\sigma_x^2(n) = \mathrm{E}\{x(n) \cdot x^*(n)\}$, where $\mathrm{E}\{\cdot\}$ is the expected value and '*' the conjugate complex operation.

- The auto-correlation of a signal is defined as $r_{xx}(n, n-l) = \mathrm{E}\{x(n) \cdot x^*(n-l)\}$. It can be represented by the vector $\mathbf{r}_{xx}(n) = [r_{xx}(n,n), r_{xx}(n,n-1), \ldots, r_{xx}(n,n-L+1)]^{\mathrm{T}}$ with the length L or the $L \times L$ auto-correlation matrix given as

$$\mathbf{R}_{xx}(n) = \begin{pmatrix} r_{xx}(n,n) & \cdots & r_{xx}(n,n-L+1) \\ \vdots & \ddots & \vdots \\ r_{xx}(n-L+1,n) & \cdots & r_{xx}(n-L+1,n-L+1) \end{pmatrix}.$$

- The cross-correlation between the signals $x(n)$ and $y(n)$ is defined as $r_{xy}(n, n-l) = \mathrm{E}\{x(n) \cdot y^*(n-l)\}$ and the vector representation is given as $\mathbf{r}_{xy}(n) = [r_{xy}(n,n), r_{xy}(n,n-1), \ldots, r_{xy}(n,n-L+1)]^{\mathrm{T}}$

The representation of linear independent systems, which are called filters throughout this thesis, are given next:

- The impulse response of a filter can be represented as a vector. Therefore, the filter is denoted as $\mathbf{f}(n) = [f_0(n), f_1(n), \ldots, f_{N-1}(n)]^{\mathrm{T}}$, where $f_j(n)$ with $j = 0, 1, \ldots, N-1$ are the filter coefficients at time index n.
- The frequency response of a filter $f(n)$ is denoted as $F(\Omega)$, which is given by the Fourier transform of the impulse response. If the frequency response of the filter is time-depended, it is defined by the short-term frequency response $F(\Omega, n)$.

The output signal $y(n)$ of a filter $f(n)$ is determined in the time domain through the convolution of the input signal $x(n)$ with $f(n)$. Hence, $y(n) = f(n) * x(n)$, where '$*$' denotes the convolution operator. For discrete signals and filters the convolution is defined as $y(n) = \sum_{k=0}^{N-1} f_k(n)x(n-k)$. In vector notation this is noted as $\mathbf{y}(n) = \mathbf{f}^{\mathrm{T}}(n)\mathbf{x}(n)$. Note in case of complex signals, one has to take the complex conjugated coefficients. Meaning, $y(n) = \sum_{k=0}^{N-1} f_k^*(n)x(n-k)$ and $\mathbf{y}(n) = \mathbf{f}^{\mathrm{H}}(n)\mathbf{x}(n)$.

Chapter 2

Overview of Acoustic Feedback Cancellation

This chapter gives an overview of the acoustic feedback cancellation. First, the acoustic feedback problem is described and conditions for howling are given in Section 2.1. After that, several methods to control, respectively, cancel the acoustic feedback are discussed in Section 2.2. This discussion leads to the conclusion to use the adaptive feedback cancellation. This method is explained in detail in Section 2.3 along with its advantages and disadvantages. Finally, in Section 2.4 the evaluation procedures for the proposed system are described. Additionally, the audio data and feedback paths used for simulations are presented.

2.1 The Acoustic Feedback Problem

The acoustic feedback problem is illustrated in Figure 2.1. The original source signal $x(n)$ is received by a microphone and then processed by the hearing aid processing represented by the forward path $G(\Omega, n)$. This results in the loudspeaker signal $u(n)$, which is played back by the loudspeaker (or receiver) of the hearing aid. Through the feedback path $F(\Omega, n)$ the sound signal travels back to the microphone. Consequently, a closed signal loop is created.
The forward path $G(\Omega, n)$ combines all signal processing methods used in a hearing aid, typically including a frequency dependent gain, automated gain control, and noise reduction. It is assumed that the forward path has a delay d_G of at least one sample. The feedback path $F(\Omega, n)$ represents along with the acoustic feedback path itself the frequency characteristics of the loudspeaker, microphone, digital-to-analog and analog-to-digital converters. Due to the processing delay of the converters and the propagation time of the sound signal from the loudspeaker to the microphone the feedback path has a delay d_F. The source signal $x(n)$ is the desired signal, i.e., without the acoustic feedback signal $v(n)$. Thus, ideally, the source signal $x(n)$ processed by the forward path $G(\Omega, n)$ results in the loudspeaker signal $u(n)$, i.e., $u(n) = \mathbf{g}(n)^{\mathrm{T}}\mathbf{x}(n)$ in the time domain or $U(\Omega, n) = G(\Omega, n)X(\Omega, n)$ in the frequency domain. However, due to the feedback path, a closed loop system is created with the frequency response

$$C(\Omega, n) = \frac{U(\Omega, n)}{X(\Omega, n)} = \frac{G(\Omega, n)}{1 - G(\Omega, n)F(\Omega, n)}. \tag{2.1}$$

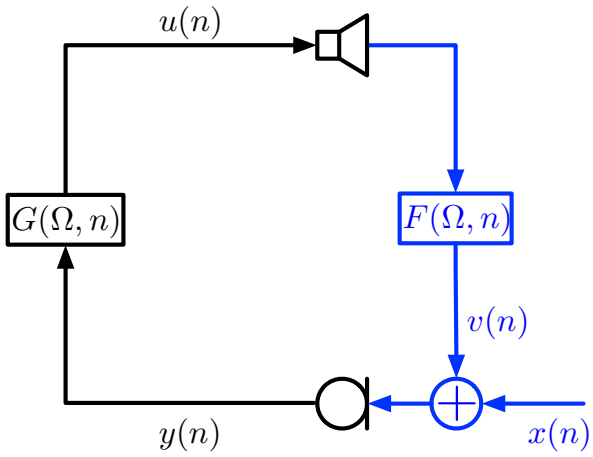


Figure 2.1. Illustration of the acoustic feedback problem. The original source signal $x(n)$ is received by a microphone and then processed by the hearing aid processing represented by the forward path $G(\Omega, n)$. This results in the output signal $u(n)$, which is played back by the loudspeaker (or receiver) of the hearing aid. Through the feedback path $F(\Omega, n)$ the sound signal is propagated back to the microphone. Consequently, a closed signal loop is created.

A closed loop system can become unstable leading to oscillations which are called howling in an acoustic system [1]. The Nyquist instability criterion [13] can be formulated as

$$\begin{cases} |F(\Omega, n)G(\Omega, n)| \geq 1 \\ \angle F(\Omega, n)G(\Omega, n) = 2\pi k, \quad k \in \mathbb{Z}. \end{cases} \tag{2.2}$$

If both conditions of Equations (2.2) are true, $C(\Omega, n)$ becomes unstable and howling occurs. The conditions are illustrated in Figure 2.2 and are essential, since all feedback cancellation methods presented in the next section have the aim to prevent either one or both conditions to be true.

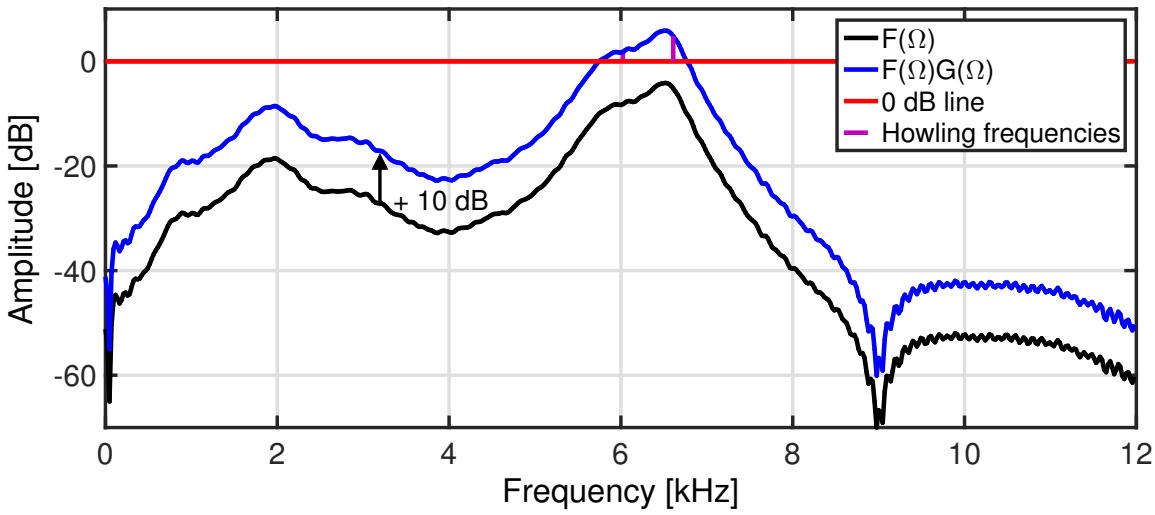


Figure 2.2. The frequency response of an exemplary feedback path. The forward path is assumed to be a 10 dB flat gain for all frequencies. $|F(\Omega)G(\Omega)|$ exceeds 0 dB at some frequencies. Howling occurs at these frequencies where both conditions of Equation (2.2) are fulfilled. These frequencies are marked with purple lines around 6 kHz.

2.2 Acoustic Feedback Control

Many solutions to the acoustic feedback problem have been proposed, cf. [1], including frequency and phase modulation, gain reduction, spatial filtering and acoustic path modeling. In this section these methods are presented along with their advantages and disadvantages:

- Frequency and Phase Modulation Methods: The frequency shift (FS) [2, 14, 15] belongs to one of the earliest acoustic feedback control approaches. The FS is applied to the audio signal between the microphone and the loudspeaker. This can be done before or after the hearing aid processing. Due to the FS the peaks of the closed loop gain and therefore the risk of howling are reduced. A drawback of the FS is the limited benefit of the approach as the peaks of $|G(\Omega, n)F(\Omega, n)|$ can at maximum be reduced by approximately 14 dB [16]. This benefit is additionally limited to avoid audible effects, like roughness for tonal signals. To limit the disturbing effects to an acceptable manner, the amount of the FS should be chosen small. This limits the reduction of the peaks even more to values below 6 dB [17].
  Other methods are phase modulation (PM) [18–20] or frequency modulation (FM) [20, 21] with the aim to prevent the second condition of Equation (2.2).

These methods suffer from the same drawbacks as the FS as they distort the audio signal and their benefit is even more limited [19–22].

- Gain Reduction Methods: An intuitive way to control acoustic feedback is to reduce the forward path gain in case of critical situations, i.e., when the closed loop gain $|G(\Omega, n)F(\Omega, n)|$ is near unity. Three different gain reduction methods exist:

  1. Automatic gain control [23, 24]: The gain is reduced equally over all frequencies.
  2. Automatic equalization [24–28]: The gain is only reduced in critical frequency regions.
  3. Notch filtering [29–33]: The gain is reduced in narrow frequency bands around critical frequencies.

  All methods need to be activated before howling occurs in order to work optimally. The drawback of the gain reduction methods is that a fast and reliable detection is very difficult to achieve. Most methods are reactive, meaning the gain is reduced after howling occurs. Although some proactive methods [28, 32, 34] exist for general feedback control, it is not desirable to reduce the gain in a hearing aid for longer time periods as it is essential to increase the loudness to compensate for hearing loss. Hence, gain reduction is used if at all as a "safety net", e.g., as last resort method if other AFC methods fail. In case of notch filtering another disadvantage is the distortion of the desired audio signal, which is unavoidable.

- Spatial Filtering Methods: The idea of spatial filtering methods [35–37] is to alter the closed loop response $G(\Omega, n)F(\Omega, n)$ by using microphone arrays and beamforming filtering. In case of hearing aids at least two microphones are normally present. The general aim is to design a beamformer with its main loop in direction of the source signal while canceling out the signals of the loudspeaker direction. Since the location of the source and the feedback path can change rapidly in a hearing aid scenario, the beamformer has to be adaptive. However, due to the closed loop the source signal and the loudspeaker signal are correlated. That means the conventional adaptive beamforming algorithms will not converge to the optimal solution. As a consequence, parts of the source signal are attenuated and parts of the feedback signal are received. Additionally, the feedback signal is not received by the microphones of the hearing aid from a single direction. Accordingly, it is very challenging to implement a spatial filtering method effectively.

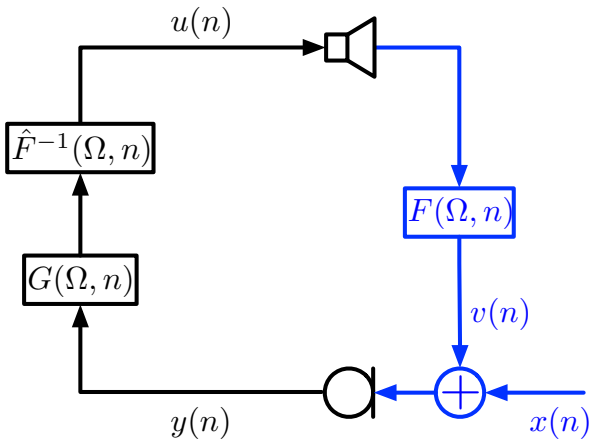


Figure 2.3. AIF scheme. The inverse model of the feedback path $\hat{F}^{-1}(\Omega, n)$ is inserted into the closed loop.

- Acoustic Path Modeling Methods: A model of the acoustic feedback path is estimated. One can differentiate between two methods of the adaptive inverse filtering (AIF) approach and the adaptive feedback canceling (AFC) approach.
  1) In the AIF approach [38, 39] an inverse model of the feedback path is determined which is then inserted into the closed loop as depicted in Figure 2.3. Hence, the closed loop gain is smoothed and the peaks are reduced. However, this method is limited as it can only achieve a peak reduction up to 10 dB [17]. Furthermore, it has a significant influence on the loudspeaker signal, which should be prohibited in hearing aid applications.
  2) The AFC approach [1–4] is conceptually similar to acoustic echo cancellation (AEC), which is used among others in hands-free telephony. The impulse response of the acoustic feedback path is modeled by an adaptive filter. The estimated impulse response is convolved with the loudspeaker signal to obtain a predicted feedback signal. This signal is then subtracted from the microphone signal. The resulting feedback-compensated signal is an estimate of the desired source signal. Therefore, ideally a complete elimination of the feedback signal is obtained. However, due to the closed loop and the correlation between source signal and loudspeaker signal the adaptive filter does not converge to the optimal solution, which can add undesired artifacts to the audio signal. This problem is the main difference between AFC and AEC. More details to AFC approach are given in Section 2.3.

In conclusion, the AFC approach is the most used acoustic feedback control method. An important reason is that with the AFC approach the largest benefit of all methods can be achieved since ideally the feedback signal is completely eliminated. However, in

reality this is not achieveable. Particularly, because of the correlation between source signal and loudspeaker signal as shown in detail in Section 2.3. In this thesis, some methods to deal with this and other problems are proposed. In the end, a suitable AFC system for real hearing aid applications is achieved.

2.3 Adaptive Feedback Cancellation

In an adaptive feedback cancellation (AFC) system the impulse response of the acoustic feedback path $\mathbf{f}(n)$ is modeled by an adaptive filter $\hat{\mathbf{f}}(n)$. The loudspeaker signal $u(n)$ is convolved with the estimated impulse response. This results in an estimate of the feedback signal. The estimated feedback signal $\hat{v}(n)$ is subtracted from the microphone signal $y(n)$. Thereby the feedback-compensated signal $e(n)$ is generated being an estimate of the source signal $x(n)$. Figure 2.4 displays the block diagram of a common AFC system.

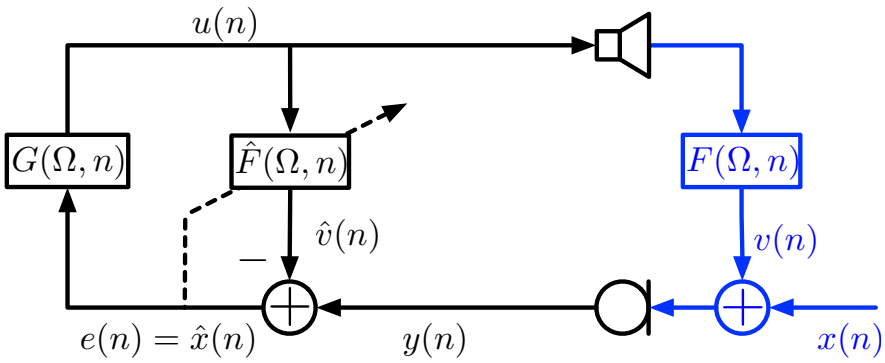


Figure 2.4. Block diagram of a common AFC system. The forward path denoted by its time dependent frequency response $G(\Omega, n)$ contains the other signal processing modules than AFC (e.g. amplification) of the hearing aid. $F(\Omega, n)$ is the acoustic feedback path between the loudspeaker and the microphone, which represents in addition to the feedback path also the frequency characteristics of the loudspeaker and microphone. The loudspeaker signal is denoted by $u(n)$ and the microphone signal by $y(n)$. The AFC algorithm estimates the feedback path $\hat{F}(\Omega, n)$ adaptively. The estimate of the feedback signal $\hat{v}(n)$ is subtracted from $y(n)$, resulting in the error signal $e(n)$, which is an estimate of the desired sound signal $x(n)$.

For the adaptive filter one can use common adaptation algorithms such as recursive least squares [40–42], affine projections [43, 44] or (normalized) least mean squares [40, 45]. Due to the trade-off between computational complexity and performance, in general, the normalized least mean squares (NLMS) algorithm is chosen as adaptive filter algorithm for AFC systems in hearing aid applications. For short filters the

adaptation speed is rather high and complex algorithms do not show a significant performance improvement. Thus, the NLMS algorithm is proposed in this thesis. The estimated impulse response is calculated by the NLMS algorithm as

$$\hat{\mathbf{f}}(n) = \hat{\mathbf{f}}(n-1) + \mu(n)\frac{e^*(n)\mathbf{u}(n)}{||\mathbf{u}(n)||^2}. \tag{2.3}$$

Note that the original complex version of the NLMS algorithm proposed by Widrow et al. [45] differs as $\mathbf{u}(n)$ is complex conjugated instead of $e(n)$. The difference occurs because of the multiple ways to define the cross-correlation. As mentioned in Section 1.3 the cross-correlation in this thesis is defined as $r_{xy}(n, n-l) = \mathrm{E}\{x(n) \cdot y^*(n-l)\}$ which results in the NLMS definition as noted in Equation (2.3). If the cross-correlation was defined as $r_{xy}(n, n-l) = \mathrm{E}\{x^*(n) \cdot y(n-l)\}$ the NLMS algorithm would be given as proposed by Widrow et al. [45]. Both versions are used in the literature.
The vector $\hat{\mathbf{f}}(n)$ in Equation (2.3) represents the N adaptive filter coefficients at time index n, the vector $\mathbf{u}(n) = [u(n), u(n-1), \ldots, u(n-N+1)]^{\mathrm{T}}$ contains the current and previous $N-1$ samples of the loudspeaker signal and the error signal is given by $e(n) = y(n) - \hat{\mathbf{f}}^{\mathrm{H}}(n-1)\mathbf{u}(n)$. The vector $\hat{\mathbf{f}}(n-1)$ contains the old estimated coefficients and the latter part of the right side is called the update term with the step size $\mu(n)$, which controls the adaptation speed. For a large and constant step size, e.g. $\mu(n) \to 1$, the adaptive filter converges faster than for a small step size, e.g. $\mu(n) \to 0$. However, with a high step size the adaptive filter strongly oscillates around the true solution while these oscillations are much smaller with a small step size. Hence, the estimation error decreases slower for small step sizes but will be smaller in the end than with a high step size. This trade-off can be avoided by using a variable step size, which increases if the difference between $f(n)$ and $\hat{f}(n)$ is large and decreases for a small difference. Several step size concepts are discussed and the one showing the best performance is proposed in Section 3.3.
Ideally, the AFC approach can eliminate the acoustic feedback completely if the adaptive filter converges to the correct solution, i.e., the impulse response of the acoustic feedback path. Nevertheless, all AFC systems have one drawback: Due to the closed loop, the source signal $x(n)$ and the loudspeaker signal $u(n)$ are correlated. Thus, a biased adaptation of $\hat{f}(n)$ occurs, c.f. [1, 5, 6]. To show this, first the *a posteriori* error signal is noted as

$$\epsilon(n) = x(n) + \sum_{l=0}^{N-1}[f_l(n) - \hat{f}_l(n)]^* u(n-l), \tag{2.4}$$

and in vector notation:

$$\epsilon(n) = x(n) + [\mathbf{f}(n) - \hat{\mathbf{f}}(n)]^{\mathrm{H}}\mathbf{u}(n). \tag{2.5}$$

The *a posteriori* error at time instance n is calculated with the updated adaptive filter coefficients $\hat{f}(n)$ while for the *a priori* error $e(n)$ the previous coefficients $\hat{f}(n-1)$ are

used. The aim of the NLMS algorithm is to minimize the mean squared error (MSE) of the *a posteriori* error signal. The MSE of $\epsilon(n)$ is given according to Equation (2.5) by

$$\begin{aligned}\mathrm{E}\{||\epsilon(n)||^2\} =& \sigma_x^2(n) + 2 \cdot \mathrm{Re}\{[\mathbf{f}(n) - \hat{\mathbf{f}}(n)]^{\mathrm{H}} \mathbf{r}_{xu}(n)\} \\ &+ [\mathbf{f}(n) - \hat{\mathbf{f}}(n)]^{\mathrm{H}} \mathbf{R}_{uu}(n) [\mathbf{f}(n) - \hat{\mathbf{f}}(n)].\end{aligned} \tag{2.6}$$

Deriving this with respect to $\hat{\mathbf{f}}(n)$ and setting the result to $\mathbf{0} = [0, \ldots, 0]^{\mathrm{T}}$ results in

$$\mathbf{r}_{xu}(n) + \mathbf{R}_{uu}(n)[\mathbf{f}(n) - \hat{\mathbf{f}}(n)] = \mathbf{0}. \tag{2.7}$$

Solving this for $\hat{\mathbf{f}}(n)$ provides the equation for the optimal estimated coefficients,

$$\hat{\mathbf{f}}_{\mathrm{opt}}(n) = \mathbf{f}(n) + \mathbf{R}_{uu}^{-1}(n)\mathbf{r}_{xu}(n), \tag{2.8}$$

for which the MSE is minimized. The bias of the adaptation given is by $\mathbf{b}(n) = \mathbf{r}_{xu}(n)\mathbf{R}_{uu}^{-1}(n)$. Due to the bias, the adaptive filter does not converge to the true impulse response of the acoustic feedback path. If $x(n)$ is spectrally colored, which is the case for all kinds of real world signals, but especially tonal signals, the bias becomes very large. Consequently, the bias can generate distortion artifacts, which are called entrainment. Entrainment occurs if the closed loop becomes unstable in consequence of the misadaptation. The most common decorrelation methods to deal with this problem are discussed and suitable methods are proposed in Section 3.2.

2.4 Evaluation Procedures and Quality Measures

In this section, the evaluation procedures and quality measures are explained, which are used to evaluate different settings and methods of the AFC system. Additionally, the audio data and feedback paths are presented which are used for the evaluation.

Quality Measures

Since the main goal of an AFC system is to avoid howling and entrainment, the best way to evaluate an AFC system is to find a measure which indicates if entrainment is present. As already discussed in Section 2.1 entrainment occurs if the closed loop system, represented by $C(\Omega)$, becomes unstable. The criteria for this are formulated in Equation (2.2) and can be modified in case of an adaptive filter application by

$$\begin{cases} |(F(\Omega, n) - \hat{F}(\Omega, n)) \cdot G(\Omega, n)| \geq 1 \\ \angle(F(\Omega, n) - \hat{F}(\Omega, n)) \cdot G(\Omega, n) = 2\pi k, \quad k \in \mathbb{Z}. \end{cases} \tag{2.9}$$

By using an AFC system one tries to avoid the first condition. Hence, in this thesis a quality measure is proposed which detects if the first condition of Equation (2.9) is fulfilled or not. This measure is called the effective closed loop gain (ECLG) and is given by

$$\begin{aligned} \mathrm{ECLG}(\Omega, n) &= |\Delta F(\Omega, n) G(\Omega, n)|, \\ &\text{with } \Delta F(\Omega, n) = F(\Omega, n) - \hat{F}(\Omega, n) \end{aligned} \tag{2.10}$$

where $G(\Omega, n)$ models the frequency response of the hearing aid gain, $F(\Omega, n)$ the frequency response of the feedback path and $\hat{F}(\Omega, n)$ is the estimated frequency response. Whenever $|\Delta F(\Omega, n) G(\Omega, n)|$ exceeds the value 1 howling or entrainment can occur. In decibel the same is true for $|\Delta F(\Omega, n) G(\Omega, n)|_{\mathrm{dB}} > 0\,\mathrm{dB}$. $G(\Omega, n)$ is assumed to be time dependent because the hearing aid amplification is realized with an automatic gain control (AGC), which adapts the hearing aid gain according to the input level. Therefore, $G(\Omega, n)$ varies depending on the loudness of the hearing aid input signal over time.
The ECLG depends on the frequency and time. Thus, one can analyze directly at which frequency and time instances entrainment occurs. This is shown in Figure 2.5 where the ECLG of an exemplary simulation is shown.

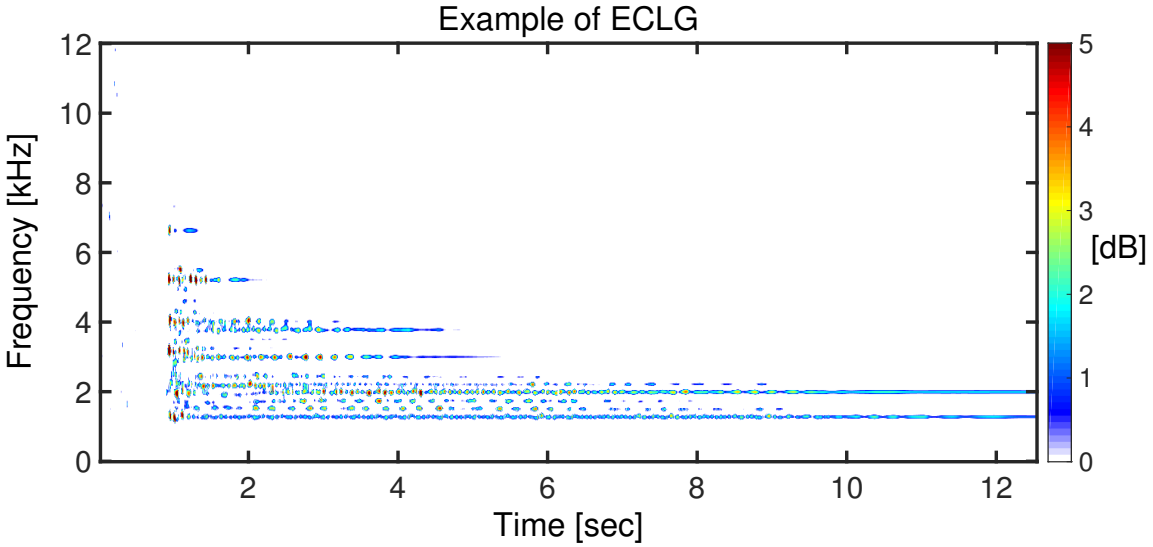


Figure 2.5. Example of the $\mathrm{ECLG}(\Omega, n)$ in dB. Only values above 0 dB are shown. Thus, one can easily see at which frequency and time instances entrainment occurs.

To evaluate if and when entrainment for a given sound signal and system setting occurs, one can take the maximum of the ECLG over all frequency, resulting in

$$\mathrm{ECLG}_{\max}(n) = \max_{\Omega} \{\mathrm{ECLG}(\Omega, n)\}. \tag{2.11}$$

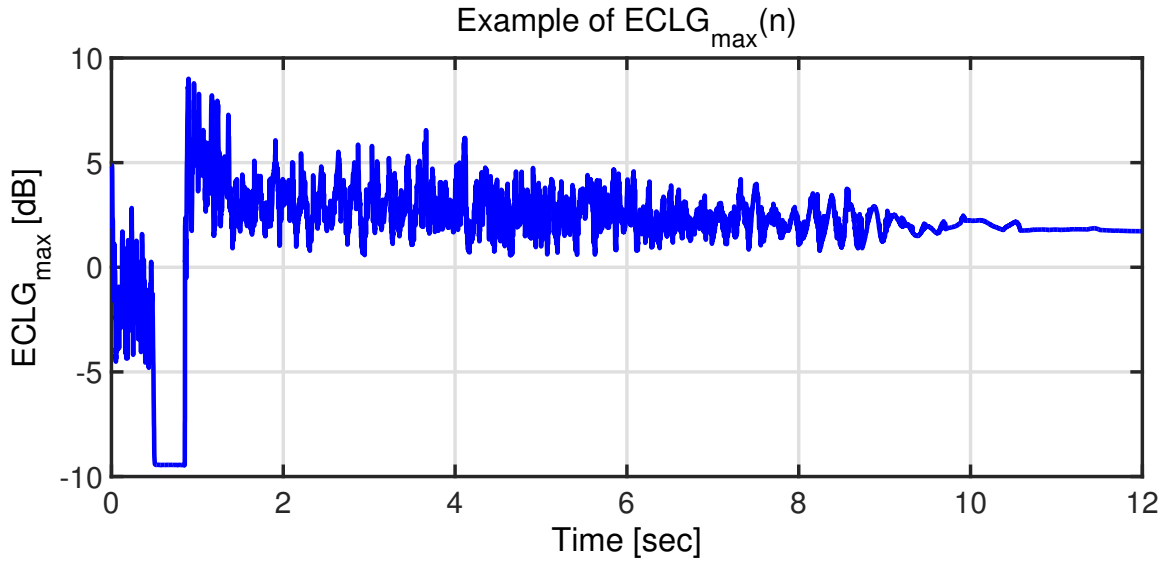


Figure 2.6. Example of the $\mathrm{ECLG}_{\max}(n)$ in dB. Here, it is easier to evaluate the amount and strength of entrainment. Also, one can efficiently compare two different methods in one plot. Note: The drop of the $\mathrm{ECLG}_{\max}(n)$ starting at 0.5 sec appears since a short silence occurs.

Figure 2.6 depicts the $\mathrm{ECLG}_{\max}(n)$ for the example given in Figure 2.5.
The presented quality measures can be used to analyze certain scenarios but does not provide an overall measure of the performance. The procedure to evaluate the overall performance is given in the following.

Simulation Procedure

The performance of an AFC system can be analyzed for different test cases:

1. Adaptation stability: The ability to eliminate and prevent entrainment for all kinds of signals, while the feedback path is fixed, e.g., does not change over time.
2. Tracking behavior: The time the systems needs to adapt in case of changing feedback paths.

First, the procedure to examine the adaptation stability is explained. Since the focus is set on the stability of the system and not yet on the tracking behavior, 0.5 seconds of white noise are added before the approximately ten minutes of audio. The white noise at the beginning allows the system to adapt to the true feedback path and it can

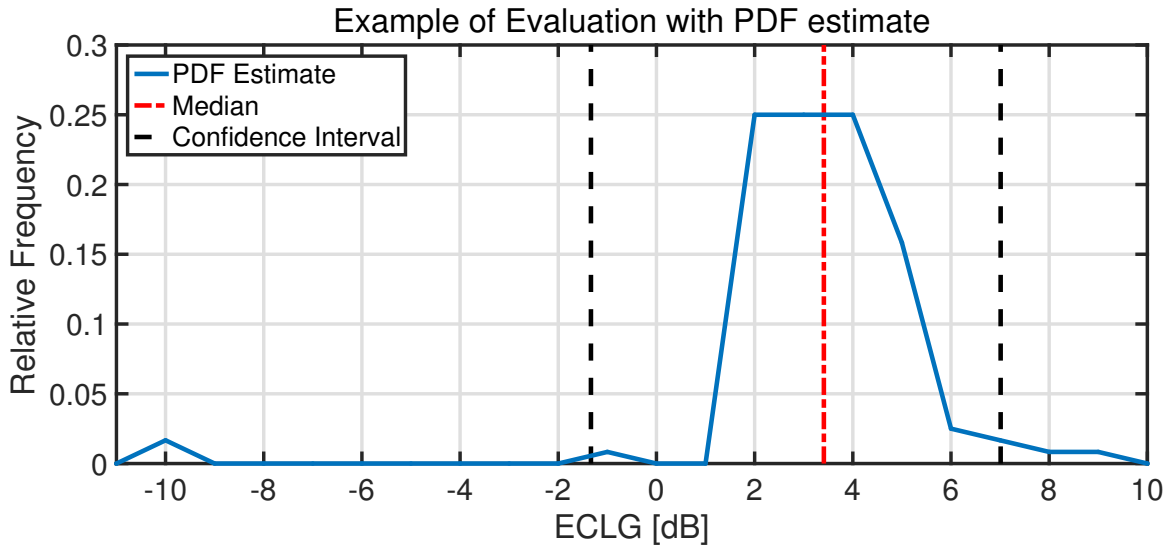


Figure 2.7. Example of the evaluation using PDF, median and CI. The PDF, which is estimated by a histogram and consists of 120 data points for this example, is displayed in blue. The red line denotes the median of all values. The two black lines indicate the lower and upper boundaries of the CI. This PDF gives a statistical evaluation of a given AFC system as one can measure for example how much entrainment occurs over the simulation time. Additionally, this evaluation can be used to compare multiple system settings in a objective and clearly structured way.

be examined if the system stays stable during the simulation or creates entrainment in consequence of the audio signal. With the generated input signal five simulations are performed separately. In each simulation one of the five different feedback paths is used. After the white noise section, $\mathrm{ECLG}_{\max}(n)$, according to Equation (2.11), is calculated for each adaptation step equal to a rate of 1 kHz. Then $\mathrm{ECLG}_{\max}(n)$ is split into intervals of 100 ms and for each interval the maximum value is calculated. This procedure is done for all five simulations. Next the probability density function (PDF) of the calculated maxima of all simulations is estimated by a histogram and called $p_{\mathrm{ECLG}}(x)$. Based on the PDF one can determine if and how probable entrainment occurs as shown in Figure 2.7. Additionally, the median and the confidence interval (CI) are calculated and depicted. The limits of the CI are determined with the method by Clopper and Pearson [46], also called the Clopper-Pearson interval. Here, the upper limit c_u is determined by $p_{\mathrm{ECLG}}(x \geq c_u) = 0.025$ and the lower limit c_l by $p_{\mathrm{ECLG}}(x \leq c_l) = 0.025$.

Second, the tracking behavior is analyzed by determining the time the system needs to adapt to changing feedback paths. The same signals as above are used. For this evaluation, though, the feedback path changes every second. After each change of the feedback path the time period that the system needs to adapt towards the new

feedback path is evaluated. Hence, the time until all values of $\mathrm{ECLG}_{\max}(n)$ are below 0 dB is measured. Then the distribution of the measured time values is analyzed by PDF. Thus, one obtains a statistical measure for the adaptation speed of the system.

Audio Data and Feedback Paths

Here, the used audio data and feedback paths are described. Approximately five minutes of speech and five minutes of music are used. For the speech recordings different speakers (male and female) and different languages (German, English and French) are considered. For the music recordings different instruments like bells, wind chimes, piano, organ, strings, etc. are chosen. All signals are calibrated for an output level of 60 dB SPL (sound pressure level).
Five different feedback paths are recorded in the audio laboratory of the Graduate School Computational Engineering (GSC CE), TU Darmstadt, using a behind-the-ear hearing aid with open fitting and a KEMAR dummy, see Figure 2.8.

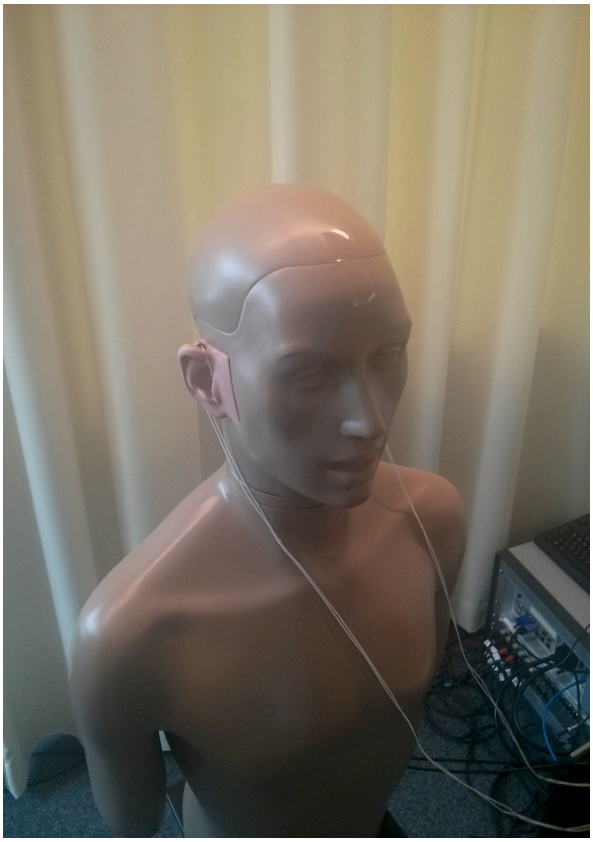


Figure 2.8. Picture of a feedback path recording scenario using a KEMAR dummy and a behind-the-ear hearing aid.

The frequency responses of the recorded feedback paths are shown in Figure 2.9. The hearing aid gain is chosen such that the closed loop gain is approximately 5 dB overcritical, meaning $\max\{|F(\Omega,n)G(\Omega,n)|_{\mathrm{dB}}\} \approx 5\,\mathrm{dB}$. This is a reasonable choice as studies show that an overcritical gain of approximately 5 dB is most challenging for an AFC system. In case of higher gains the system can easier adapt as the feedback signal power is stronger and thus the influence of the original sound signal $x(n)$, which is a disturbance for the adaptive filters, is less. Due to the adaptation in sub-bands and the shape of $F(\Omega,n)G(\Omega,n)$ the system has to deal with lower magnitudes of the closed loop gain $|F(\Omega,n)G(\Omega,n)|$ also. Figure 2.9 shows for each feedback path a closed loop frequency response $|F(\Omega,n)G(\Omega,n)|$.

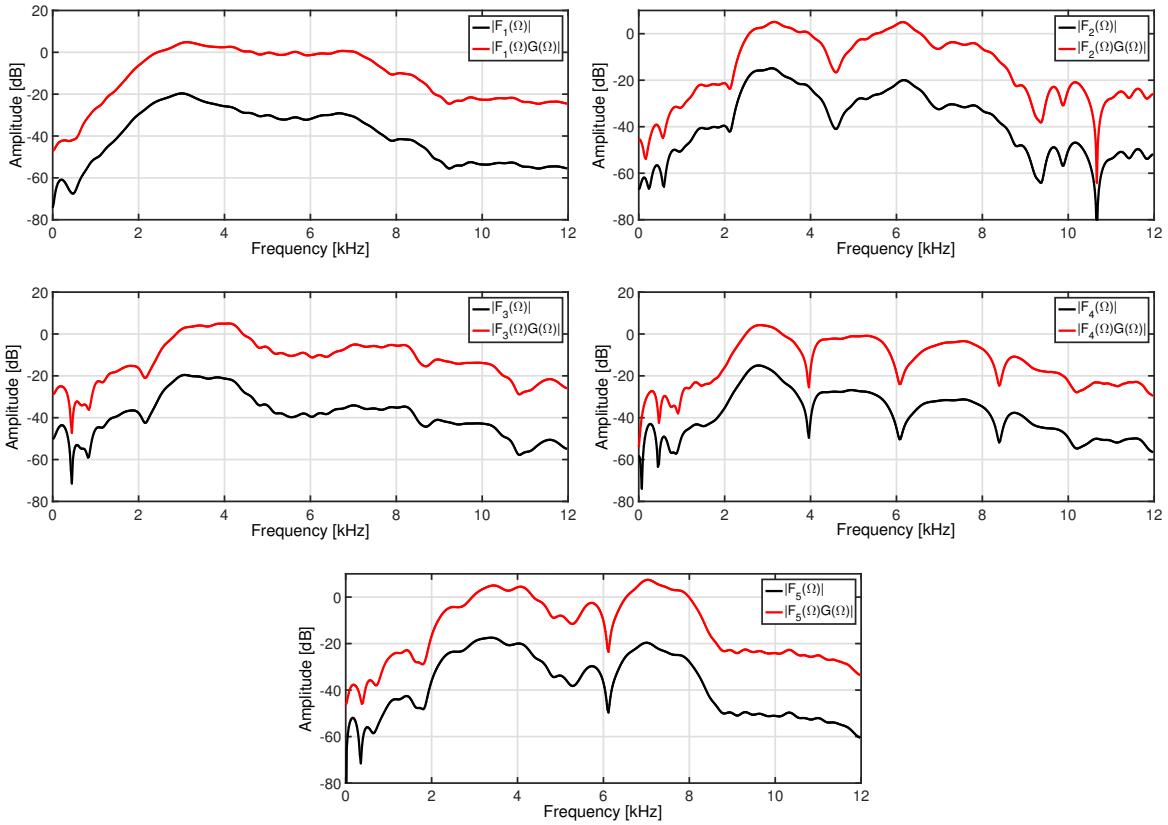


Figure 2.9. The frequency responses of all five feedback paths are depicted as well as frequency responses of the closed loop with an overcritical gain equal to 5 dB. $F_1(\Omega)$ corresponds to the scenario shown in Figure 2.8. The other frequency responses represent situations in which a hand is placed around the ear of the KEMAR. Due to the different placement of the hand the frequency responses vary strongly. Exemplary, at frequency 6 kHz a dynamic of about 25 dB ($F_2(6\,\mathrm{kHz}) \approx 5\,\mathrm{dB}$; $F_4(6\,\mathrm{kHz}) \approx -20\,\mathrm{dB}$) is observed. Similar observation can be made for other frequency regions. However, the region between 2 kHz and 3 kHz is critical in all cases. Additionally, the frequency dependence of the hearing aid gain is depicted, e.g., for all cases a difference of about 7 dB between 2 kHz and 10 kHz is observable.

Chapter 3

Adaptive Feedback Cancellation: Proposed System

In this chapter, the proposed AFC system is presented in detail. First, the basic setup is explained in Section 3.1. It consists of adaptive finite impulse response (FIR) filters applied in uniform frequency sub-bands. In Section 3.2 two decorrelation methods are introduced to reduce the bias of the adaptation, which depends on the correlation between the source signal and the loudspeaker signal. To control the adaptation the use of a variable step size (VSS) is proposed. Therefore, several VSS methods are discussed and the superior one is determined in Section 3.3. In Section 3.4, additional adaptation control methods are explained. The first two are required when dealing with real conditions. The third is a novel approach to detect the correlation between the source and loudspeaker signals.

3.1 Basic System

In this section the basic system, depicted in Figure 3.1, is introduced. It consists of a loudspeaker (or receiver) and a microphone as hardware components. The blocks in between show different signal processing components.

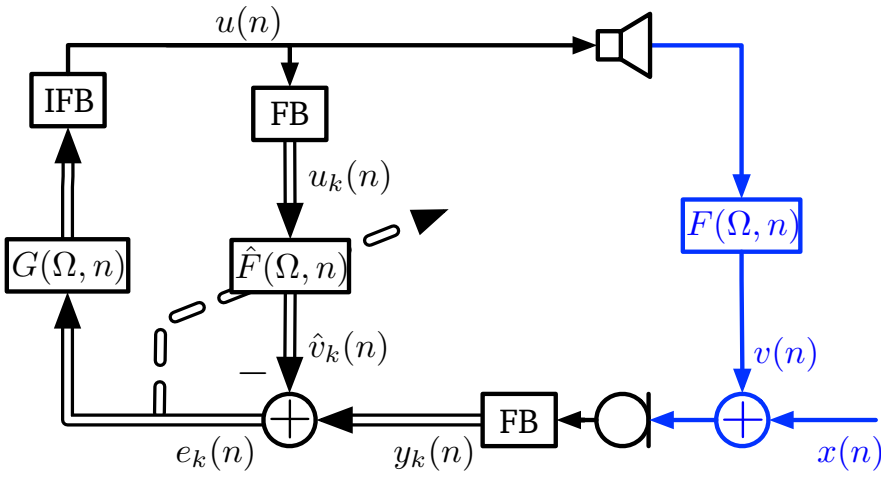


Figure 3.1. Scheme of the basic AFC concept.

The signal processing setup is as follows: The microphone input signal of the hearing

aid is decomposed into several sub-bands with a filter bank (FB). This allows a flexible compensation of frequency dependent hearing losses and the utilization of other frequency domain signal processing techniques, such as noise reduction. For further theoretical analyses of the AFC system, all these components are combined in one block with a linear frequency response $G(\Omega, n)$. Nevertheless, practical investigations of the AFC performance in this thesis, which are based on simulations and real-time investigations, are done with a non-linear automatic gain control (AGC) component for hearing loss compensation as they are used by real hearing aids. An inverse filter bank (IFB) is applied before the output signal is forwarded to the hearing aid loudspeaker. The AGC reduces the hearing aid gain to ensure that the loudspeaker output does not exceed the uncomfortable level of the hearing aid user. Therefore, the amplification is reduced above a knee point c_k by applying a compression ratio c_r. Figure 3.2 shows the applied gain using an AGC depending on the input power level with $c_k = 65\,\text{dB SPL}$ and $c_r = 2$.

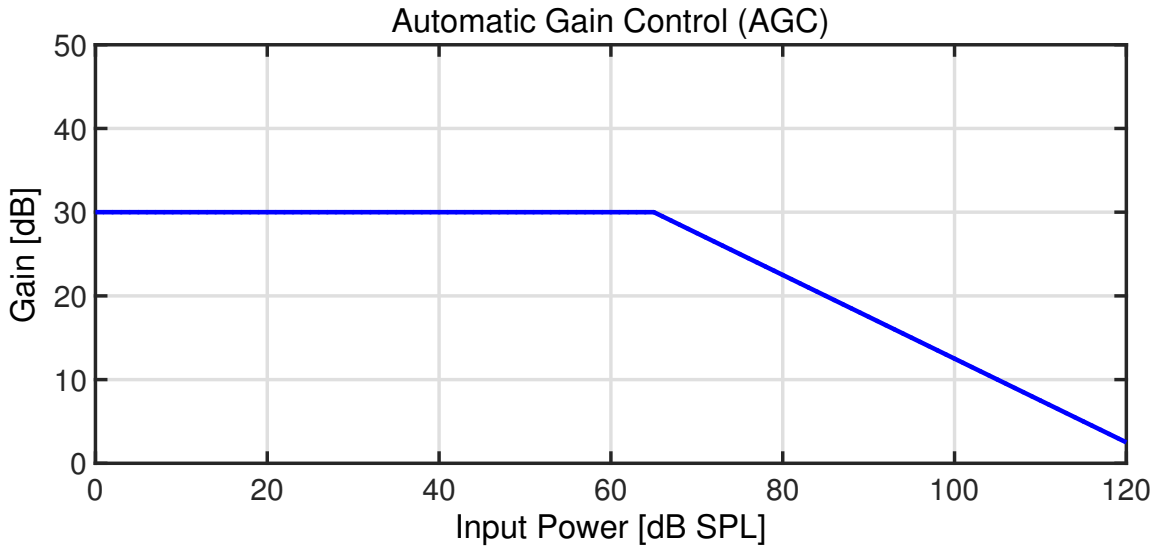


Figure 3.2. The figure shows the applied gain using an AGC depending on the input power level. For a power level below the knee point of 65 dB SPL an exemplary gain of 30 dB is used. Above the knee point the gain is reduced by the compression ratio of $c_r = 2$.

The feedback cancellation is realized in sub-bands by a frequency decomposition of the loudspeaker signal and the application of independent adaptive filters in each sub-band to model the feedback path. The output signals of the adaptive filters are subtracted from the sub-bands of the input signal to cancel the feedback.
The sub-band processing reduces the computational complexity and allows a frequency selective adaptation and control. Furthermore, due to the sub-band processing the performance of the NLMS algorithm is increased significantly, cf. [47, 48]. To decompose

the signals into 48 effective sub-bands, each with a bandwidth of 250 Hz, a polyphase filter bank with a prototype low-pass filter, c.f. Section 3.1.1, is used. However, no adaptation is performed in the first six sub-bands, which is equal to 1500 Hz. The reason is that the feedback paths show a bandpass characteristic with low feedback risk below 1500 Hz. An adaptation in these bands has the risk to create entrainment without any benefit.
For the adaptation a normalized least mean squares (NLMS) adaptation rule [40] is formulated:

$$\hat{\mathbf{f}}_k(n) = \hat{\mathbf{f}}_k(n-1) + \mu_k(n)\frac{e_k^*(n)\mathbf{u}_k(n)}{||\mathbf{u}_k(n)||^2}, \tag{3.1}$$

where the index k denotes the k-th sub-band. The vector $\hat{\mathbf{f}}_k(n)$ represents the N adaptive filter coefficients at time index n, the vector $\mathbf{u}_k(n)$ contains the current and the previous $N-1$ samples of the loudspeaker signal and the error $e_k(n)$ is given by $e_k(n) = y_k(n) - \hat{\mathbf{f}}_k(n-1)^{\mathrm{H}}(n)\mathbf{u}_k$. The scalar $\mu_k(n)$ is known as step size and controls the adaptation speed. The number of coefficients for each adaptive filter is set to $N = 3$. More coefficients do not increase the performance of the AFC system significantly, see Section 3.1.2, but the computational complexity increases. To prevent values of the norm $||\mathbf{u}_k(n)||^2$ in Equation (3.1) close to zero a regularization value equal to 0 dB SPL is added to the norm.

3.1.1 Polyphase Filter Bank

The polyphase filter bank, [49], is an efficient way to decompose a signal into the sub-bands. The structure of an efficiently implemented polyphase filter bank is depicted in Figure 3.3.
In addition to the reduced computational costs the polyphase filter bank has another advantage compared to the standard discrete fourier transform (DFT). The base-band signal $y(n)$ is divided into $K = 96$ sub-bands by the convolution with K bandpass filters g_k, $k \in \{1, 2, \ldots, K-1\}$, with length N_{FB}. The mathematical formulation for one sub-band is expressed as follows

$$\overline{y}_k(n) = \sum_{i=0}^{\infty} \hat{g}_{k,i} y(n-i). \tag{3.2}$$

The frequency resolution can be improved by choosing N_{FB} larger than the length of the DFT. The bandpass filters are based on a frequency shifted prototype lowpass filter g_0 and are given by

$$g_{k,i} = g_{0,i} \cdot e^{j2\pi i \frac{k}{K}}. \tag{3.3}$$

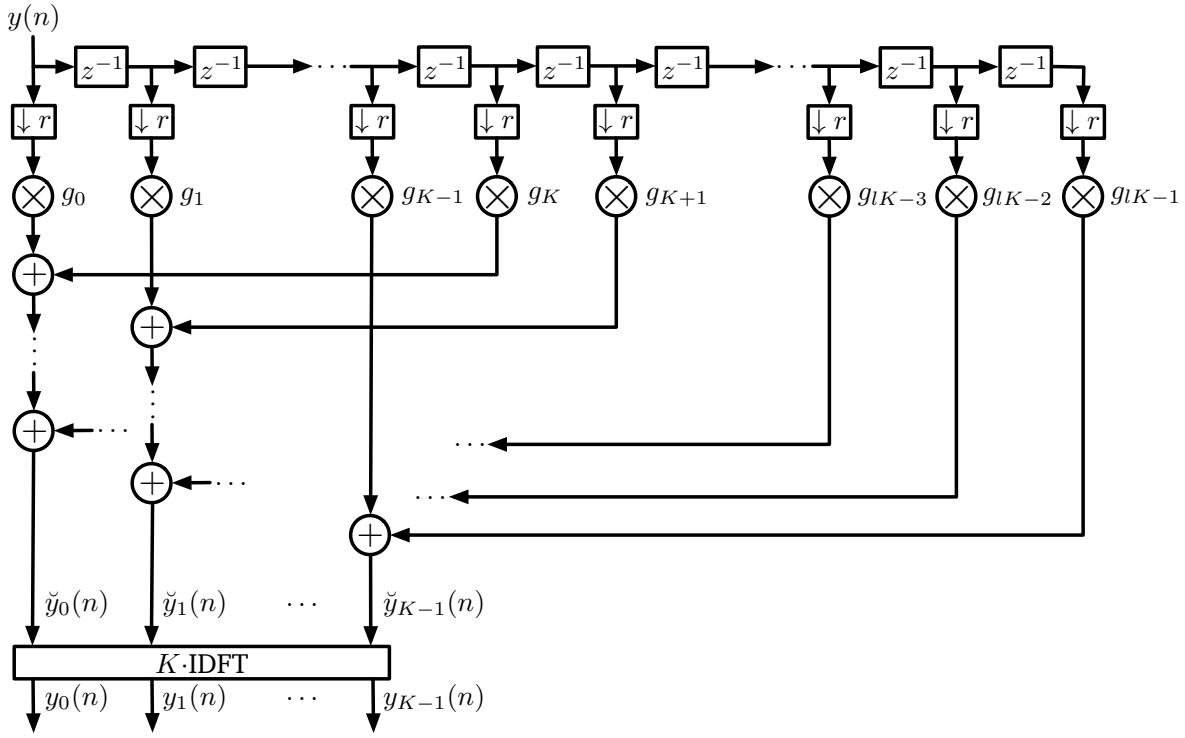


Figure 3.3. Scheme of the analysis polyphase filter bank.

Since the bandwidth for each sub-band is smaller than for the baseband, the signals $\overline{y}_k(n)$ can be downsampled by the factor r. Resulting in

$$y_k(n) = \overline{y}_k(rn). \tag{3.4}$$

The sampling frequency is reduced for each sub-band to f_s/r. For ideal bandbass filters the downsampling factor can be choosen as $r = K$. However, for realistic applications r has to be small enough to prevent aliasing. Here, r is set to 24, i.e., an oversampling of $K/r = 4$ is applied.
Now the filter length is chosen to be an integer multiple of K. Thus,

$$N_{\mathrm{FB}} = LK \text{ with } L \in \mathbb{N}. \tag{3.5}$$

This can be easily achieved by adding zeros to the impulse response of $g_{0,i}$. Then the index i can be replaced by $i = lK+\nu$ with $l \in \{0, 1, \ldots, L-1\}$ and $\nu \in \{0, 1, \ldots, K-1\}$. This results in

$$y_k(n) = \sum_{l=0}^{L-1} \sum_{\nu=0}^{K-1} y(rn - lK - \nu) g_{0,lK-\nu} e^{j2\pi \frac{lK+\nu}{K} k} \tag{3.6}$$

$$= \sum_{l=0}^{L-1} \underbrace{\sum_{\nu=0}^{K-1} y(rn - lK - \nu) g_{0,lK-\nu}}_{=\breve{y}_\nu(n)} \underbrace{e^{j2\pi lk}}_{=1} e^{j2\pi \frac{\nu}{K} k}, \tag{3.7}$$

which corresponds to the inverse discrete fourier transform (IDFT). Hence,

$$\mathbf{y}(n) = K \cdot \mathrm{IDFT}\{\breve{\mathbf{y}}(n)\} \tag{3.8}$$

with $\mathbf{y}(n) = [y_0(n), y_1(n), \ldots, y_{K-1}(n)]$ and $\breve{\mathbf{y}}(n) = [\breve{y}_0(n)], \breve{y}_1(n), \ldots, \breve{y}_{K-1}(n)$. This is the mathematical formulation of the analysis filterbank. The synthesis filterbank is given analogously as depicted in Figure 3.4.

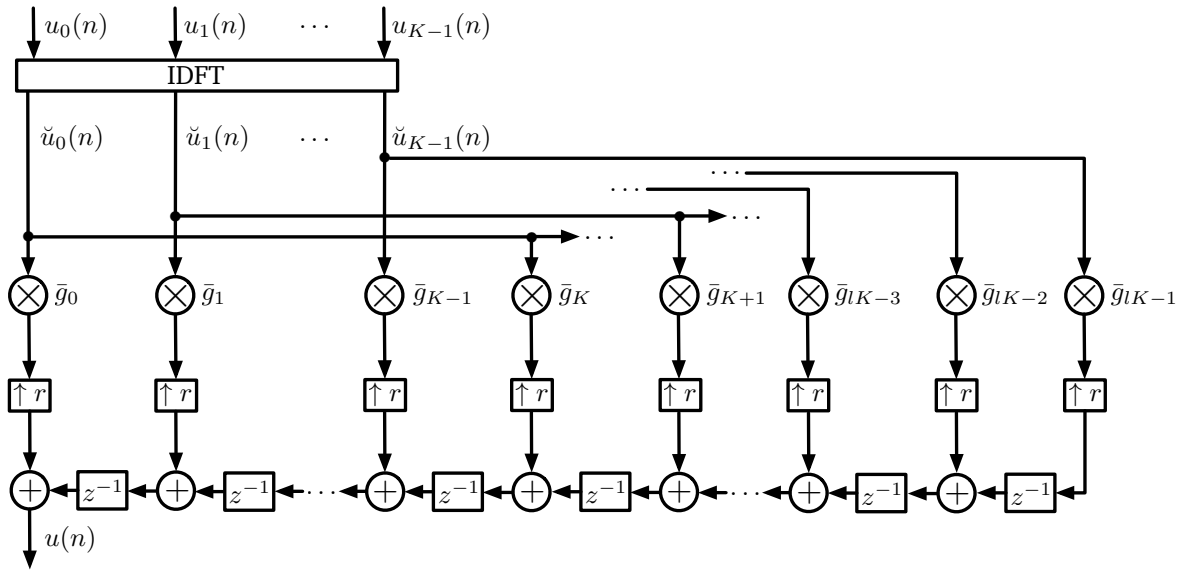


Figure 3.4. Scheme of the synthesis polyphase filter bank.

Note that the resulting sub-band signals are complex valued.
If the baseband signal $y(n)$ and the filter coefficients of the prototype lowpass filter g_0 are real valued, the sub-band signals are conjugate-complex symmetric. Meaning,

$$y_k(n) = y^*_{K-k}(n). \tag{3.9}$$

Thus, if K is an even number, only the sub-bands up to $K/2$ have to be considered. Furthermore, one can assume that the sub-band signal $y_{K/2}(n)$ is sufficiently small and can be neglected. This reduces the effective number of sub-bands to $K/2$.
Finally, the sub-band signals are frequency shifted to the baseband,

$$y_{k,\mathrm{base}}(n) = y_k(n) e^{-j\pi(2r/K+1)kn}, \tag{3.10}$$

to get sub-band signals centered around zero.
The computational complexity is reduced by using the sub-band processing because the lower sampling rate of f_s/r allows to reduce the number of filter coefficients per sub-band by the factor r. Additionally, the calculation frequency is reduced by the same factor.

3.1.2 Evaluation of the Basic System

In this section the suitable number of adaptive filter coefficients N is determined and the sub-band AFC system is compared to the baseband AFC system.
Both evaluations are done by using the added stable gain (ASG). The ASG denotes the increase of the maximum possible gain which could be added by the hearing aid without creating howling. Mathematically, the ASG in dB is given as

$$\mathrm{ASG}_{\mathrm{dB}}(n) = -20\log_{10}\left(\frac{\max_{\Omega}|G(\Omega)(F(\Omega)-\hat{F}(\Omega))|}{\max_{\Omega}|G(\Omega)F(\Omega)|}\right).$$

Here, $G(\Omega)$ is assumed to be constant over all frequencies. This assumption leads to

$$\mathrm{ASG}_{\mathrm{dB}}(n) = -20\log_{10}\left(\frac{|(F(\Omega)-\hat{F}(\Omega))|}{|F(\Omega)|}\right).$$

To determine the suitable number of coefficients N the basic AFC system, Figure 3.1, with a constant step size of $\mu = 2^{-4}$ is applied to obtain stable results. As source signal white noise with the length of about two seconds is used. 600 different recorded feedback paths are used and N is varied from 1 to 10. For each feedback path and choice of N the ASG is calculated over time. Then, the median values of the ASG over the 600 feedback paths are calculated for each N. The results for $N = [1, 5]$ are shown in Figure 3.5. The results for $N = [6, 10]$ are excluded, because they do not supply any additional information w.r.t. $N = [1, 5]$.

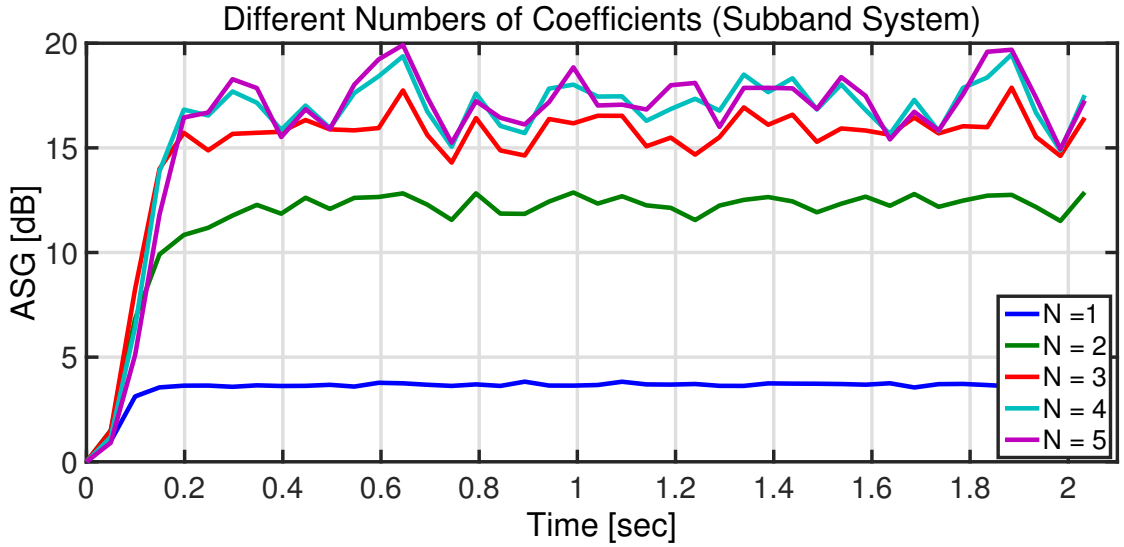


Figure 3.5. Median values of the ASG over 600 feedback paths for $N = [1, 5]$.

Between $N = 1$ and $N = 3$ a significant increase of the ASG can be seen. For $N > 3$

it increases only slightly. Since an increase of the number of coefficients leads to an increase of the computational costs, three coefficients provide the best trade-off between performance and computational costs for the sub-band AFC system.
To compare the sub-band AFC system with the baseband system the same evaluation procedure is used. For both systems a fixed number of coefficients is applied. For the sub-band AFC system $N_{\mathrm{sub-band}}$ is chosen as 3, according to Figure 3.5, and for the baseband system N_{baseband} is set to $r \cdot N_{\mathrm{sub-band}} = 72$ to ensure the same precondition in terms of filter length. The comparison of the two systems according to the median values of the ASG is illustrated in Figure 3.6.

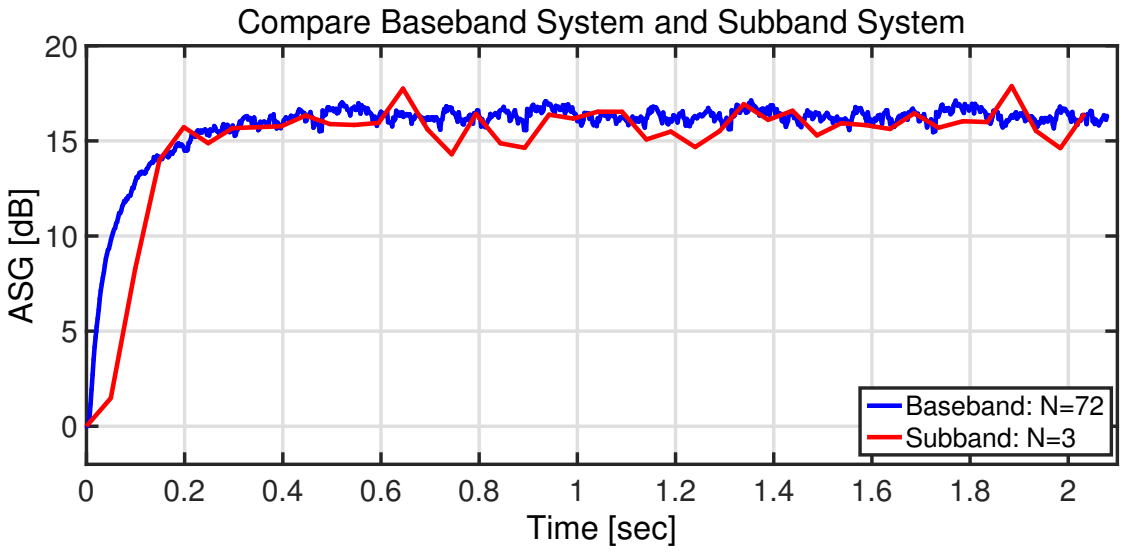


Figure 3.6. Comparison between baseband AFC and sub-band AFC. Depicted are the median values of the ASG over 600 feedback paths for $N_{\mathrm{baseband}} = 72$ and $N_{\mathrm{sub-band}} = 3$.

The performance of both systems in terms of the ASG is similar. The reason is that white noise as an input is used. Since the values of white noise are not correlated which each other, the problem of the bias, described in Section 2.3, does not occur. However, the simulation shows that the sub-band AFC system performs as well as the baseband AFC system under these conditions with less computational costs.

3.1.3 Parameter Setting

The following parameters are chosen for the basic AFC system:

- Sampling frequency $f_s = 24\,\text{kHz}$.
- Number of sub-bands $K = 96$. Resulting in 48 effective sub-bands.
- Downsampling factor $r = 24$. Resulting in a sampling rate of 1 kHz per sub-band.
- Number of adaptive filter coefficients $N = 3$.
- A regularization value equal to 0 dB SPL is added to the norm $||\mathbf{u}_k(n)||^2$.
- AGC: knee point $c_k = 65\,\text{dB SPL}$ and compression ratio $cr = 2$.

3.2 Decorrelation Methods

As stated in Section 2.3 the main problem, which occurs by using an AFC system, is the bias of the adaptation. The amount of this bias is characterized by the cross-correlation of the source signal $x(n)$ and the loudspeaker signal $u(n)$. Therefore, it is obvious that the decorrelation of these signals will reduce the bias of the adaptation. Decorrelation of the signals can be achieved by several approaches [1]:

- Noise injection [4, 39, 50]: A white noise signal is added to the loudspeaker signal. The result can be used as an input signal of the adaptive filter. With the addition of the noise, the correlation between the source signal and the loudspeaker signal is decreased. Therefore, the bias will be reduced. Alternatively, only the noise signal can be used as input of the adaptive filter. By this procedure the bias will be completely eliminated. The drawback is that the convergence of the adaptive filter will be rather slow since the source signal components of the feedback signal are considered as additional disturbance by the adaptive filter. In both cases the sound quality decreases drastically in order to reduce the bias satisfactorily. However, a decrease of the sound quality should be prohibited in hearing aids.
- Forward path delay [6]: A processing delay is added in the forward path. Due to the delay the correlation between the source signal and the loudspeaker signal can be reduced. However, this works only sufficiently if the autocorrelation functions

of the source signals decay rapidly. This is not the case for voiced speech or music. Furthermore, adding a processing delay in the forward path is critical in hearing aids since the overall tolerable delay from microphone to loudspeaker is limited. Otherwise it would become noticeable to the hearing aid user. Hence, additional delays should be avoided if possible.

- Decorrelating prefilters [51–53]: The idea is to reduce the autocorrelation of the microphone and loudspeaker signals. These prefiltered signals are used in the adaptation to estimate the feedback path instead of the original signals. The inverse of the estimated source signal model can be used as such a prefilter. The common choice in AFC systems is the prediction error filter (PEF). For further details on PEF see Section 3.2.1. This approach can effectively reduce the bias and does not influence the loudspeaker signal. Thus the sound quality is not affected and it is highly suited to be used in an AFC system for hearing aids.

- Time-varying processing [1, 14, 15]: A periodically time-varying filter is used as a decorrelation filter in the forward path. A common choice is a frequency shift (FS), which can also be considered as a non-linear operation. In Section 3.2.2 the FS is discussed in detail. The FS can add roughness to the loudspeaker signal. However, this can be minimized by using a small FS, which can still reduce the bias significantly [15].

The PEF and FS are the best suited methods to use in the context of AFC in hearing aids. In the next sections both methods are discussed in detail and their ability to reduce the bias is shown. Additionally, their benefit to the adaptation is evaluated and a combination of both is proposed. Despite of their frequently use in AFC systems the combination of both was not suggested in the literature until [54, 55].

3.2.1 Prediction Error Filter

The prediction error filter (PEF), cf. [51–53], is used to estimate the inverse frequency response of the signal model. The idea behind the signal model is that the source signal $x(n)$ can be interpreted as an output of a linear time-independent (LTI) system $H(\Omega, n)$ with a white signal $w(n)$ as input. This is depicted in Figure 3.7.
If the source signal model $H(\Omega, n)$ is known one can use the inverse source signal model to pre-whiten the microphone $y(n)$ and the loudspeaker signal $u(n)$ and therefore eliminate the correlation between the source signal $x(n)$ and $u(n)$.

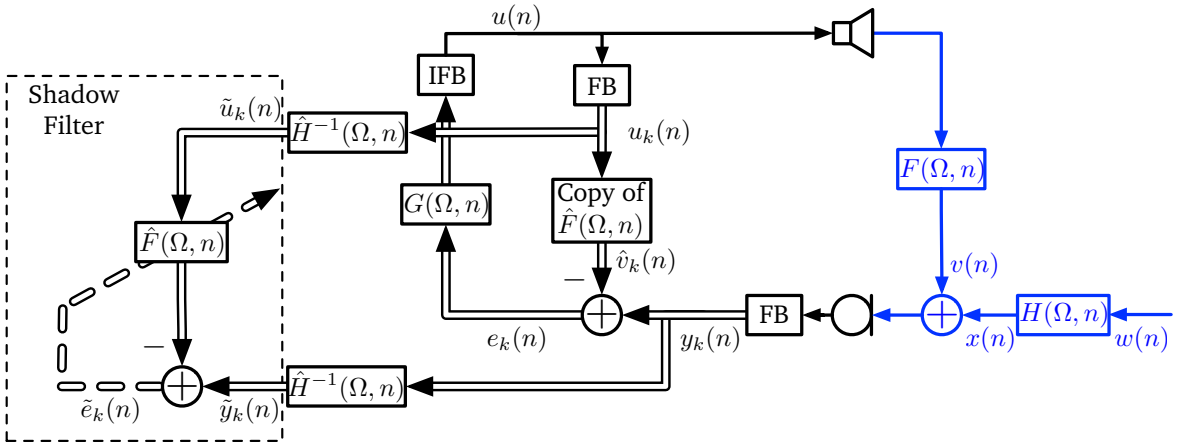


Figure 3.7. Sub-band AFC system with integrated prediction error filter (PEF) $\hat{H}^{-1}(\Omega,n)$. The PEF pre-whitens the signals $y(n)$ and $u(n)$. The pre-whitened signals $\tilde{y}(n)$ and $\tilde{u}(n)$ are used for the adaptation of the filter $\hat{F}(\Omega,n)$ also called shadow filter in this context. The original loudspeaker signal $u(n)$ is filtered with the copy of $\hat{F}(\Omega,n)$ to get the estimated feedback signal $\hat{v}(n)$.

To estimate the inverse source signal model a PEF denoted by $\hat{H}^{-1}(\Omega,n)$ is chosen. The pre-whitened signals $\tilde{y}(n)$ and $\tilde{u}(n)$ are used for the adaptation of the adaptive filter $\hat{F}(\Omega,n)$ which is called shadow filter in this context. The original loudspeaker signal $u(n)$ is filtered with the copy of $\hat{F}(\Omega,n)$ to get the estimated feedback signal $\hat{v}(n)$.

The concept of the PEF is based on the fact that a white signal consists of a series of uncorrelated samples. Due to the delay introduced by the hearing aid processing the output $u(n)$ is at least one tap delayed in relation to the white signal $w(n)$. Consequently, the correlation $r_{wu}(n,l)$ will be 0. By using the PEF the bias of the shadow filter depends on $r_{wu}(n,l)$. Thus the bias in the adaptation of $\hat{F}(\Omega,n)$ is canceled out. The PEF is based on the prediction filter (PF), which estimates the current sample of a signal based on the previous samples of the signal. This is done by an FIR filter structure, which is given by

$$\hat{x}(n) = \sum_{m=1}^{M} \hat{h}_m(n) x(n-m), \tag{3.11}$$

where $\hat{h}_m(n)$ are the estimated coefficients and M the order of the PF. The difference between the true signal $x(n)$ and the estimated signal $\hat{x}(n)$ denotes the output of the PEF.

In the context of AFC a shadow filter for each sub-band is used to calculate the adaptive filter of this sub-band. As shown in Figure 3.7, the inputs of the shadow filters are

pre-whitened by the same PEFs and are given by the following equations

$$\tilde{u}_k(n) = u_k(n) - \hat{\mathbf{h}}_k^{\mathrm{H}}(n)\mathbf{u}_k(n-1) \text{ and} \tag{3.12}$$

$$\tilde{y}_k(n) = y_k(n) - \hat{\mathbf{h}}_k^{\mathrm{H}}(n)\mathbf{y}_k(n-1), \tag{3.13}$$

where $\hat{\mathbf{h}}_k(n)$ denotes the PF for the k-th sub-band. As Spriet et al. [51] proposed the calculation of the PF coefficients is based on $e_k(n)$. Alternatively $y_k(n)$ can be used. But it contains more information of the feedback signal than $e_k(n)$. By calculating the PF coefficients based on $y_k(n)$ not only the correlation of $x_k(n)$ is modeled but also the correlation due to the feedback. Thus, the PEF would model the feedback signal and therefore, the adaptation speed would be reduced. The formula to calculate the optimal PF coefficients is

$$\hat{\mathbf{h}}_k(n) = \mathbf{R}_{e_k e_k}^{-1}(n)\mathbf{r}_{e_k e_k,1}(n). \tag{3.14}$$

Note that $\mathbf{r}_{e_k e_k,1}(n) = [r_{e_k e_k}(n, n-1), r_{e_k e_k}(n, n-2), \ldots, r_{e_k e_k}(n, n-M)]^{\mathrm{T}}$.
Since the PEF is applied in the sub-band setup, it is suitable to use a PEF with $M = 1$ in each sub-band, which reduces the computational complexity significantly. Consequently, the PF coefficient for each PEF can be calculated by the simplified equation

$$\hat{h}_k(n) = \frac{r_{e_k e_k}(n, n-1)}{r_{e_k e_k}(n, n)}. \tag{3.15}$$

The consideration of a higher number of coefficients does not increase the performance significantly. The reduction of the bias and the improved adaptation performance due to the PEF are evaluated in Section 3.2.3. Another proposed decorrelation method, frequency shift, is explained next.

3.2.2 Frequency Shift

Another method to reduce the correlation between $x(n)$ and $u(n)$ is the frequency shift (FS), cf. [1,14,15]. The correlation and therefore the bias in the adaptation is reduced because the FS is a non-linear operation.
An FS can be modeled as a periodically time-varying filter

$$h_{\mathrm{FS}}(n) = e^{j2\pi\frac{f_m}{f_s}n}, \tag{3.16}$$

where f_m is the amount of the FS. The FS is applied on the analytical representation of the signal $d(n)$, given by

$$d_a(n) = d(n) + jd_{\mathrm{H}}(n), \tag{3.17}$$

where $d_{\mathrm{H}}(n)$ denotes the Hilbert transform of $d(n)$. The Hilbert transform is realized by wave digital filters [56]. The output of the FS is then given by

$$u(n) = \mathrm{Re}\{d_a(n) h_{\mathrm{FS}}(n)\}. \tag{3.18}$$

With (3.16) and (3.17) this results in

$$u(n) = d(n)\cos(2\pi f_m n) - d_{\mathrm{H}}(n)\sin(2\pi f_m n). \tag{3.19}$$

The FS is applied to the output of the hearing aid processing $d(n)$ as depicted in Figure 3.8.

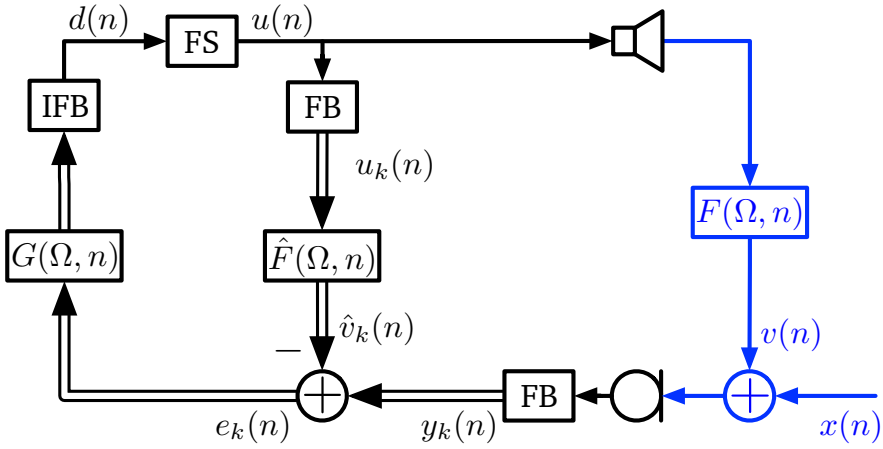


Figure 3.8. Sub-band AFC system with integrated frequency shift (FS). The FS is applied to the output signal of the hearing aid processing $d(n)$.

In contrast to the PEF, the FS is implemented in the baseband domain, but it is only used for frequencies higher than 1 kHz. There are two reasons for that: 1) No adaption is performed by the system at lower frequencies, c.f. Section 3.1, thus there is no need to shift these frequencies. 2) The FS can add roughness to a signal and the relative influence of the FS has a higher impact at lower frequencies. Hence, the negative effects of the FS is more noticeable for low frequencies. Consequently, the FS is not applied at low frequencies. To apply the FS only for frequencies higher than 1 kHz a split band filter is implemented before the actual FS. The FS is then applied to the signal with the high frequency components, while the signal with the low frequency components is not processed.

As mentioned the FS adds roughness to the signals, especially for tonal signals. Therefore, it is advisable to use a small FS, which obviously limits the benefit of the method. In this work, a FS of $f_m = 12$ Hz is proposed, which is a good compromise between reducing the correlation and adding roughness [15].

Note that the adaptation is performed for frequencies higher than 1.5 kHz as stated in

Section 3.1, while the FS is applied for frequencies higher than 1 kHz. The reason is the transition band of the split band filter between the low pass components without frequency shift and the high pass components with frequency shift, which cannot be infinitely small. Hence, a cut-off frequency for the FS of 1 kHz guarantees that all signal components above 1.5 kHz are frequency shifted.
In the next section, the effect on the bias and the adaptation of PEF and FS is shown.

3.2.3 Effect on the Bias and the Adaptation

In this section the effects of the decorrelation methods, PEF and FS, are discussed. Therefore, the influence on the bias of the adaptation and on the adaptation performance when using a constant step size is evaluated. This evaluation is performed for each method and for the combination of the two. It will be shown that a combination of both is highly advisable.

Effect on the Bias

First the influence of the decorrelation methods on the bias is examined. Therefore, an open loop structure of the system, i.e., without feedback path and adaptive filter, is used as depicted in Figure 3.9.

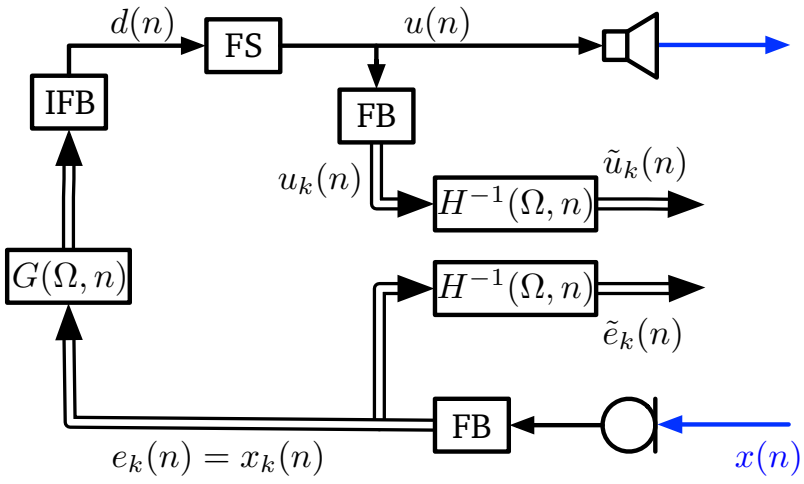


Figure 3.9. Open-loop structure to evaluate the effect of PEF and FS on the bias of the adaptation. $\hat{H}^{-1}(\Omega)$ denotes the PEF.

Using this structure is equivalent to assume a perfect estimation of the feedback path by the adaptive filter, which includes a complete elimination of the feedback signal from the microphone signal.
The source signal consists of ten minutes audio data as introduced in Section 2.4. On the basis of this data the bias is calculated for each time step. The bias vector for a sub-band k is given by

$$\mathbf{b}_k(n) = \mathbf{R}^{-1}_{u_k u_k}(n)\mathbf{r}_{x_k u_k}(n) \tag{3.20}$$

with the length $N = 3$, the derivation is stated in Section 2.1. The autocorrelation matrix and the cross-correlation vector are estimated by

$$\hat{\mathbf{R}}_{u_k u_k}(n) = \alpha \cdot \mathbf{u}_k(n)\mathbf{u}_k^{\mathrm{H}}(n) + (1-\alpha) \cdot \hat{\mathbf{R}}_{u_k u_k}(n-1) \text{ and} \tag{3.21}$$

$$\hat{\mathbf{r}}_{x_k u_k}(n) = \alpha \cdot x_k(n)\mathbf{u}_k^{*}(n) + (1-\alpha) \cdot \hat{\mathbf{r}}_{x_k u_k}(n-1) \tag{3.22}$$

with $\alpha = 0.01$ and $x_k(n) = e_k(n)$. The bias is calculated for four settings:

1. Both decorrelation methods are turned off.
2. Only the FS is used.
3. Only the PEF is used.
4. Both methods are applied.

Note, that if the PEF is applied instead of $u_k(n)$ and $e_k(n)$ the prewhitened signals $\tilde{u}_k(n)$ and $\tilde{e}_k(n)$ are used to calculate the bias.
Figure 3.10 shows the spectrogram of an exemplary music part together with the magnitude of the first coefficient of the bias vector over frequency and time for each setting.

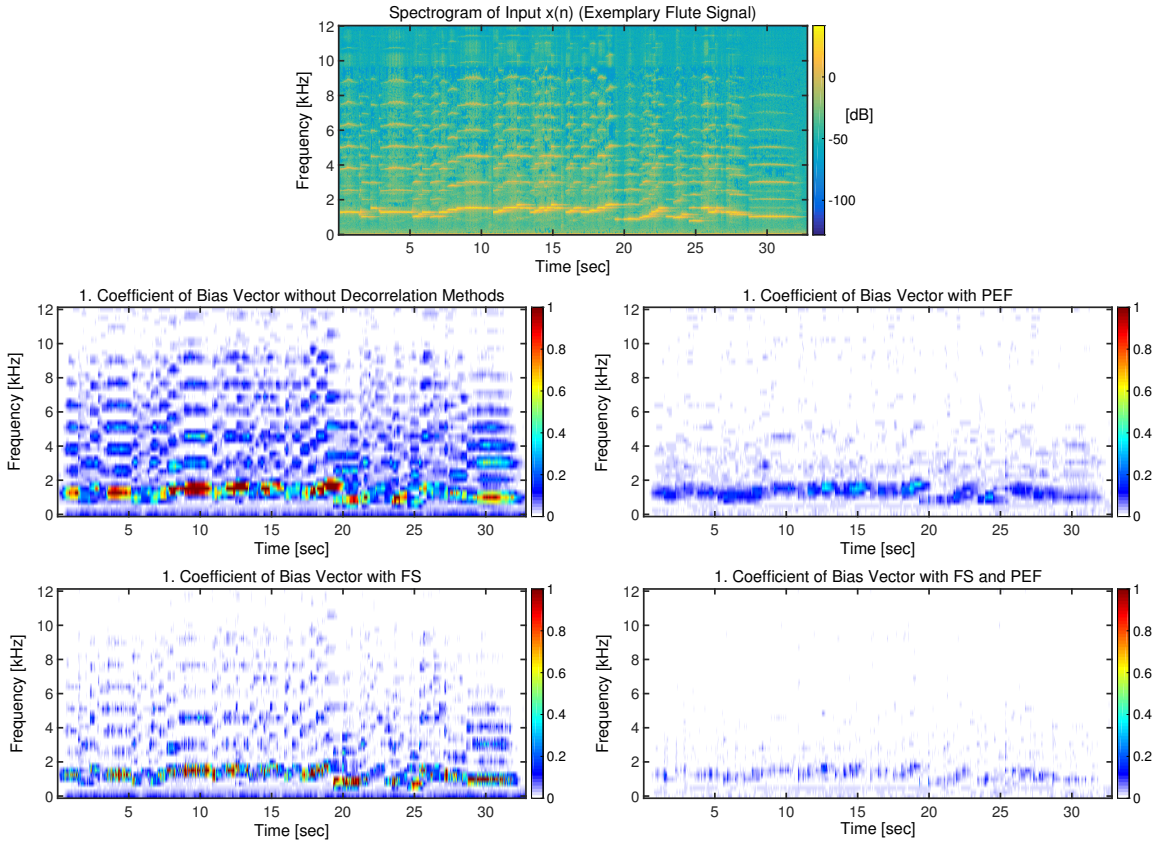


Figure 3.10. Example of the effect of the decorrelation methods on the bias. On the top the spectrogram of an exemplary music part is depicted. The other plots show the magnitude of the first coefficient of the bias vector for each of the 48 sub-bands over frequency and time for each setting.

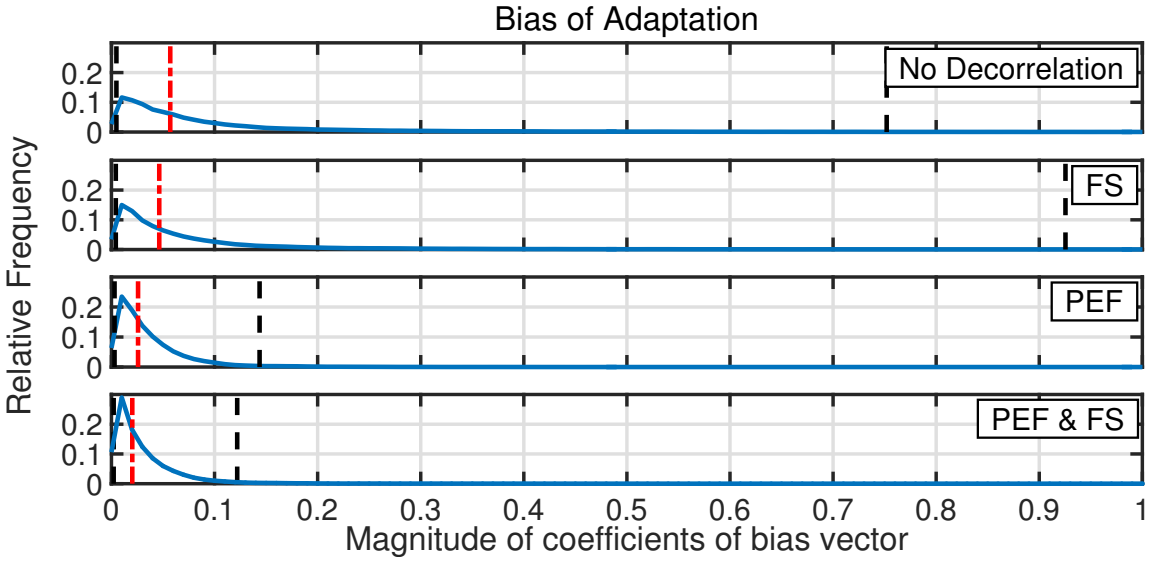


Figure 3.11. The evaluation with PDFs show the effect of the decorrelation methods on the bias. In each figure the combined frequency of the magnitudes of all three coefficients of the bias vector are depicted. The benefit of using the PEF is significant while the use of the FS seems not notable at first. However, in combination the benefit of the FS is clearly seen. This motivates the use of both methods.

It is shown that both PEF and FS reduce the bias. For the shown example, PEF reduces the bias significantly more compared to FS. However, the highest bias reduction is observed when both PEF and FS are used jointly.
For a complete evaluation of the decorrelation methods regarding the bias, a simulation for each setting with five minutes of speech and five minutes of music as source signal are performed. Then the autocorrelation vector of the source signal $x(n)$ is estimated by

$$\hat{\mathbf{r}}_{x_k x_k}(n) = \alpha \cdot x_k(n)\mathbf{x}_k^*(n) + (1-\alpha) \cdot \hat{\mathbf{r}}_{x_k x_k}(n-1) \tag{3.23}$$

with $\alpha = 0.01$. For each simulation the bias vector $\mathbf{b}_k(n)$ is calculated with Equation (3.20). The values higher than 55 dB SPL for each coefficient of $\hat{\mathbf{r}}_{x_k x_k}(n)$ are determined and the corresponding coefficients of $\mathbf{b}_k(n)$ are collected. The magnitudes of these coefficients of the bias vector for each setting are displayed together by PDF estimates. These PDFs are presented in Figure 3.11. As described in Section 2.4 the median and the CI are depicted as well. With the PEF the bias decreases greatly which is noticeable due to the shape of the PDF, the median and especially the upper bound of the CI. The median decreases from 0.06 to 0.02 with the use of the PEF. The upper bound of the CI even from 0.75 to 0.14. In contrast, with the FS the median decreases only slightly, while the upper bound of the CI even increases. Nevertheless, the combination of PEF and FS shows that the additional use of the FS improves the

results significantly. This is most observable due to the shape of the PDF and the upper bound if the CI, which is located at 0.12. In conclusion, the evaluation shows that the combination of PEF and FS clearly reduces the bias and this results in a improved adaptation performance [55]. In the following, this is examined in more detail.

Effect on the Adaptation

Here, the effect of the two decorrelation methods on the adaptation performance is shown. As quality measure the ECLG is used and four different settings are compared:

1. A constant step size (CSS) $\mu_{\text{const}} = 0.1$ is used without any additional methods.
2. The same CSS is applied with the FS.
3. The same CSS is applied with the PEF.
4. The combination of CSS, FS and PEF is used.

First the ECLG is displayed in Figure 3.12 with respect to frequency and time for the example also used in Figure 3.10. For the adaptation without using decorrelation methods a huge amount of entrainment occurs. The entrainment is reduced by using the decorrelation methods as expected.
The complete evaluation with multiple audio signals and feedback paths is illustrated with PDFs as described in Section 2.4. In Figure 3.13 the results are shown.
For the CSS setting without decorrelation methods approximately 79 % of the values are above 0 dB. Thus, during 79 % of the simulation time entrainment occurs. The results of the settings using the decorrelation methods show clearly the benefit which is provided by the decorrelation methods. It is also visible that the PEF has the larger influence in comparison with the FS. With the PEF approximately 42 % of the values are above 0 dB while for the setting with the FS still 68 % are larger than 0 dB. However, the best results are obtained by the combination of both methods (26 % above 0 dB). The evaluations using the ECLG are conform with the analysis regarding the bias.

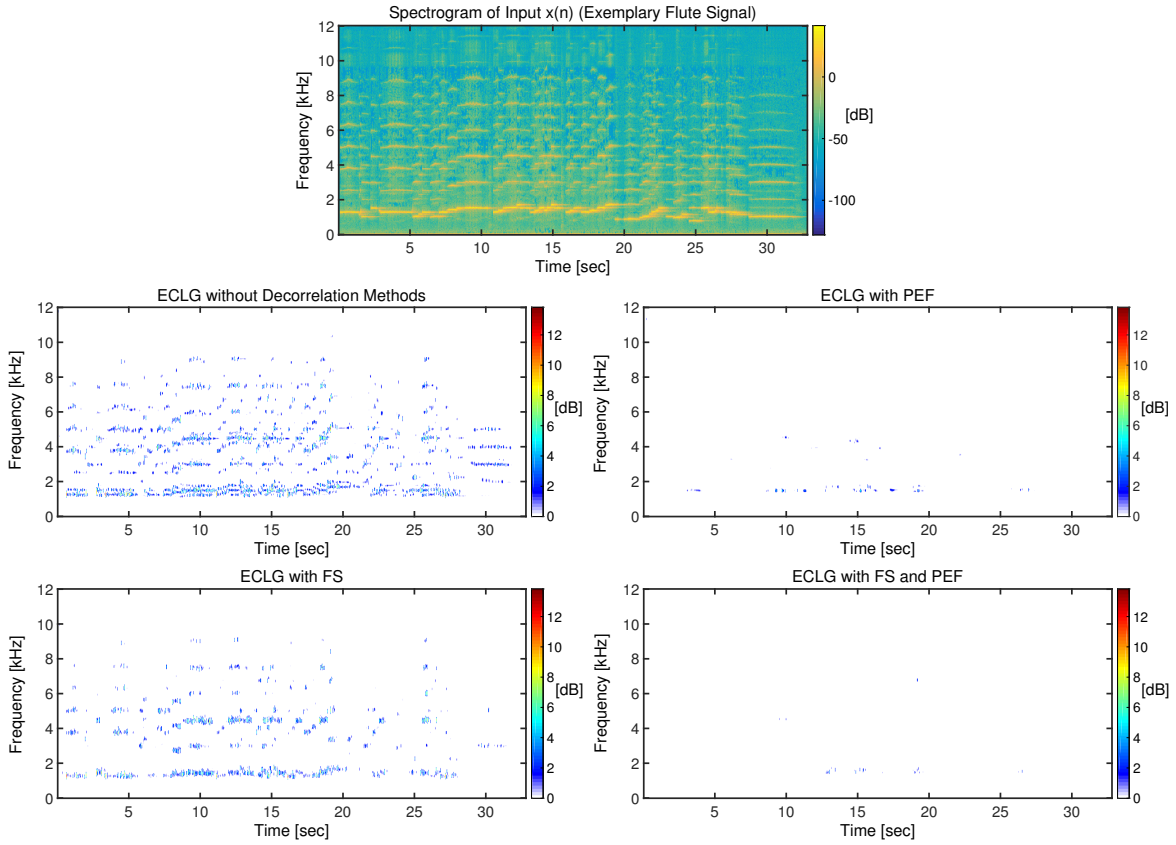


Figure 3.12. Example of the adaptation performance by using decorrelation methods. On the top the spectrogram of an exemplary music part is depicted. The other plots show the effective close loop gain (ECLG) over frequency and time for each adaptation setting.

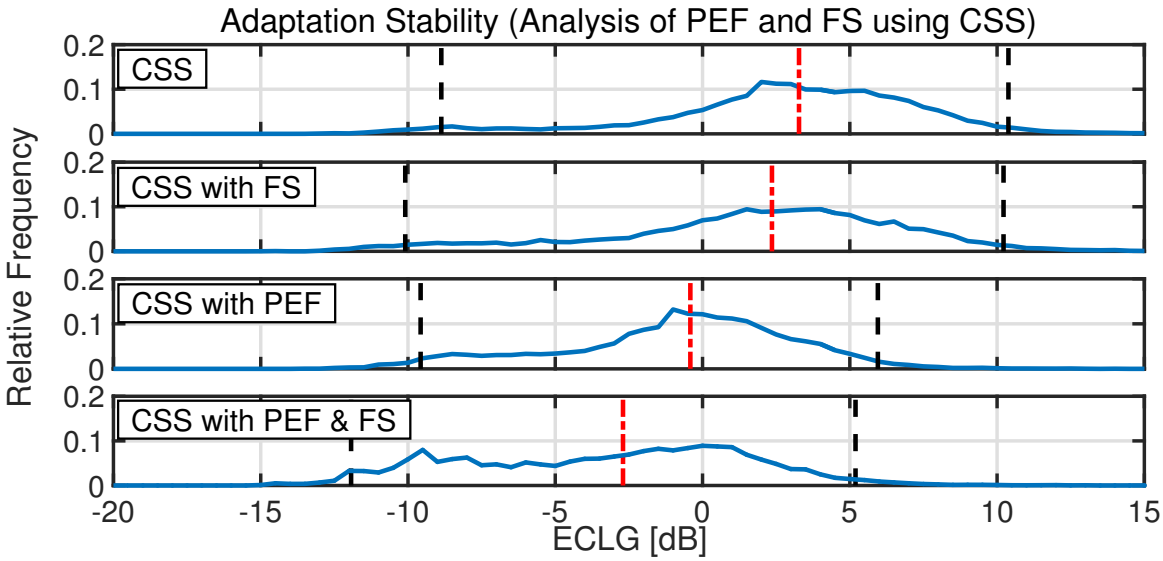


Figure 3.13. Here the effect of the decorrelation methods on the adaptation performance is shown. The quality measure is the ECLG represented by PDFs, c.f. Section 2.4. The first PDF shows the results using a CSS of 0.1. For the second PDF the same CSS with the FS is used. The third PDF displays the result of the CSS with the PEF. For the last PDF the CSS, FS and PEF are combined. The results are similar to the results shown in Figure 3.11 as the influence of the PEF is much higher compared to the FS. However, the result of the combination shows that the use of the FS is beneficial.

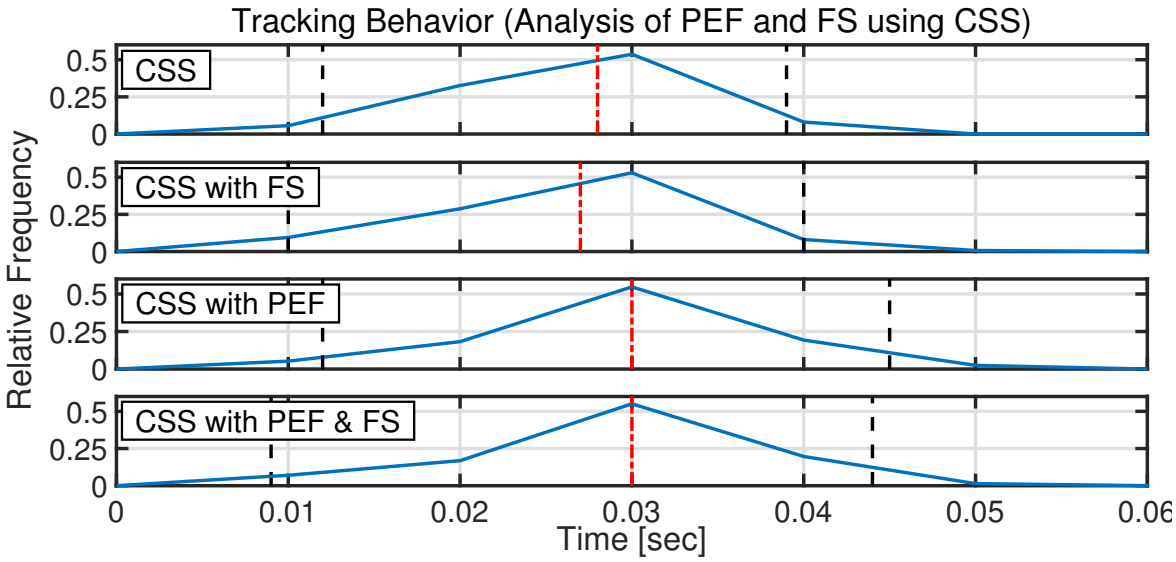


Figure 3.14. Here the effect of the decorrelation methods on the tracking behavior is shown. The adaptation times based on the ECLG are represented by PDFs, c.f. Section 2.4. The first PDF shows the results using a CSS of 0.1. For the second PDF the same CSS with the FS is used. The third PDF displays the result of the CSS with the PEF. For the last PDF the CSS, FS and PEF are combined. The results show that the FS has no significant effect on the tracking behavior while the tracking gets slower with the use of the PEF. This is of course the same for the combination of FS and PEF.

As mentioned in Section 2.4 the system is evaluated not only in terms of the adaptation stability but also based on the tracking ability. In order to evaluate the influence of the decorrelation methods on the tracking ability only white noise as input signal instead of speech and music is used. The reason is that the bad adaptation stability achieved with only the CSS influences the tracking behavior in a negative way. Hence, white noise is used to get a fair comparison between the four settings. Except of the audio signal the evaluation of the tracking behavior is done as described in Section 2.4. The results are depicted in Figure 3.14.

As assumed the FS has no real effect on the tracking behavior. Whereas the PEF has an influence as the system reacts slower to feedback path changes. This effect is explainable by the fact that the PF models the correlation given by the feedback which is present in $e(n)$ during the time the system needs to adapt to a new feedback path additionally to the correlation of the source signal. Consequently, the system needs longer to adapt since this information of the feedback path is deleted by the PEF. As the simulations show the tracking gets slower by 2 ms (median value). Nevertheless, the benefit of using PEF clearly outweighs the disadvantage.

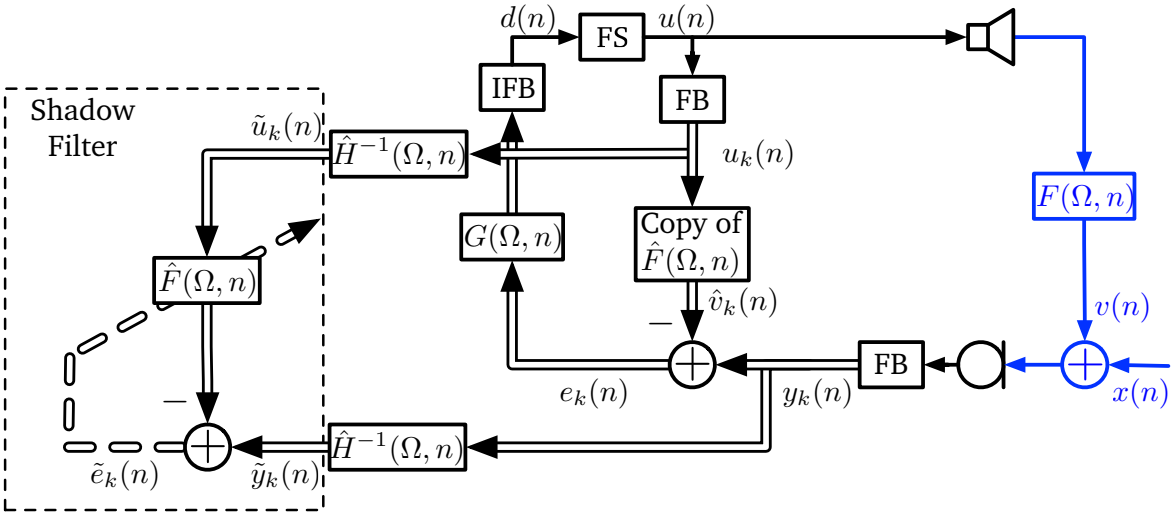


Figure 3.15. Proposed sub-band AFC system with integrated PEF $\hat{H}^{-1}(\Omega)$ and FS.

The combined setting with PEF and FS is depicted in Figure 3.15, which leads to a new NLMS adaptation equation:

$$\hat{\mathbf{f}}_k(n) = \hat{\mathbf{f}}_k(n-1) + \mu_k(n)\frac{\tilde{e}_k^*(n)\tilde{\mathbf{u}}_k(n)}{||\tilde{\mathbf{u}}_k(n)||^2}. \tag{3.24}$$

Even though the benefit of using both decorrelation methods is clearly observable, it is obvious that even for the combined setting many values ($\approx 26\,\%$) of the maxima distributions are above 0 dB. This means that still significant entrainment occurs during the corresponding simulations. Hence, the additional use of a VSS method is recommended. Several VSS methods are introduced and discussed in the following.

3.3 Variable Step Sizes

The step size in the NLMS update equation can be constant or time-varying. For the latter case $\mu(n)$ is called a variable step size (VSS). A VSS is used to control the adaptation as a factor of the update term. It determines how much the filter coefficients of the adaptive filter change for time index n.
In the case of AFC the desired behavior of the VSS is that it approaches zero if the system distance $|f(n) - \hat{f}(n)|$ is small to prevent a false adaptation, which can lead to entrainment. If the system distance is large the VSS should also be large to allow a

fast adaptation. A fast adaptation is important when the feedback path is changing because the faster the system adapts the less howling occurs. However, a fast adapting systems has a high risks of reacting not only to feedback path changes but also to loud source signals if the relation of source signal and feedback signal is high enough. As a consequence, entrainment might occur.
In Section 3.3.1 several step size estimation methods are introduced and analyzed based in their behavior. Also due to the practical implementation the amount of tuneable parameters is considered. Two promising step size estimation methods are introduced in detail and are compared in Sections 3.3.2 and 3.3.3. As a result, it is proposed to utilize the non-parametric variable step size (NPVSS) by Benesty et al. in the presented AFC system. To implement the NPVSS the power of the source signal has to be estimated. Therefore an approach is presented in Section 3.3.3.

3.3.1 Overview

Several VSS methods are introduced, here. Some are general methods for LMS or NLMS algorithms others were derived in the context of AFC or the related acoustic echo cancellation (AEC). All these methods will be analyzed based on their behavior. Therefore, two extreme cases are defined:

1. $|f(n) - \hat{f}(n)| = 0$: The AFC system has adapted perfectly to the feedback path. This results in $e(n) = x(n)$. Consequently, an optimal VSS should tend to 0.
2. $|f(n) - \hat{f}(n)| \rightarrow |f(n)|$: The AFC system has not adapted, yet. Hence, an optimal VSS should be as high as possible.

Another criterion, which will be a factor to find the most suitable step size, is the number of parameters needed to calculate the step size. Since the AFC system should work for a general setup, one would like to use the lowest number of parameters as possible to ensure generality.

Variable Step Sizes

1. VSSLMS by Kwong and Johnston [57]: This step size is derived for the general LMS algorithm without a specific application in mind.
   The VSSLMS by Kwong is calculated by

   $$\mu(n+1) = \alpha\mu(n) + \gamma|e(n)|^2,$$

   with $0 < \alpha < 1$ and $\gamma > 0$. The step size at time index $n+1$ depends on the previous calculated step size at n and the power of the error signal $e(n)$. The authors state that the step size should be upper limited with the condition $\mu_{\max} \leq \frac{2}{3||\mathbf{u}(n)||^2}$. Without this boundary the LMS algorithm would be highly sensitive to the powers of the signals $e(n)$ and $u(n)$.
   Behavior of VSSLMS:

   - $|f(n) - \hat{f}(n)| = 0$:
     Then the VSSLMS can be stated as

     $$\mu(n+1) = \alpha\mu(n) + \gamma|x(n)|^2.$$

     The step size depends on the power of $x(n)$ if the system distance is zero. Hence, for a loud source signal the step size increases, which is not desirable.
   - $|f(n) - \hat{f}(n)| \to |f(n)|$:
     In this case the power of $e(n)$ is large, which will lead to a high step size as intended.

   The VSSLMS depends on the power of $x(n)$ if the AFC system is well adapted. It increases when the power increases. This can lead to misadaptation and in consequence to entrainment.

2. MVSS by Aboulnasr and Mayyas [58]: The MVSS is very similar to the VSSLMS:

   $$\mu(n+1) = \alpha\mu(n) + \gamma r_{ee}(n, n-1)$$

   with $0 < \alpha < 1$, $\gamma > 0$ and $\mu_{\max} \leq \frac{2}{3||\mathbf{u}(n)||^2}$. Instead of the power of $e(n)$ the MVSS uses the autocorrelation $r_{ee}(n, n-1)$. However, the result of the behavior analysis is nearly the same:

   - $|f(n) - \hat{f}(n)| = 0$:
     The error signal is equal to the source signal. Therefore,

     $$\mu(n+1) = \alpha\mu(n) + \gamma r_{xx}(n, n-1).$$

     This results in a high step size if the source signal is strongly correlated.

- $|f(n) - \hat{f}(n)| \rightarrow |f(n)|$:
  Now, the step size will increase since $r_{ee}(n, n-1)$ will be high.

Unlike the VSSLMS, the MVSS uses the autocorrelation $r_{ee}(n, n-1)$. Nevertheless, this leads to the same problem if the source signal is highly correlated which is the case for speech and music signals. Additionally, both step sizes need two parameters, which have to be optimized. Thus, it is not proposed to use the VSSLMS or the MVSS.

3. VS-NLMS of Shin et al. [59]: The VS-NLMS is derived as a general step size for the NLMS algorithm. It is given by

$$\mu(n) = \mu_{\max} \frac{||\mathbf{p}(n)||^2}{||\mathbf{p}(n)||^2 + C},$$
$$\mathbf{p}(n) = \beta \mathbf{p}(n-1) + (1-\beta) \frac{\mathbf{u}(n) e^*(n)}{||\mathbf{u}(n)||^2},$$

with $0 < \mu_{\max} < 2$, $0 < \beta < 1$ and $C > 0$. Proposed values for $C > 0$ are, e.g., 0.005, 0.01 and 0.05. The vector $p(n)$ is the average of the NLMS update vector. The arbitrarily chosen $\mu_{\max}$ is multiplied by the quotient $\frac{||\mathbf{p}(n)||^2}{||\mathbf{p}(n)||^2 + C}$. We analyze this quotient to draw conclusions for the behavior of VS-NLMS:

- $|f(n) - \hat{f}(n)| = 0$:
  $||\mathbf{p}(n)||^2$ should be much smaller than C because $e(n)$ is small. Hence, $\frac{||\mathbf{p}(n)||^2}{||\mathbf{p}(n)||^2 + C} \rightarrow 0$ and $\mu(n) \rightarrow 0$. However, this is only true if the power of $x(n)$ is small in comparison to C. Since for $|f(n) - \hat{f}(n)| = 0$ the error signal $e(n)$ is equal to the source signal $x(n)$ and $u(n)$ also depends on $x(n)$ the VS-NLMS is highly sensitive to loud source signals and the choice of C.
- $|f(n) - \hat{f}(n)| \rightarrow |f(n)|$: The fraction $\frac{||\mathbf{p}(n)||^2}{||\mathbf{p}(n)||^2 + C} \rightarrow 1$ because $||\mathbf{p}(n)||^2$ is much larger than C, supposed C is chosen small enough. This results in the maximum step size, which is desirable.

Due to the sensitivity of the VS-LMS to the relation between the source signal $x(n)$ and the parameter C, this step size is not recommended to be used in an AFC system.

4. GSER by Lee et al. [60]: The GSER is again a generally derived VSS for the NLMS algorithm. It is determined as

$$\mu(n) = \mu_{\max} \frac{\sigma_e^2(n)}{\sigma_e^2(n) + C}$$

with $0 < \mu_{\max} < 2$ and $C > 0$. Obviously, the GSER is similar to the VS-NLMS. Instead of the average update vector it uses the signal power of $e(n)$.

Nevertheless, it has the same disadvantages as the step sizes before because it is very sensitive to the source signal power and the parameter C. Therefore, the GSER is also not proposed to be chosen for an AFC system.

5. VSS-MDNLMS by Rotaru et al. [7]: The VSS-MDNLMS is a VSS that is particularly derived for an AFC system. MDNLMS stands for modified decorrelated NLMS algorithm. It uses pre-whitened signals to calculate the update vector as explained in Section 3.2.1. The proposed VSS is given by

$$\mu(n) = \left| 1 - \frac{\sigma_y(n)}{\sigma_e(n)} \right|.$$

In comparison to the other step sizes, introduced so far, the VSS-MDNLMS originally needs no parameters. However, Rotaru et al. propose to use a $\mu_{\max}$ and a $\mu_{\min}$ to set a range for the step size. They choose $\mu_{\max} = 0.005$ and $\mu_{\min} = 0.0005$, which are very small choices for the step size and will result in a very slow adaptation speed. The behavior analysis of the VSS-MDNLMS shows why the bounds have to be chosen as proposed:

- $|f(n) - \hat{f}(n)| = 0$:
  In this case $y(n) = x(n) + \mathbf{f}^{\mathrm{H}}(n)\mathbf{u}(n)$ and $e(n) = x(n)$. This leads to $\frac{\sigma_y}{\sigma_e} > 1$ or in case of very strong feedback $\frac{\sigma_y}{\sigma_e} \gg 1$. The step size increases as the feedback signal increases even when the system estimates the feedback path perfectly. Hence, the upper bound guarantees that the adaptive filter does not diverge.
- $|f(n) - \hat{f}(n)| \to |f(n)|$:
  Now $e(n) = x(n) + |\mathbf{f}(n) - \hat{\mathbf{f}}(n)|^{\mathrm{H}}\mathbf{u}(n)$ with $|\mathbf{f}(n) - \hat{\mathbf{f}}(n)|$ being very large. Consequently, $\frac{\sigma_y(n)}{\sigma_e(n)} \to 1$ and the step size increases to its maximum. This behavior is intended.

Nevertheless, to get a stable AFC system one has to chose a very small upper bound which results in a very slow adaptation speed. Therefore, the VSS-MDNLMS can not be used in an AFC system with a fast changing feedback path as it is the case for hearing aid applications.

6. OSS by Mader et al. [61]: The OSS is derived to control the adaptation of an echo cancellation system. It is calculated by

$$\mu(n) = \frac{\mathrm{E}\{e_{\mathrm{u}}^2(n)\}}{\mathrm{E}\{e^2(n)\}} = \frac{\sigma_{e_{\mathrm{u}}}^2(n)}{\sigma_e^2(n)},$$

where e_{u} is called the undisturbed error and is given by $e_{\mathrm{u}}(n) = |\mathbf{f}(n) - \hat{\mathbf{f}}(n)|^{\mathrm{H}}\mathbf{u}(n)$. The behavior analysis of the OSS:

- $|f(n) - \hat{f}(n)| = 0$:
  The power of the undisturbed error $\sigma^2_{e_\mathrm{u}}(n)$ is zero while $e(n) = x(n)$. Thus, the step size goes to zero which is the ideal behavior.
- $|f(n) - \hat{f}(n)| \to |f(n)|$:
  In this case $|\mathbf{f}(n) - \hat{\mathbf{f}}(n)|^\mathrm{H}\mathbf{u}(n) \gg x(n)$ and this results in a step size of 1 since $\sigma_{e_\mathrm{u}}(n) \approx \sigma_e(n)$. This is also desirable.

The OSS shows a perfect behavior for an AFC system and further it does not use additional parameters. However, the signal e_u is not accessible in an AFC system because the system distance is not known. Despite some proposed solutions for this problem the OSS is not applicable for the proposed sub-band AFC system. Further comments on that and on the OSS itself are given in Section 3.3.2.

7. NPVSS by Benesty et al. [12]: The non-parametric variable step size (NPVSS) is derived as a general VSS for the NLMS algorithm. The equation to calculate the NPVSS is
$$\mu(n) = 1 - \frac{\sigma_x(n)}{\sigma_e(n)},$$
where $\sigma_{x_k}(n)$ and $\sigma_{e_k}(n)$ denote the standard deviations of $x_k(n)$ and $e_k(n)$, respectively. The results of the behavior analysis are:

   - $|f(n) - \hat{f}(n)| = 0$:
     The error signal $e(n)$ is equal to the source signal $x(n)$. Hence, $\frac{\sigma_x(n)}{\sigma_e(n)} = 1$ and this means that the step size is 0 as desired.
   - $|f(n) - \hat{f}(n)| \to |f(n)|$:
     Here, $\sigma_e(n)$ is much larger than $\sigma_x(n)$. Thus, the step size increases to 1 which is the perfect behavior.

   Such as the OSS, the NPVSS shows an ideal behavior, does not need parameters and is very simple to calculate. However, it depends on a signal which is not accessible in real applications. For the NPVSS it is the power $\sigma^2_x(n)$ of the source signal $x(n)$. In Section 3.3.3 a solution for this problem will be proposed and the NPVSS will be discussed in detail.

8. VSS by Huang and Lee [62]: The VSS by Huang and Lee is a general VSS for the NLMS algorithm without a specific application and is given by
$$\mu(n) = \alpha\mu(n-1) + (1-\alpha)\frac{\sigma^2_e(n)}{\beta\sigma^2_x(n)}$$
with $0 < \alpha < 1$ and $\beta > 0$. It is similar to the NPVSS which results in a similar behavior. However, it uses more parameters than the NPVSS. Despite

this the VSS by Huang and Lee has the same drawback as the NPVSS since it also depends on $\sigma_x^2(n)$. In contrast to Benesty et al., Huang and Lee propose a method to estimate the power of the source signal:

$$\hat{\sigma}_x^2(n) = \hat{\sigma}_e^2(n) - \frac{1}{\sigma_u^2(n)} \hat{\mathbf{r}}_{eu}^{\mathrm{H}}(n) \hat{\mathbf{r}}_{eu}(n). \tag{3.25}$$

One idea is to use this estimation for the NPVSS calculation since the NPVSS needs less parameters and is computationally less demanding. However, in Section 3.3.3 it is shown that the estimate given in Equation (3.25) is not useable for an AFC application since several assumptions within the derivation do not hold [54].

The conclusion of this overview of several approaches for calculating an optimal step size is that the OSS and the NPVSS seem to be applicable for the proposed AFC system. Hence, in the following both approaches are discussed in detail.

3.3.2 Optimal Step Size

The optimal step size (OSS) by Mader et al. [61] was derived in 2000 for acoustic echo cancellation (AEC) which is related to AFC. The OSS is optimal in terms of the maximum decrease of the mean system mismatch. In the following the derivation is shown in detail. Then the drawbacks of the OSS for an AFC system are discussed. These drawbacks lead to the conclusion not to use the OSS for the proposed AFC system.

Theoretical Derivation

In this section the theoretical derivation of the OSS is given. For simplicity, the subband index k is neglected. First the mismatch vector

$$\mathbf{m}(n) = \mathbf{f}(n) - \hat{\mathbf{f}}(n) \tag{3.26}$$

is defined. With the NLMS update equation

$$\hat{\mathbf{f}}(n+1) = \hat{\mathbf{f}}(n) + \mu(n) \frac{e^*(n)\mathbf{u}(n)}{||\mathbf{u}(n)||^2}$$

the mismatch vector can be stated as

$$\mathbf{m}(n+1) = \mathbf{m}(n) - \mu(n) \frac{e^*(n)\mathbf{u}(n)}{||\mathbf{u}(n)||^2}. \tag{3.27}$$

Since with each adaptation step the mismatch should decrease one can define a decreasing cost function as follows

$$\mathrm{E}\{||\mathbf{m}(n+1)||^2\} - \mathrm{E}\{||\mathbf{m}(n)||^2\} \overset{!}{<} 0, \tag{3.28}$$

where $\mathrm{E}\{||\mathbf{m}(n)||^2\}$ is called the system mismatch. With Equation (3.27) and the undisturbed error $e_\mathrm{u}(n) = \mathbf{m}^\mathrm{H}(n)\mathbf{u}(n)$ the cost function is rewritten as

$$\mathrm{E}\{||\mathbf{m}(n+1)||^2\} - \mathrm{E}\{||\mathbf{m}(n)||^2\} \tag{3.29}$$

$$= \mu^2(n)\mathrm{E}\left\{\frac{e^2(n)}{||\mathbf{u}(n)||^2}\right\} - 2\mu(n)\mathrm{E}\left\{\frac{e(n)e_\mathrm{u}(n)}{||\mathbf{u}(n)||^2}\right\} \overset{!}{<} 0. \tag{3.30}$$

This leads to the condition

$$0 < \mu(n) < 2\frac{\mathrm{E}\left\{\frac{e(n)e_\mathrm{u}(n)}{||\mathbf{u}(n)||^2}\right\}}{\mathrm{E}\left\{\frac{e^2(n)}{||\mathbf{u}(n)||^2}\right\}}, \tag{3.31}$$

which the step size has to fulfill in order to guarantee in mean a decreasing system mismatch. The derivation of Equation (3.29) shows that the cost function decreases at most with

$$\mu_\mathrm{opt}(n) = \frac{\mathrm{E}\left\{\frac{e(n)e_\mathrm{u}(n)}{||\mathbf{u}(n)||^2}\right\}}{\mathrm{E}\left\{\frac{e^2(n)}{||\mathbf{u}(n)||^2}\right\}}. \tag{3.32}$$

With $e(n) = e_\mathrm{u}(n) + x(n)$ and the approximation that $||\mathbf{u}(n)||^2$ is constant one gets

$$\mu_\mathrm{opt}(n) = \frac{\mathrm{E}\{e_\mathrm{u}^2(n)\} + \mathrm{E}\{x(n) \cdot e_\mathrm{u}(n)\}}{\mathrm{E}\{e^2(n)\}}. \tag{3.33}$$

Now, it is assumed that $x(n)$ and $e_\mathrm{u}(n)$ are uncorrelated which results in the OSS:

$$\mu_\mathrm{OSS}(n) = \frac{\mathrm{E}\{e_\mathrm{u}^2(n)\}}{\mathrm{E}\{e^2(n)\}}. \tag{3.34}$$

For the implementation in an AFC system two problems have to be solved. The first is the assumption that $x(n)$ and $e_\mathrm{u}(n)$ are uncorrelated. Since $e_\mathrm{u}(n) = \mathbf{u}^\mathrm{H}(n)\mathbf{m}(n)$ this is only true if $x(n)$ and $u(n)$ are uncorrelated which is the case for an AEC system but not for an AFC system. Nevertheless, by using the decorrelation method proposed in Section 3.2 the assumption is reasonable as shown in Section 3.3.4. The other problem is that the OSS depends on the system mismatch unknown in real systems. The proposed method by Mader et al. [61] to estimate the system mismatch can not be used for the proposed AFC system in this thesis. Before the reason for that is discussed the estimation procedure is explained.

Estimation of System Mismatch

The system mismatch $||\mathbf{m}||^2 = ||\mathbf{f}(n) - \hat{\mathbf{f}}(n)||^2$ can not be calculated because the real impulse response $\mathbf{f}(n)$ is unknown in real systems. Mader et al. propose to use an additional artificial delay equal to N_D filter coefficients before the real impulse response. This delay is also modeled by the adaptive filter and the first N_D coefficients of the new $\mathbf{h}(n)$ are known to be zero. Hence, the system mismatch vector coefficients corresponding to the artificial delay are given as

$$m_i(n) = -\hat{h}_i(n),\ i \in [0, \ldots, N_\mathrm{D} - 1].$$

This can be extrapolated to the estimated system distance

$$b_\mathrm{D} = \frac{N}{N_\mathrm{D}} \sum_{i=0}^{N_\mathrm{D}-1} c_i^2(n). \tag{3.35}$$

The proposed method is not suitable for the AFC system proposed in this contribution due to the use of filter banks. Since one sample in each sub-band equals 1 msec and one sample is not representative, one would need multiple milliseconds as delay. However, if the delay becomes too large it is noticeable to the hearing aid user. Hence, an additional delay before the loudspeaker should be avoided. Therefore, despite its very good theoretical behavior, which is also conformed by the simulation results in Section 3.3.4, the OSS can not be implemented because of the unknown system mismatch in this AFC application.

3.3.3 Non-Parametric Variable Step Size

The non-parametric variable step size (NPVSS) was introduced in 2006 by Benesty et al. [12] and is derived as a general VSS for the standard NLMS algorithms. Non-parametric means here that in theory no parameters are needed to calculate the step size. However, for practical applications some parameters are needed as described in Section 3.3.4. In the following, the theoretical derivation of the NPVSS is shown. Then a method is proposed to estimate the power of the feedback-free source signal $x(n)$ [55]. Since Benesty at al. do not give a solution for this problem.

Theoretical Derivation

Here the derivation of the NPVSS is shown. For simplicity, the sub-band index k is neglected in the following. The NPVSS is based on the criterion

$$\mathrm{E}\{|\epsilon(n)|^2\} \stackrel{!}{=} \mathrm{E}\{|x(n)|^2\}\ \forall n, \tag{3.36}$$

where $\epsilon(n)$ is the *a posteriori* error signal. Equation (3.36) is deduced from the desired behavior of a feedback canceler where $\epsilon(n) \stackrel{!}{=} x(n)$ if $\hat{f}(n) = f(n)$. The *a posteriori* error signal is given by

$$\epsilon(n) = y(n) - \hat{\mathbf{f}}^{\mathrm{H}}(n)\mathbf{u}(n). \tag{3.37}$$

With the use of the NLMS update equation,

$$\hat{\mathbf{f}}(n) = \hat{\mathbf{f}}(n-1) + \mu(n)\frac{e^*(n)\mathbf{u}(n)}{||\mathbf{u}(n)||^2}, \tag{3.38}$$

Equation (3.37) can be rewritten as

$$\epsilon(n) = y(n) - \left(\hat{\mathbf{f}}(n-1) + \mu(n)\frac{e^*(n)\mathbf{u}(n)}{||\mathbf{u}(n)||^2}\right)^{\mathrm{H}}\mathbf{u}(n) \tag{3.39}$$

$$= \underbrace{y(n) - \hat{\mathbf{f}}(n-1)^{\mathrm{H}}\mathbf{u}(n)}_{e(n)} - \mu(n)\frac{e(n)\overbrace{\mathbf{u}(n)^{\mathrm{H}}\mathbf{u}(n)}^{||\mathbf{u}(n)||^2}}{||\mathbf{u}(n)||^2} \tag{3.40}$$

$$= e(n)(1-\mu(n)). \tag{3.41}$$

Hence,

$$\mathrm{E}\{|\epsilon(n)|^2\} = \sigma_e^2(n)(1-\mu(n))^2 \tag{3.42}$$

and according to Equation (3.36)

$$\sigma_x^2(n) \stackrel{!}{=} \sigma_e^2(n)(1-\mu(n))^2. \tag{3.43}$$

This can be rewritten as

$$\mu^2(n) - 2\mu(n) + \left(1 - \frac{\sigma_x^2(n)}{\sigma_e^2(n)}\right) = 0, \tag{3.44}$$

which can be solved for $\mu(n)$. Resulting in the NPVSS:

$$\mu_{\mathrm{NPVSS}}(n) = 1 - \frac{\sigma_x(n)}{\sigma_e(n)}. \tag{3.45}$$

The NPVSS can be used for each sub-band separately as follows

$$\mu_{\mathrm{NPVSS},k}(n) = 1 - \frac{\sigma_{x_k}(n)}{\sigma_{e_k}(n)},\ k \in [0,\ldots,K-1]. \tag{3.46}$$

The use of NPVSS, in theory, leads to an optimal behavior of the adaptive filters in all cases including the AFC application [54, 55]. The reasons are that the NPVSS is derived based on Equation (3.36) and during the derivation almost no assumptions are made. The only one is that the power of the loudspeaker signal $u(n)$ stays constant over N samples. N is the length of the adaptive filters. Since in each sub-band $N = 3$ (equals 3 milliseconds) is used, this is a reasonable assumption also for practical applications with non-stationary signals. This statement will be validated with simulation results in Section 3.3.4.
To obtain the optimal behavior one has to assume to know $\sigma_{x_k}(n)$. However, the signal $x_k(n)$ is only accessible in simulations and not in real applications. Furthermore, it is not trivial to estimate $\sigma_{x_k}(n)$ in the AFC context. The problem of estimating $\sigma_{x_k}(n)$ will be addressed in the following.

Estimation of Source Signal Power

In this section, a solution to estimate the power $\sigma_x^2(n)$ without having direct access to $x(n)$ is derived. As before the sub-band index is neglected for the derivation.
At first, an expression for $\Delta\mathbf{f}(n) = \mathbf{f}(n) - \hat{\mathbf{f}}(n)$ is needed without using the in real systems unknown $\mathbf{f}(n)$. Therefore, we use the cross-correlation of $e(n)$ and $u(n)$, which is given as

$$r_{eu}(n, n-l) = \mathrm{E}\{e(n) \cdot u^*(n-l)\}. \tag{3.47}$$

The error signal $e(n)$ can be expressed as

$$e(n) = \Delta v(n) + x(n), \tag{3.48}$$

with $\Delta v(n) = v(n) - \hat{v}(n) = \sum_j \Delta f_j^*(n) u(n-j)$. Thus,

$$e(n) = \sum_j \Delta f_j^*(n) u(n-j) + x(n). \tag{3.49}$$

Hence, $r_{eu}(n, n-l)$ can be noted as

$$\begin{aligned} r_{eu}(n, n-l) =& \mathrm{E}\Bigg\{ \Bigg(\sum_j \Delta f_j^*(n) u(n-j) + x(n) \Bigg) \\ & \cdot u^*(n-l) \Bigg\} \end{aligned} \tag{3.50}$$

$$\begin{aligned} =& \sum_j \Delta f_j^*(n) r_{uu}(n-j, n-l) \\ & + r_{xu}(n, n-l). \end{aligned} \tag{3.51}$$

This can be written in a matrix-vector notation:

$$\mathbf{r}_{eu}(n) = \mathbf{R}_{uu}(n)\Delta\mathbf{f}^*(n) + \mathbf{r}_{xu}(n). \tag{3.52}$$

Equation (3.52) can be solved for $\Delta\mathbf{f}^*(n)$. This leads to the following expression for $\Delta\mathbf{f}(n)^*$,

$$\Delta\mathbf{f}^*(n) = \mathbf{R}_{uu}^{-1}(n)(\mathbf{r}_{eu}(n) - \mathbf{r}_{xu}(n)). \tag{3.53}$$

Now, an estimate for the power of $x(n)$ can be derived. The power of $x(n)$ is defined as

$$\sigma_x^2(n) = r_{xx}(n,n) = \mathrm{E}\{x(n)\cdot x^*(n)\}. \tag{3.54}$$

With $x(n) = e(n) - \Delta v(n)$ one obtains

$$\begin{aligned}
\sigma_x^2(n) =& \mathrm{E}\{(e(n) - \Delta v(n))(e(n) - \Delta v(n))^*\} && (3.55)\\
=& \mathrm{E}\{|e(n)|^2\} + \mathrm{E}\{|\Delta v(n)|^2\} && (3.56)\\
&- \mathrm{E}\{e(n)\cdot\Delta v^*(n)\} - \mathrm{E}\{e^*(n)\cdot\Delta v(n)\}\\
=& \sigma_e^2(n) + \sigma_{\Delta v}^2(n) - (\mathrm{E}\{e(n)\cdot\Delta v^*(n)\} && (3.57)\\
&+ \mathrm{E}\{e^*(n)\cdot\Delta v(n)\})\\
=& \sigma_e^2(n) + \sigma_{\Delta v}^2(n) && (3.58)\\
&- 2\cdot\mathrm{Re}\{\mathrm{E}\{e(n)\cdot\Delta v^*(n)\}\},
\end{aligned}$$

where $\mathrm{Re}\{\cdot\}$ denotes the real component. The signal $e(n)$ is accessible in real applications and can be used directly to estimate $\sigma_e^2(n)$. Since $\Delta v(n)$ is not accessible it is more complicated to estimate $\sigma_{\Delta v}^2(n)$ than $\sigma_e^2(n)$:

$$\begin{aligned}
\sigma_{\Delta v}^2(n) =& \mathrm{E}\{\Delta v(n)\cdot\Delta v^*(n)\} && (3.59)\\
=& \mathrm{E}\Big\{\sum_j\sum_m \Delta f_j^*(n)\Delta f_m(n) && (3.60)\\
&\cdot u(n-j)u^*(n-m)\Big\}\\
=& \sum_j\sum_m \Delta f_j^*(n)\Delta f_m(n) && (3.61)\\
&\cdot r_{uu}(n-j, n-m).
\end{aligned}$$

In matrix-vector notation this results in

$$\sigma_{\Delta v}^2(n) = \Delta\mathbf{f}^{\mathrm{T}}(n)\mathbf{R}_{uu}(n)\Delta\mathbf{f}^*(n). \tag{3.62}$$

Using Equation (3.53), one obtains:

$$\begin{aligned}\sigma^2_{\Delta v}(n) =&(\mathbf{R}^{-1}_{uu}(n))^*(\mathbf{r}_{eu}(n)-\mathbf{r}_{xu}(n))^*)^{\mathrm{T}} \qquad (3.63)\\ &\cdot \mathbf{R}_{uu}(n)(\mathbf{R}^{-1}_{uu}(n)(\mathbf{r}_{eu}(n)-\mathbf{r}_{xu}(n)))\\ =&(\mathbf{r}_{eu}(n)-\mathbf{r}_{xu}(n))^{\mathrm{H}}(\mathbf{R}^{-1}_{uu}(n))^{\mathrm{H}} \qquad (3.64)\\ &\cdot (\mathbf{r}_{eu}(n)-\mathbf{r}_{xu}(n)).\end{aligned}$$

The estimation of the magnitude $\mathrm{E}\{e(n)\cdot\Delta v^*(n)\}$ of Equation (3.58) is derived as follows

$$\begin{aligned}\mathrm{E}\{e(n)\cdot\Delta v^*(n)\} &= \mathrm{E}\{e(n)\cdot\sum_j \Delta f_j(n)u^*(n-j)\} \qquad (3.65)\\ &= \sum_j \Delta f_j(n)\cdot r_{eu}(n,n-j). \qquad (3.66)\end{aligned}$$

In matrix-vector notation and with Equation (3.53):

$$\begin{aligned}\mathrm{E}\{e(n)\cdot\Delta v^*(n)\} &= \Delta\mathbf{f}^{\mathrm{T}}(n)\cdot\mathbf{r}_{eu}(n) \qquad (3.67)\\ &= (\mathbf{r}_{eu}(n)-\mathbf{r}_{xu}(n))^{\mathrm{H}}(\mathbf{R}^{-1}_{uu}(n))^{\mathrm{H}}\mathbf{r}_{eu}(n).\end{aligned}$$

With Equations (3.64) and (3.67) we obtain for Equation (3.58)

$$\begin{aligned}\sigma^2_x(n) =&\sigma^2_e(n)+(\mathbf{r}_{eu}(n)-\mathbf{r}_{xu}(n))^{\mathrm{H}}(\mathbf{R}^{-1}_{uu}(n))^{\mathrm{H}} \qquad (3.68)\\ &\cdot(\mathbf{r}_{eu}(n)-\mathbf{r}_{xu}(n))-2\mathrm{Re}\{(\mathbf{r}_{eu}(n)-\mathbf{r}_{xu}(n))^{\mathrm{H}}\\ &\cdot(\mathbf{R}^{-1}_{uu}(n))^{\mathrm{H}}\mathbf{r}_{eu}(n)\}.\end{aligned}$$

Equation (3.68) provides an optimal estimation of $\sigma^2_x(n)$ and in [55] and in Section 3.3.4 it is shown that the system performs almost identically to the system which uses $x(n)$ to estimate $\sigma^2_x(n)$. Nevertheless, the cross-correlation between $x(n)$ and $u(n)$ is necessary. Since $x(n)$ is not accessible in realistic applications, the cross-correlation vector $\mathbf{r}_{xu}(n)$ can not be calculated easily. However, since the PEF and FS, c.f. Section 3.2, which decorrelate the signals $x(n)$ and $u(n)$, are already applied to the system, the cross-correlation vector $\mathbf{r}_{xu}(n)$ can be omitted from Equation (3.68). Thus, a method which estimates the power of $x_k(n)$ in each sub-band and can be used in real AFC applications is obtained by,

$$\begin{aligned}\hat{\sigma}^2_{x_k}(n) =&\hat{\sigma}^2_{\tilde{e}_k}(n)+\mathbf{r}^{\mathrm{H}}_{\tilde{e}_k\tilde{u}_k}(n)(\mathbf{R}^{-1}_{\tilde{u}_k\tilde{u}_k}(n))^{\mathrm{H}}\mathbf{r}_{\tilde{e}_k\tilde{u}_k}(n) \qquad (3.69)\\ &-2\mathrm{Re}\{(\mathbf{r}^{\mathrm{H}}_{\tilde{e}_k\tilde{u}_k}(n)(\mathbf{R}^{-1}_{\tilde{u}_k\tilde{u}_k}(n))^{\mathrm{H}}\mathbf{r}_{\tilde{e}_k\tilde{u}_k}(n)\}\end{aligned}$$

according to Figure 3.15. By using Equation (3.69) the NPVSS is applicable in the proposed AFC system [55].

3.3.4 Evaluation of Variable Step Sizes

In this section the OSS and NPVSS are evaluated by simulations. Therefore the ECLG is expressed using PDFs, c.f. Section 2.4, and additionally the median and CI are depicted. First, it is shown that the AFC system with the OSS only performs satisfactorily if the PEF and FS are additionally used [54]. The reason for this is the assumption made during the derivation of the OSS as mentioned in Section 3.3.2. For the simulation the OSS is applied for each sub-band k as

$$\mu_{\mathrm{OSS},k}(n) = a \cdot \left(\frac{\hat{\sigma}^2_{e_{\mathrm{u}},k}(n)}{\hat{\sigma}^2_{e,k}(n)} \right), \tag{3.70}$$

where a is called attenuation parameter, which is used to increase the stability of the system. The parameter a is chosen in simulations equal to 0.25. The power values are estimated by

$$\begin{aligned} \hat{\sigma}^2_{e_{\mathrm{u}},k}(n) &= \alpha_1 \cdot |e_{\mathrm{u},k}(n)|^2 + (1-\alpha_1) \cdot \hat{\sigma}^2_{e_{\mathrm{u}},k}(n-1), \\ \hat{\sigma}^2_{e_k}(n) &= \alpha_1 \cdot |e(n)|^2 + (1-\alpha_1) \cdot \hat{\sigma}^2_{e,k}(n-1) \end{aligned}$$

with $\alpha_1 = 0.01$. The undisturbed error signal $e_{\mathrm{u},k}(n)$ is calculated by

$$e_{\mathrm{u},k} = |\mathbf{f}_k(n) - \hat{\mathbf{f}}_k(n)|^{\mathrm{H}} \mathbf{u}_k(n),$$

which is only possible in simulations and not in real applications.
In Figure 3.16 two PDFs depicting the ECLG for two settings are shown. For the first setting only the OSS is used to control the adaptation. For the second setting the PEF and FS are also applied.
Clearly the system which only uses the OSS is not stable which can be observed since much entrainment occurs. By adding the decorrelation methods, the performance increases significantly. However, due to the lack of an estimation method for the system mismatch this setting is not applicable in a real system.

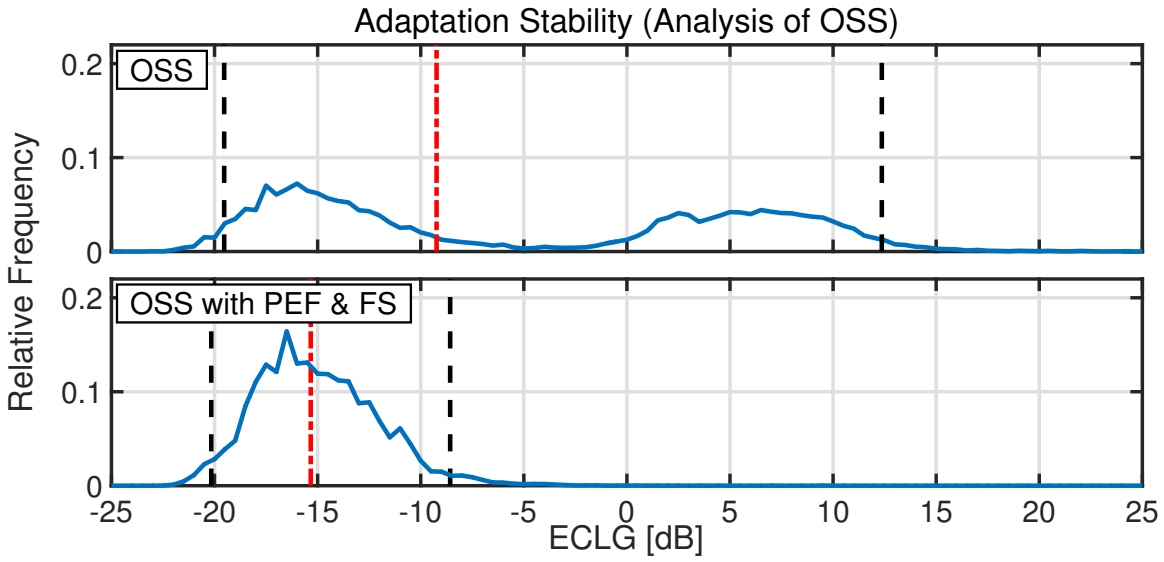


Figure 3.16. The two PDFs depict the ECLG values for the AFC system only using the OSS (upper plot) and using PEF and FS additionally (lower plot). The system using the OSS without modifications is not stable since much entrainment occurs. By adding the decorrelation methods the performance increases significantly.

Now, the NPVSS is evaluated by the same simulation method. The NPVSS for subband k is given by

$$\mu_{\text{NPVSS},k}(n) = a \cdot \left(1 - \frac{\hat{\sigma}_{x_k}(n)}{\hat{\sigma}_{e_k}(n)}\right) \tag{3.71}$$

with the same attenuation parameter a as mentioned in Equation (3.70). For the simulation the NPVSS is directly used by assuming to have access to the decomposed source signal $x_k(n)$. Hence, the power values $\sigma^2_{x_k}(n)$ and $\sigma^2_{e_k}(n)$ can be estimated by

$$\begin{aligned}\hat{\sigma}^2_{\text{true},x_k}(n) &= \alpha_1 \cdot |x_k(n)|^2 + (1-\alpha_1) \cdot \hat{\sigma}^2_{x_k}(n-1),\\ \hat{\sigma}^2_{e_k}(n) &= \alpha_1 \cdot |e_k(n)|^2 + (1-\alpha_1) \cdot \hat{\sigma}^2_{e_k}(n-1).\end{aligned}$$

The smoothing parameter α_1 is set to 0.01 again. In Figure 3.17 it is shown that the AFC system performs very well by using the NPVSS as no values of the ECLG exceeds 0 dB. Consequently, no entrainment occurs. This proves the statement that by using the NPVSS without additional decorrelation method in the AFC system one can reduce the feedback and no artifacts are generated [54, 55]. This is a novel discovery.

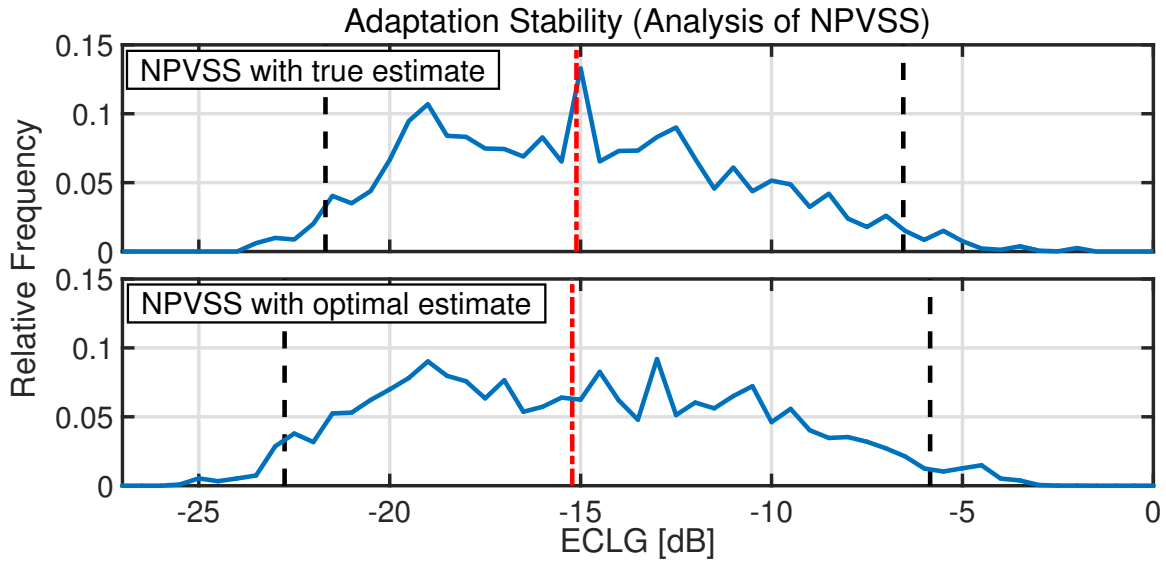


Figure 3.17. The ECLG values for the AFC system with NPVSS assuming to have access to $x(n)$ (upper plot) and with the NPVSS using the optimal estimate of $\sigma_x^2(n)$ (lower plot). For both cases no entrainment occurs as all values are below 0 dB.

Nevertheless, for the calculation of the NPVSS the power of $x_k(n)$, which is not accessible in real systems, is needed. Therefore, in Section 3.3.3 a method to estimate $\sigma_{x_k}^2(n)$ is proposed. The optimal estimate of $\sigma_{x_k}^2$ is given by

$$\begin{aligned}\hat{\sigma}_{\text{opt},x_k}^2(n) =& \hat{\sigma}_{e_k}^2(n) + (\hat{\mathbf{r}}_{e_k u_k}(n) - \hat{\mathbf{r}}_{x_k u_k}(n))^{\mathrm{H}} (\hat{\mathbf{R}}_{u_k u_k}^{-1}(n))^{\mathrm{H}} \\ & \cdot (\hat{\mathbf{r}}_{e_k u_k}(n) - \hat{\mathbf{r}}_{x_k u_k}(n)) - 2\mathrm{Re}\{(\hat{\mathbf{r}}_{e_k u_k}(n) - \hat{\mathbf{r}}_{x_k u_k}(n))^{\mathrm{H}} \\ & \cdot (\hat{\mathbf{R}}_{u_k u_k}^{-1}(n))^{\mathrm{H}} \hat{\mathbf{r}}_{e_k u_k}(n)\}.\end{aligned} \tag{3.72}$$

The magnitudes can be estimated by

$$\begin{aligned}\hat{\sigma}_{e_k}^2(n) &= \alpha_1 \cdot |e_k(n)|^2 + (1-\alpha_1) \cdot \hat{\sigma}_{e_k}^2(n-1), \\ \hat{\mathbf{R}}_{u_k u_k}(n) &= \alpha_1 \cdot \mathbf{u}_k(n)\mathbf{u}_k^{\mathrm{H}}(n) + (1-\alpha_1) \cdot \hat{\mathbf{R}}_{u_k u_k}(n-1), \\ \hat{\mathbf{r}}_{e_k u_k}(n) &= \alpha_1 \cdot e_k(n)\mathbf{u}_k^{*}(n) + (1-\alpha_1) \cdot \hat{\mathbf{r}}_{e_k u_k}(n-1), \\ \hat{\mathbf{r}}_{x_k u_k}(n) &= \alpha_1 \cdot x_k(n)\mathbf{u}_k^{*}(n) + (1-\alpha_1) \cdot \hat{\mathbf{r}}_{x_k u_k}(n-1).\end{aligned}$$

The result in Figure 3.17 shows that by using the optimal estimate the system performs as well as with the true estimate. However, both settings can only be used in theory, because the first assumes that the signal $x(n)$ is known and the second that the correlation between $x(n)$ and $u(n)$ is known. To get a system which works under authentic conditions the two decorrelation methods, PEF and FS, have to be additionally used [55]. By this, the magnitude $\hat{\mathbf{r}}_{x_k u_k}(n)$ can be neglected in Equation (3.72),

resulting in

$$\begin{aligned}\hat{\sigma}^2_{x_k}(n) =& \hat{\sigma}^2_{\tilde{e}_k}(n) + \mathbf{r}^{\mathrm{H}}_{\tilde{e}_k\tilde{u}_k}(n)(\mathbf{R}^{-1}_{\tilde{u}_k\tilde{u}_k}(n))^{\mathrm{H}}\mathbf{r}_{\tilde{e}_k\tilde{u}_k}(n) \\ & - 2\mathrm{Re}\{(\mathbf{r}^{\mathrm{H}}_{\tilde{e}_k\tilde{u}_k}(n)(\mathbf{R}^{-1}_{\tilde{u}_k\tilde{u}_k}(n))^{\mathrm{H}}\mathbf{r}_{\tilde{e}_k\tilde{u}_k}(n)\},\end{aligned} \tag{3.73}$$

which is applicable in real systems and not only in simulations. The magnitudes are estimated as stated above. Before evaluating the overall performance of the NPVSS with the proposed estimate, the advantage of the proposed estimation procedure in comparison to the estimate proposed by Huang and Lee given as

$$\hat{\sigma}^2_{\mathrm{Huang},x_k}(n) = \hat{\sigma}^2_{\tilde{e}_k}(n) - \frac{1}{\sigma^2_{\tilde{u}_k}(n)}\hat{\mathbf{r}}^{\mathrm{H}}_{\tilde{e}_k\tilde{u}_k}(n)\hat{\mathbf{r}}_{\tilde{e}_k\tilde{u}_k}(n) \tag{3.74}$$

is examined. The proposed estimate in Equation (3.73) is more general than the estimate by Huang and Lee. To achieve the result of Equation (3.74) one has to assume that the signals $x_k(n)$ and $\Delta v_k(n)$ are uncorrelated and that the loudspeaker signal $u_k(n)$ is white. Both assumption do not hold in case of a hearing aid application. Hence, the estimate proposed by Huang and Lee is not applicable to an AFC system for hearing aids [54]. To confirm this statement an exemplary simulation result is shown. Therefore, a bell ringing of approximately 12 seconds and the feedback path $F_1(\Omega)$, Figure 3.18, are used. The hearing aid gain is such that $\max|F_1(\Omega)G(\Omega)| \approx 5\,\mathrm{dB}$.

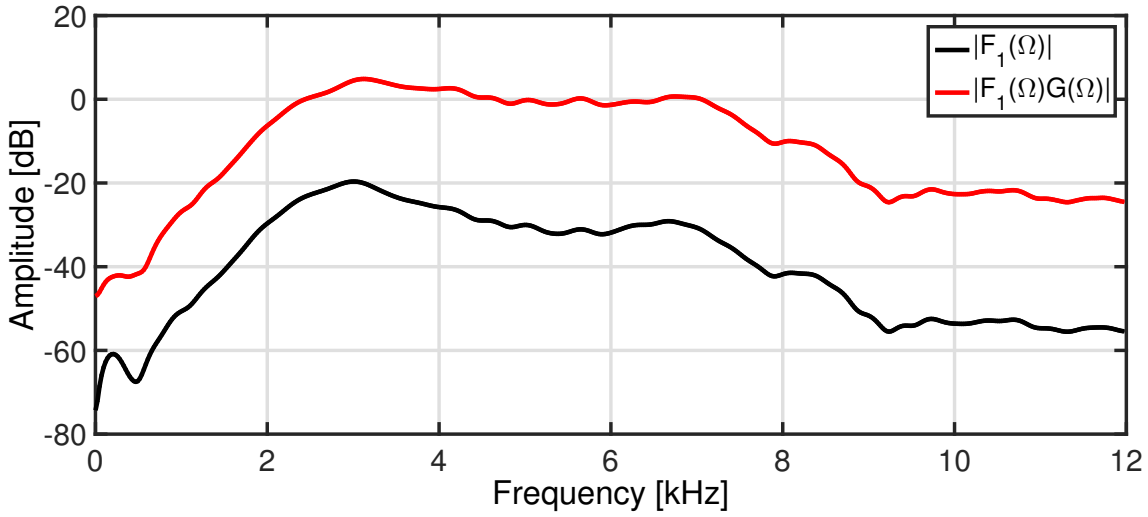


Figure 3.18. Feedback path $F_1(\Omega)$ with $G(\Omega)$. The hearing aid gain is chosen that $\max|F_1(\Omega)G(\Omega)| \approx 5\,\mathrm{dB}$.

No adaptation during the simulation is performed. Figure 3.19 shows the results of three different estimation methods for the source signal power $\sigma^2_{x_k}(n)$:

1. Power is estimated by smoothing $|\tilde{x}_k(n)|^2$ (requires direct access to $\tilde{x}_k(n)$; only accessible in simulations).
2. Power is estimated by Equation (3.74).
3. Power is estimated by the proposed new method, c.f. Equation (3.73).

Note, to calculate the power with the first method the original source signal $x_k(n)$ is processed by the same PEF as $u_k(n)$ and $e_k(n)$ to obtain $\tilde{x}_k(n)$. This is necessary for a comparison between all three estimation methods since the other two rely on signals which are pre-whitened by the PEF. The estimate by Huang and Lee strongly overestimates the power at the frequencies with strong feedback, while the proposed method shows a significant better estimate.

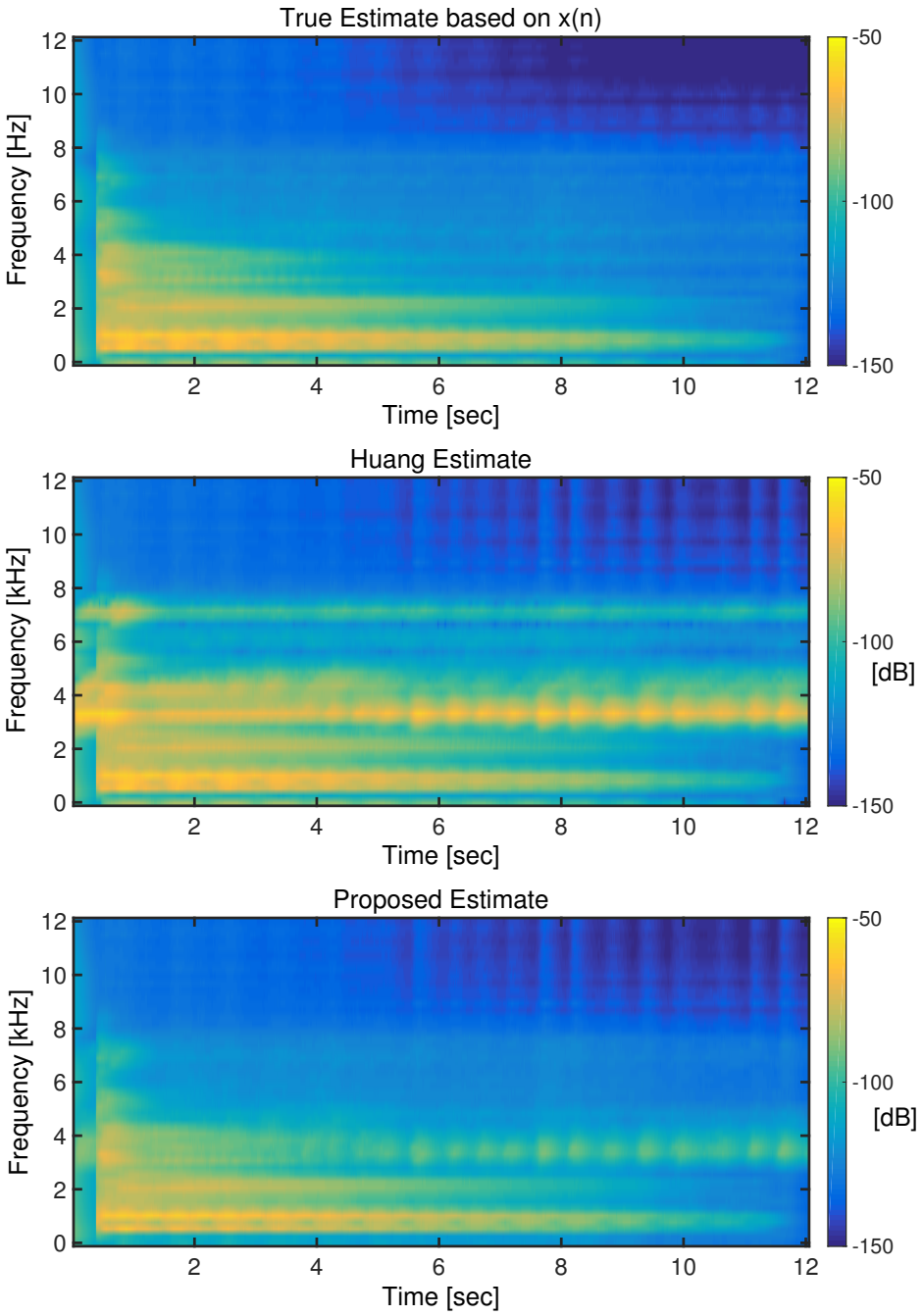


Figure 3.19. The figure depicts three different estimates of the source signal power: Power is estimated by smoothing $|x_k(n)|^2$ (upper plot) for reference, power is estimated by Equation (3.74) (middle plot) and power is estimated by the proposed method, c.f. Equation (3.73) (lower plot). The estimate by Huang and Lee strongly overestimates at the frequencies with strong feedback, while the proposed method shows a significant better estimate since it is close to the true reference of the upper plot.

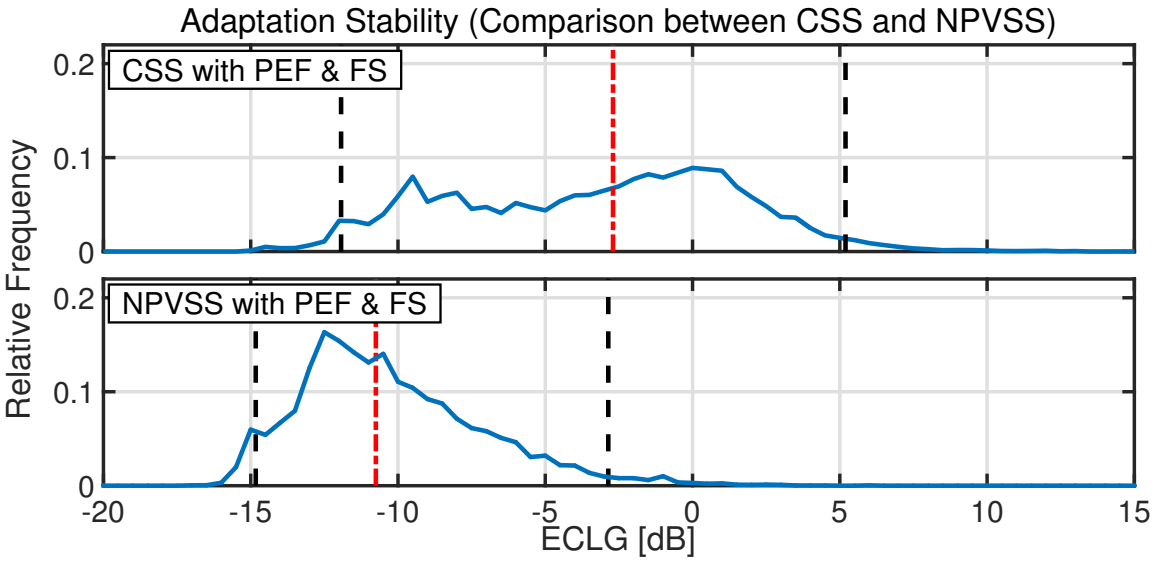


Figure 3.20. The ECLG values for the AFC system with CSS $\mu_{\text{CSS}} = 0.1$ (upper plot) and NPVSS using the proposed estimate (lower plot). For both settings the decorrelation methods, PEF and FS, are additionally used. The benefit of using the NPVSS is clearly observable.

Finally, it should be shown that the AFC system using the NPVSS with the proposed estimate and the decorrelation methods additionally performs satisfactorily. Therefore, Figure 3.20 depicts the result for the AFC system with NPVSS, PEF, FS and using the proposed estimate for $\sigma^2_{x_k}(n)$. For comparison the result for the system with CSS ($\mu_{\text{CSS}} = 0.1$), PEF and FS is also shown.
The benefit of using this applicable version of NPVSS is clearly observable. The median for example is reduced by 8.5 dB and by using the NPVSS only 0.9 % of the values are above 0 dB compared to 29 % when using a CSS of 0.1.

3.3.5 Conclusion: Proposed Step Size Control

Based on the analyses and evaluations in the previous sections the NPVSS with the addition of the PEF and FS is proposed to control the adaptation. The NPVSS shows in theory an optimal behavior even without decorrelation methods. In order to make the NPVSS applicable for a realistic AFC system an estimate of the source signal power is derived. This estimate is only reliable in combination with the previously proposed and already implemented decorrelation methods, PEF and FS. Hence, the decorrelation methods do not only decrease the adaptation bias but also ensure that the NPVSS performs as intended.

The AFC system controlled by the combination of NPVSS, PEF, and FS performs very well as depicted by evaluations above and in Chapter 5.

3.4 Additional Control Methods

In this section, three additional control methods are presented. The first two are important when dealing with realistic signals and conditions. The last one is a new approach to measure the amount of correlation of the source signal. The latter one is used to control the adaptation additionally with the benefit of increasing the tracking ability of the adaptation without decreasing the stability.

3.4.1 Impulse Detection

The first additional control method is called impulse detection. It is used to detect wideband impulses of the source signal, e.g., clapping or knocking. Problems occur because the update term in Equation (3.1), given as $\frac{e_k^*(n)\mathbf{u}_k(n)}{||\mathbf{u}_k(n)||^2}$, increases strongly when impulses occur. The reason is the time delay between $e_k(n)$ and $u_k(n)$. For sudden power increases due to impulses the power of $e_k(n)$ increases rapidly while the norm $||\mathbf{u}_k(n)||^2$ lags behind and can not compensate the power increase of the error signal immediately. Thus, the NLMS algorithm reacts to impulses. This leads to misadaptations which can create entrainment. The solution is to detect these impulses and freeze the adaptation when an impulse is detected. To freeze the adaptation the step sizes for all sub-bands are set to 0. It is important to detect only wideband impulses since otherwise one would also detect howling, which occurs if the feedback path changes suddenly for narrowband frequency regions. Thus, a freezing of the adaptation in these situations would result in extended howling.
The concept to detect impulses is straight forward. The loudspeaker's sub-band signals $y_k(n)$, $k \in [0, \ldots, K-1]$, are smoothed in two different ways. A fast smoothing, which follows the signal power fast:

$$\overline{y}_{\text{fast},k}(n) = \begin{cases} \beta_r \cdot |y_k(n)|^2 + (1-\beta_r) \cdot \overline{y}_{\text{fast},k}(n-1), \ \overline{y}_{\text{fast},k}(n-1) \geq y_k(n) \\ \beta_f \cdot |y_k(n)|^2 + (1-\beta_f) \cdot \overline{y}_{\text{fast},k}(n-1), \ \overline{y}_{\text{fast},k}(n-1) < y_k(n) \end{cases} . \quad (3.75)$$

Here, two smoothing parameters β_r in case of increasing power and β_f in case of decreasing power are used. The second smoothing is declared as slow smoothing, meaning the smoothed value follows changes of the power slowly:

$$\overline{y}_{\text{slow},k}(n) = \gamma \cdot |y_k(n)|^2 + (1-\gamma) \cdot \overline{y}_{\text{slow},k}(n-1). \quad (3.76)$$

This time only one smoothing parameter is used. Evaluations show that this is sufficient. With the two smoothed power values an indicator function $I_k(n)$ is calculated for each sub-band by

$$I_k(n) = 10 \log_{10} \left(\frac{\overline{y}_{\text{fast},k}(n)}{\overline{y}_{\text{slow},k}(n)} \right). \tag{3.77}$$

This indicator function is then averaged over all sub-bands resulting in

$$I_{\text{comp}}(n) = \frac{1}{K} \sum_{k=0}^{K-1} I_k(n). \tag{3.78}$$

By taking the average over the sub-bands one guarantees that only wideband impulses are detected. If $I_{\text{comp}}(n)$ exceeds an arbitrarily chosen threshold the step sizes of all sub-bands are set to 0 to freeze the adaptation. It is crucial to use an average indicator function over all sub-bands and therefore, to detect only wideband impulses, as howling increases the power in some sub-bands. Hence, howling would be interpreted as an impulse in these sub-bands and the freezing of the adaptation would lead to much longer howling.
Three constraints for the smoothing parameters of Equations (3.75) and (3.76) are defined:

1. $\beta_r > \gamma$. This constraint ensures that $\overline{y}_{\text{fast},k}(n)$ increases faster than $\overline{y}_{\text{slow},k}(n)$ if impulses occur.

2. $\beta_f \leq \gamma$. This constraint results in a slower (or equal) decrease of $\overline{y}_{\text{fast},k}(n)$ compared with $\overline{y}_{\text{slow},k}(n)$. Hence, it is guaranteed that $I_k(n)$ does not decrease too fast. This is crucial for repeated impulses.

3. $\beta_r > \beta_f$. This is the logical conclusion of the previous two constraints.

Optimal Smoothing Parameters

The optimal smoothing parameters are found by simulations. Therefore, audio data consisting of speech and music are used. Additionally real howling and impulses, like clapping and knocking, are recorded in the audio laboratory. The impulses are tagged by "hand", the rest of the audio data is labeled as non-impulses. Based on the complete audio data the indicator function $I_{\text{comp}}(n)$ is calculated. The smoothing parameters are varied with respect to the constraints given above. The resulting values of the indicator function $I_{\text{comp}}(n)$ are splitted into two groups, impulses and non-impulses, based on the labels of the input data. By using histograms the PDFs for both groups,

$p_{\text{impulses}}(I_{\text{comp}})$ and $p_{\text{no-impulses}}(I_{\text{comp}})$, are estimated. The threshold T_{ID} is set to the first intersection of the PDFs. Then the error probability is calculated. The error probability is given by

$$p_{\text{error}} = \sum_{I_{\text{comp}}=-\infty}^{T_{\text{ID}}} (p_{\text{impulses}}(I_{\text{comp}})) + \sum_{I_{\text{comp}}=-T_{\text{ID}}}^{\infty} (p_{\text{no-impulses}}(I_{\text{comp}})). \tag{3.79}$$

This is done for each parameter setting and the optimal parameter setting is found minimizing the error probability, which is about 6%. The corresponding parameters are: $\beta_r = 0.05$, $\beta_f = 0.005$, and $\gamma = 0.005$. The threshold T_{ID} is set to 4 dB. Figure 3.21 shows the PDFs resulting with these parameters.

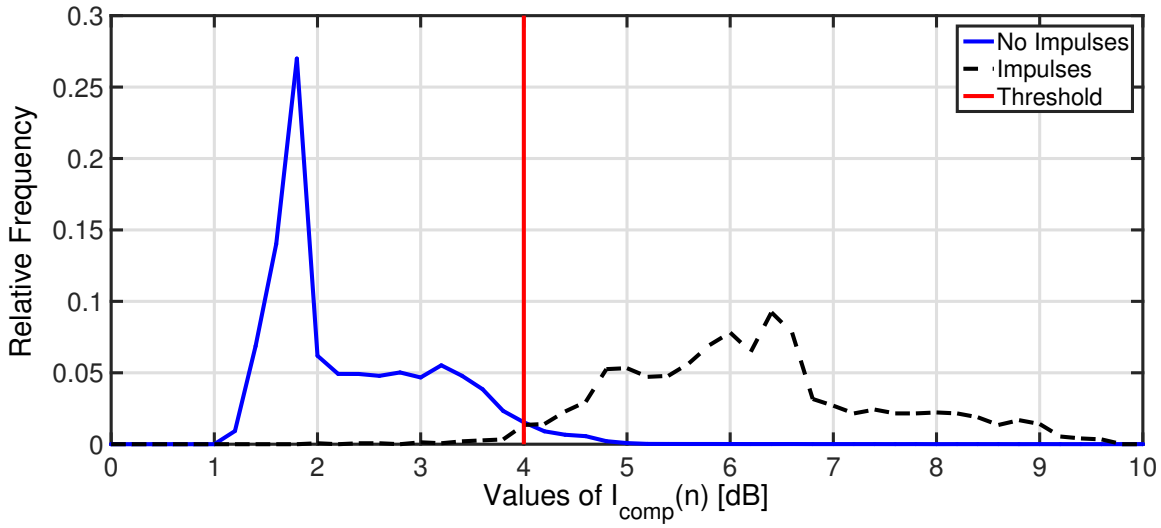


Figure 3.21. The PDFs $p_{\text{no-impulses}}(I_{\text{comp}})$ (blue) and $p_{\text{impulses}}(I_{\text{comp}})$ (black). Used parameters: $\beta_r = 0.05$, $\beta_f = 0.005$, and $\gamma = 0.005$. The threshold T_{ID}, equal to 4 dB, is depicted by the red vertical line.

Impulse Detection Example

For a more detailed analysis of the proposed impulse detection procedure an example is given in Figure 3.22. It shows three recorded impulses in the time domain and also the corresponding spectrograms. Exemplary the tenth sub-band is chosen, for which $\overline{y}_{\text{fast},k}(n)$, $\overline{y}_{\text{slow},k}(n)$, and $I_k(n)$ are depicted. Additionally, $I_k(n)$ for all sub-bands and $I_{\text{comp}}(n)$ over time are shown. One observes that the impulses are detected for the proposed threshold of 4 dB.

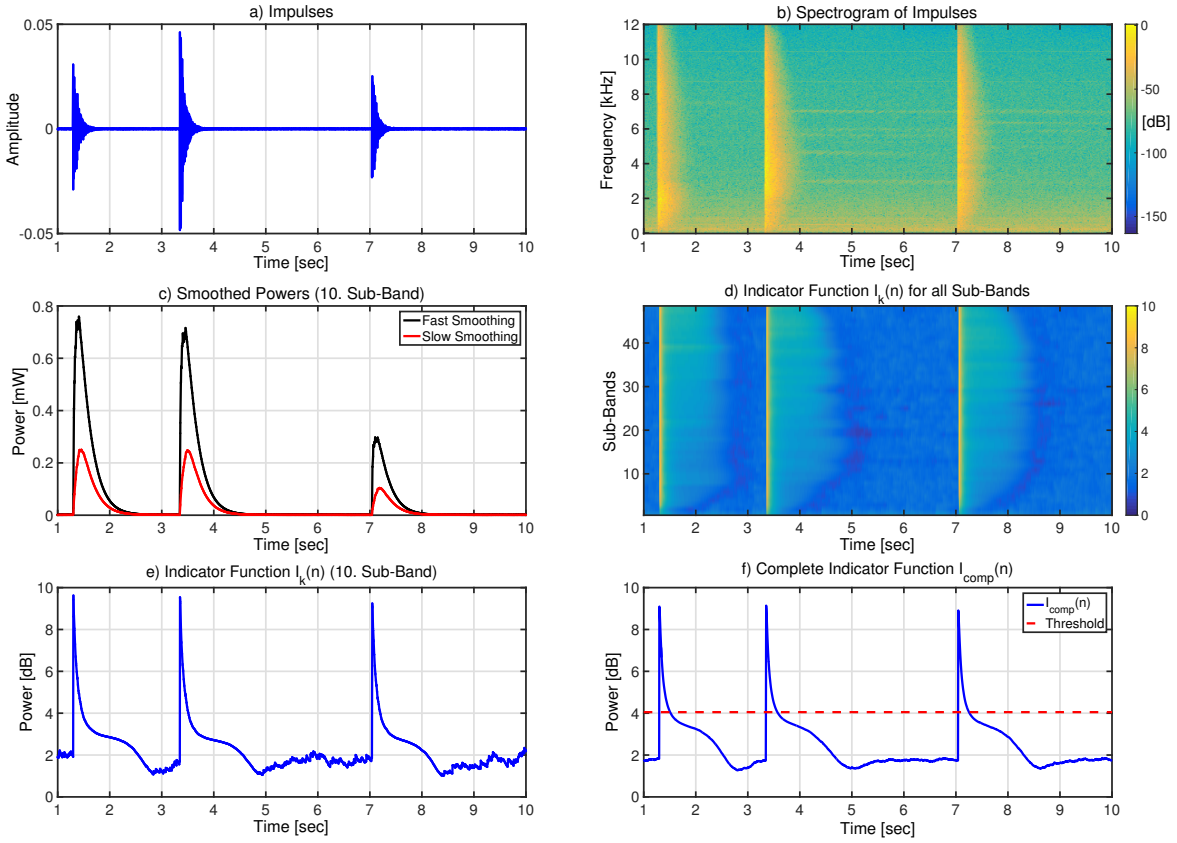


Figure 3.22. An example of the impulse detection is depicted. The different plots show: a) Three recorded impulses in the time domain. b) The three impulses in a time-frequency representation. c) The two smoothed power estimates for sub-band 10. d) The resulting indicator function $I_k(n)$ for sub-band 10. e) $I_k(n)$ for all sub-bands. f) The complete indicator function $I_{\text{comp}}(n)$. By setting the threshold to 4 dB all three impulses are detected.

Conclusion

The impulse detection is developed to handle wideband impulses, which leads to an increased update term and in consequence, to misadaptation and entrainment. For the detection fast and slow power estimates of each sub-band are compared. Based on the comparison an indicator function for each sub-band is calculated. By the average of each indicator function an overall indicator function is obtained. A threshold is used to decide if a impulse is present. If an impulse is detected the step sizes of all sub-bands are set to zero. Hence, the adaptation is frozen and misadaptation is prevented. In Section 5.5.1 the benefit of the impulse detection is shown based on the real-time system and in Section 5.4.1 based on simulation.

3.4.2 Modified Estimate of Source Signal Power

In Section 3.3.4 it was shown that the combination with NPVSS, PEF and FS works very well. Nevertheless, there are still problems and room for improvement if practical aspects, e.g., a realistic hearing aid processing unit, are considered. In this section, a modification to for estimation procedure for $\sigma^2_{x_k}(n)$, c.f. Equation (3.73), is proposed. This modification is needed in case of fast increases of the source signal power, e.g., jarring voices, which are no impulses. In this case the estimate of $\sigma^2_{x_k}(n)$ becomes inaccurate and entrainment can occur. The reason is explained in the following and a solution is proposed.

Since for this thesis a hearing aid processing with automatic gain control and frequency domain processing is considered, similar to a real hearing aid processing unit, $u_k(n)$ is slightly delayed compared to $e_k(n)$. Because of this delay the estimate of $\sigma^2_{x_k}(n)$ based on Equation (3.69) increases too slowly in case of fast increases of the true source signal power. This leads to high peaks in the variable step size which can lead to entrainment [55]. In Figure 3.23 an example is depicted.

To solve this problem one has to modify the estimation of $\sigma^2_{x_k}(n)$ given by Equation (3.69). The idea, which is empirically found, is to combine the autocorrelation matrices of $u_k(n)$ and $e_k(n)$. Hence, the autocorrelation matrix $\mathbf{R}_{u_k u_k}(n)$ in Equation (3.69) is replaced by a modified matrix $\tilde{\mathbf{R}}_{u_k u_k}(n)$ given as

$$\tilde{\mathbf{R}}_{u_k u_k}(n) = \frac{1}{2}\left(\mathbf{R}_{u_k u_k}(n) + \mathbf{R}_{e_k e_k}(n)\frac{\sigma^2_{u_k}(n)}{\sigma^2_{e_k}(n)}\right). \tag{3.80}$$

Since the non-delayed signal $e_k(n)$ is integrated into the calculation of the matrix the influence of the delay is reduced and the estimate $\hat{\sigma}^2_{x_k}(n)$ lags less behind the true power $\sigma^2_{x_k}(n)$ if the power increases rapidly. The factor $\sigma^2_{u_k}(n)/\sigma^2_{e_k}(n)$ compensates the power difference between $e_k(n)$ and $u_k(n)$, which allows the replacement of $\mathbf{R}_{u_k u_k}(n)$ by $\tilde{\mathbf{R}}_{u_k u_k}(n)$. If the signal power stays constant or changes slowly, $\tilde{\mathbf{R}}_{u_k u_k}(n)$ is approximately equal to $\mathbf{R}_{u_k u_k}(n)$. Only in case of fast power changes the difference occurs, which is desired.

The quantities of Equation (3.80) can be estimated as follows

$$\hat{\mathbf{R}}_{u_k u_K}(n) = \alpha_1 \cdot \mathbf{u}_k(n)\mathbf{u}^{\mathrm{H}}_k(n) + (1-\alpha_1)\cdot \hat{\mathbf{R}}_{u_k u_k}(n-1), \tag{3.81}$$

$$\hat{\mathbf{R}}_{e_k e_k}(n) = \alpha_2 \cdot \mathbf{e}_k(n)\mathbf{e}^{\mathrm{H}}_K(n) + (1-\alpha_2)\cdot \hat{\mathbf{R}}_{e_k e_K}(n-1), \tag{3.82}$$

$$\hat{\sigma}^2_{e_k}(n) = \alpha_1 \cdot |e_k(n)|^2 + (1-\alpha_1)\cdot \hat{\sigma}^2_{e_k}(n-1), \tag{3.83}$$

$$\hat{\sigma}^2_{u_k}(n) = \alpha_1 \cdot |u_k(n)|^2 + (1-\alpha_1)\cdot \hat{\sigma}^2_{u_k}(n-1). \tag{3.84}$$

Hence, the modified estimate is given as

$$\begin{aligned}\hat{\sigma}^2_{x_k,\mathrm{mod}}(n) =& \hat{\sigma}^2_{e_k}(n) + \mathbf{r}^{\mathrm{H}}_{e_k u_k}(n)(\tilde{\mathbf{R}}^{-1}_{u_k u_k}(n))^{\mathrm{H}}\mathbf{r}_{e_k u_k}(n) \\ &- 2\mathrm{Re}\{(\mathbf{r}^{\mathrm{H}}_{e_k u_k}(n)(\tilde{\mathbf{R}}^{-1}_{u_k u_k}(n))^{\mathrm{H}}\mathbf{r}_{e_k u_k}(n)\}.\end{aligned} \tag{3.85}$$

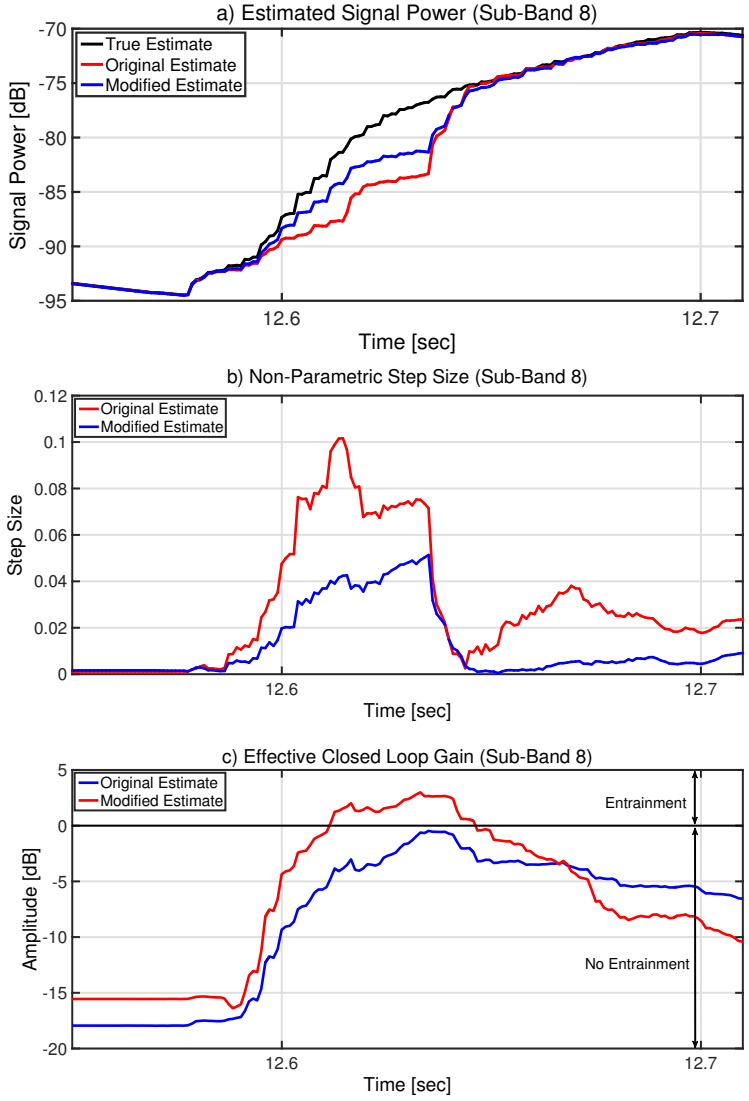


Figure 3.23. An example to illustrate the problem which occurs if the signal power of $x_k(n)$ increases fast. Additionally, the positive impact of the modification, c.f. Equation (3.80), is shown. For this example only the eighth sub-band is considered. Plot a): The true power of the sub-band signal $x_8(n)$ increases by 24 dB in a timespan of 0.12 seconds. The original estimate of the power given by Equation (3.69) does not increase as fast as the true power, because of the delay between $u_8(n)$ and $e_8(n)$. The modified estimate, see Equation (3.85), follows the true power faster. Plot b): The underestimation of $\sigma^2_{x_8}(n)$ leads to an increased step size, though the step size should stay low since the feedback path is not changing. With the modified estimate the underestimation is reduced and therefore, the step size is also reduced in comparison to the step size calculated with the original estimate. Plot c): The ECLG shows the benefit of using the modified estimate. For the original estimate the high step size leads to a misadaptation and this leads to entrainment in this case. Using the modified estimate the step size is reduced which leads to less misadaptation and no entrainment occurs, i.e., the ECLG is below 0 dB.

Note that Equation (3.85) works only sufficiently when the decorrelation methods are used. Hence, $e_k(n)$ and $u_k(n)$ are replaced by $\tilde{e}_k(n)$ and $\tilde{u}_k(n)$ respectively.
Due to the proposed modification, $\hat{\sigma}^2_{x_k,\mathrm{mod}}(n)$ increases faster if the true signal power $\sigma^2_{x_k}(n)$ increases rapidly which reduces the step size and minimizes the generation of entrainment [55]. The smoothing parameter α_2 is set to 0.05 while $\alpha_1 = 0.01$. Both choices are motivated in Section 5.2. Figure 3.23 illustrates the positive effect on the example given before. A complete analysis is given in Section 5.4.2.

Conclusion

The source signal power is modified to deal with sudden power increases in one or few sub-bands. These power increases happen in case of, e.g., jarring voices and lead to a incorrect estimate of the source signal power. Hence, entrainment can occur. This is prevented with the modification of the autocorrelation matrix $\mathbf{R}_{u_k u_k}(n)$, which is used to calculate the estimate of the source signal power. Simulations in Section 5.4.2 demonstrate the increased performance of the AFC system.

3.4.3 Correlation Detection

In this section a new approach is proposed to detect whether the source signal is highly correlated or not. With this detection the adaptation can be additionally controlled.
The AFC system consisting of NPVSS, PEF and FS works well for all kinds of signals and especially shows a very good stability with respect to entrainment, c.f. Figure 3.20. However, one has to deal with the trade-off between adaptation stability and speed. The attenuation factor a of the NPVSS is chosen in favor of stability which is at the expense of the tracking behavior. This is illustrated by the optimization of a in Section 5.2.1. The proposed correlation detection is able to distinguish between correlated input signals and correlation resulting from feedback path changes. This allows us to solve this trade-off. Hence, the adaptation speed is increased without influencing the stability negatively [66].
The correlation detection is based on the concept to measure the effect of the decorrelation methods, PEF and FS. In case of a correlated input signal these methods reduce the correlation significantly whereas a correlation caused by feedback is not affected. To measure the efficiency of decorrelation, two cross-correlation values are estimated. For the first case the signals $e_k(n)$ and $d_k(n)$ are used, which are not influenced by the decorrelation methods, see Figure 3.24. The second one is based on the signals after the applications of decorrelation, i.e., $\tilde{e}_k(n)$ and $\tilde{u}_k(n)$. To measure the difference

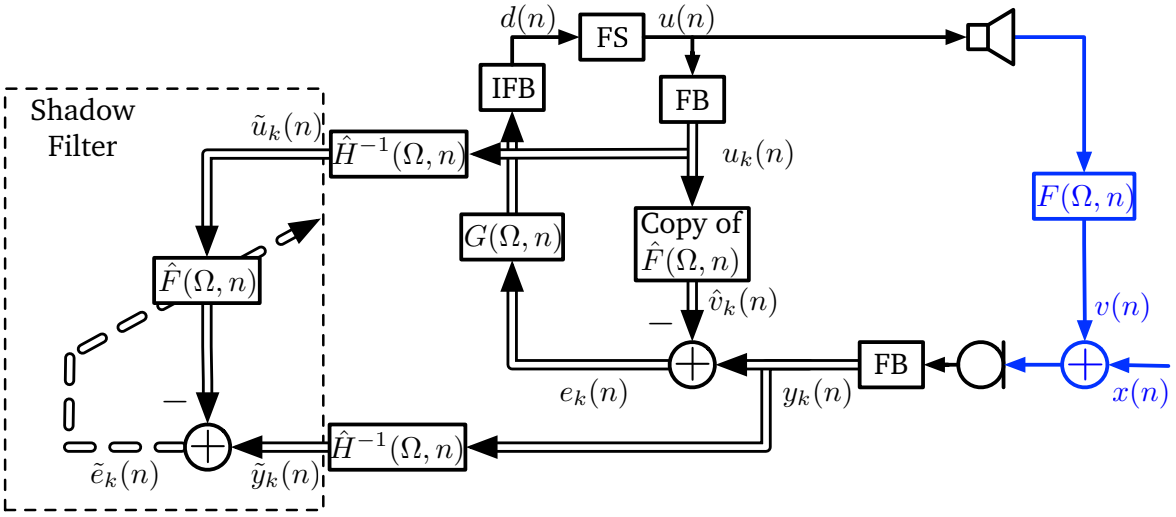


Figure 3.24. Proposed AFC System.

between the two estimates an indicator function for each sub-band is defined:

$$C_k(n) = 10\log_{10}\left(\frac{||\mathbf{r}_{e_k d_k}(n)||^2}{||\mathbf{r}_{\tilde{e}_k \tilde{u}_k}(n)||^2}\right). \tag{3.86}$$

Based on the comparison of this indicator function with a threshold it is detected if the source signal is highly correlated or not. To explain why this method works reliably a closer look on the cross-correlation is provided. First the definition of the cross-correlation is noted as

$$r_{e_k u_k}(n, n-l) = \mathrm{E}\{e_k(n) \cdot u_k^*(n-l)\}. \tag{3.87}$$

The signal $e_k(n)$ is given by

$$e_k(n) = \Delta v_k(n) + x_k(n) = \sum_{j=0}^{N-1} \Delta f_{k,j}^*(n) u_k(n-j) + x_k(n), \tag{3.88}$$

where $\Delta f_{k,j}(n)$ is the difference between the true feedback path coefficients and the estimated coefficients, i.e., $\Delta f_{k,j}(n) = f_{k,j}(n) - \hat{f}_{k,j}(n)$. Hence, $r_{e_k u_k}(n, n-l)$ can be rewritten as

$$r_{e_k u_k}(n, n-l) = \mathrm{E}\left\{\left(\sum_{j=0}^{N-1} \Delta f_{k,j}^*(n) u_k(n-j) + x_k(n)\right) \cdot u_k^*(n-l)\right\} \tag{3.89}$$

$$= \sum_{j=0}^{N-1} \Delta f_{k,j}^*(n) r_{u_k u_k}(n-j, n-l) + r_{x_k u_k}(n, n-l). \tag{3.90}$$

This can be also written in matrix-vector notation as

$$\mathbf{r}_{e_k u_k}(n) = \mathbf{R}_{u_k u_k}(n) \Delta \mathbf{f}_k^*(n) + \mathbf{r}_{x_k u_k}(n). \tag{3.91}$$

Based on that the two cross-correlation vectors given in Equation (3.86) can be noted for each sub-band as

$$\mathbf{r}_{e_k d_k}(n) = \mathbf{R}_{u_k d_k}(n)\Delta\mathbf{f}_k^*(n) + \mathbf{r}_{x_k d_k}(n) \text{ and} \tag{3.92}$$

$$\mathbf{r}_{\tilde{e}_k \tilde{u}_k}(n) = \mathbf{R}_{\tilde{u}_k \tilde{u}_k}(n)\Delta\mathbf{f}_k^*(n) + \mathbf{r}_{\tilde{x}_k \tilde{u}_k}(n). \tag{3.93}$$

Now, three cases are discussed:

1. In the first one the feedback path is optimally estimated, meaning $\Delta\mathbf{f}_k(n) \approx \mathbf{0}$, and the input signal is correlated. Consequently, the indicator function is reduced to

$$C_k(n) \approx 10\log_{10}\left(\frac{||\mathbf{r}_{x_k d_k}(n)||^2}{||\mathbf{r}_{\tilde{x}_k \tilde{u}_k}(n)||^2}\right). \tag{3.94}$$

   Since the input signal is highly correlated, the norm $||\mathbf{r}_{x_k d_k}(n)||^2$ will be high as well, while the norm $||\mathbf{r}_{\tilde{x}_k \tilde{u}_k}(n)||^2$ tends to zero. This results in high values of $C_k(n)$, i.e., $C_k(n) \gg 0\,\text{dB}$.

2. In the second case the input signal is uncorrelated and the feedback path is still optimally estimated. Again $C_k(n)$ can be reduced to Equation (3.94). However, this time the difference between the two correlation vectors is small, which means $C_k(n) \approx 0\,\text{dB}$.

3. In the third case the feedback path changes. Hence, $\Delta\mathbf{f}_k(n)$ increases and the first term of Equation (3.92) and (3.93) dominates. Consequently, $C_k(n)$ can be expressed as

$$C_k(n) \approx 10\log_{10}\left(\frac{||\mathbf{R}_{u_k d_k}(n)\Delta\mathbf{f}_k^*(n)||^2}{||\mathbf{R}_{\tilde{u}_k \tilde{u}_k}(n)\Delta\mathbf{f}_k^*(n)||^2}\right). \tag{3.95}$$

   First only the FS is considered, i.e., $\mathbf{R}_{\tilde{u}_k \tilde{u}_k}(n) \rightarrow \mathbf{R}_{u_k u_k}(n)$, to allow an easier analysis. Due to the use of the FS the magnitudes of the elements of $\mathbf{R}_{u_k d_k}(n)$ are smaller than the magnitudes of the element of $\mathbf{R}_{u_k u_k}(n)$. Hence, the indicator function $C_k(n)$ is small, i.e., $C_k(n) < 0\,\text{dB}$.
   If the PEF is additionally considered and ideal decorrelation is assumed, the two matrices can be expressed as $\mathbf{R}_{u_k d_k}(n) = \mathbf{0}$ and $\mathbf{R}_{\tilde{u}_k \tilde{u}_k}(n) = \sigma^2_{\tilde{u}_k}(n)\mathbf{I}$. This leads again to $C_k(n) < 0\,\text{dB}$.
   This behavior is independent of correlated hearing aid input signals. Thus, it is also true if howling occurs due to sudden feedback path changes.

The three cases show that the proposed indicator function can be used for the correlation detection in an AFC system. Since the indicator function $C_k(n)$ is only high for correlated input signals and a stable feedback path, the adaptation speed can be reduced if a high correlation is detected to prevent entrainment and it can be increased if the feedback path changes.
In the following, examples are shown to confirm the statements above and illustrate the behavior of $C_k(n)$ in practice.

Correlation Detection Examples

In practice, the two cross-correlation vectors can be estimated by recursive smoothing

$$\mathbf{r}_{e_k d_k}(n) = \alpha \cdot e_k(n)\mathbf{d}_k^*(n) + (1-\alpha) \cdot \mathbf{r}_{e_k d_k}(n-1) \text{ and} \tag{3.96}$$

$$\mathbf{r}_{\tilde{e}_k \tilde{u}_k}(n) = \alpha \cdot \tilde{e}_k(n)\tilde{\mathbf{u}}_k^*(n) + (1-\alpha) \cdot \mathbf{r}_{\tilde{e}_k \tilde{u}_k}(n-1), \tag{3.97}$$

with $\alpha = 0.01$. Figure 3.25 shows the first example of the correlation detection. Therefore, 0.5 seconds of white noise followed by a bell ringing as audio data are used. As feedback path $F_1(\Omega)$ presented in Section 2.4 is applied. The spectrogram of the signal and the indicator function $C_k(n)$ are depicted. The indicator function has high values for highly correlated signal components while it is lower for signals with less correlation like for the white noise at the beginning.
A second example is depicted in Figure 3.26. This time an audio signal starting with 1 second of white noise followed by 4 seconds of string instruments is used. The spectrogram of the signal and the indicator functions $C_k(n)$ for two scenarios are depicted. For the first scenario the feedback path is constant while for the second it is changed from $F_1(\Omega)$ to $F_3(\Omega)$, c.f. Section 2.4, at about 2.5 seconds. Again $C_k(n)$ has high values for highly correlated signal components while it is lower for signals with less correlation. When the feedback path is changed, $C_k(n)$ decreases rapidly until the systems adapts to the new feedback path. That is even the case at the frequencies around 5 kHz where the two feedback paths differ with about 8 dB and howling occurs due to the rapid feedback path change as depicted by the ECLG.
Both examples demonstrate the reliability of the indicator function to be used for the correlation detection.

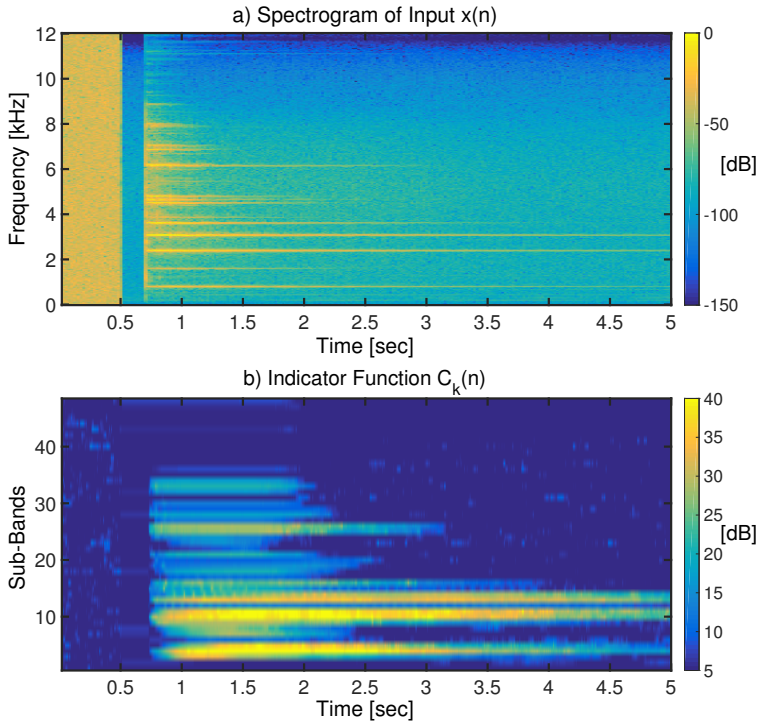


Figure 3.25. An example of the correlation detection is shown. First, the spectrogram of the audio signal, which consists of 0.5 seconds of white noise followed by a bell ringing, is depicted in a). Then, in b) the indicator function $C_k(n)$ is given. The indicator function has high values for highly correlated signal components while it is lower for signals with less correlation like for the white noise at the beginning.

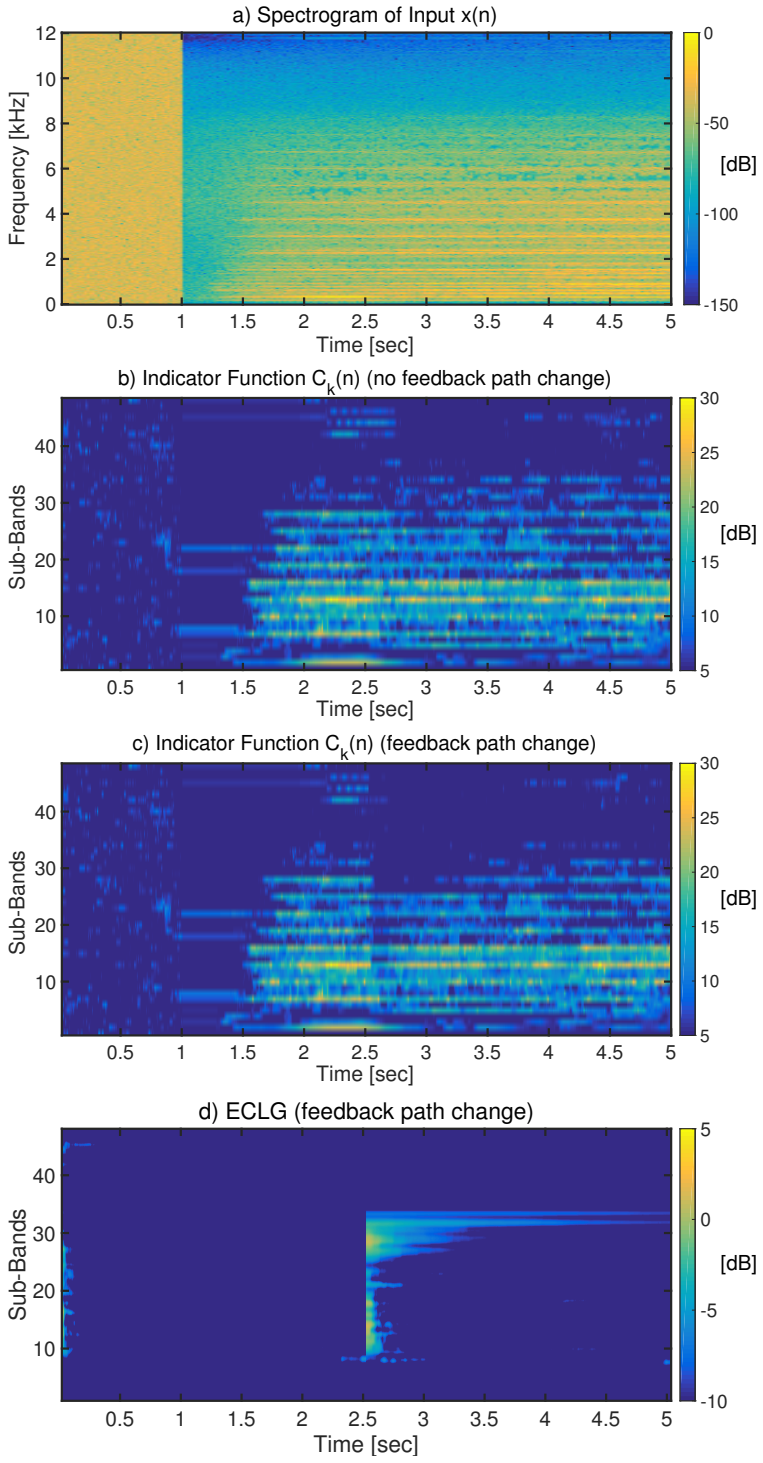


Figure 3.26. Example of the correlation detection. a) Spectrogram of an audio signal consisting of 1 second of white noise followed by 4 seconds of string instruments. b) Indicator function $C_k(n)$ for constant feedback path. c) Indicator function $C_k(n)$ for feedback path change at about 2.5 seconds. d) ECLG(Ω, n) to illustrate the adaptation of the system. $C_k(n)$ has high values for highly correlated signal components while it is lower for signals with less correlation. When the feedback path is changed, $C_k(n)$ decreases rapidly until the systems adapts to the new feedback path. That is also the case at the frequencies around 5 kHz where the two feedback paths differ with about 8 dB and howling occurs due to the rapid feedback path change as depicted in d).

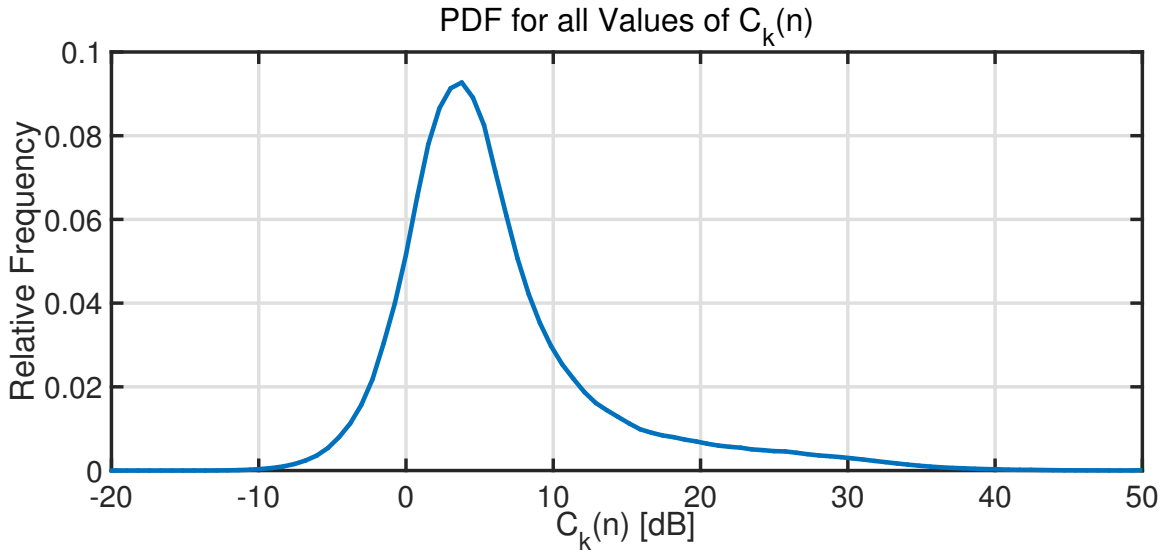


Figure 3.27. PDF $p(x)$ for all values over time and frequency of $C_k(n)$. Almost all values (99.7 %) of $C_k(n)$ are between -10 dB and 40 dB. The PDF has one peak at about 4 dB. However, $p(\mathbf{x})$ is not symmetric as the right tail is longer than the left tail. The right tail is presumably related to highly correlated components of $C_k(n)$.

Finding Optimal Threshold T_{CD}

To decide whether a sub-band signal is highly correlated at time index n the indicator function $C_k(n)$, c.f. Equation (3.86), is compared with a threshold T_{CD}, which is chosen identically for all sub-bands. To find the optimal value for this threshold a simulation with ten minutes of audio data (music and speech) as excitation signal of the AFC is performed. During the simulation the indicator function $C_k(n)$ for each sub-band is calculated. All calculated values of $C_k(n)$ for all sub-bands over time are collected. Based on these values a PDF of $C_k(n)$ is estimated by a histogram. The PDF is called $p(x)$, where x stands for all values of $C_k(n)$ over time and frequency. Note, that the audio data used as input is not the same data as described in Section 2.4. Hence, training and evaluation are separated. This guarantees the reliability of the method for all kinds of audio signals. Since the chosen feedback path has low influence on the values of $C_k(n)$, a typical feedback path as $F_1(\Omega)$, c.f. Section 2.4, is used for this simulation. Figure 3.27 shows the PDF $p(x)$.
Almost all values (99.7 %) of $C_k(n)$ are between -10 dB and 40 dB. The PDF has one peak at about 4 dB. However, $p(x)$ is not symmetric as the right tail is longer than the left tail. The right tail is presumably related to highly correlated components of $C_k(n)$. The frequency of these highly correlated parts is low compared to the less correlated parts. This is due to the nature of most audio signals. Most audio signals excite only a few sub-bands while the rest is covered by uncorrelated background noise.

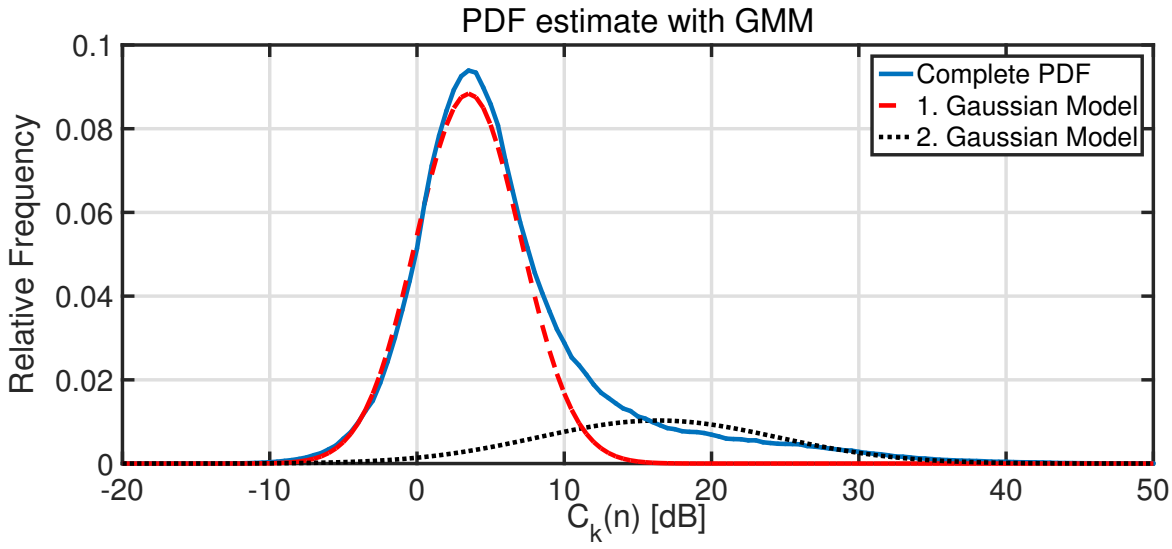


Figure 3.28. PDF $p(x)$ with GMM. $C_k(n)$ is depicted as blue line while the two models of the GMM are given as red and black line. The two models of the GMM are separated satisfactorily. The first model represents the uncorrelated components of the sub-band signals. The second model represents the correlated components.

As consequence, the highly correlated parts are not present in the PDF as a second peak but as a longer right tail.
In order to distinguish the less correlated parts and the highly correlated ones, $p(x)$ is modeled by a Gaussian mixture model (GMM) [63]. The GMM $\hat{p}(x)$ has one dimension and consists of two Gaussians. It is described mathematically by

$$\hat{p}(x) = \sum_{m=0}^{1} g_m \mathcal{N}(x|\mu_m, \sigma_m), \tag{3.98}$$

where μ_m is the mean value of the $m-$th Gaussian, σ_m is the corresponding standard deviation and g_m is called the weight. These parameters are calculated by the expectation-maximization (EM) algorithm [63,64]. Since the EM algorithm is an iterative procedure, the parameters have to be initialized. This is done by the Linde-Buzo-Gray (LBG) algorithm [65]. Here, a codebook with two codebook vectors is trained based on all data values. The codebook vectors are used as initialization for the mean values μ_0 and μ_1. The initial standard deviations are given by the standard deviations of the data allocated to the respective codebook vector. The weights are initialized by the number of elements allocated each codebook divided by the number of all elements. The resulting GMM is displayed together with the PDF $p(x)$ in Figure 3.28.
The two models of the GMM are separated satisfactorily. Hence, the first model can be used to represent the less correlated components of the sub-band signals while the second model can represent the highly correlated components. To verify this, a simu-

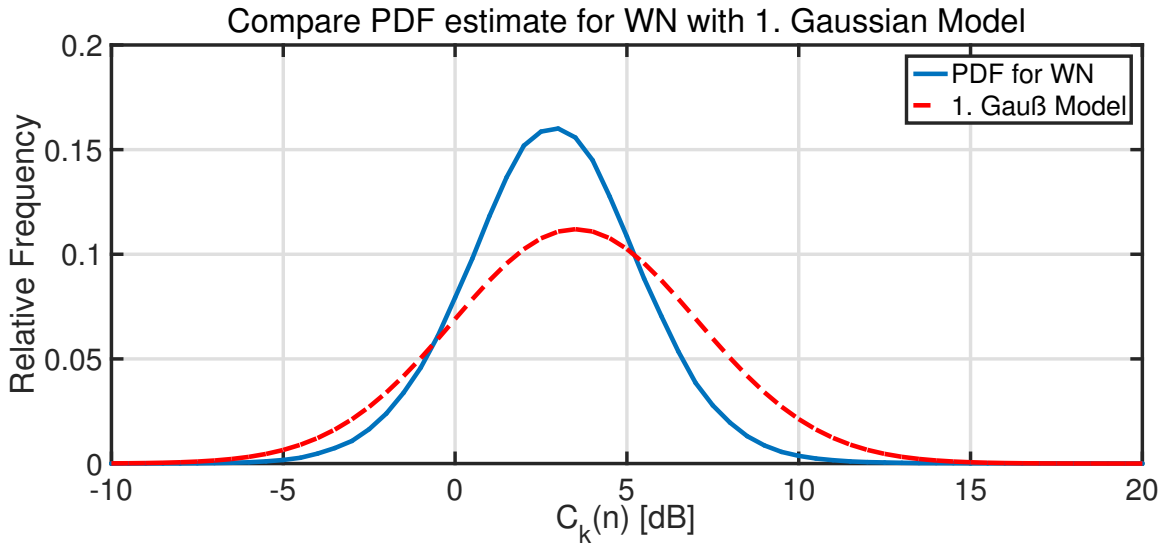


Figure 3.29. Verification of first Gaussian model using the PDF of $C_k(n)$ for a white noise excitation of the AFC system. The PDF $p_{WN}(x)$ is Gaussian distributed and has its peak at 3 dB, which is only slightly less than the peak of the first Gaussian model at 3.5 dB. However, $p_{WN}(x)$ has a smaller variance. Nevertheless, the first Gaussian model is suitable to represent the less correlated components.

lation with only white noise is performed. The PDF of the resulting values of $C_k(n)$ is called $p_{WN}(x)$ and is shown together with the first Gaussian model in Figure 3.29.
The PDF $p_{WN}(x)$ is Gaussian distributed and has its peak at 3 dB, which is only slightly less than the peak of the first Gaussian model at 3.5 dB. However, $p_{WN}(x)$ has a smaller variance. Nevertheless, the first Gaussian model is suitable to represent the less correlated parts. These parts do not only consist of white signal parts but also on slightly correlated parts. Therefore, it is reasonable that the first Gaussian model covers also higher values of $C_k(n)$ than $p_{WN}(x)$.
Based on the two Gaussian models the threshold T_{CD} is determined. The threshold is set to the value for which the smallest error rate occurs for a Bayesian detection. This means, the error rate is determined based on the sum of false alarm and missed detection probabilities. Figure 3.30 shows the error rate for thresholds from -10 dB to 50 dB.
The smallest error rate, which is 6.8 %, is obtained for $T_{\mathrm{CD}} = 11$ dB. Hence, based on the given assumptions this is the optimal value for the threshold. In Figure 3.31 the threshold and the GMM are shown together.
With the optimal threshold the correlation detection can be used in the proposed AFC system. In the next section the idea how to apply the correlation detection is described.

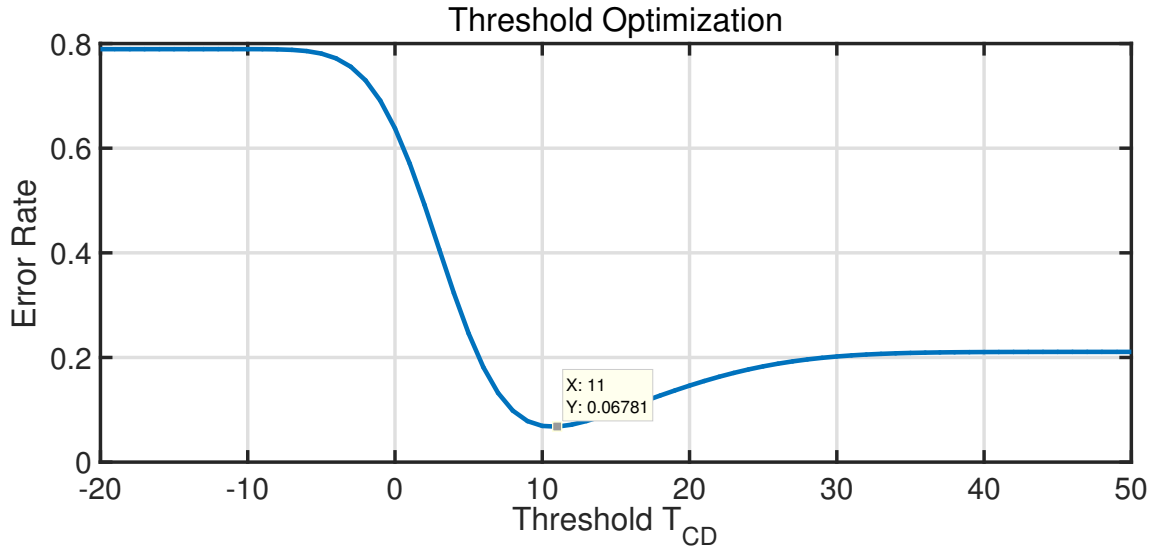


Figure 3.30. Optimization of Threshold T_{CD}. The error rate for thresholds from $-10\,\mathrm{dB}$ to $-50\,\mathrm{dB}$ is depicted. The smallest error rate (6.8 %) is obtained for $T_{\mathrm{CD}} = 11\,\mathrm{dB}$.

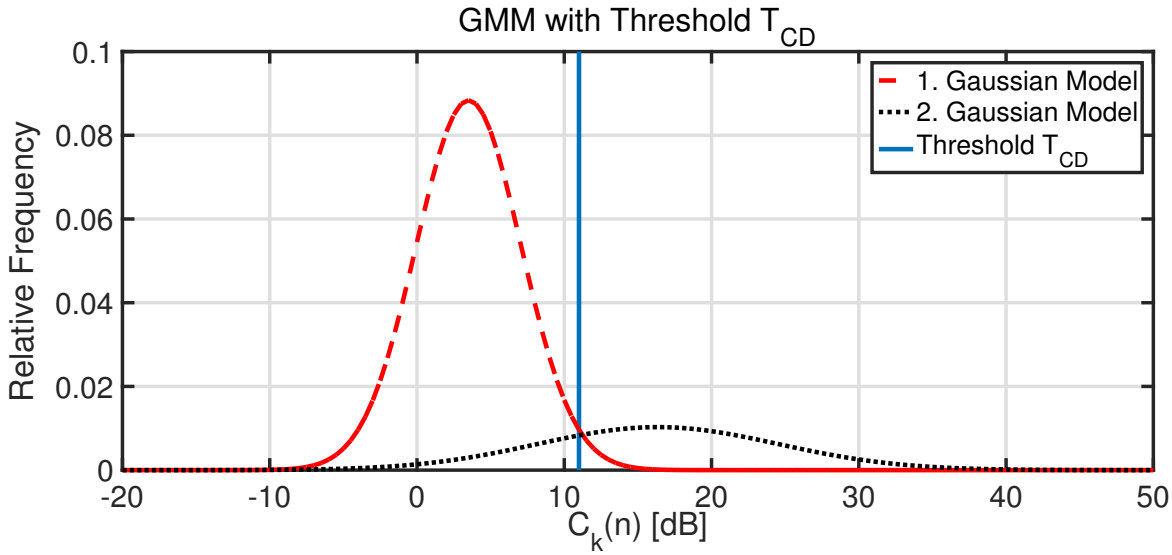


Figure 3.31. GMM with Threshold T_{CD}.

Application of Correlation Detection

The concept is to control the adaptation additionally by the correlation detection [66]. This is done by using different attenuation values for the calculation of the step size. If a high correlation is detected a stronger attenuation is used to prevent entrainment. If the correlation is less the attenuation is reduced to allow a faster adaptation. Mathematically this is expressed as

$$\mu_{\mathrm{NPVSS},k}(n) = \begin{cases} a_1 \cdot \left(1 - \frac{\sigma_{x_k}(n)}{\sigma_{e_k}(n)}\right), & C_k(n) \geq T_{\mathrm{CD}} \\ a_2 \cdot \left(1 - \frac{\sigma_{x_k}(n)}{\sigma_{e_k}(n)}\right), & C_k(n) < T_{\mathrm{CD}}, \end{cases}$$

where the threshold T_{CD} is chosen as 11 dB and the attenuation parameters as $a_1 = 0.25$ and $a_2 = 0.5$. By choosing $a_1 = a = 0.25$, it is guaranteed that the system will not adapt slower with the additional correlation detection method than without it. With $a_2 = 0.5$ a faster adaptation is obtained when the source signal is less correlated. Hence, the correlation detection increases the tracking behavior while the adaptation stability is not much effected [66].
The parameters are optimized in Section 5.2.2 and a complete investigation on the impact of the correlation detection is given in Section 5.4.3.

Conclusion

The correlation based control approach is proposed to enhance the trade-off between adaptation stability and speed. The correlation detection measures the effect of the decorrelation methods by estimating two different cross-correlation values. Based on these values, indicator functions for each sub-band are calculated. With the help of a threshold it is decided if the correlation of the source signal is high or not. The threshold is optimized based on two PDFs, one representing uncorrelated source signals and the other correlated source signals. If a high correlation is detected, a low attenuation factor is chosen, which leads to a slow adaptation. For a low correlation a high attenuation factor is used, which means the adaptation speed is increased. In consequence, the tracking speed is increased while the stability is not influenced. The benefit of this additional adaptation control approach is demonstrated in Section 5.5.1 by the real-time system and in Section 5.4.3 by simulations.

3.5 Conclusion

The complete proposed AFC system, which is developed for a realistic hearing aid setup, was described in this chapter. The NLMS algorithm was chosen to perform the adaptation of the estimated feedback path. In order to provide a stable adaptation, the NPVSS was proposed for an adaptation control in combination with the decorrelation methods, PEF and FS. For the application of the NPVSS in a realistic hearing aid environment, a new estimate of the source signal power was derived.
Furthermore, three additional control methods were developed and implemented in the AFC system: an impulse detection, a modified estimation method of the source signal power and a correlation detection.
The impulse detection ensures that the adaptation is stable in case of wideband impulses, e.g., clapping. The modified estimate increases the stability for sudden power increases in few sub-bands, e.g., jarring voices. With the correlation detection the trade-off between adaptation stability and tracking behavior is improved.
With the combination of all these methods an AFC system with excellent adaptation stability and speed was developed as demonstrated through real-time evaluations in Chapter 4 and offline simulations in Chapter 5.

Chapter 4

Adaptive Feedback Cancellation: Real-time Implementation

This chapter addresses the real-time system which is used in this thesis. The real-time system allows a fast and realistic examination of the proposed AFC system.
Fast refers to the immediate results one gets by using the real-time system. Offline simulations require a certain amount of computational effort. Hence, they take time and computing capacity. This is also the case for the evaluation based on the ECLG, which is proposed in this thesis, c.f. Section 2.4, as it has to be calculated at every adaptation step. With the real-time system one is able to receive immediate feedback regarding the performance of the analyzed AFC system.
Furthermore, the system can be tested with a large variety of realistic scenarios. This includes different audio signals and effectively an unlimited number of feedback paths, which can be varied in a realistic manner during the analyses.
In the following section the experimental setup of the real-time system is discussed, which emphasizes the benefit of using it. Next the graphical user interface (GUI) is described. The GUI is used to observe and control the AFC system.

4.1 Experimental Setup

In this section the experimental setup of the real-time system is described. The main components are a performance real-time target machine called Speedgoat system, a head and torso simulator called KEMAR, and a personal computer (PC). Figure 4.1 illustrates the scheme of the setup.
Real-time simulations are performed as follows: First the proposed AFC system is realized with Matlab Simulink. The resulting Simulink model can easily be compiled into C-Code using a Matlab toolbox. This C-Code is then transfered to the Speedgoat system where it is run in real-time. This is the main reason why Simulink is used. Additionally, it allows a clear structure of all proposed methods.
The Speedgoat system, c.f. Figure 4.2, provides a high performing test environment. It includes a 24-bit input module with 12 differential channels, which guarantees a very high input resolution. The module uses a sigma-delta A/D converter for sample rate up to 200 ksamples/s and provides a signal-to-noise ratio (SNR) up to 93 dB. The output module consists of 8 channels with a resolution of 16 bits. Each channel has

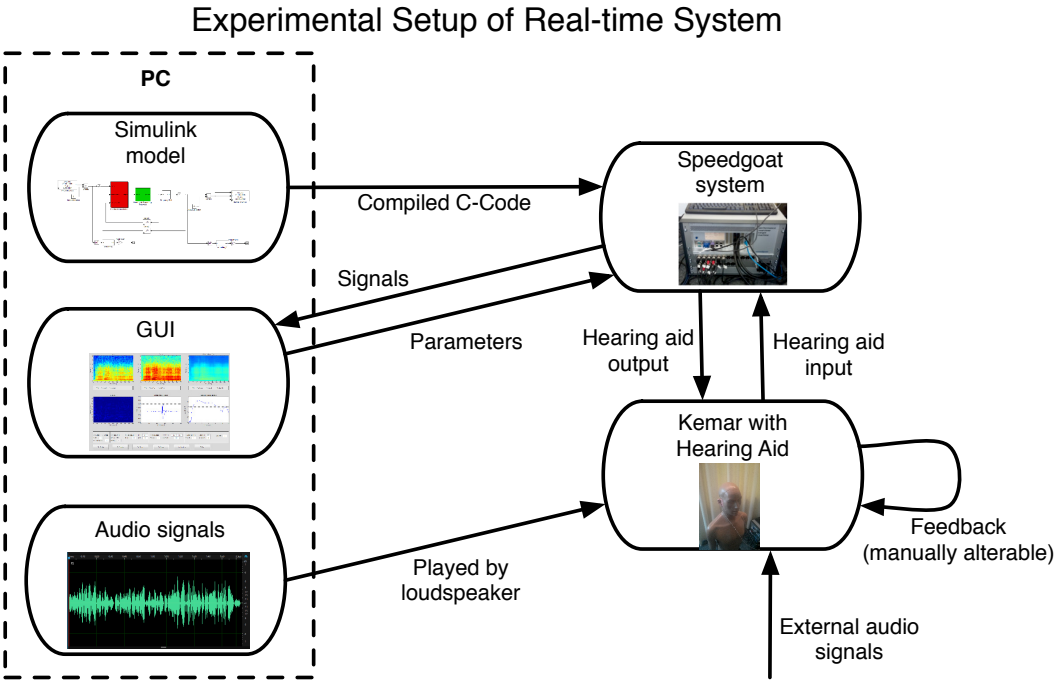


Figure 4.1. Scheme to illustrate the experimental setup of the real-time system.

a dedicated D/A converter with an update time of 2.2 μs. Both modules ensure that high numbers of inputs and outputs are available. Hence, several different experimental setups can be applied. Due to the high-end hardware built in the Speedgoat system a high performance is achieved resulting in a low overall latency equal to the latency in a real hearing aid. This is very important for feedback cancellation and concrete hearing aid applications in general as a realistic closed loop delay is required for meaningful evaluations with a real-time system. Additionally, the Speedgoat system allows with the help of specialized tools from Mathworks the opportunity to change the values of parameters and to turn specific features on or off while the AFC system is running.
To control the AFC system, i.e., change parameter values, enable procedures, and observe important signals, a GUI is used. The GUI is implemented in Matlab and runs on the PC during the analysis. In Section 4.2 the GUI is described in detail.
The Speedgoat system is connected to a hearing aid dummy with microphones and a loudspeaker, i.e., receiver, which is placed in the ear model of the KEMAR. The KEMAR, c.f. Figure 4.3, represents the hearing aid user. It models the human body, head, and ear. Furthermore, it consists of a pinna simulator, an ear canal extension, and an ear simulator to resemble the acoustic characteristics of the human ear. The ear simulator is designed according to the IEC 60318-4 standard with a coupler volume of 1260 mm^3 at 500 Hz. All the mentioned components are important in order to obtain a realistic feedback path, which can be changed instantaneously, c.f. Figure 4.3.

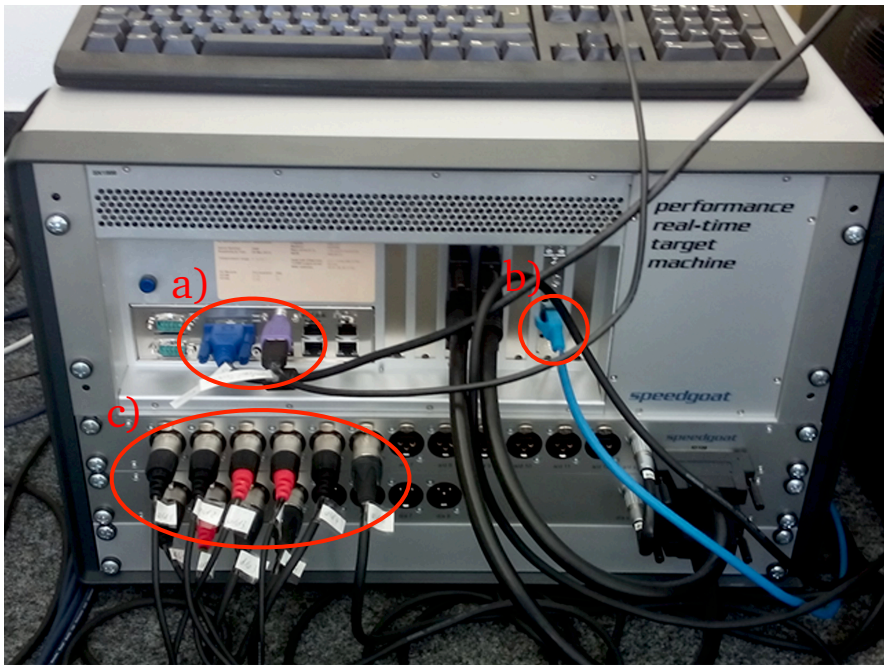


Figure 4.2. The front of the Speedgoat system is depicted. a) shows the connections to a monitor and keyboard, which are used to configure the Speedgaot system. b) displays the local area network cable to connect the system with the PC. c) indicates the 12 inputs (upper row) and the 8 outputs (lower row).

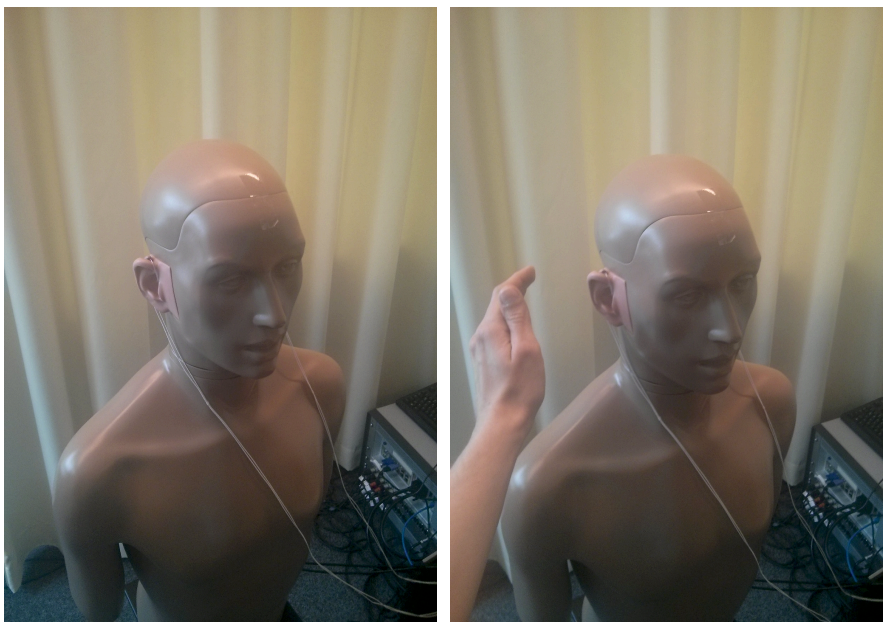


Figure 4.3. The KEMAR with a hearing aid dummy is shown. The left and right figure illustrate two different feedback scenarios.

There are two possibilities to create audio examples in the presented real-time system. The first possibility is to play audio signals from the PC with a loudspeaker. For the second possibility the audio signals are created externally meaning by speaking, clapping, etc. Obviously both methods can be combined.
Due to the possibility to vary the feedback path manually and use different audio examples, one is able to test the AFC system with a large variety of realistic scenarios.

4.2 Graphical User Interface

The GUI is used to control and observe the proposed AFC system on the Speedgoat system in real-time. With the GUI one is able to observe several signals of the system. Additionally, it is possible to change parameter values and to enable or disable certain methods. The GUI is implemented in Matlab. Figure 4.4 shows a screenshot of the GUI while the AFC system is running to give an overview.

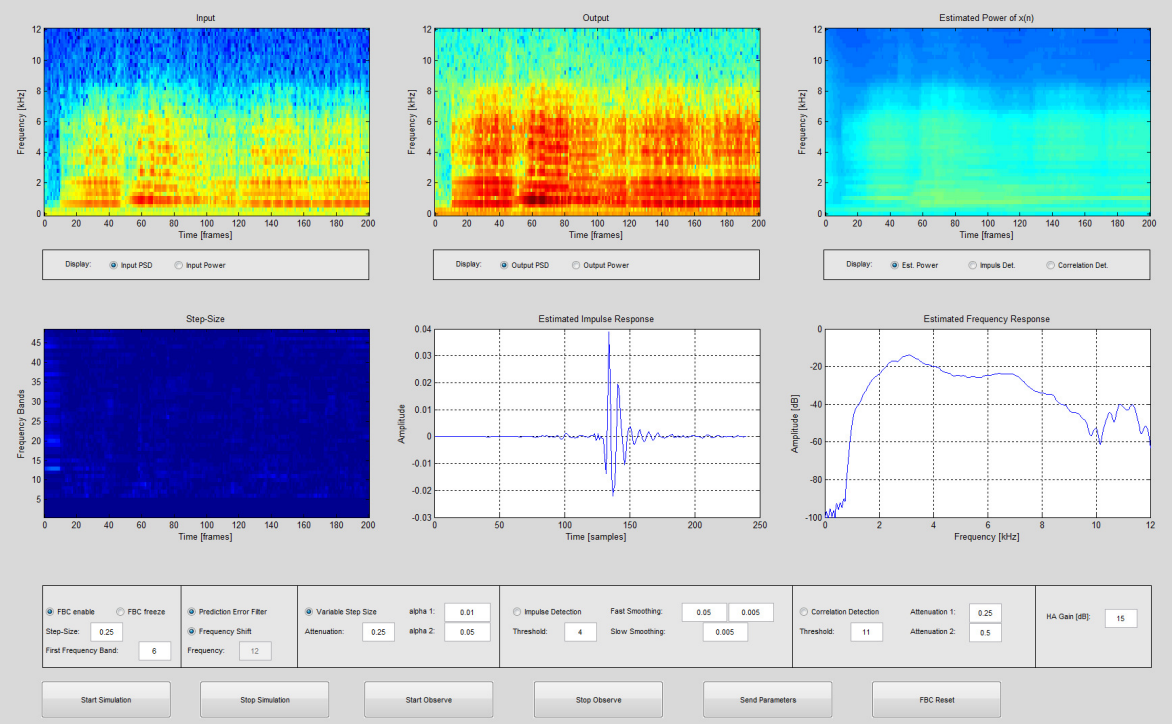


Figure 4.4. Screenshot of GUI while AFC system is running.

The GUI consists of six plots. The upper left plot shows the input power spectral density (PSD), i.e., the PSD of the microphone signal of the hearing aid. The upper middle plot displays the output PSD. Output stands for the loudspeaker signal of the hearing aid. Instead of PSDs the plots can also show the broadband signal power of the

input or output signal. This can be changed by radio buttons below the corresponding plots. These two plots are used to observe the input and output signals over time, i.e., showing spectrogram plots.
For the upper right plot one has the choice for the analysis of three signals. The first one is the estimated PSD of the source signal $x(n)$, which is active in the depicted screenshot and indicates that the source signal power is estimated correctly. Furthermore, the indicator function of the impulse detection and the indicator function of the correlation detection can be displayed. Both help to evaluate the impulse, respectively the correlation detection and optimize the respective parameters as the immediate impact on the methods can be observed by the indicator functions.
The left plot in the lower row shows step size values applied to each sub-band. This plot helps to analyze if the system adapts as indented during the real-time simulations, e.g., when the feedback path is changed an increase of the step sizes should be noticeable while the step size values should remain low if the feedback path is constant independently of the surrounding sounds.
In the next two plots the estimated impulse and the corresponding frequency response are depicted. With these two plots one is able to analyze the estimated feedback path over time. Note, that the impulse and frequency responses have to be transformed to the baseband based on the 48 sub-band impulse responses. This is done on the PC with Matlab procedures in real-time in order to display them in the GUI. This is due to the fact that the adaptation is performed in sub-bands and hence, no baseband frequency response is estimated by the proposed system. The procedure to calculate the impulse response is displayed in Figure 4.5.

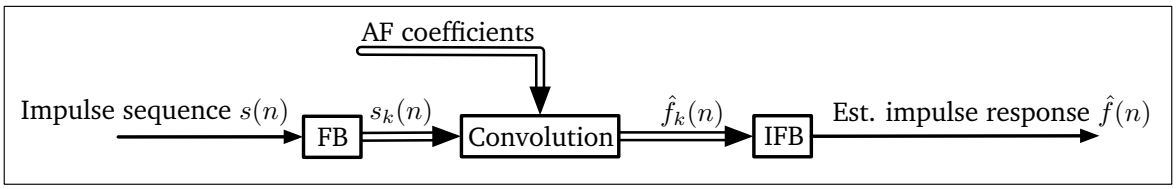


Figure 4.5. Scheme of the procedure to calculate the estimated impulse response based on the adaptive filter coefficients.

An impulse is decomposed into the 48 sub-bands. Then, the decomposed signals are filtered with the estimated filter coefficients. The resulting signals are transformed into the full band with a synthesis FB and finally, the impulse response is obtained. Based on the impulse response the frequency response is calculated with a 1024-point fast Fourier transform (FFT).
Below the plots several radio buttons and text boxes are placed in the GUI. The radio buttons are used to enable different procedures of the system. If 'FBC', which stands for feedback cancellation, is disabled no feedback cancellation is performed. By acti-

vating 'FBC freeze' the adaptation is disabled, meaning the current estimated feedback path is no more updated. With the text boxes important parameters of the methods and the broadband hearing aid gain can be changed by the user. This allows the user to test the impact of the parameters when changing them. Eventually, the user can optimize the parameters based on the plots, e.g., indicator functions and step size values, and based on listening to the output signal of the hearing aid, where howling and artifacts can be detected easily.
Multiple buttons are located at the button of the GUI. With the first two buttons one can start and stop the simulation. The next two buttons start and stop the observation, meaning the visualization of the signals. The button 'Send Parameters' has to be used to send changed parameters to the Speedgoat system. With 'FBC reset' the adaptive filter coefficients are all set to zero.
With the GUI one can evaluate the overall performance of the system. Therefore, the lower three plots are important as they show where and how strongly the systems adapts and with that the estimated feedback path changes. Depending on the created situation an adaptation is necessary and desired or it represents a misadaptation. In Section 5.5 some analyses to illustrate the evaluation with the real-time system are described with the help of the GUI.

Chapter 5

Parameter Optimization and Evaluation

This chapter presents several simulations and evaluations of the overall system and specific methods. Mostly the simulations are based on evaluation procedures explained in Section 2.4. The used audio signals and feedback paths are presented in Section 2.4 as well.
Before the actual simulations are covered, the overall system and each proposed control method are summarized in Section 5.1. Several settings, which consist of different combinations of control methods, are defined. These settings are used for different simulations. Additionally, the standard values for all parameters are presented in a clearly structured way.
In Section 5.2 the values of the most important parameters, which are not determined throughout the previous chapters, are then optimized based on their effect on the system performance. Therefore, the NPVSS with the addition of the decorrelation methods are used. The NPVSS alone is not usable in realistic applications as the proposed method to estimate the source signal power is only valid with the use of decorrelation methods. Hence, the decorrelation methods FS and PEF are applied for the optimization of the parameters.
After the optimization of the parameters, the effects of the decorrelation methods in combination with NPVSS are examined in Section 5.3. These effects are evaluated based on the estimation error, which occurs by the estimation of source signal power, and the ECLG. It is shown that during the simulations entrainment occurs regularly when the decorrelation methods are not applied. This is due to high values of the estimation error, which occur without decorrelation.
In Section 5.4 the effects of the three additional control methods, described in Section 3.4, are discussed based on the adaptation stability and tracking behavior.
Finally, in Section 5.5 all proposed methods and the complete AFC system are evaluated with the real-time system. Therefore, specific situations are analyzed by the GUI and a simulation regarding the estimation error of the source signal power is performed.

5.1 Overview: Overall System and Parameters

In this section an overview of the complete proposed system is given. This overview includes a recap of all applied methods along with the parameter configurations, which are finally chosen. Several simulation settings of the system are defined. These settings are different combinations of the presented methods and are used to explain clearly which methods are active during a specific simulation.
First of all, the basic scheme of the proposed sub-band AFC system is once more presented in Figure 5.1.

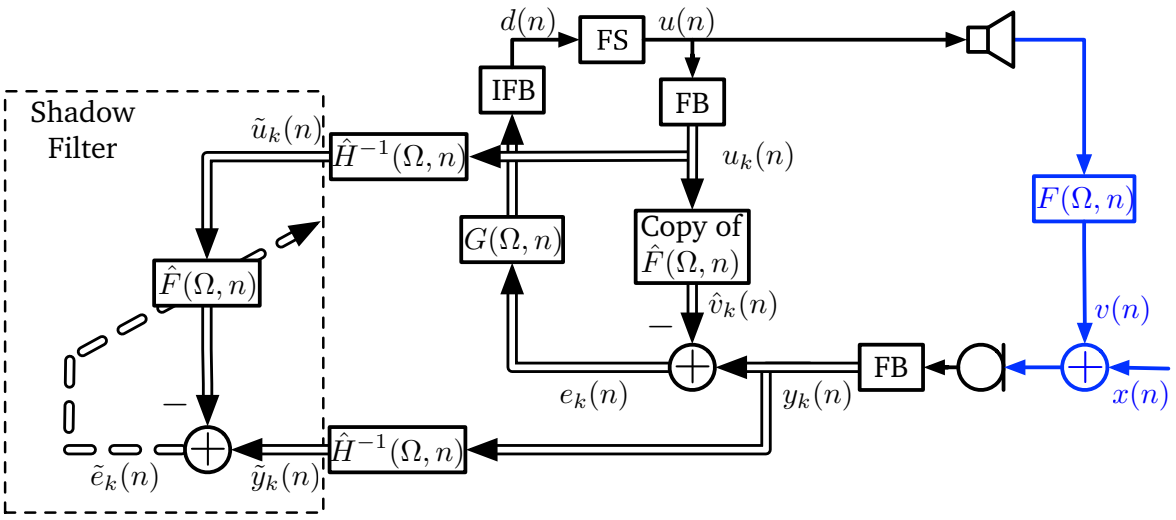


Figure 5.1. Basic scheme of the proposed sub-band AFC system.

5.1.1 Simulation Settings

Before describing the simulation settings the basic system, presented in detail in Section 3.1, is briefly discussed. It consists of the microphone and loudspeaker of the hearing aid, the filter banks (FBs) and the inverse filter bank (IFB) as well as the hearing aid processing unit represented by $G(\Omega, n)$. The processing unit includes the hearing aid gain and the AGC. The basic system is the basis for the following simulation settings.

- Setting 1: This setting consists of the combination of the NPVSS, the PEF, and the FS. No additional control methods are applied. The adaptive filters are calculated in the shadow domain as illustrated in Figure 5.1 with the NLMS adaptation rule for a sub-band k

$$\hat{\mathbf{f}}_k(n) = \hat{\mathbf{f}}_k(n-1) + \mu_{\mathrm{NPVSS},k}(n)\frac{\tilde{e}_k^*(n)\tilde{\mathbf{u}}_k(n)}{||\tilde{\mathbf{u}}_k(n)||^2}. \tag{5.1}$$

The NPVSS $\mu_{\mathrm{NPVSS},k}(n)$ is calculated by

$$\mu_{\mathrm{NPVSS},k}(n) = a \cdot \left(1 - \frac{\sigma_{\tilde{x}_k}(n)}{\sigma_{\tilde{e}_k}(n)}\right), \tag{5.2}$$

where the estimates of $\sigma_{\tilde{x}_k}(n)$ and $\sigma_{\tilde{e}_k}(n)$ are given by

$$\begin{aligned}\hat{\sigma}^2_{\tilde{x}_k}(n) =& \hat{\sigma}^2_{\tilde{e}_k}(n) + \mathbf{r}^{\mathrm{H}}_{\tilde{e}_k\tilde{u}_k}(n)(\mathbf{R}^{-1}_{\tilde{u}_k\tilde{u}_k}(n))^{\mathrm{H}}\mathbf{r}_{\tilde{e}_k\tilde{u}_k}(n) \\ &- 2\mathrm{Re}\{(\mathbf{r}^{\mathrm{H}}_{\tilde{e}_k\tilde{u}_k}(n)(\mathbf{R}^{-1}_{\tilde{u}_k\tilde{u}_k}(n))^{\mathrm{H}}\mathbf{r}_{\tilde{e}_k\tilde{u}_k}(n)\}\end{aligned} \tag{5.3}$$

and

$$\hat{\sigma}^2_{\tilde{e}_k}(n) = \alpha_1 \cdot |\tilde{e}_k(n)|^2 + (1-\alpha_1) \cdot \hat{\sigma}^2_{\tilde{e}_k}(n-1), \tag{5.4}$$

which is also used in Equation (5.3). The correlation values of Equation (5.3) are estimated by

$$\hat{\mathbf{R}}_{\tilde{u}_k\tilde{u}_k}(n) = \alpha_1 \cdot \tilde{\mathbf{u}}_k(n)\tilde{\mathbf{u}}^{\mathrm{H}}_k(n) + (1-\alpha_1) \cdot \hat{\mathbf{R}}_{\tilde{u}_k\tilde{u}_k}(n-1), \tag{5.5}$$

$$\hat{\mathbf{r}}_{\tilde{e}_k\tilde{u}_k}(n) = \alpha_1 \cdot \tilde{e}_k(n)\tilde{\mathbf{u}}^{*}_k(n) + (1-\alpha_1) \cdot \hat{\mathbf{r}}_{\tilde{e}_k\tilde{u}_k}(n-1). \tag{5.6}$$

The NPVSS is described in detail in Section 3.3.3, the PEF in Section 3.2.1 and the FS in Section 3.2.2. Setting 1 is used in Section 5.2.1 to optimize parameters used to calculate the NPVSS, in Section 5.3 to examine the effects of the decorrelation methods and in Section 5.5.2 to evaluate the estimation error of the source signal power with the real-time system.

- Setting 2: Additionally to Setting 1 the impulse detection described in Section 3.4.1 is used. The indicator function of the impulse detection is given by

$$I_{\mathrm{comp}}(n) = \frac{1}{K}\sum_{k=0}^{K-1} I_k(n) \tag{5.7}$$

with $I_k(n)$ calculated in each sub-band by

$$I_k(n) = 10\log_{10}\left(\frac{\overline{y}_{\mathrm{fast},k}(n)}{\overline{y}_{\mathrm{slow},k}(n)}\right). \tag{5.8}$$

The smoothed values $\overline{y}_{\mathrm{fast},k}(n)$ and $\overline{y}_{\mathrm{slow},k}(n)$ are obtained by

$$\overline{y}_{\mathrm{fast},k}(n) = \begin{cases} \beta_r \cdot |y_k(n)|^2 + (1-\beta_r) \cdot \overline{y}_{\mathrm{fast},k}(n-1), & \overline{y}_{\mathrm{fast},k}(n-1) \geq |y_k(n)|^2 \\ \beta_f \cdot |y_k(n)|^2 + (1-\beta_f) \cdot \overline{y}_{\mathrm{fast},k}(n-1), & \overline{y}_{\mathrm{fast},k}(n-1) < |y_k(n)|^2 \end{cases}$$

and

$$\overline{y}_{\mathrm{slow},k}(n) = \gamma \cdot |y_k(n)|^2 + (1-\gamma) \cdot \overline{y}_{\mathrm{slow},k}(n-1).$$

If $I_{\mathrm{comp}}(n)$ exceeds the threshold T_{ID} an impulse is detected and the adaptation is stopped by $\mu_{\mathrm{NPVSS},k}(n) = 0\ \forall k$. Setting 2 is used in Section 5.4.1 to evaluate the effects of the impulse response.

- Setting 3: In contrast to Setting 1 the modified estimate of the source signal power is used. Meaning the autocorrelation matrix $\mathbf{R}_{\tilde{u}_k\tilde{u}_k}(n)$ in Equation (5.3) is replaced by

$$\tilde{\mathbf{R}}_{\tilde{u}_k\tilde{u}_k}(n) = \frac{1}{2}\left(\mathbf{R}_{\tilde{u}_k\tilde{u}_k}(n) + \mathbf{R}_{\tilde{e}_k\tilde{e}_k}(n)\frac{\sigma^2_{\tilde{u}_k}(n)}{\sigma^2_{\tilde{e}_k}(n)}\right) \tag{5.9}$$

  with the estimated quantities:

$$\hat{\mathbf{R}}_{\tilde{u}_k\tilde{u}_k}(n) = \alpha_1 \cdot \tilde{\mathbf{u}}_k(n)\tilde{\mathbf{u}}_k^{\mathrm{H}}(n) + (1-\alpha_1)\cdot\hat{\mathbf{R}}_{\tilde{u}_k\tilde{u}_k}(n-1), \tag{5.10}$$

$$\hat{\mathbf{R}}_{\tilde{e}_k\tilde{e}_k}(n) = \alpha_2 \cdot \tilde{\mathbf{e}}_k(n)\tilde{\mathbf{e}}_k^{\mathrm{H}}(n) + (1-\alpha_2)\cdot\hat{\mathbf{R}}_{\tilde{e}_k\tilde{e}_k}(n-1), \tag{5.11}$$

$$\hat{\sigma}^2_{\tilde{e}_k}(n) = \alpha_1 \cdot |\tilde{e}_k(n)|^2 + (1-\alpha_1)\cdot\hat{\sigma}^2_{\tilde{e}_k}(n-1), \tag{5.12}$$

$$\hat{\sigma}^2_{\tilde{u}_k}(n) = \alpha_1 \cdot |\tilde{u}_k(n)|^2 + (1-\alpha_1)\cdot\hat{\sigma}^2_{\tilde{u}_k}(n-1). \tag{5.13}$$

  The idea and explanation of the modified estimate is given extensively in Section 3.4.2. For the reminder of this section the NPVSS with using this modified estimate is called $\mathrm{NPVSS_{mod}}$. Setting 3 is used for the optimization of α_2 in Section 5.2.2 and for the evaluation of $\mathrm{NPVSS_{mod}}$ in Section 5.4.2.

- Setting 4: The correlation detection is added to Setting 1. Hence, the step sizes are calculated by

$$\mu_{\mathrm{NPVSS},k}(n) = \begin{cases} a_1 \cdot \left(1 - \frac{\sigma_{\tilde{x}_k}(n)}{\sigma_{\tilde{e}_k}(n)}\right), & C_k(n) \geq T_{\mathrm{CD}} \\ a_2 \cdot \left(1 - \frac{\sigma_{\tilde{x}_k}(n)}{\sigma_{\tilde{e}_k}(n)}\right), & C_k(n) < T_{\mathrm{CD}}, \end{cases} \tag{5.14}$$

  where the indicator function $C_k(n)$ is given as

$$C_k(n) = 10\log_{10}\left(\frac{||\mathbf{r}_{e_kd_k}(n)||^2}{||\mathbf{r}_{\tilde{e}_k\tilde{u}_k}(n)||^2}\right). \tag{5.15}$$

  For detailed information and derivation of the correlation detection see Section 3.4.3. With Setting 4 the parameter a_2 is optimized in Section 5.2.3 and the performance of the correlation detection is examined in Section 5.4.3.

In the following section the standard parameter configuration is given.

5.1.2 Parameters

The following parameter settings are used for all simulations if not noted separately. For a better overview the parameters are sorted based on for which methods they are used:

- General settings:
  - Sampling frequency: $f_\mathrm{s} = 24\,\mathrm{kHz}$
  - Number of sub-bands: $K = 96 \Rightarrow 48$ effective sub-bands; frequency resolution is $250\,\mathrm{Hz}$
  - Downsampling factor: $r = 24 \Rightarrow$ sampling frequency $f_{\mathrm{s},k} = 1\,\mathrm{kHz}$ in each sub-band
  - The adaptation is not performed in the first 5 sub-bands
  - Number of adaptive filter coefficients: $N = 3$ (in each sub-band)
  - A regularization value equal to $0\,\mathrm{dB}$ SPL is added to the norm $||\tilde{\mathbf{u}}_k(n)||^2$ of Equation (5.1) .
- AGC:
  - Knee point: $c_k = 65\,\mathrm{dB}$ SPL
  - Compression ratio: $c_r = 2$
- FS:
  - Amount of frequency shift: $f_m = 12\,\mathrm{Hz}$
  - Only applied for frequency above $1\,\mathrm{kHz}$
- PEF:
  - PEF order: $M = 1$ (in each sub-band)
- NPVSS:
  - Attenuation: $a = 0.25$
  - Smoothing parameter: $\alpha_1 = 0.1$
- Impulse detection:
  - Fast smoothing parameters: $\beta_r = 0.05$, $\beta_f = 0.005$
  - Slow smoothing parameter: $\gamma = 0.005$
  - Threshold: $T_\mathrm{ID} = 4\,\mathrm{dB}$
- Modified estimate:
  - Smoothing parameter: $\alpha_2 = 0.05$
- Correlation detection:
  - Threshold: $T_\mathrm{CD} = 11\,\mathrm{dB}$
  - Attenuation parameters: $a_1 = 0.25$, $a_2 = 0.5$

5.2 Parameter Optimization

Before the actual evaluation is done, the choice of important parameters, which are given in Section 5.1.2, is motivated in this section. This is done by optimization of the parameters based on the evaluation of the adaptation stability and tracking behavior, which are described in Section 2.4. Hence, the aim is to obtain a good stability and fast tracking at the same time. Since most of the time the parameters influence the adaptation stability and tracking behavior contrarily, the chosen values are a trade-off to get satisfactory results for both.
The name of the non-parametric variable step size (NPVSS) success that no parameters are needed to calculate it. This is true in theory. However, for a realistic application of the NPVSS some parameters are needed. In this thesis, the smoothing parameter α_1 to estimate the signal powers, c.f. Equations (5.4)-(5.6), and the attenuation parameter a, c.f. Equation (5.2), to improve the stability are proposed. Both parameters are optimized in Section 5.2.1. After that the choice of the smoothing parameter α_2, which is needed to calculate $\text{NPVSS}_{\text{mod}}$, c.f. Equation (5.11), is optimized. Finally, the attenuation parameter a_2 used for the correlation detection, c.f. Equation (5.14), is selected.
All parameters are optimized separately of each other. Meaning, one parameter is optimized while the other parameters are set to their fixed standard values. Only the two parameters to calculate the NPVSS α_1 and a are optimized together as they interact. The parameters used for the modified estimate and the correlation detection are influenced by the choice of α_1 and a. However, both methods are proposed as additional methods. Thus, it is reasonable that the choices of α_2 and a_2 are based on the choices of α_1 and a.

5.2.1 Parameters for Variable Step Size

First the parameters to calculate the NPVSS are optimized. These are the smoothing parameter α_1 and the attenuation parameter a. The smoothing parameter α_1 is needed to estimate all quantities of Equation (5.3) as described by Equations (5.4)-(5.6). The attenuation parameter a is multiplied with the step size used in Equation (5.2) to improve the stability of the adaptation. A combination of both parameters is optimized since they influence each other. For the optimization Setting 1 described in Section 5.1.1 is used, which consists of the combination of NPVSS, PEF and FS.
Three different values for a are chosen: 0.15, 0.25 and 0.35. For each value, α_1 is varied between 0.005 and 0.04. All other parameters are chosen as defined in Section

5.1.2. The results of the adaptation and tracking behavior for the different values of a and α_1 are displayed in Figures 5.2-5.4.

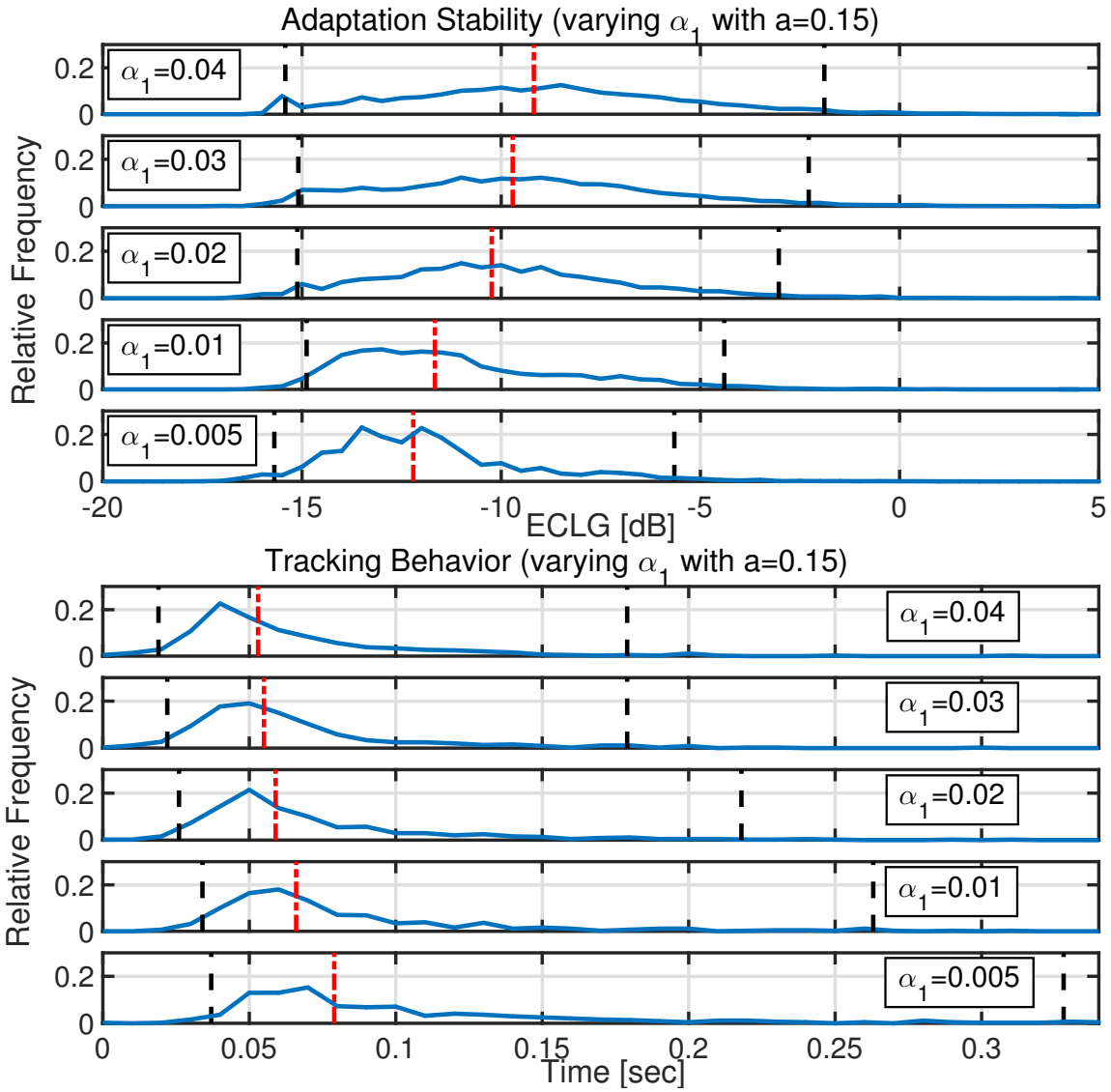


Figure 5.2. The parameter α_1 is varied from 0.005 to 0.04 while $a = 0.15$.

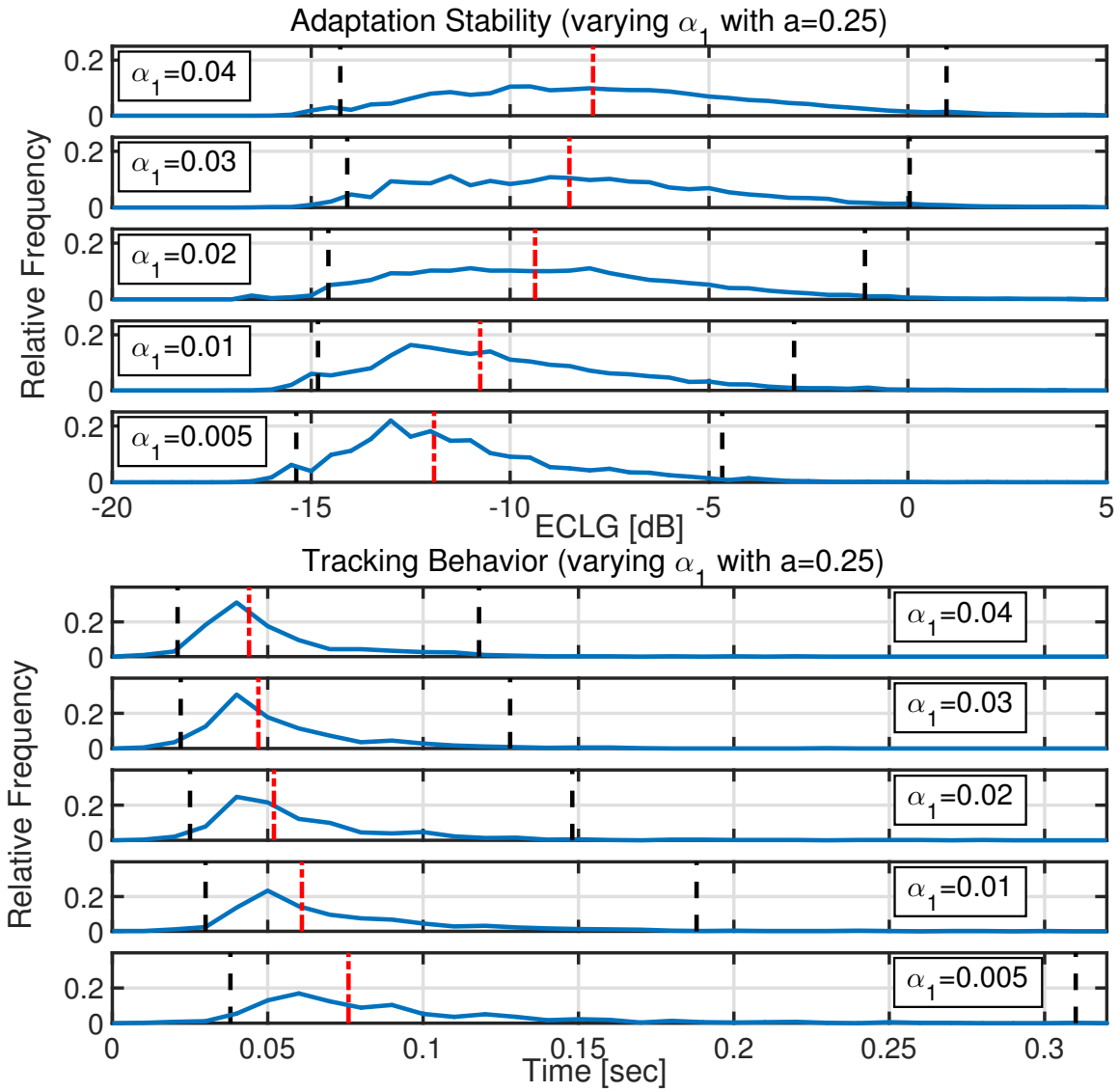


Figure 5.3. The parameter α_1 is varied from 0.005 to 0.04 while $a = 0.25$.

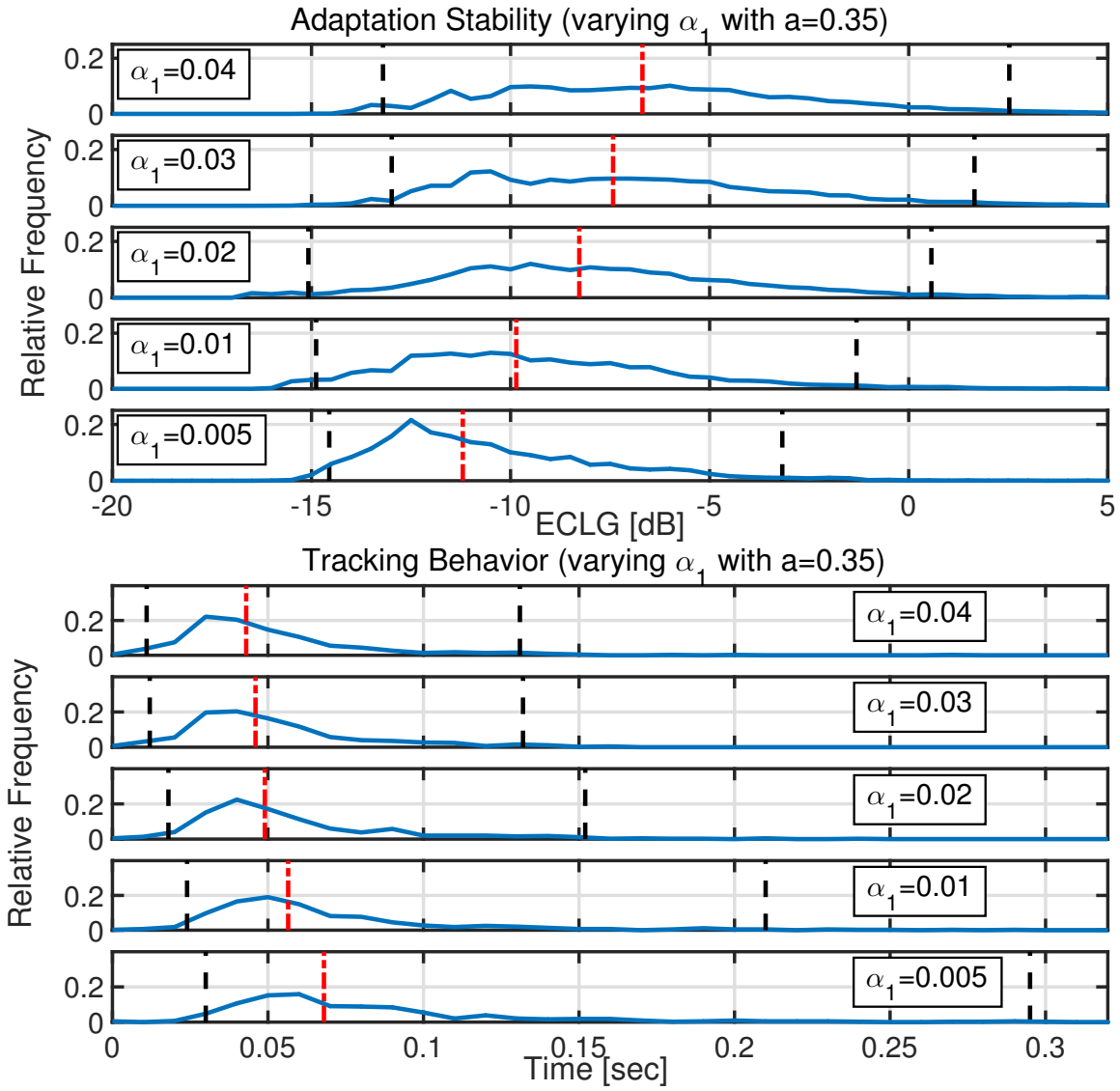


Figure 5.4. The parameter α_1 is varied from 0.005 to 0.04 while $a = 0.35$.

In all three figures, it is noticeable that a lower α_1 leads to a better ability of the system to prevent entrainment. This is because the lower is α_1 the stronger is the smoothing. With stronger smoothing the estimation error is decreased. However, in parallel, the time the systems needs to track a feedback path change increases with decreasing α_1. With a small α_1 the estimate of $\sigma^2_{\tilde{e}_k}(n)$ reacts slowly to changes of the true power of $\tilde{e}_k(n)$ resulting in a decreased adaptation speed.
To ensure that almost no entrainment occurs only the parameter combinations for which less than 1 % of $\mathrm{ECLG}_{\mathrm{max}}$ are above 0 dB should be considered. Since higher values of α_1 increase the adaptation speed, the highest α_1 for each value of a is chosen. Three combinations of the parameters are found:

- 1. Combination: $a = 0.15$ and $\alpha_1 = 0.02$
- 2. Combination: $a = 0.25$ and $\alpha_1 = 0.01$
- 3. Combination: $a = 0.35$ and $\alpha_1 = 0.005$

These combinations are again compared in terms of the resulting adaptations stability and tracking behavior. For a better overview the results for all three combinations are depicted in Figure 5.5.
In terms of the adaptation stability the third combination is the best. However, the differences are only slightly visible due to the median and the tracking behavior of the third combination is significantly worse than for the other two combinations. For the tracking behavior the median values of the first and second combination are almost equal but the upper bound of the CI is higher for the first combination. This suggests that in some cases the adaptation speed is reduced for the first combination. Furthermore, the adaptation stability is also slightly better for the second combination. Hence, the two parameters of the second combination ($a = 0.25$ and $\alpha_1 = 0.01$) are chosen for the proposed AFC system as they provide the best trade-off between adaptation stability and tracking behavior.

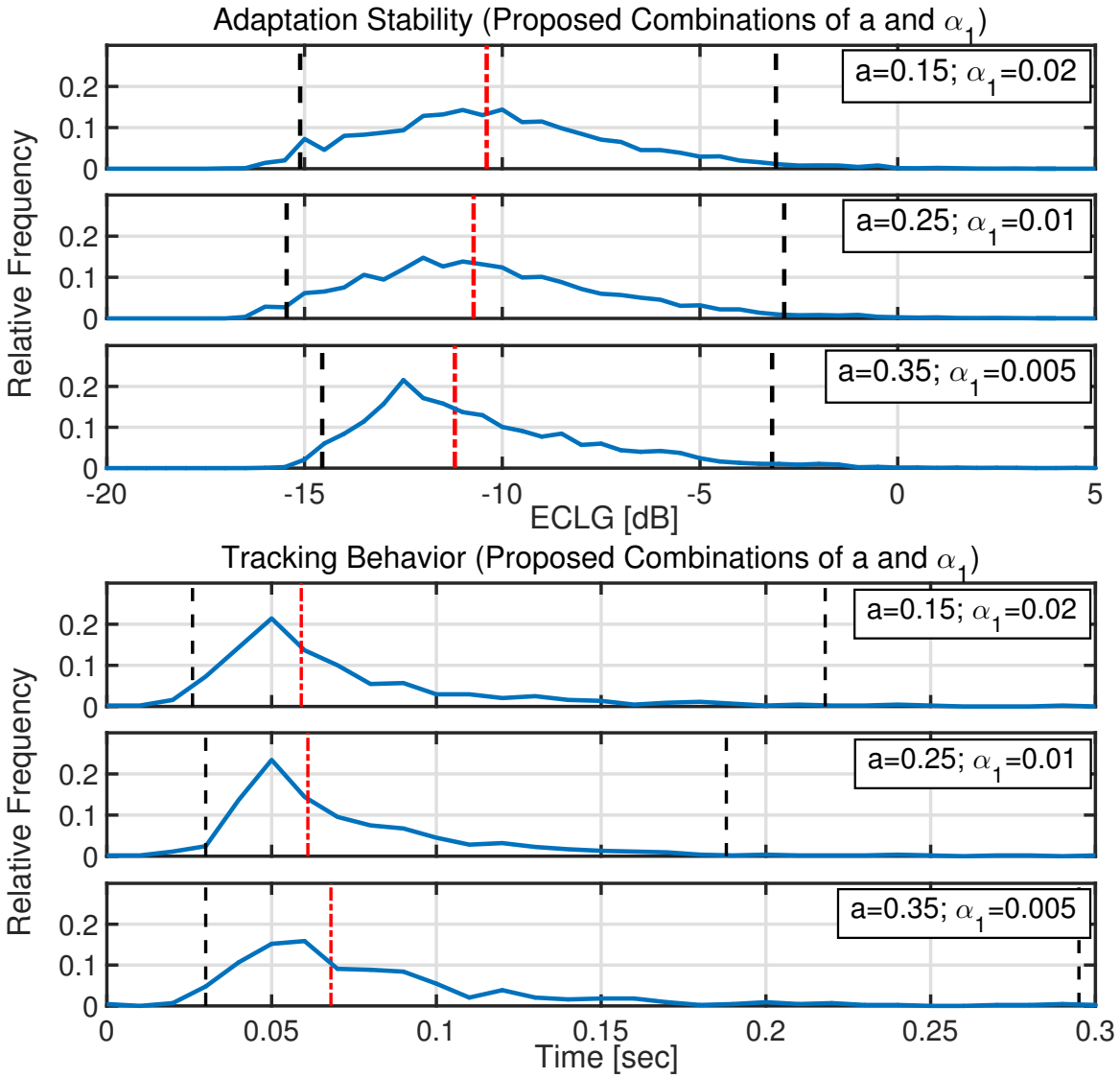


Figure 5.5. Comparison of the 1. combination ($a = 0.15$ and $\alpha_1 = 0.02$), 2. combination ($a = 0.25$ and $\alpha_1 = 0.01$), and 3. combination ($a = 0.35$ and $\alpha_1 = 0.005$). The second combination is chosen as it provides the best trade-off between adaptation stability and tracking behavior.

Additional Investigation of Parameter a

Here, the influence of the parameter a should be examined. Therefore, a is varied between 0.05 and 0.45 while α_1 is set to 0.01. Figure 5.6 shows the results for the adaptation stability and the tracking behavior. As intended a lower value of a results in an improved adaptation stability. This is not surprising as the parameter a attenuates the actual calculated step sizes related with the reduction of the complete update term of each sub-band adaptive filter. Hence, the adaptive filters are less sensitive to correlated input signals. However, the adaptive filters do react slower to feedback path changes and this results in a slower tracking behavior. This is shown by the simulations as well. For the proposed choice of $a = 0.25$ the tracking behavior is good and higher values of a only increase the tracking speed slightly.

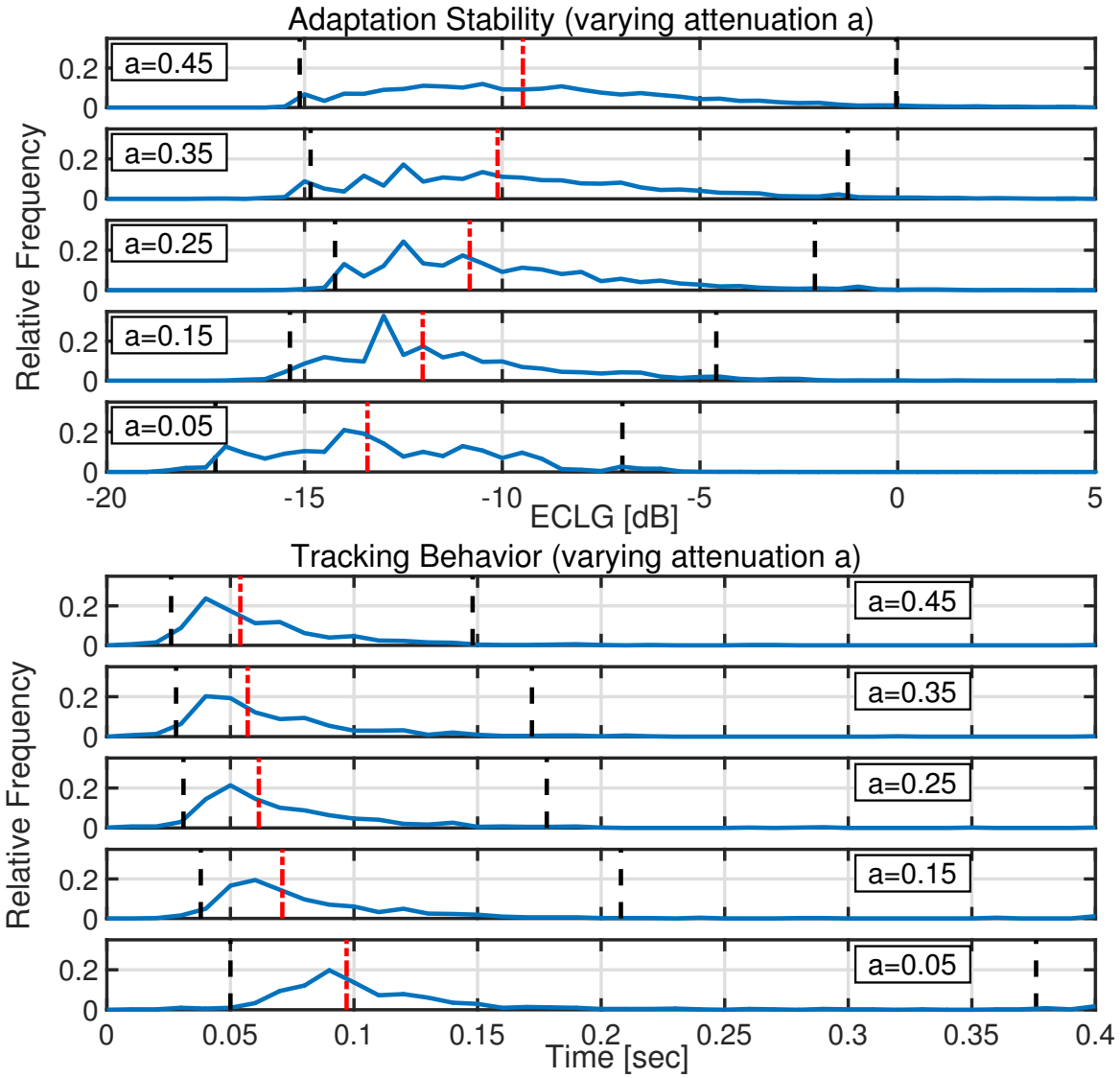


Figure 5.6. The parameter a is varied from 0.05 to 0.45 while $\alpha_1 = 0.01$. A lower a leads to a better adaptation stability, but the ability to track feedback path changes is reduced in this case. The choice $a = 0.25$ provides a very good adaptation stability and an acceptable tracking behavior.

5.2.2 Parameter for Modified Estimate

In the following, the parameter α_2 which is needed to calculate the modified estimate is optimized. Therefore, Setting 3, which includes $\text{NPVSS}_{\text{mod}}$ and is described in Section 5.1.1, is applied.

Smoothing Parameter α_2

The smoothing parameter α_2 is used to calculate the estimate $\hat{\mathbf{R}}_{ee}(n)$, c.f. Equation (5.9). The parameter α_2 is varied between 0.03 and 0.07. Note, that α_2 has to be greater than α_1, which is set to 0.01, in order to have the indented effect on the estimate. Figure 5.7 shows the simulation results for varying α_2. With increasing α_2 the median of the maxima distribution decreases, which indicates an improved adaptation stability. However, for $\alpha_2 \geq 0.05$ the medians are nearly the same. Thus, the improvement of the adaptation stability saturates. The adaptation speed decreases from $\alpha_2 = 0.01$ to $\alpha_2 = 0.06$ after that it increases only slightly. These observations lead to the choice $\alpha_2 = 0.05$ as the entrainment is minimized and the slower adaptation speed is acceptable.

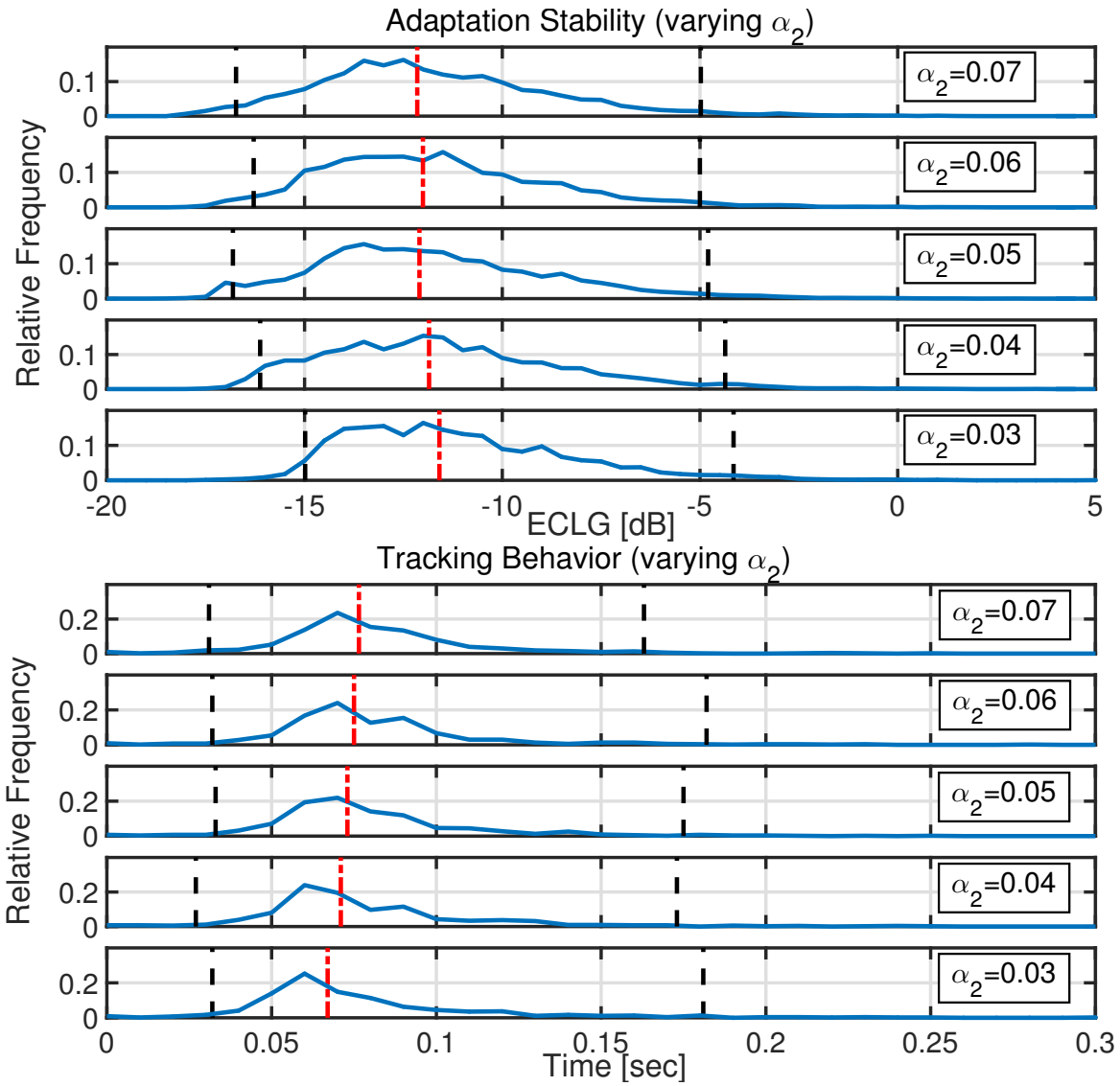


Figure 5.7. Parameter optimization of α_2. The parameter α_2 is varied between 0.03 and 0.07. An increasing α_2 improves the adaptation stability. However, the adaptation stability saturates for $\alpha_2 \geq 0.05$. The tracking of the system gets slower with increasing α_2, but also saturates at $\alpha_2 = 0.06$. Thus, α_2 is set to 0.05.

5.2.3 Parameter for Correlation Detection

Here, one parameter of the correlation detection is optimized, namely the attenuation factor a_2. Therefore, Setting 4, c.f. 5.1.1, is used. Note that in Section 5.1.2 three parameters (a_1, a_2 and T_{CD}) are mentioned. The threshold T_{CD} is optimized in Section 3.4. The attenuation factor a_1 is not optimized since it is chosen to be equal with a, which was optimized in Section 5.2.1. In case of $a_1 < a$ the system would react slower to feedback path changes, but the aim of the correlation detection is to increase the tracking ability. A higher a_1 would of course lead to a faster tracking but also to less adaptation stability similar to increasing a, c.f. Figure 5.6.

Attenuation Factor a_2

The attenuation factor a_2 is multiplied with the NPVSS if the correlation detection decides that the corresponding sub-band signals are not highly correlated. In order to see the influence of a_2 on the adaptation stability and the tracking behavior in simulations it is set to three different values: 0.35, 0.5, and 0.75. Figure 5.8 depicts the results of these simulations. For comparison the results of the simulations using Setting 1 (no correlation detection) and $a = 0.25$ or $a = 1$ are also shown.
A significant increase of the tracking speed can be noticed when the correlation detection is used. However, for values of a_2 equal or above 0.5 the specific value has only a small influence on the tracking behavior. As a drawback of the correlation detection the adaptation stability is influenced negatively. Setting a_2 to 0.5 leads to the best tracking behavior and an excellent adaptation stability as only 2.5 % of the time entrainment occurs. This results in the obvious choice to set $a_2 = 0.5$.

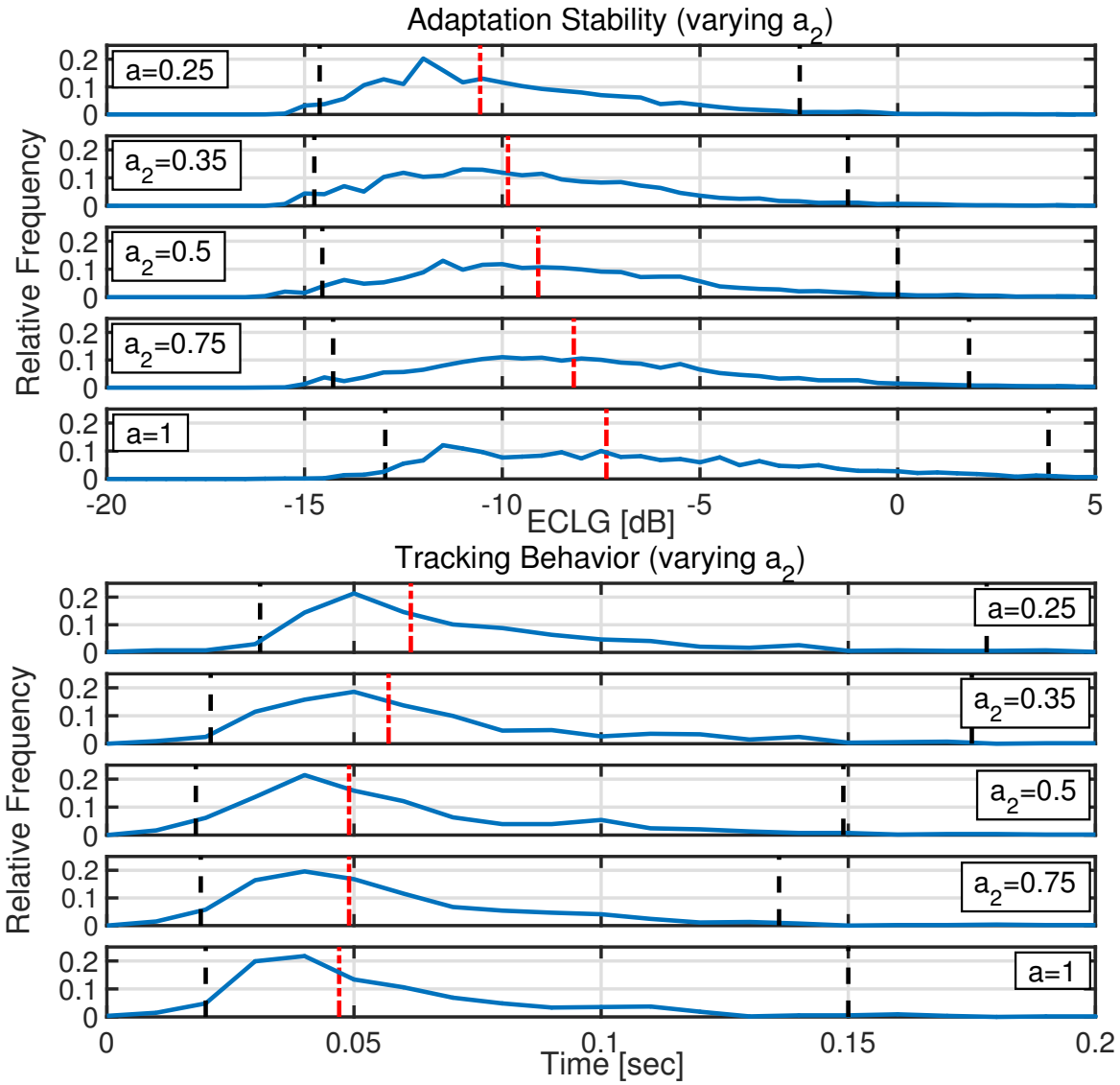


Figure 5.8. Parameter optimization of a_2. The parameter a_2 is varied from 0.35 to 0.75. As comparison the results for $a = 0.25$ and $a = 1$ with correlation detection off are shown as well. A significant increase of the tracking speed can be noticed when the correlation detection is used. However, the tracking speed saturates for $a_2 \geq 0.5$. As a drawback of the correlation detection the adaptation stability is influenced negatively. Setting a_2 to 0.5 leads to the best tracking behavior and a good adaptation stability as only 2.5 % of the time entrainment occurs. Hence, the choice $a_2 = 0.5$ is obvious.

5.2.4 Conclusion

In this section four parameters were optimized based on their effect on the adaptation stability and tracking behavior of the overall system. All other applied parameters mentioned in Section 5.1.2 could be optimized and chosen based on other criteria as described throughout this thesis.
First, the smoothing parameter α_1 and the attenuation parameter a are optimized, which are needed to calculate the NPVSS. The combined optimization of both is necessary to deal with the trade-off between the adaptation stability and the tracking behavior. In consequence, the chosen parameter setting leads to an overall good performance with a focus on the adaptation stability.
The other two parameters, α_2 and a_2, are used for two additional control methods, the modified estimate and the correlation detection. Both optimizations lead to clear favored values.
The investigations in this section allowed to find a complete parameter setting which is used for the following evaluations.

5.3 Performance Analysis of Decorrelation Methods

In this section, the effects of using the decorrelation methods, PEF and FS, in combination with the NPVSS are discussed. The examination in Section 5.3.1 is based on the estimation error of $\sigma^2_{x_k}(n)$. This estimation is necessary in order to make the NPVSS applicable. An optimal estimate leads to an optimal adaptation without any entrainment as demonstrated in Section 3.3.4. That statement is even true without using the decorrelation methods. However, the optimal estimate cannot be calculated in concrete applications as $x_k(n)$ is inaccessible. Therefore, a new estimate with the application of the PEF and FS is derived in Section 3.3.3, which is stated in Equation (5.3). This new estimate is applicable without having access to $x_k(n)$. Of course an estimation error occurs, which leads to a non-optimal behavior of the system. The smaller this error is the better the system performs. Hence, the evaluation of the estimation error is of particular interest.
In section 5.3.2 the effects based on the system performance are evaluated. The system performance is measured by the ECLG. It is shown that the amount of estimation error influences directly the overall system performance.
In both sections Setting 1, c.f. Section 5.1.1, is used but slightly modified as one or

both decorrelation methods are disabled.

5.3.1 Evaluation based on Estimation Error of $\hat{\sigma}_x^2$

The estimation error is defined as $|10\log_{10}(|\hat{\sigma}^2_{\text{true},x_k}(n)|) - 10\log_{10}(|\hat{\sigma}^2_{x_k}(n)|)|$, which is the difference between the true estimate and the proposed estimate of this thesis. To calculate the true estimate the feedback-free input signal $x(n)$ accessible in simulations only, which is divided into sub-band signals, is used. The estimate is given as

$$\hat{\sigma}^2_{\text{true},x_k}(n) = \alpha_1 \cdot |x_k(n)|^2 + (1-\alpha_1) \cdot \hat{\sigma}^2_{x_k}(n-1). \tag{5.16}$$

The proposed estimate $\hat{\sigma}^2_{x_k}(n)$, which does not rely on $x_k(n)$, is defined in Equation (5.3). Note, when the PEF is applied the signals $\tilde{x}_k(n)$ are considered instead of $x_k(n)$. For the evaluation of the estimation error the signals and feedback path models described in Section 2.4 are applied. Also, Setting 1, which is described in Section 5.1.1, is used but the decorrelation methods are disabled. For each simulation, the estimation error is calculated in sub-bands.
All resulting estimation errors of all simulations, which are collected over time and all frequencies, are divided into two groups. The two groups are:

1. Entrainment: Consists of all estimation errors for which entrainment at the same time and frequency, meaning $\text{ECLG}_{\text{dB}}(\Omega, n) > 0\,\text{dB}$, occurs.

2. No Entrainment: Consists of all estimation errors for which no entrainment at the same time and frequency, meaning $\text{ECLG}_{\text{dB}}(\Omega, n) \leq 0\,\text{dB}$, occurs.

The values in dB of the estimation errors for each group are depicted in Figure 5.9 by PDFs together with the median and confidence intervals. The distribution analysis is used for a classification of non-critical errors:
One can clearly see that the estimation errors are differently distributed for both groups. If no entrainment occurs, most values are small, e.g., the 75th percentiles of the corresponding distribution is at 2.3 dB. Most values for which entrainment occurs are significantly higher. Hence, estimation errors below 2.3 dB are defined as non-critical errors.

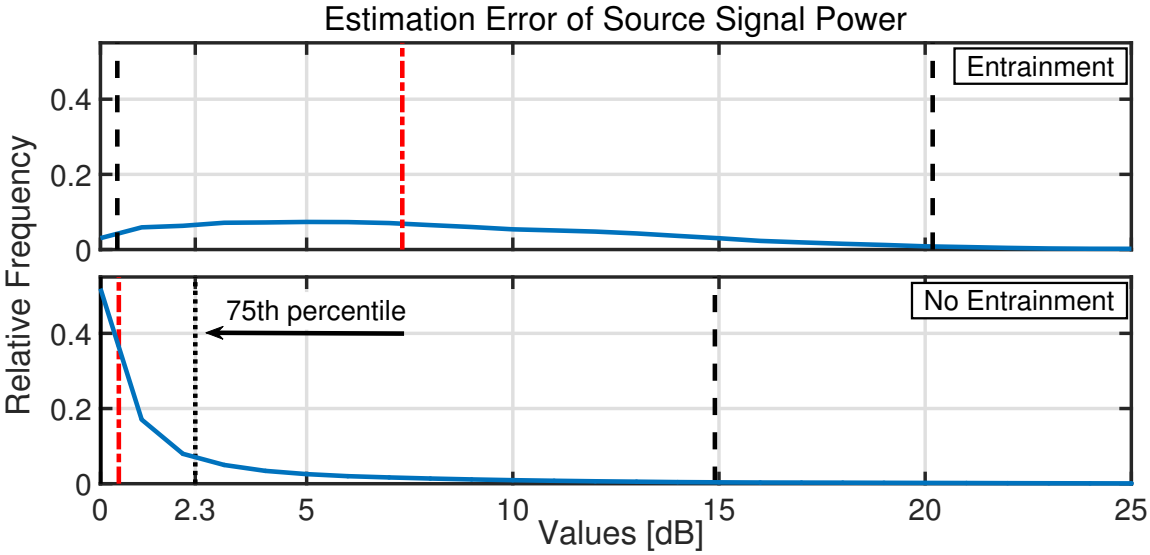


Figure 5.9. Estimation errors are divided into two groups, 'entrainment' and 'no entrainment', and the distributions are illustrated by PDFs. If no entrainment occurs, most values are small, while most values for which entrainment occurs are significantly higher. Hence, estimation errors below 2.3 dB, which is equal to the 75th percentile of the lower PDF, are defined as non-critical errors.

In the following, the percentage of these non-critical errors for different setups of Setting 1 are examined. The different setups concern the decorrelation methods:

1. The NPVSS is used without FS and PEF.
2. The NPVSS is applied with FS.
3. The NPVSS is applied with PEF.
4. The combination of NPVSS, FS and PEF is used.

For each setup the percentage of the non-critical errors for all simulations are calculated and depicted in Figure 5.10.
If the decorrelation methods are turned off non-critical errors occur with about 84 %. When the FS is additional used a percentage of 95 % can be achieved, while it is 90 % for the PEF. With the combination of NPVSS, PEF and FS almost no critical errors are present as 97 % are reached. The FS decreases the estimation error more than the PEF. However, both methods combined with the NPVSS lead to the lowest number of errors.

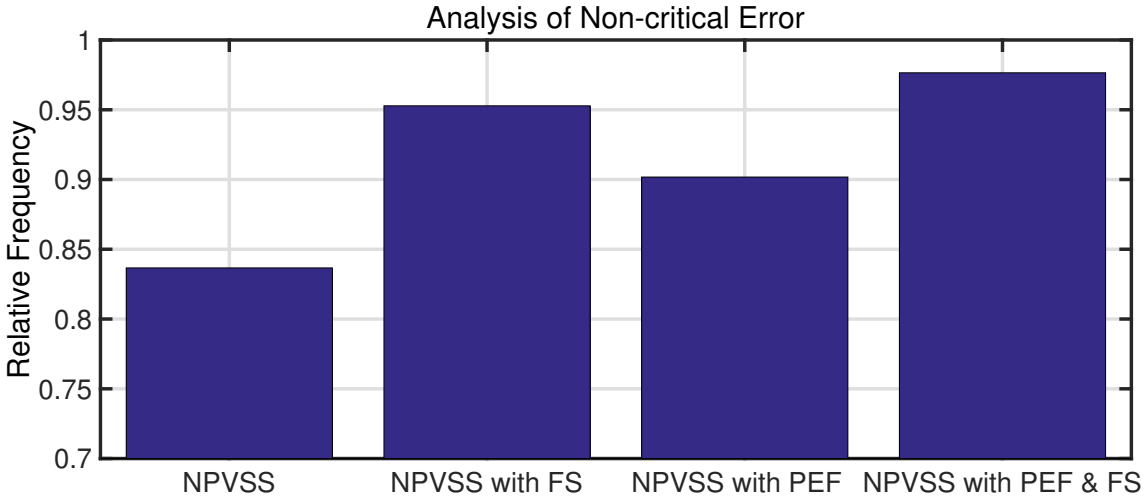


Figure 5.10. The relative frequencies of non-critical errors are shown for four different setups. If the decorrelation methods are turned off non-critical estimation errors of the source signal power occur with about 84 %. When the FS is additional used a percentage of 95 % can be achieved, while it is 90 % for the PEF. With the combination of NPVSS, PEF and FS almost no critical errors are present as 97 % are reached.

In the next section it is shown that this can be directly translated to the system performance as the use of the FS leads to a better adaptations stability than with the PEF. However, with the combination of both methods the stability is maximized.
Since the evaluation of the estimation error behaves qualitatively similar to the evaluation based on the ECLG it can be used to evaluate the real-time system. This is described in Section 5.5.2.

5.3.2 Evaluation based on Effective Closed Loop Gain

In this section the effects of the decorrelation methods on the system performance are evaluated based on the ECLG, c.f. Section 2.4. For the simulations, four different setups of Setting 1 are used:

1. The NPVSS is used without FS and PEF.
2. The NPVSS is applied with FS.
3. The NPVSS is applied with PEF.
4. The combination of NPVSS, FS and PEF is used.

The results of all four settings are combined in Figure 5.11.
Note that for the simulations considering the tracking behavior only white noise is used, since the relatively bad adaptation stability, which is the result when using the NPVSS alone, would bias the tracking behavior in this case. The simulations show that the FS has nearly any effect on the tracking behavior while the PEF makes the tracking slower. This behavior can be explained by the fact that the PF not only models the correlation of the source signal but also the correlation due to the feedback, which is present in $e(n)$ during the time the system needs to adapt to a new feedback path. The same observation is made in Section 3.2.3 when using a CSS.
The analysis in terms of the adaptation stability is more interesting. It shows that the positive influence of the FS is higher than it is the case when using the PEF. This observation is, in contrast to the results obtained in Section 3.2.3, by using CSS. When using CSS the influence of the PEF is much higher compared to the FS. This behavior is traceable to the fact that with the PEF the bias of the adaptation is eliminated more efficiently than with the FS, which is also discussed in Section 3.2.3. However, the simulations of the estimation error of the source signal power, c.f. Section 5.3.1, show that with the FS a better estimate of the source signal power is obtained. With a better estimate a more stable adaptation is obtained as demonstrated by the results in Figures 5.10 and 5.11. Both analyses motivate the combination of PEF and FS as it provides the best behavior. Furthermore, due to the shown connection between estimation error and adaptation performance the analysis of the non-critical error, see Section 5.3.1, can be used to evaluate the performance of the AFC system. This is done in Section 5.5.2 to evaluate the real-time system.

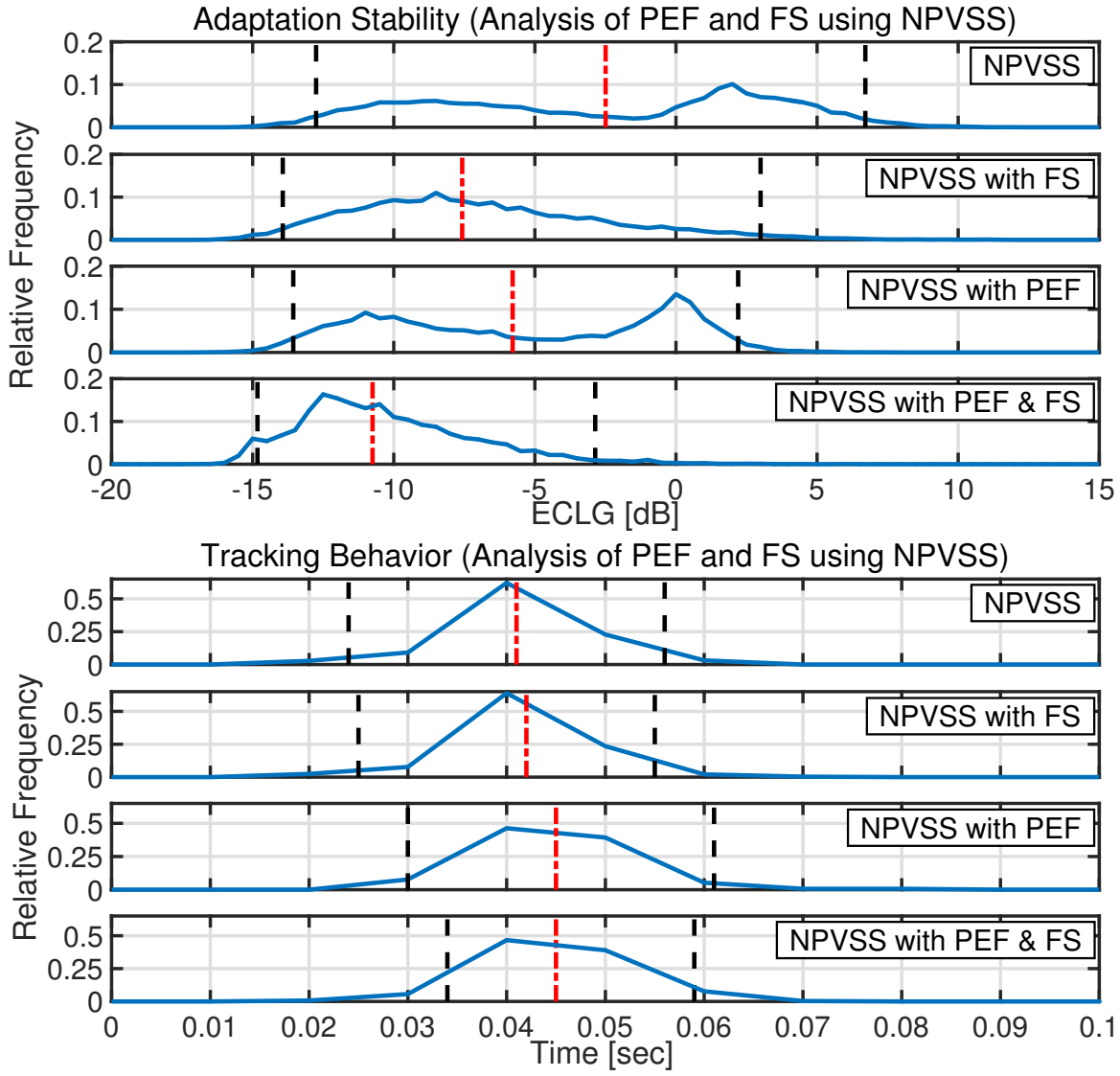


Figure 5.11. Analysis on the influence of PEF and FS in combination with the NPVSS in terms of the adaptation stability and the tracking behavior. The first PDFs of each plot show the NPVSS without FS and PEF. For the second the FS is added. In the third PDFs the NPVSS and PEF are applied. For the last ones the combination of NPVSS, FS and PEF is used. Both decorrelation methods improve the adaptation stability. In contrast to the simulations using CSS, see Figure 3.13 in Section 3.2.3, the benefit of using the FS is higher in comparison to the PEF. However, the combination of both methods shows the best and a very good performance. The tracking behavior is only effected by the PEF as it was the case for the simulations using CSS in Section 3.2.3. But the increase of the tracking speed is only minor.

5.4 Performance Analysis of Additional Control Methods

This section illustrates the effects of the additional control methods, introduced in Section 3.4, based on the evaluation of the adaptations stability and the tracking behavior. The additional control methods are the impulse detection, the modified estimate of the source signal power, and the correlation detection.

5.4.1 Impulse Detection

To evaluate the effect of the impulse detection the applied input signals are slightly modified. A sequence consisting of impulses with an interval of one second between each other is added to the speech and music signals, which are described in Section 2.4. These impulses are recorded at the audio laboratory at the GSC CE and contain for example clapping, knocking on a table, or clapping on a cup with a spoon. The superposition with the impulses is necessary in order to check the effect of the impulse detection. Setting 2 is applied for the simulations and the results are depicted in Figure 5.12.
The analysis based on ECLG shows that the system performs better in terms of the adaptation stability if the impulse detection is used. If the impulse detection is disabled the median of the ECLG distribution and the upper bound of the CI rise. Most important 1.6 % of the values are above 0 dB. With the impulse detection the percentage is decreased to 0.4 %. Furthermore, the result is even slightly improved compared to the results of the system without superposition with the impulses, c.f. Figure 5.11. This indicates that the impulses are detected and the impulse detection ensures that the adaptation stays stable when impulses occur. The impulse detection has a small influence on the tracking behavior. The reduction of the tracking speed occurs if howling, which is present because of abrupt feedback path changes, is misinterpreted as an impulse which happens in very few cases. Nevertheless, the analysis shows the benefit of the impulse detection.

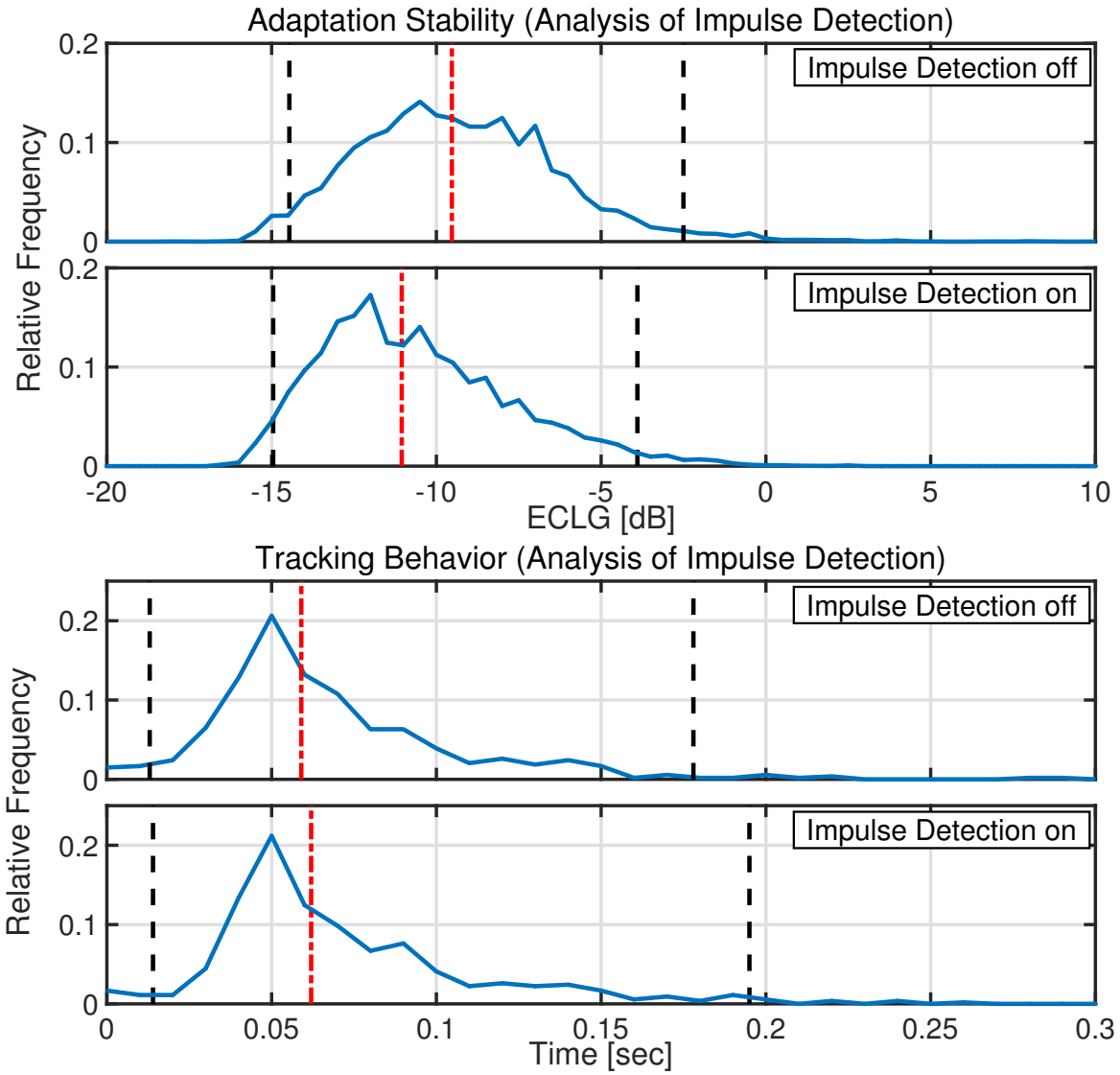


Figure 5.12. Analysis of the effect using the impulse detection. For the simulations recorded impulses are added to the input signals used for the other simulations. The system performs better in terms of stability with activated impulse detection and only a minor influence on the tracking behavior is observed.

5.4.2 Modified Estimate of Source Signal Power

For the evaluation of $\text{NPVSS}_{\text{mod}}$ Setting 3, c.f. Section 5.1.1, is chosen. In Figure 5.13 the results for the adaptation stability and tracking behavior are depicted.
The analysis proves the benefit of the modification. It is shown that with $\text{NPVSS}_{\text{mod}}$ the system performs overall better and the values of the maxima distribution above 0 dB are reduced. As a consequence, also the entrainment is minimized. Certainly, still some values are above 0 dB, but this concerns only 45 of 28170 values in total. Hence, only 0.16% of the time entrainment occurs which is obviously a very low number. A small drawback is the slower tracking behavior, which is observable in the second plot. However, this is only a minor effect since the median is increased by only approximately 6.5 ms and can be tolerated since the improved stability outweighs it.

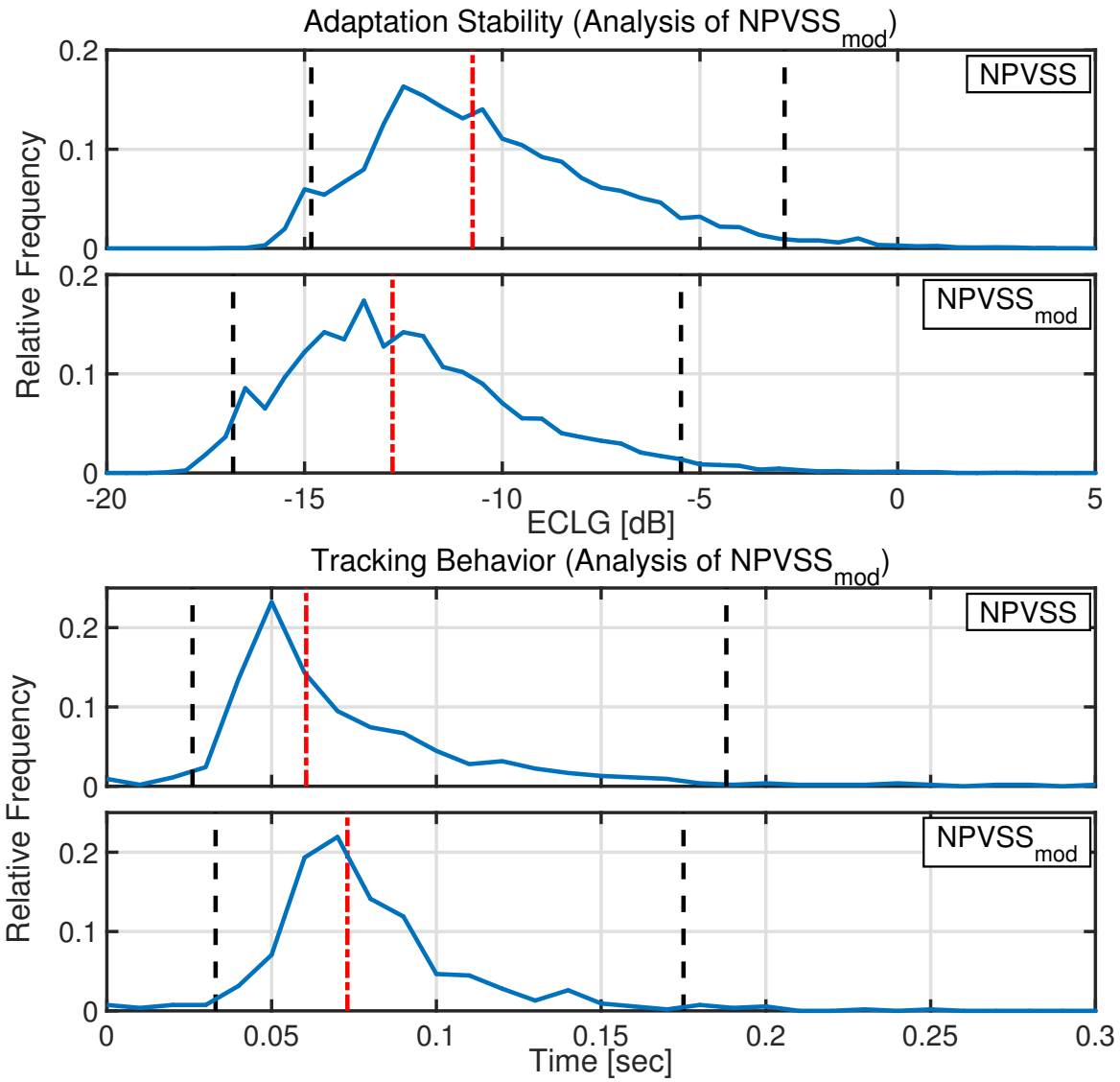


Figure 5.13. Analysis of the effect using $\text{NPVSS}_{\text{mod}}$. In the upper plot, which shows the results for the adaptation stability, it is clear to see that with the modified estimate the stability is increased. However, the lower plot shows that the tracking gets a bit slower. Since the tracking speed is only slightly reduced, it is acceptable.

5.4.3 Correlation Detection

The effect of the correlation detection is evaluated with Setting 4, c.f. Section 5.1.1, and the results are shown in Figure 5.14.
With the correlation detection the systems adapts significantly faster to feedback path changes, which was intended by the implementation of the correlation detection. This can be seen in the lower plot. A side effect is a worse adaptation stability. However, it is still acceptable as only 2.5 % of the values are above 0 dB.
In Section 5.2.1 it is shown that by increasing the attenuation parameter a of Setting 1 a similar behavior is noticeable. A higher choice of a leads to a faster tracking but worse adaptation stability as depicted in Figure 5.6. To demonstrate the difference between varying a and using the correlation detection, in Figure 5.15 the setting without correlation detection (Setting 1) and $a = 0.45$ is compared to the setting with correlation detection (Setting 4).
For Setting 1 with $a = 0.45$ and Setting 4 the adaptation stability is identical. However, the system with the correlation detection adapts faster to feedback path changes. This is noticeable by the mean value which is reduced by 5 ms. Furthermore, the lower bound of the CI is 8 ms lower. The improved tracking ability motivates to use the correlation detection instead of increasing the attenuation factor a.

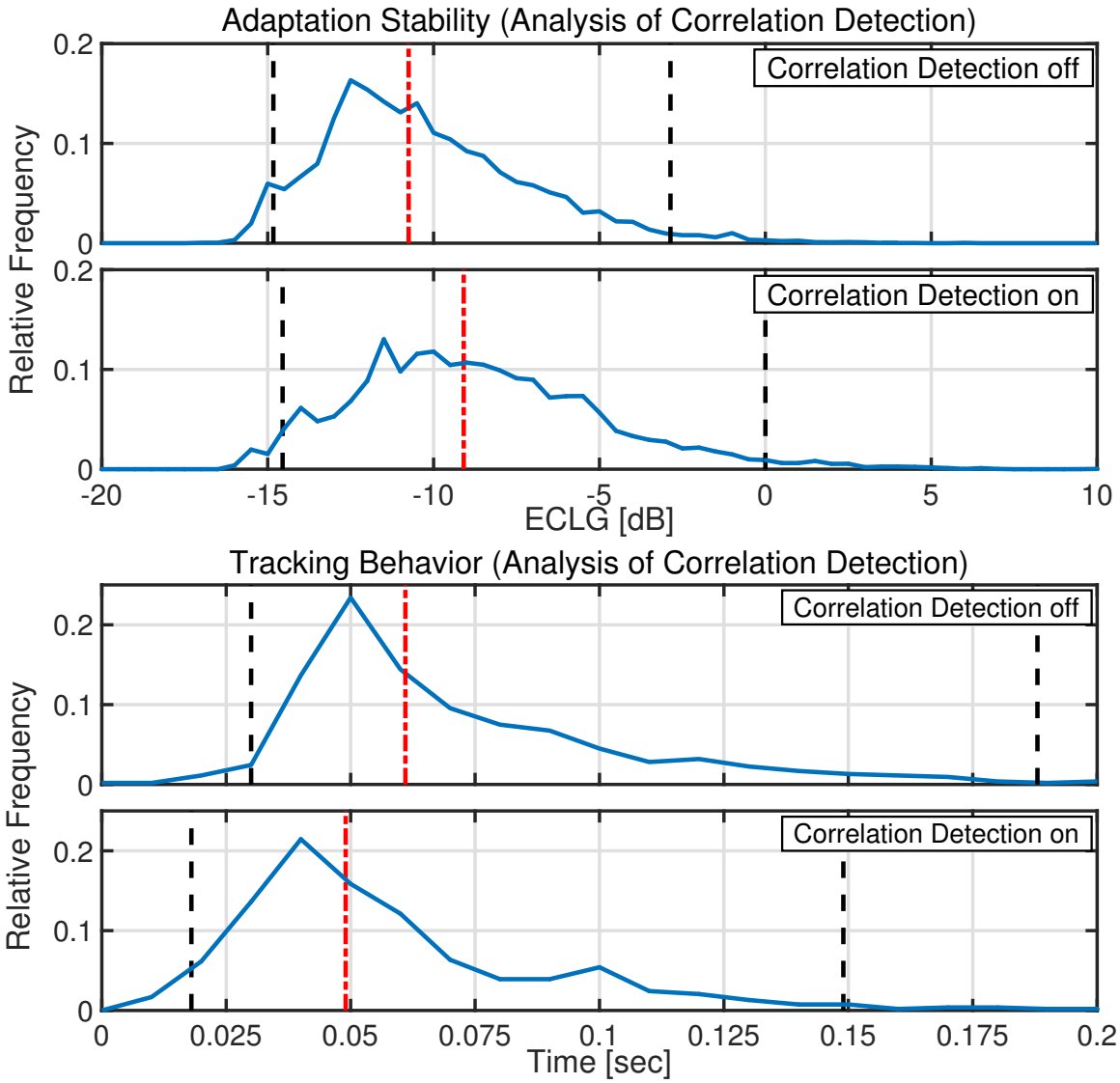


Figure 5.14. Analysis of the effect when applying the correlation detection. The lower plot shows that by using the correlation detection the system adapts significantly faster to feedback path changes. A side effect is a worse adaptation stability, which is still acceptable as only 2.5 % of the values are above 0 dB.

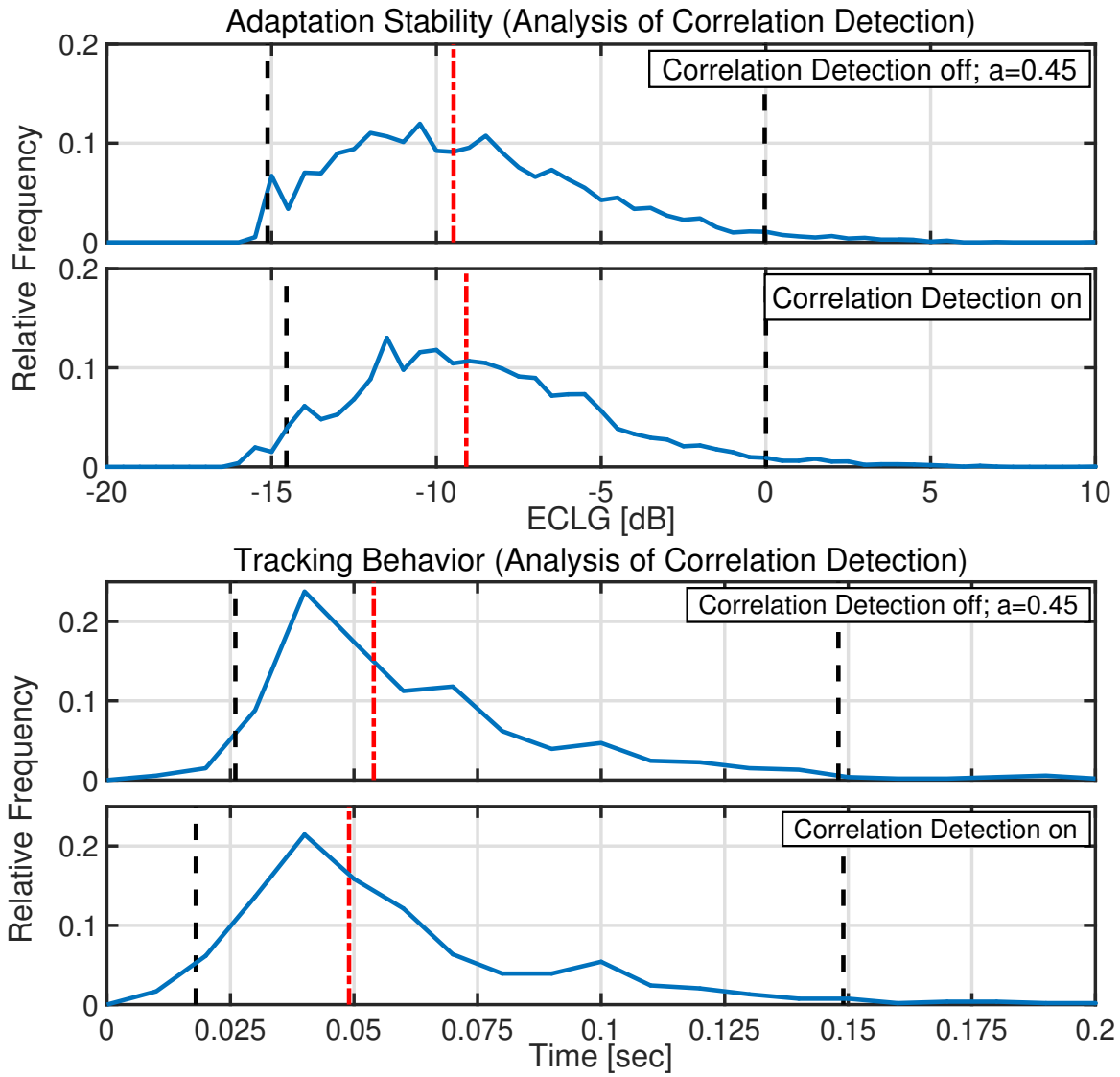


Figure 5.15. Comparison of setting without correlation detection and attenuation parameter $a = 0.45$ and setting with correlation detection. For both settings the adaptations stability is identical. However, the system with the correlation detection adapts faster to feedback path changes. This is noticeable by the mean value which is reduced by 5 ms. Furthermore, the lower bound of the CI is 8 ms lower.

5.5 Evaluation using Real-time System

In the previous sections all proposed methods are evaluated by extensive offline simulations. Here, the methods and the overall AFC system is evaluated on the real-time system, which is described in detail in Chapter 4 along with its benefits compared to offline simulations. These benefits include the possibility to receive immediate results and the large variety of realistic scenarios which can be tested.
In the following some analyses of specific scenarios are demonstrated. After that, the evaluation of the effects of the decorrelation methods based on the estimation error, which occur by estimating the source signal power, is presented. This is similar to the evaluation done with offline simulations in Section 5.3.1.

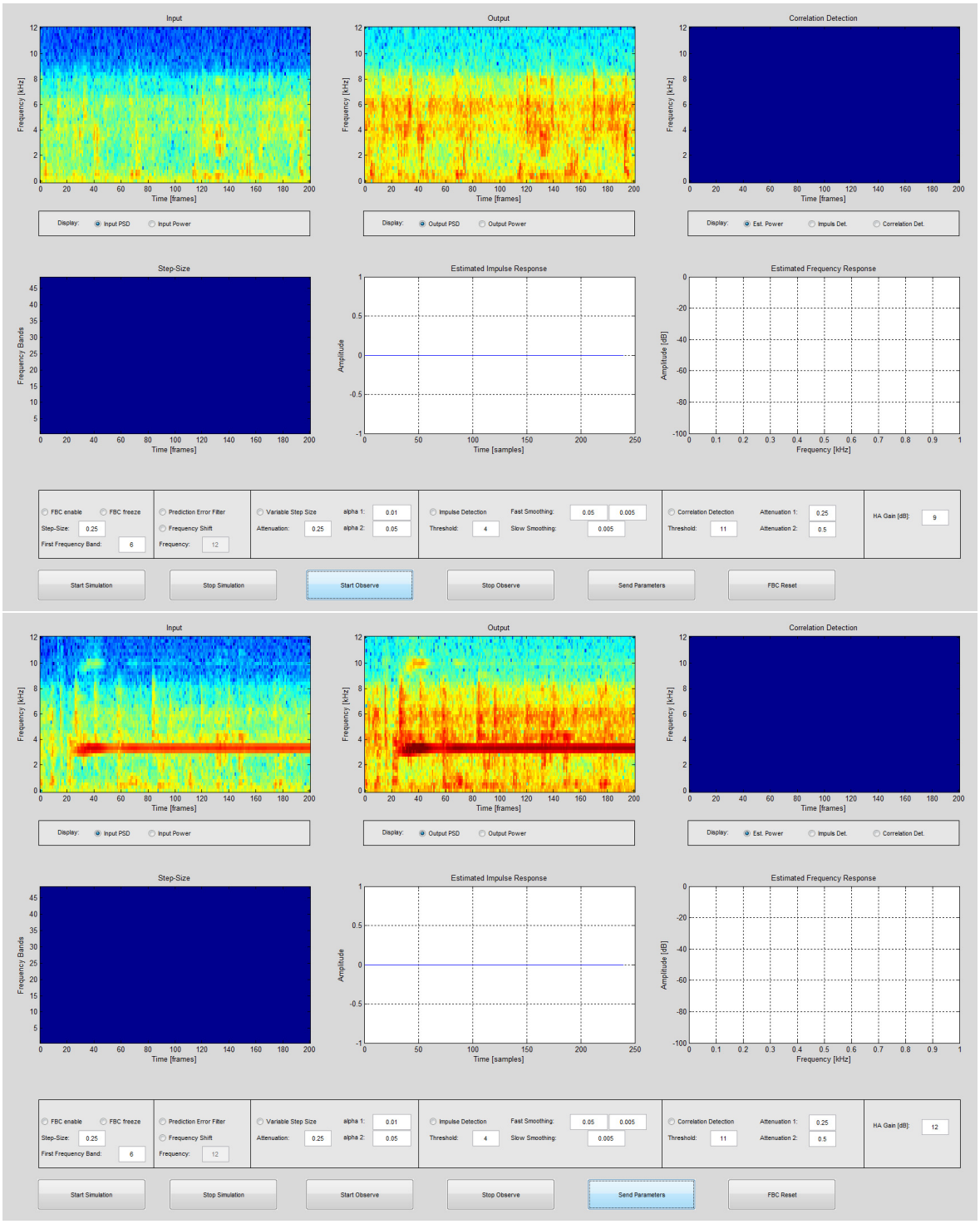
5.5.1 Analyses of Specific Situations

In this section some analyses with the real-time system are presented to demonstrate how the real-time system can be used to evaluate the performance of the AFC system. Therefore, screenshots of the GUI, see Figure 4.4, are illustrated to show specific situations of the analyses. These situations are exemplary chosen and confirm conclusions which are drawn with offline simulations in the previous sections. Hence, it is shown that the performance of the system quantified by offline simulations is also obtained by real-time simulations for a larger variety of realistic test scenarios. For some analyses a sequence of screenshots is shown to reveal the behavior over time. In this case the screenshots are taken in intervals of one second.
In the following the situations and the resulting conclusions are explained.

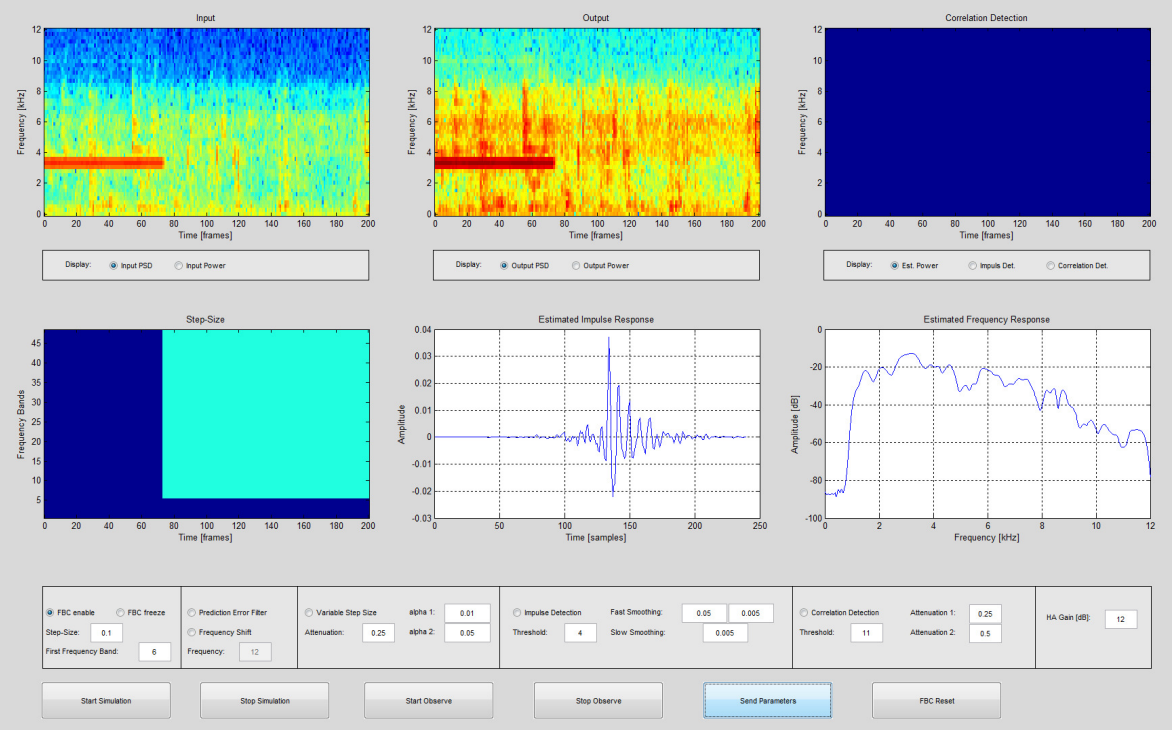
Howling

Firstly, the general functionality of the AFC system to prevent howling is shown by three pictures in Figure 5.16.
For the first picture the FBC is turned off and the hearing aid gain is chosen such that the closed loop gain is non-critical and no howling occurs. During the observation speech is played by the loudspeaker. To generate howling the hearing aid gain is increased by 3 dB. The howling becomes visible in the second picture by the input and output PSDs from about 3 kHz to 4 kHz. To eliminate howling the FBC with a CSS of 0.25 is activated. The plot, showing the step sizes, of the third picture illustrates this. At time frame 73 (no indication in seconds possible since no guaranteed readout speed

of the system) the adaptation is activated and the step sizes of the sub-bands from 6 to 48 increase. No adaptation is performed in the first five sub-bands, c.f. Section 3.1. At the same time the howling disappears, observable by the input and output PSDs. Furthermore, the estimated impulse and frequency response are illustrated.



a) Screenshot 1 and 2

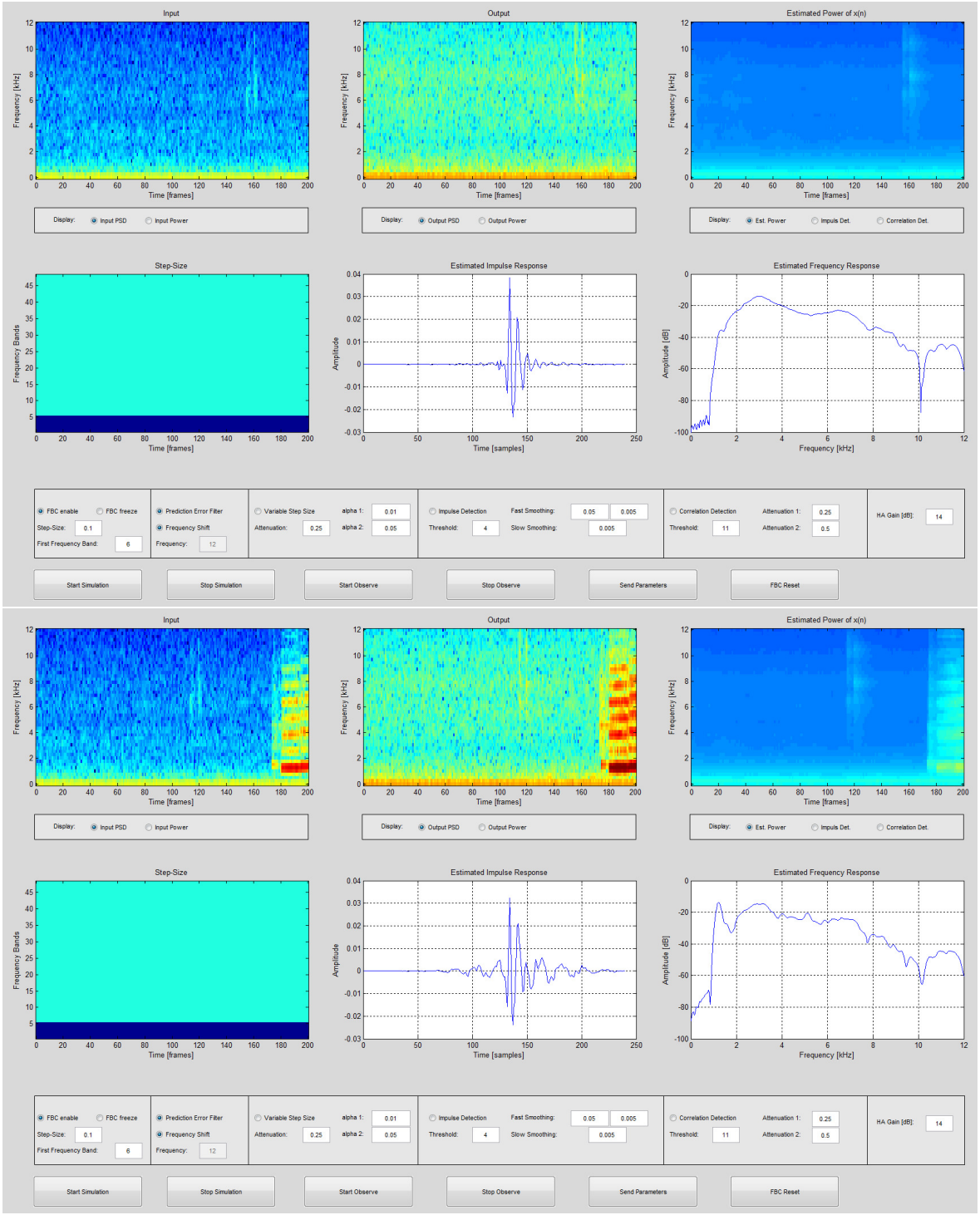


b) Screenshot 3

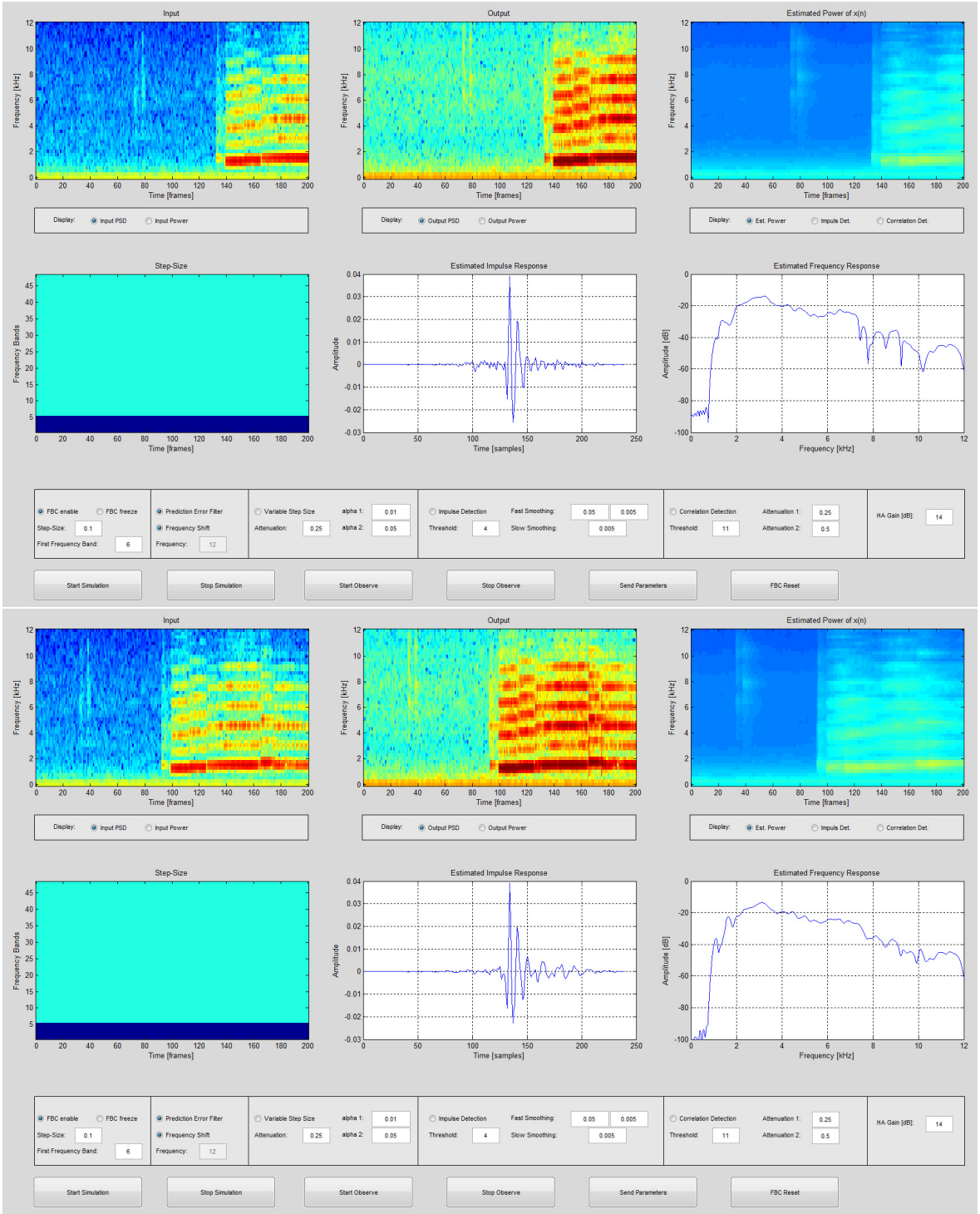
Figure 5.16. Analysis to demonstrate how howling can occur and the AFC system eliminates it. During the observation speech is played. For the first picture the FBC is turned off and the hearing aid gain is too low to generate howling. For the second picture the gain is increased and howling occurs, which is observable by the input and output PSDs. To avoid howling the FBC with a CSS of 0.25 is activated at time frame 73 of the third picture.

Comparison between CSS and NPVSS

This analysis shows the benefit of applying the NPVSS instead of a CSS. In both cases the decorrelation methods PEF and FS are used additionally. To show the difference a flute signal is played by the loudspeaker while the true feedback path is constant.
In Figure 5.17 the analysis using a CSS equal to 0.1 is presented by four screenshots, which show the behavior of the system over time. The estimated feedback path is changing strongly although the original feedback path is constant. The adaptive filters react too strongly to the sound signal, which is an indicator for misadaptation. The misadaptation leads to entrainment which is proven later on.



a) Screenshot 1 and 2

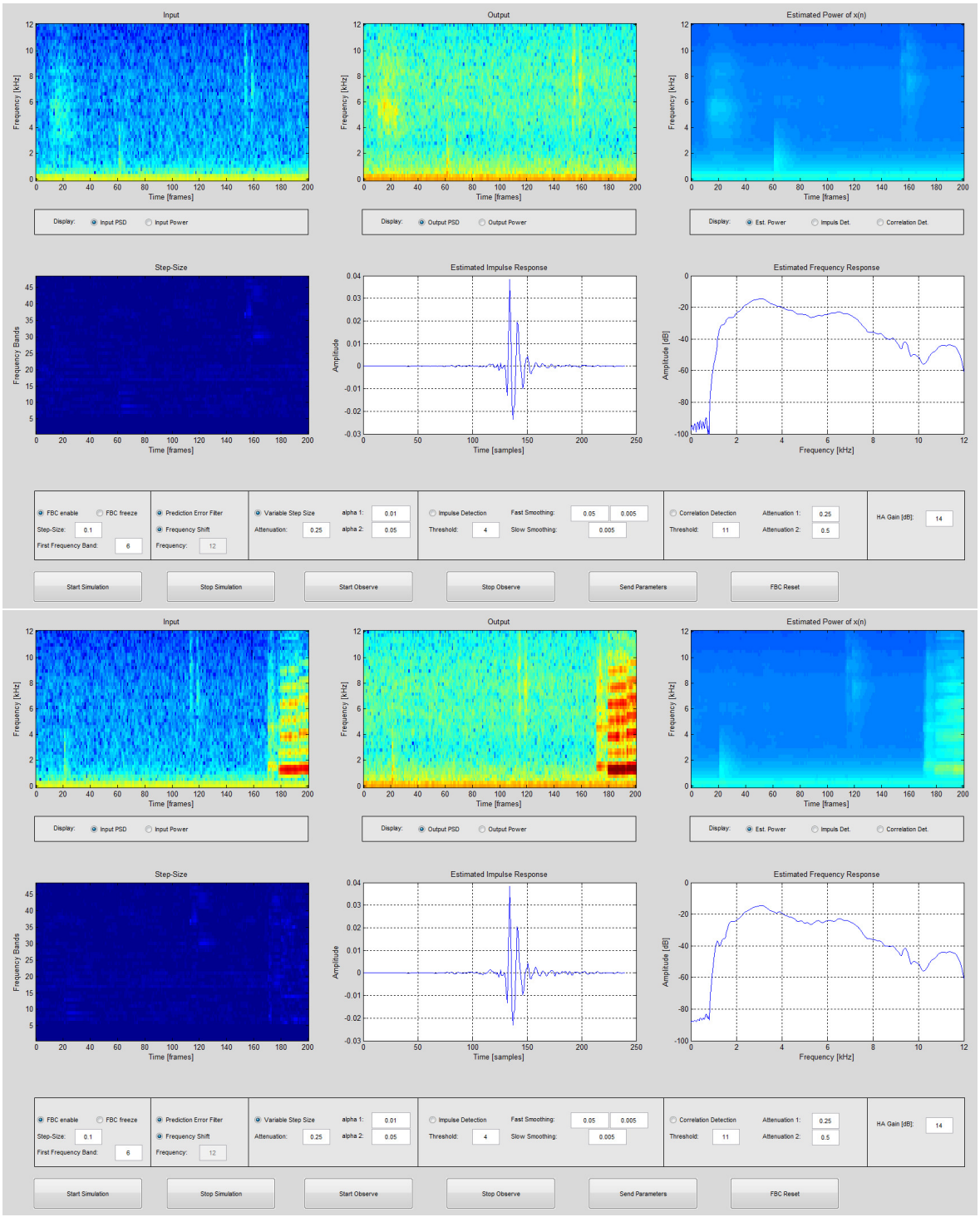


b) Screenshot 3 and 4

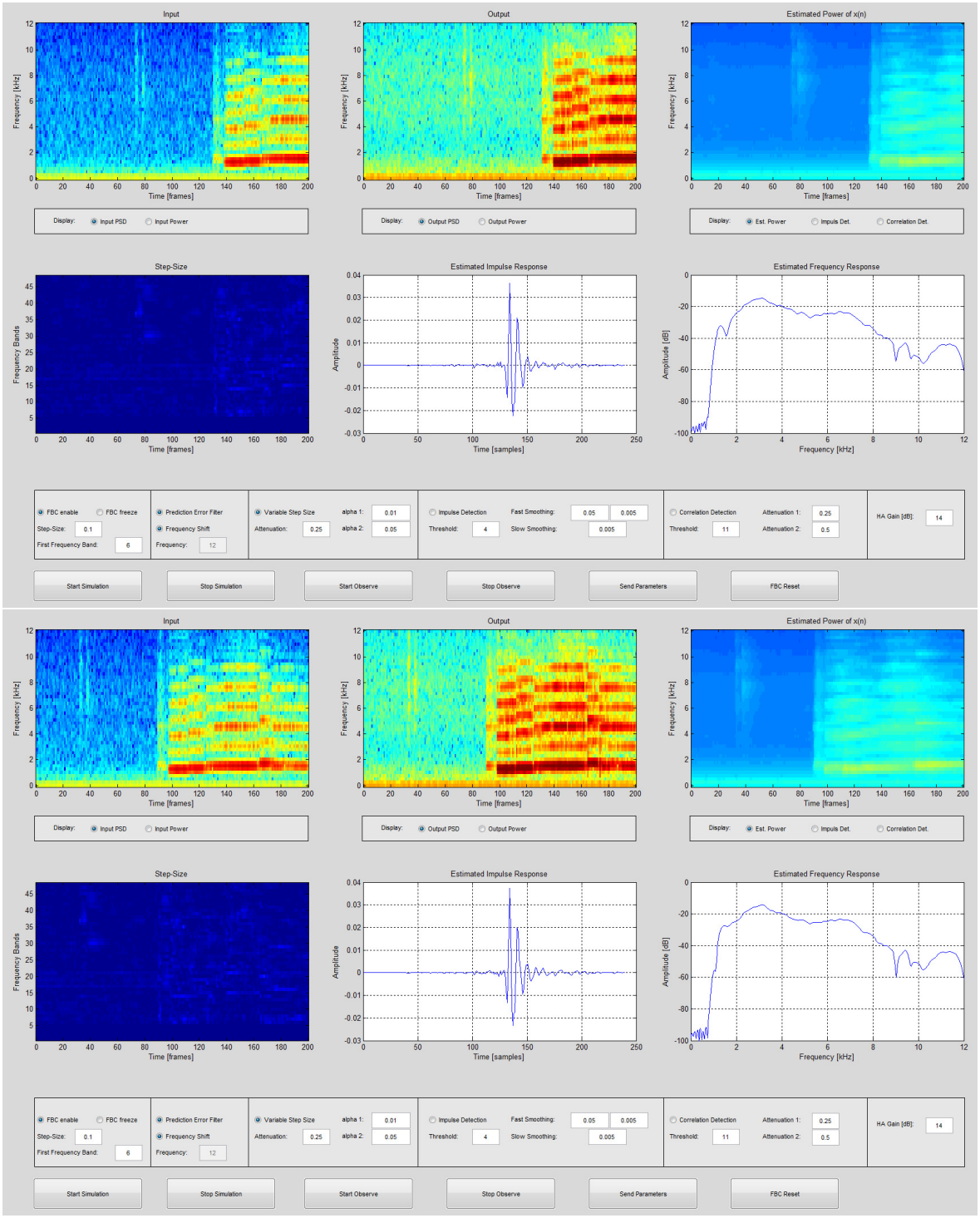
Figure 5.17. Analysis of using a CSS of 0.1 while a flute signal is played and feedback path is constant. The estimated frequency response is changing strongly. This is an indicator for misadaptation, which leads to entrainment.

The screenshots of the analysis for the NPVSS are shown in Figure 5.18. With the help of the NPVSS the estimated frequency response is almost as constant as it should

be. Hence, the adaptive filters stay at the correct estimate of the feedback path. This is also visible by observing the step sizes as their values are almost zero during the analysis. Consequently, no entrainment occurs.



a) Screenshot 1 and 2



b) Screenshot 3 and 4

Figure 5.18. Analysis of using the NPVSS while a flute signal is played and feedback path is constant. The estimated frequency response is almost constant. Consequently, no entrainment occurs.

Since entrainment can be described as artifacts in the audio signal, it becomes visible through differences in the output PSDs of the two analyses. Figure 5.19 shows the output PSD for the analysis with CSS on the left and for the analysis with NPVSS on the right. Some differences representing entrainment are observable in this example. Two obvious ones are marked by black circles.

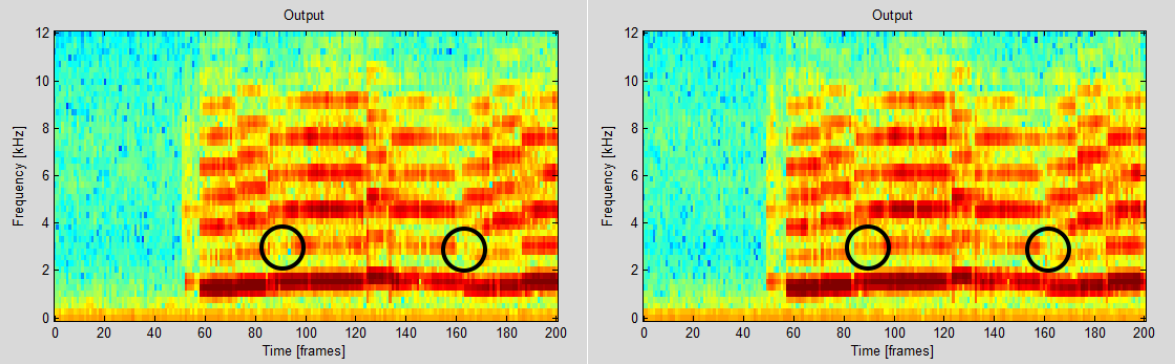


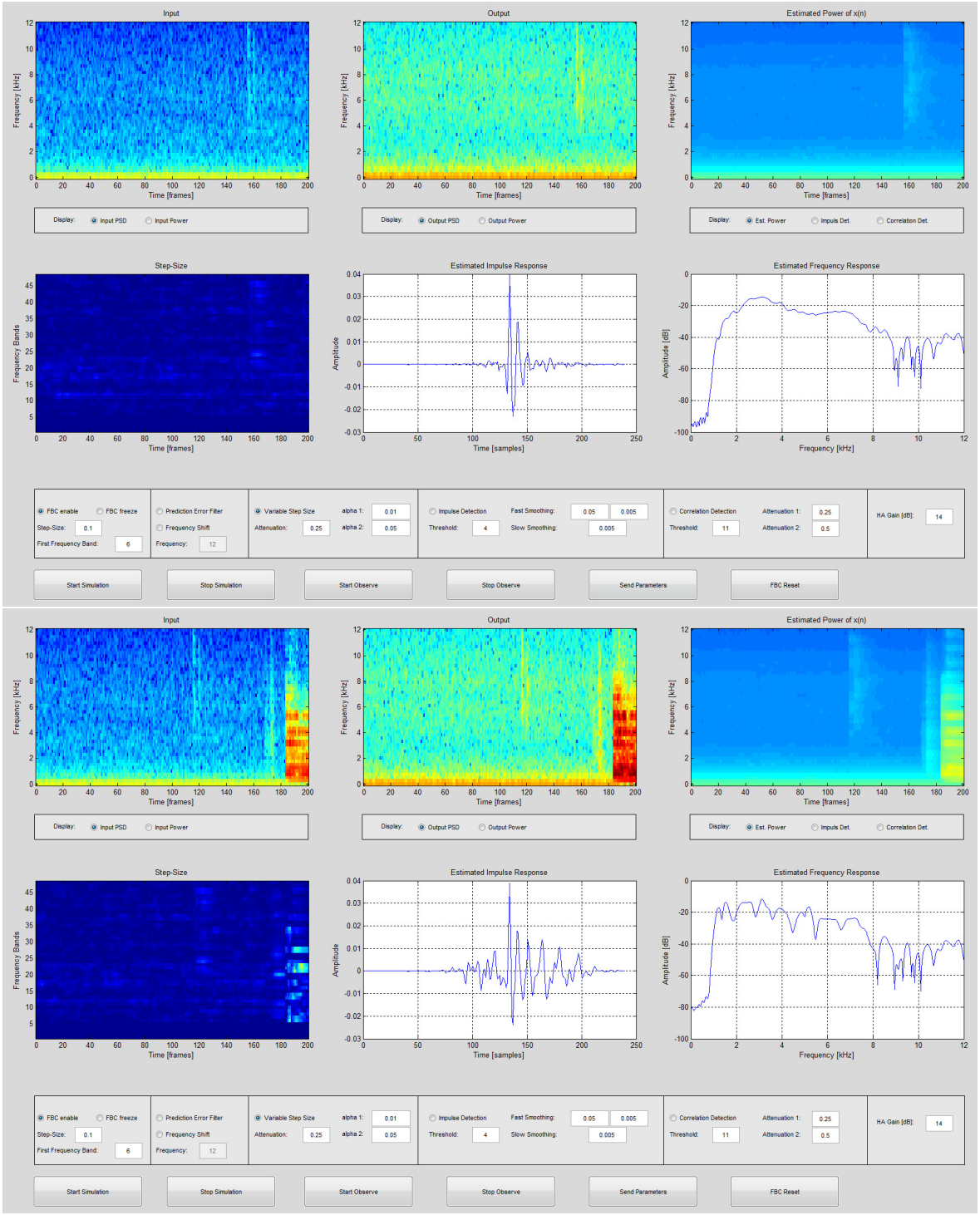
Figure 5.19. Analysis to show entrainment which occurs when a CSS is used. The left picture shows the output PSD for the analysis with CSS. The right picture the output PSD for the analysis with NPVSS. Some differences representing entrainment are observable. The most obvious ones are marked by black circles.

This analysis with the real-time system is equivalent to the simulation depicted by Figure 3.20 in Section 3.3.4. Figure 3.20 demonstrates based on the ECLG that with the NPVSS far less entrainment occurs with the NPVSS than with the CSS. This is confirmed by the example shown with the real-time system in this section.

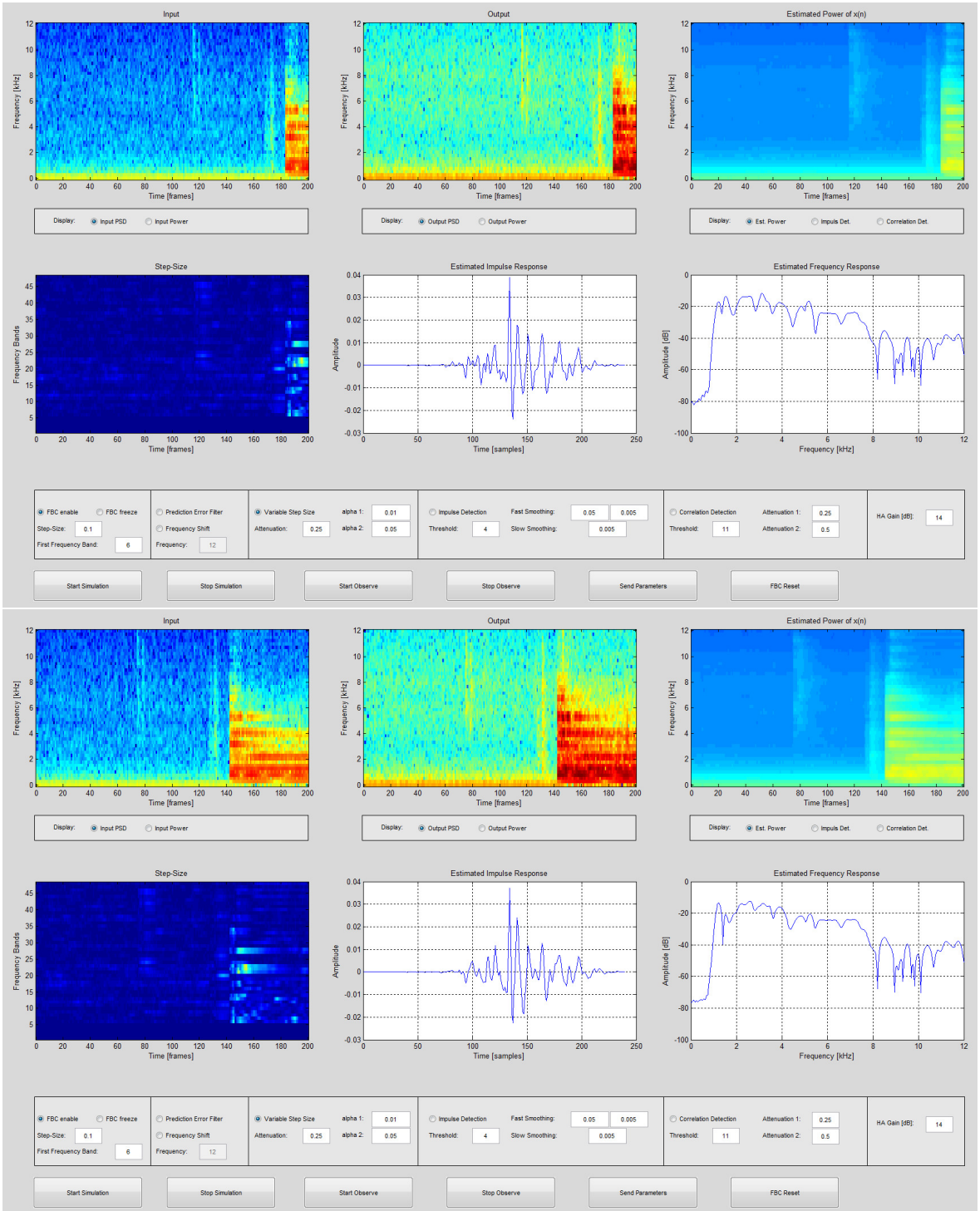
Effects of Decorrelation Methods

In this section, the effects of the decorrelation methods, PEF and FS, are examined with the real-time system. Therefore, the AFC system with the NPVSS is used to cancel the feedback while a bell ringing is played and the original feedback path is kept constant. First the PEF and FS are disabled, the observations are depicted in Figure 5.20, then the signal is played again and this time the decorrelation methods are applied, the result can be observed in Figure 5.21.

In Figures 5.20 four screenshots of the analysis are shown when the decorrelation methods are disabled. By observing the estimated impulse and frequency responses one sees that they are changing significantly over time while the original feedback path stays constant. Hence, the adaptive filters are adapting to the wrong solution and this leads to entrainment. The misadaptation occurs since the step sizes increase rapidly at certain frequencies, which can be observed also in Figure 5.20. The step sizes reach high values since significant errors occur by estimating the source signal power. This is the case as the proposed estimate of the source signal power, described in Section 3.3.3, is only reasonable if the decorrelation methods are used.



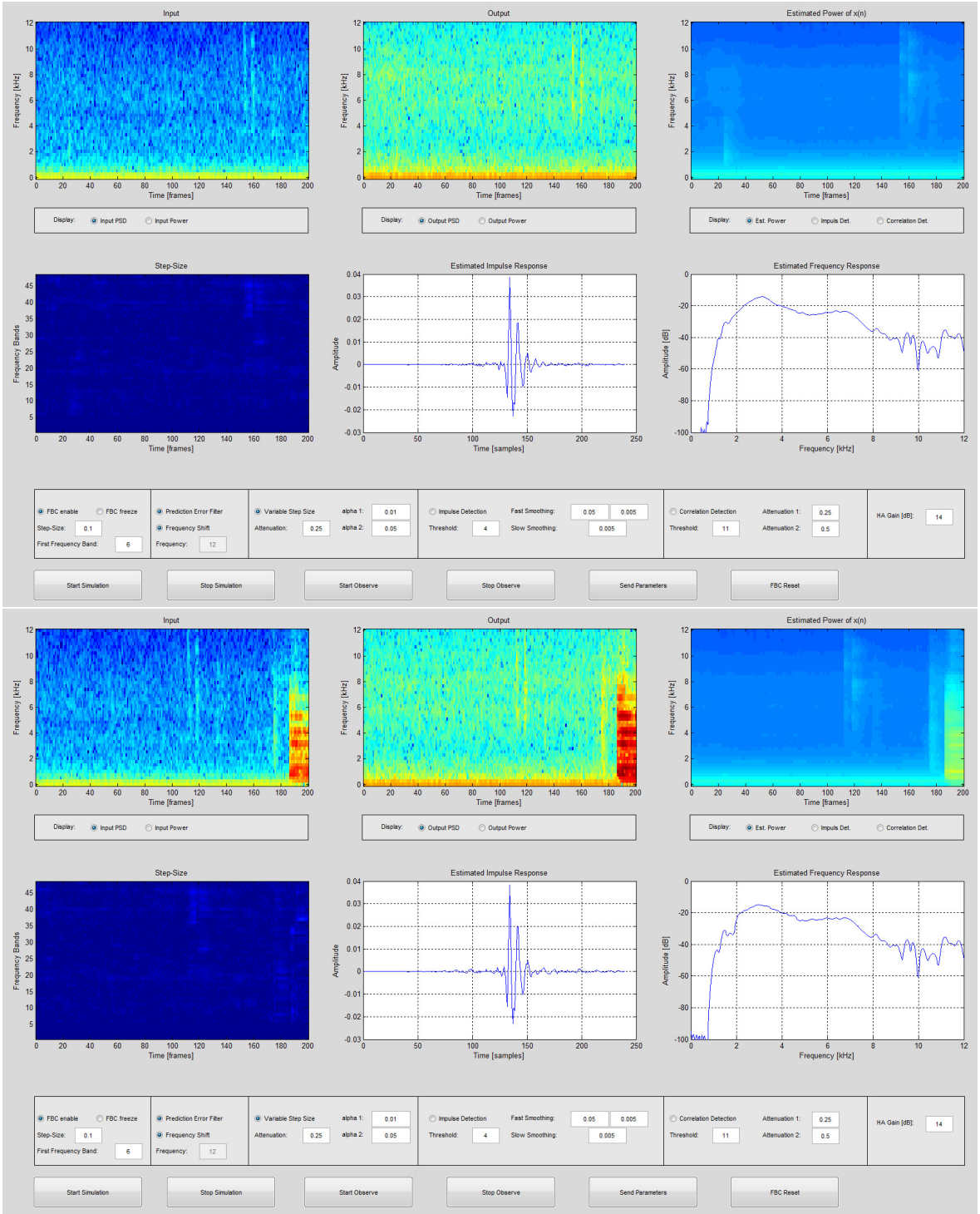
a) Screenshot 1 and 2



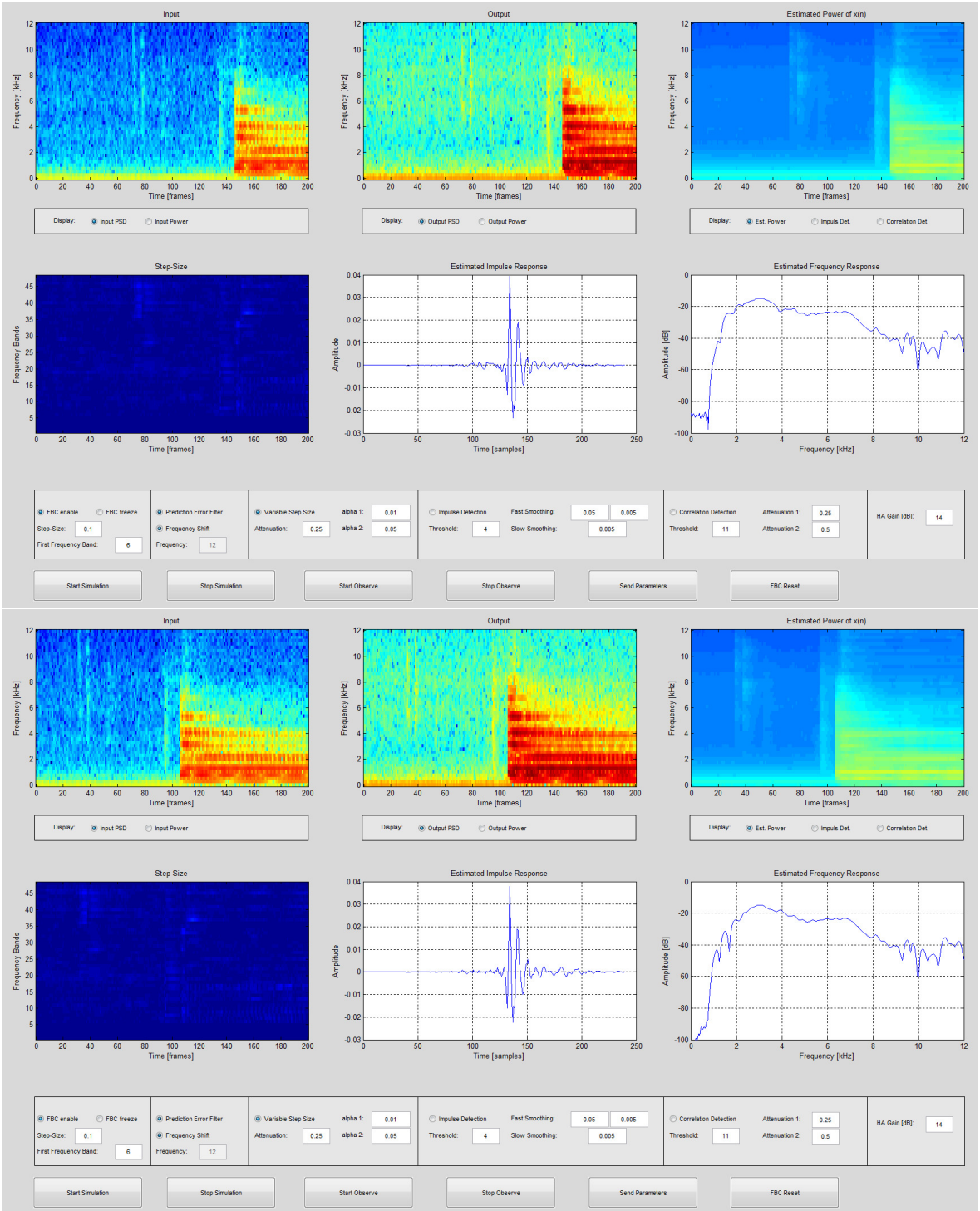
b) Screenshot 3 and 4

Figure 5.20. Analysis of using the NPVSS without PEF and FS while a bell ringing is played and the feedback path is constant. The estimated impulse and frequency responses are changing significantly over time. Hence, the adaptive filters are adapting to the wrong solution and this leads to entrainment.

Figure 5.21 confirms the improved performance with the PEF and FS as they are applied for this analysis. In this case, the step sizes only take small values and consequently, the estimated impulse and frequency responses are almost constant during the analysis. Consequently, no entrainment occurs.



a) Screenshot 1 and 2



b) Screenshot 3 and 4

Figure 5.21. Analysis of using the NPVSS with PEF and FS while a bell ringing is played and the feedback path is constant. In this case, the step sizes only take small values and consequently, the estimated impulse and frequency responses are almost constant during the analysis. Consequently, no entrainment occurs.

As for the analysis before the entrainment is visible in the output PSDs. In Figure 5.22 the output PSD without decorrelation methods on the left is compared to the output PSD with decorrelation methods on the right.

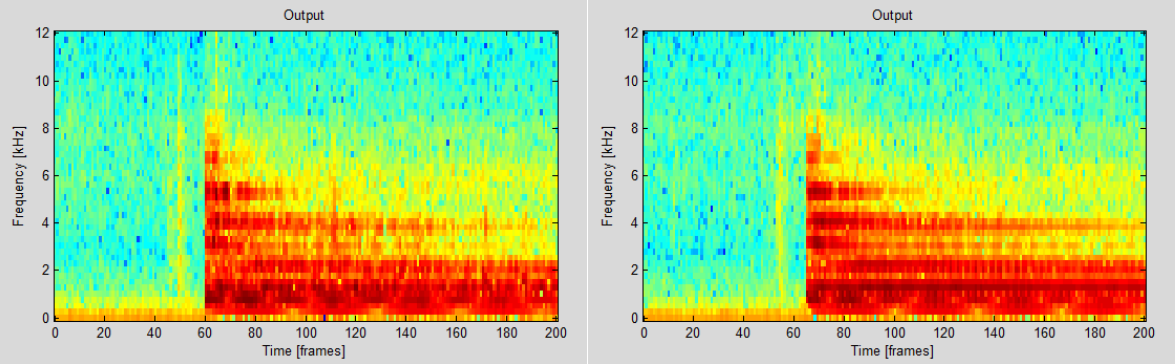


Figure 5.22. Analysis to show entrainment which occurs when no decorrelation methods are applied. The left picture shows the output PSD for the analysis without decorrelation methods. The right picture the output PSD for the analysis with decorrelation methods. For this analysis the entrainment is clearly visible.

For this analysis the entrainment is clearly visible. The results of the corresponding simulations, which reveal the benefit of using the decorrelation methods based on the ECLG, are shown in Section 5.3.

Effect of Impulse Detection

The effect of the impulse detection is shown in this section. While the real-time system is running impulses are generated by clapping. The feedback path is estimated by the use of NPVSS, PEF and FS. First the impulse detection is disabled. Then the analysis is repeated with enabled impulse detection. Figure 5.23 shows screenshots for both analyses. Here, the influence function of the impulse detection along with the threshold are displayed in the right upper plot of each screenshot.

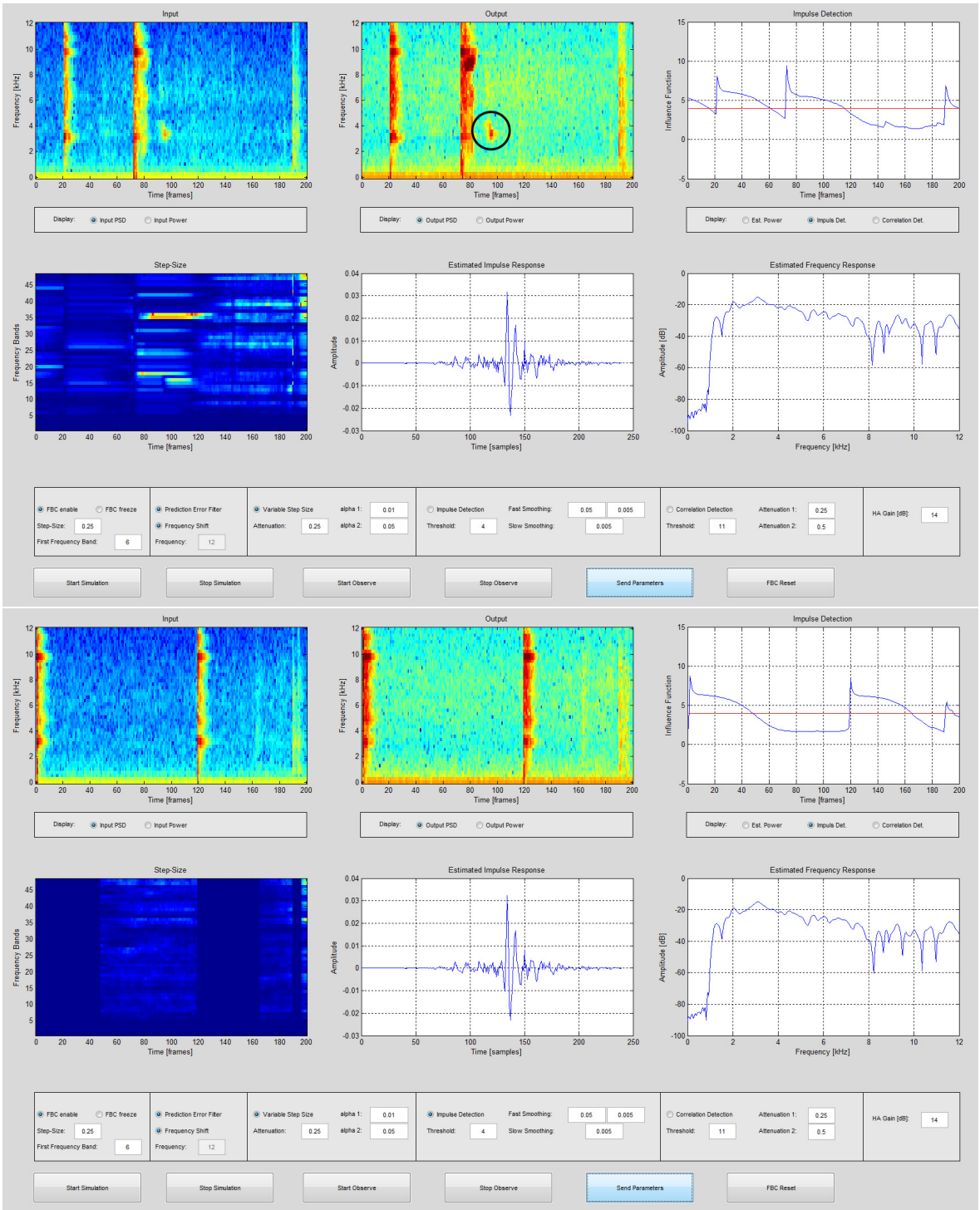


Figure 5.23. Analysis of the benefit of the impulse detection. Therefore, impulses are generated. Upper picture illustrates the analysis without the impulse detection. The step sizes reach high values immediately after the impulses and entrainment occurs, which can be seen in the output PSD (marked with a black circle). For the lower picture the impulse detection is enabled. Here, the step sizes are set to zero at the moment the impulses occur and with that the entrainment is avoided.

The difference between both pictures is clearly noticeable. The upper picture illustrates the analysis without the impulse detection. Therefore, the step sizes reach high values immediately after the impulses. Due to the high step sizes the system adapts to the wrong solution and entrainment occurs, which can be seen in the output PSD. By using the impulse detection, which is done for the analysis shown by the lower picture, the step sizes are set to zero at the moment the impulses occur and stay there long enough so that the impulses have no influence on the adaptation. Thus, no entrainment occurs. Furthermore, in both pictures the indicator function $I_{\text{comp}}(n)$ of the impulse detection is shown in the upper right plot. The indicator function increases strongly and rises above the threshold when the impulses occur. After the impulses the indicator function decreases slowly. This is important to insure the stability as the high values of the step sizes after an impulse can exceed the time of the impulse, which can be observed in the upper picture.
In Section 5.4.1 the impulse detection is analyzed with offline simulations. These simulations confirm the observation described in this section.

Analysis of Correlation Detection

The following analyses examine the functionality of the correlation detection. More precisely, the performance of the indicator function is analyzed. This is done by two tests: The first one shows how the indicator function can differentiate between highly correlated signal components and non-correlated signal components. The second one indicates that a feedback path change decreases the indicator function.
Both properties are very important for the functionality of the correlation detection. As described in detail in Section 3.4.3 the aim of this method is to increase the tracking speed without a negative influence on the adaptation stability. Therefore, the attenuation factor used to calculate the NPVSS is increased when less correlated components of the source signal are present and the risk of entrainment is lower. For highly correlated signal components this attenuation factor is reduced to prevent entrainment. The attenuation factor should not be alternated if the feedback path changes. This is ensured by low values of the indicator function in case the feedback path is changing.
To illustrate the functionality of the indicator function to differentiate between highly correlated signal components and non-correlated signal components white noise is played during the analysis followed by a flute signal. The observation is shown in Figure 5.24.

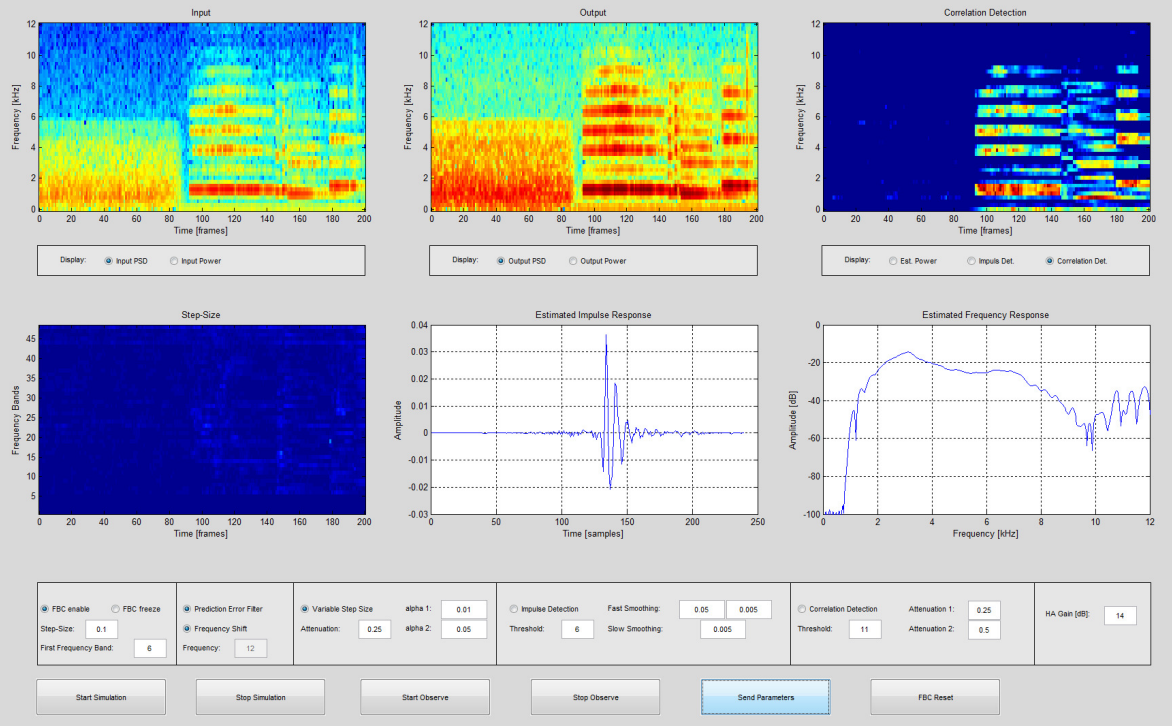


Figure 5.24. Analysis to demonstrate how the indicator function differentiates between highly correlated signal parts and non-correlated signal components. Therefore, white noise is played followed by a flute signal. The right upper plot shows the values of the indicator function. All values below the threshold of 11 dB are illustrated in dark blue. It is clear to see that the indicator function takes only values above 11 dB when the source signal is highly correlated.

The right upper plot shows the values of the indicator function for each sub-band over time. All values below the threshold of 11 dB are illustrated in dark blue. It is clear to see that the indicator function takes only values above 11 dB when the source signal is highly correlated.

To verify that the indicator function decreases in case the feedback path changes the same audio signal, a wind chime, is played twice. The first time the feedback path is kept stable and the second time the feedback path is changed. Then the differences of the indicator function of both analyses are examined. The result is illustrated in Figure 5.25.

The upper picture shows the observation of the analyses without a feedback path change. For the lower one the feedback path is varied, which is visible by the increased step sizes during the analysis. The comparison of the upper and lower pictures reveals that the values of the influence function decrease when the feedback path is changing. Hence, the two tests demonstrate that the influence function performs as intended and can be used for the correlation detection. In the next section, the effect of the correlation detection on the tracking behavior is indicated.

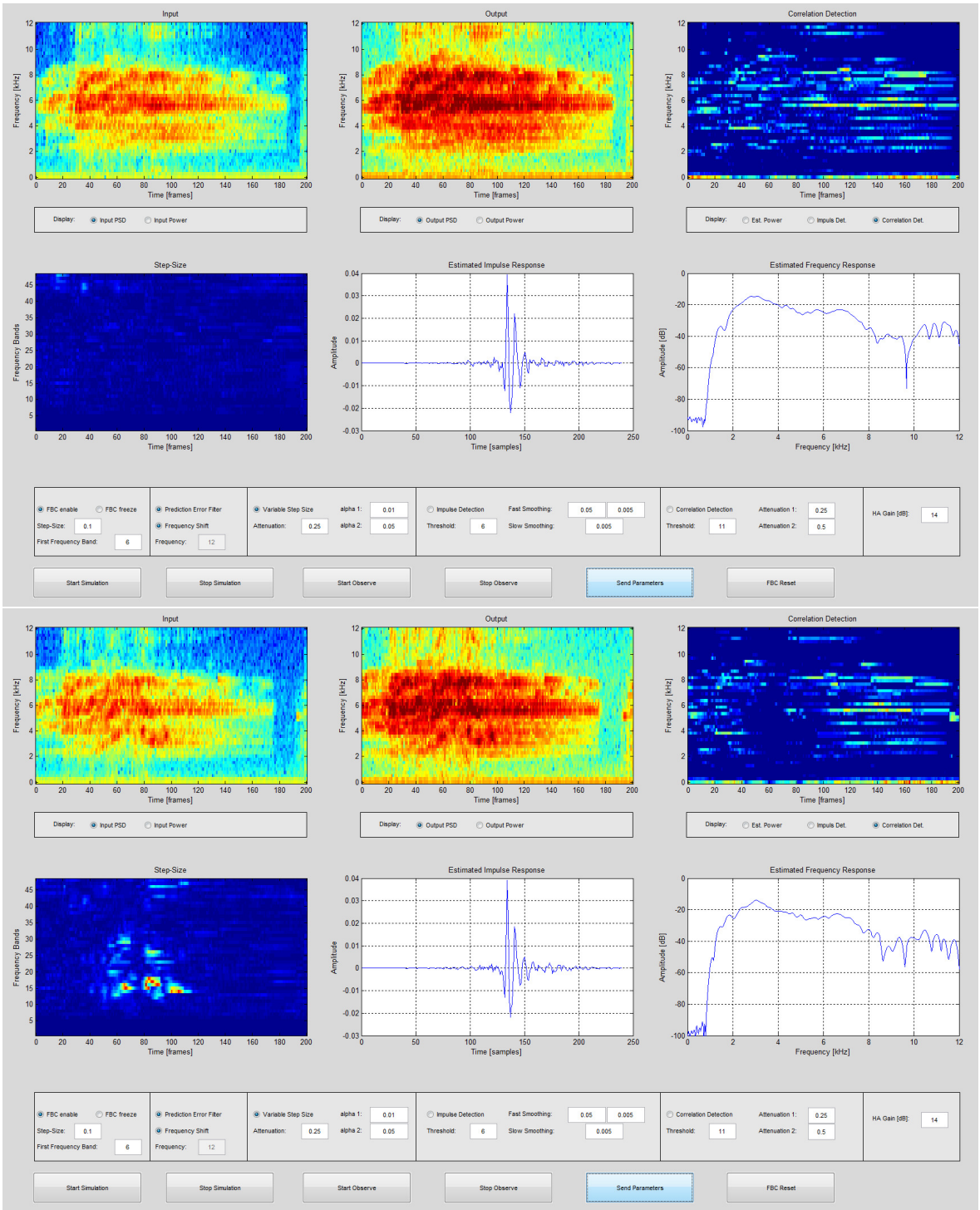


Figure 5.25. Analysis to examine the behavior of the influence function when the feedback path is changed. Therefore, a wind chime signal is played twice. The first time the feedback path is kept stable and the second time the feedback path is changed. For the upper picture the feedback path is constant while it is changed for the lower picture, which is visible by the increased step sizes. The comparison reveals that the values of influence function decreases when the feedback path is changing.

Effect of Correlation Detection

In this section the effect of the correlation detection on the tracking behavior is discussed. It is not possible to illustrate the tracking behavior based on the GUI directly. However, the values of the step sizes can indicate how fast the system adapts if the feedback path has changed. The higher the values of the step sizes are the faster the system can react to a changing feedback path. Consequently, by using the correlation detection the step sizes should take higher values for source signal parts which are not highly correlated.
To demonstrate this behavior, an audio example, consisting of white noise followed by a flute signal, is played twice by the loudspeaker. The first time the correlation detection is disabled while it is applied for the second time. Figure 5.26 shows the corresponding screenshots.
On the upper picture the correlation detection is disabled while for the lower picture it is applied. When the influence function of the correlation detection is low, the step sizes in the lower picture reach higher values compared to the upper picture. Therefore, the system could react faster to a feedback path change. The improved ability to track feedback path changes is also demonstrated by simulations in Section 5.4.3.

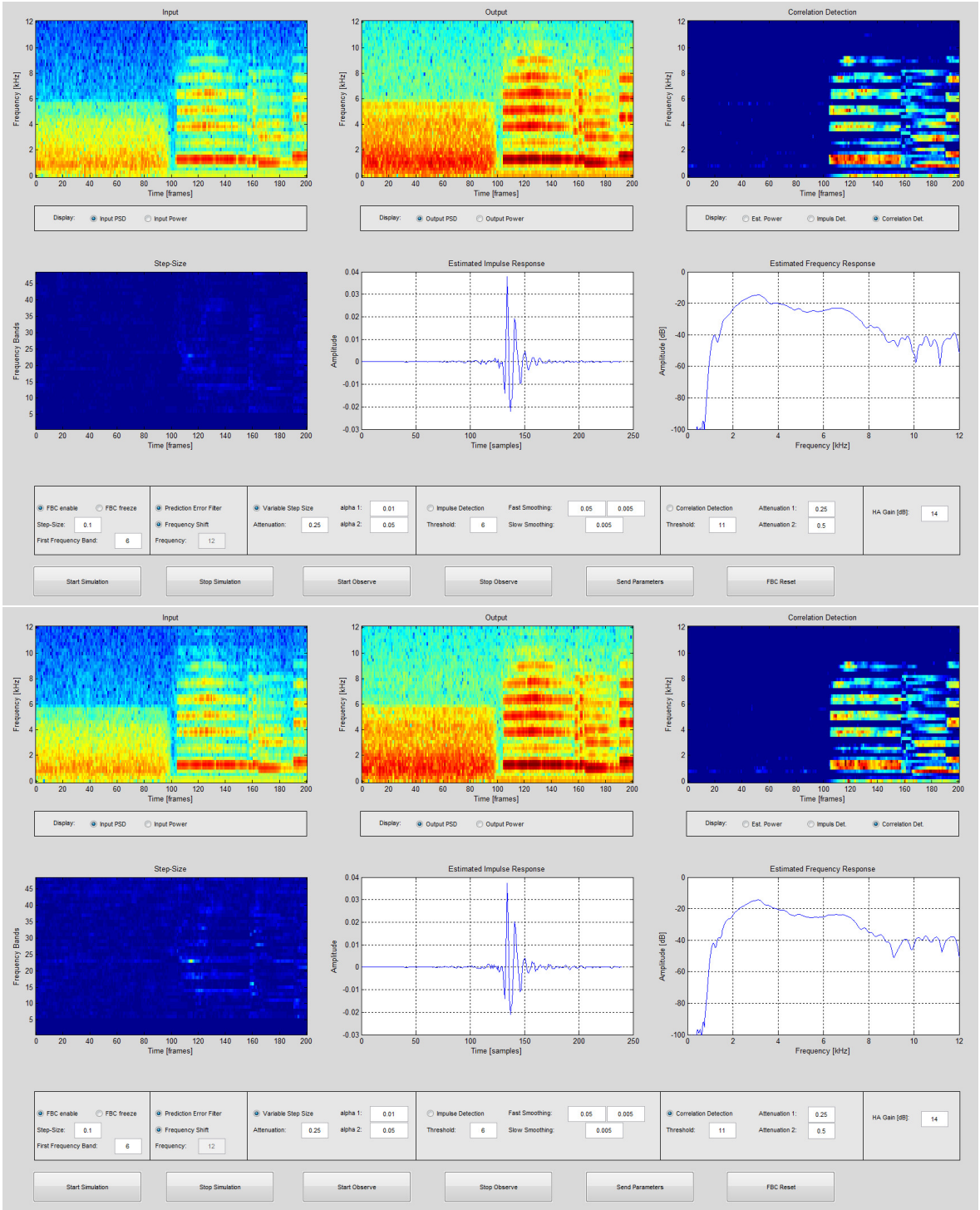
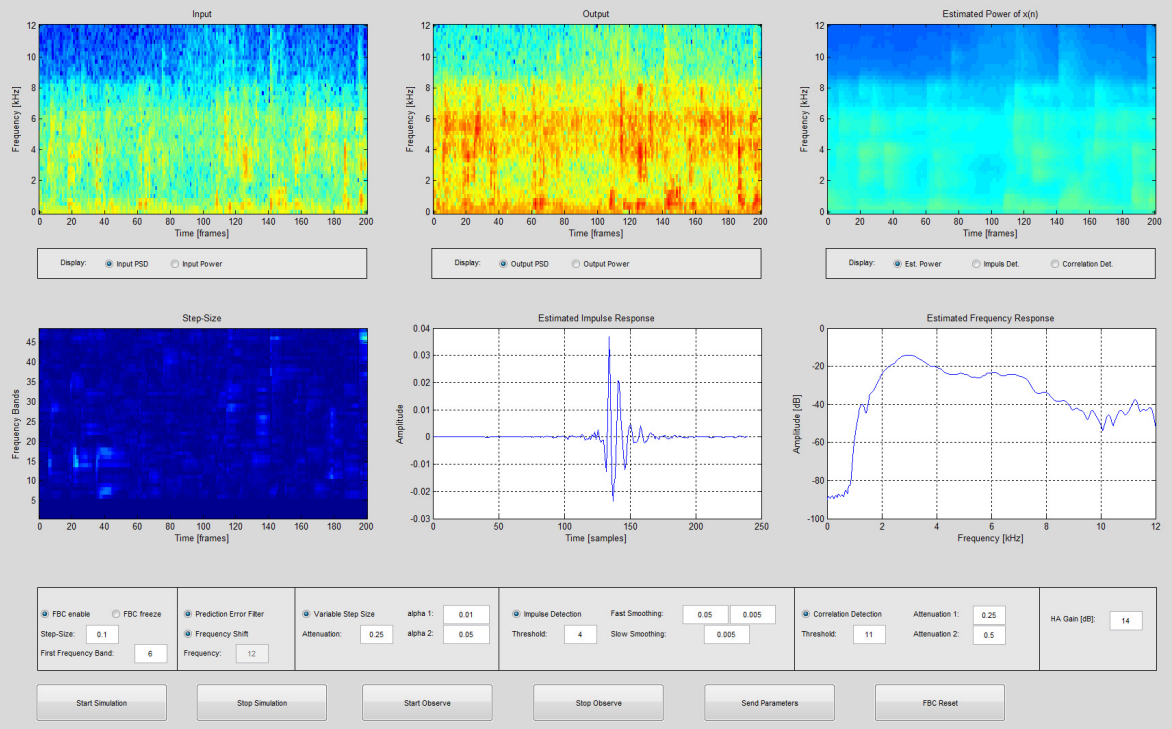


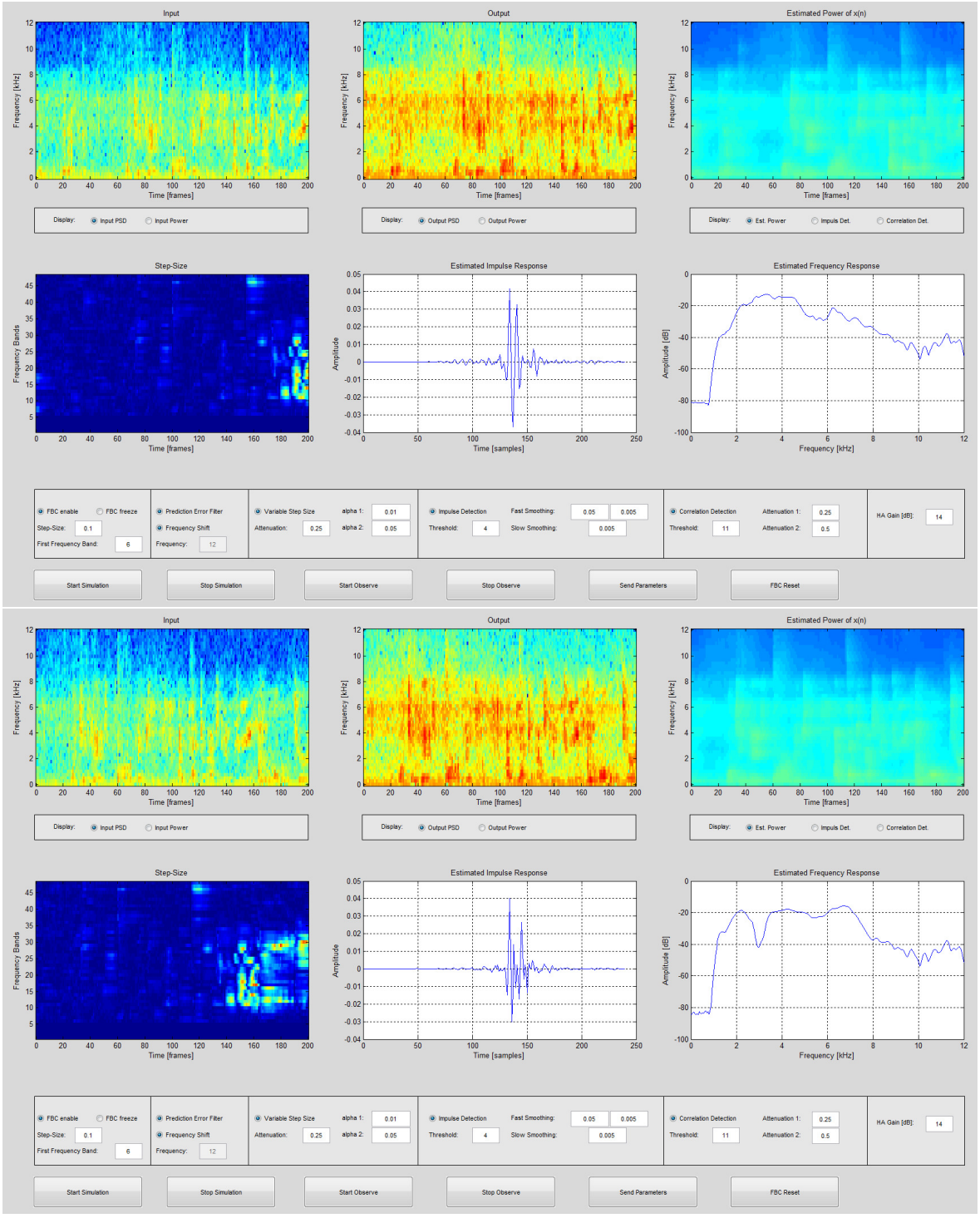
Figure 5.26. Analysis on the effect of the correlation detection. White noise followed by a flute signal is played. First the correlation detection is disabled. The corresponding observation is shown by the upper picture. Then the correlation detection is enabled. The analysis can be observed by the lower picture. When the influence function of the correlation detection is low, the step sizes in the lower picture reach higher values compared to the upper picture. This suggests that the system reacts faster to a feedback path change.

Tracking Feedback Paths

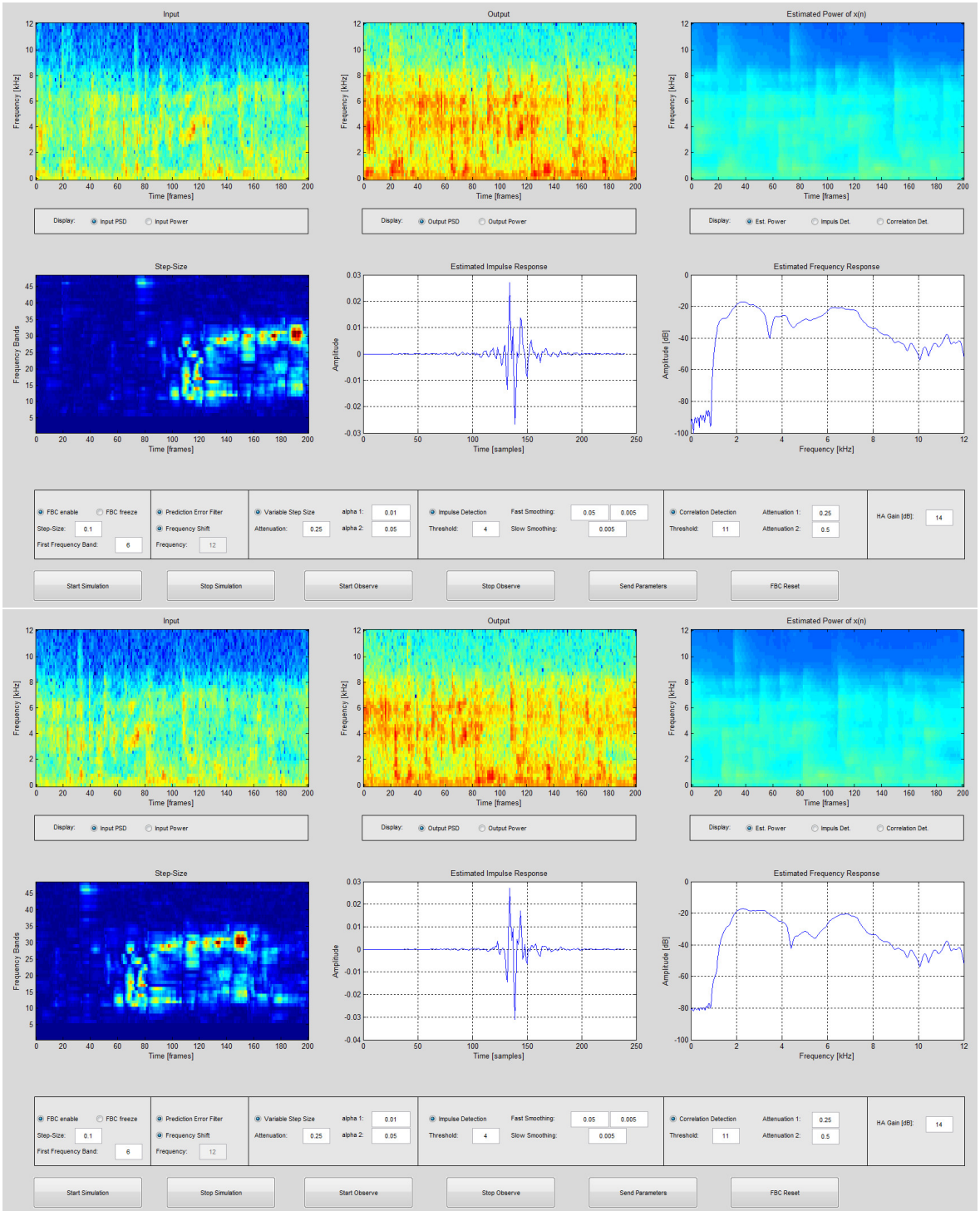
Finally, the tracking ability of the AFC system is demonstrated with the real-time system. Therefore, a speech signal is played while the feedback path is varied by moving a hand around the ear. For this analysis all methods of the the proposed AFC system are enabled. Figure 5.26 shows five screenshots of the observation.



a) Screenshot 1



b) Screenshot 2 and 3



c) Screenshot 4 and 5

Figure 5.26. Analysis of the tracking behavior. Therefore, a speech signal is played while the feedback path is varied by moving a hand around the ear. The estimated impulse and frequency responses are changing with time according to the changed feedback path. This is also indicated by the step size values.

It is clear to see that the estimated impulse and frequency responses are changing with time according to the changed feedback path. This is also indicated by the increased step size values. Qualitative evaluations can not be shown with the GUI as differences of the tracking behavior are only hear able. However, detailed evaluations of the tracking ability were discussed in Sections 5.2 to 5.4 by extensive offline simulations.
All the shown analyzes with the real-time system reflect the results obtained by the offline simulations. While the offline simulations provide quantification of the results the real-time simulations show that these results are obtained for a large variety of realistic test scenarios. Hence, both evaluation procedures validate and complement each other.

5.5.2 Evaluation based on Estimation Error of $\hat{\sigma}_x^2$

As depicted in the previous section, the evaluation with the real-time system is mostly done by observing the figures showed in the GUI and simultaneously, listening to the hearing aid signals while playing test signals, changing the feedback path manually, turning methods on and off, and changing parameters.
These evaluations show a very good performance for the complete system with each presented method comparable to the simulation results. However, as the true feedback path is unknown when using the real-time system, evaluations based on the ECLG are not possible. To give an indication of the system performance based on simulations the evaluation of the estimation error, c.f. Section 5.3.1, is applied to the real-time system. Therefore, Setting 1, c.f. Section 5.1.1, and the signals described in Section 2.4 are used. To obtain the true signal power for all signals, the signals are played by a loudspeaker, received at the hearing aid microphone, and the signal power is calculated. This is done with the loudspeaker signal of the hearing aid turned off by setting the hearing aid gain to zero. Hence, no feedback is present. In this way the true signal power can be obtained. Based on this the estimation error is calculated and the same evaluation as in Section 5.3.1 is done for the real-time system. The results are shown in Figure 5.27.
The results show that the influence of the decorrelation methods are similar for the real-time simulations and for the offline simulations. By using the FS alone a higher percentage (86 %) of non-critical errors are gained compared to only applying the PEF (slightly less than 85 %). However, the combination of both gives the best result 90 %. Compared to the results obtained with the offline simulations a bias of 5 % to 9 % is noticeable. The reason is that the estimation of the source signal power is more difficult under realistic conditions, e.g., constant slight changes of the feedback path.

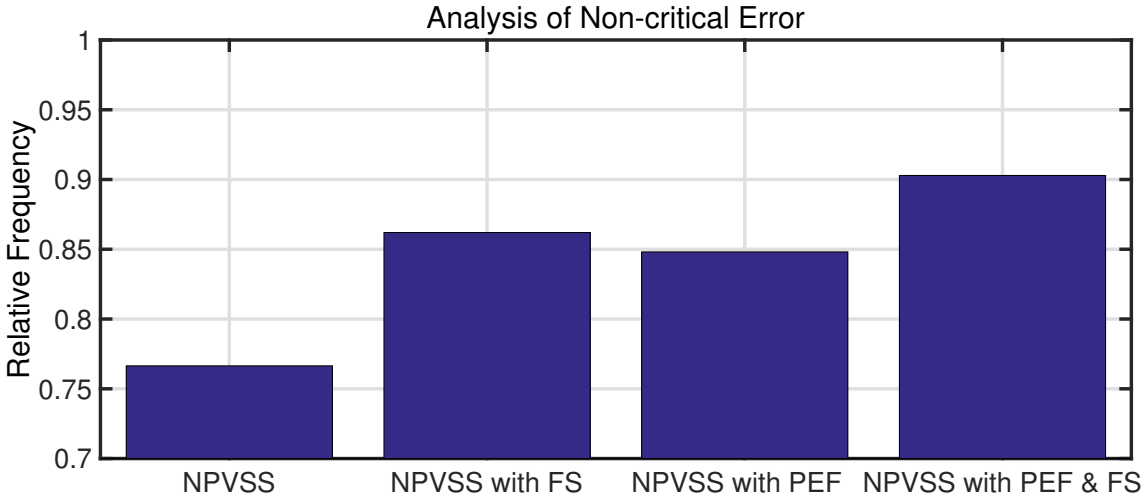


Figure 5.27. Non-critical errors are shown for four different setups of the real-time system. If the decorrelation methods are turned off non-critical errors occur with about 77 %. When the FS is additionally used a percentage of 86 % can be achieved, while it is slightly less than 85 % for the PEF. With the combination of NPVSS, PEF and FS almost no critical errors are present as 90 % are reached.

Nevertheless, these results demonstrate that the real-time system behaves similarly to the offline system, which emphasizes the practical relevance of the presented methods in this thesis.

5.6 Conclusion

In this chapter the simulations to evaluate all proposed methods and the complete AFC system were presented. Therefore, several simulations settings, which combine multiple methods, were introduced. With these different settings the evaluation and optimization of the corresponding methods is performed. Additionally, an overview of the complete parameter setting was given.
First, four parameters, two needed to calculate the NPVSS and two for additional control methods, were optimized based on their effect on the performance of the overall system. All other applied parameters could be optimized and chosen because of other criteria as described throughout this thesis. All parameter choices were used for all evaluations and are proposed to be applied in the presented AFC system independent on the conditions and input signals.
Then, the effect of the proposed methods in this thesis were evaluated intensively. The

decorrelation methods, PEF and FS, were motivated extensively in Section 3.2 by theoretic analyses as well as by simulation results. The same was done for the NPVSS in Section 3.3.3. However, the benefit to use the decorrelation methods in combination with the NPVSS was finally quantified in this chapter. This was done based on the estimation error of the source signal power and on the ECLG, which indicates the performance of the overall AFC system in terms of the adaptation stability and tracking behavior.

After that, the three additional control methods, i.e., the impulse detection, the modified estimate, and the correlation detection, were evaluated. It was shown that all methods increase the performance of the system significantly.

In addition to the offline simulations, the evaluation with the real-time system was presented as well. This was done by demonstrating the behavior of the AFC system for specific scenarios and the evaluation of the estimation error of the source signal power on the real-time system.

The offline simulations and the evaluation on the real-time systems demonstrate the benefit of each proposed method, i.e., NPVSS, PEF, FS, impulse detection, modified estimate, and correlation detection, and the excellent performance of the AFC system if all these methods are combined. In consequence, the complete AFC system combining all these methods is proposed in this thesis.

Chapter 6
Summary, Conclusions and Future Work

6.1 Summary and Conclusion

In this doctoral project a new adaptive feedback cancellation (AFC) system for hearing aid applications was derived, optimized, and evaluated extensively.
The proposed system was setup to satisfy several requirements: First, the AFC system is realized in a realistic hearing aid environment. Therefore, a realistic automatic gain control as it is applied in hearing aids is used as well as several feedback paths which are measured under realistic conditions. Additionally, the system is successfully implemented and evaluated on the real-time system Speedgoat. This guarantees a realistic test environment and unlimited test cases. Other requirements concern the adaptation stability and tracking speed, which typically need to be balanced as a trade-off. For both, the system shows an excellent performance. Furthermore, the computational costs of the system should not exceed the capability of a hearing aid. Hence, no methods with high computational demands were applied.

In the course of developing and optimizing the AFC system several contributions were made.
First of all, an objective and realistic evaluation procedure to evaluate all proposed methods and the overall system was developed. Therefore, the effective closed loop gain (ECLG) was proposed, which is based on the Nyquist stability criterion. It is used to examine the performance of different system settings and all proposed methods. These simulations evaluate the adaptation stability and the tracking behavior. In order to obtain objective results, a complete simulation procedure was implemented using different types of audio signals, i.e., speech and music, and various feedback paths, which are recorded in realistic conditions. The results are visualized by figures and plots.

As starting point for the final optimized system a basic AFC system was set up. The most important components of this system are the AFC method, the normalized least mean squares (NLMS) algorithm, as adaptation procedure, and a sub-band processing. Many approaches dealing with acoustic feedback have been proposed in the past. However, the AFC method shows the largest benefit compared to other approaches as in

ideal cases it allows to eliminate the feedback signal completely.
The NLMS algorithm is preferred as adaptation procedure, in general and in this thesis, due to the trade-off between computational complexity and adaptation performance.
The AFC is realized in sub-bands by using a polyphase filter bank. The sub-band processing reduces the computational complexity and allows a frequency selective adaptation and control. Additionally, the performance of the NLMS algorithm is increased significantly.
However, all AFC approaches regardless of the adaptation procedure show one major drawback. The adaptive filter adapts to a biased solution due to the closed feedback loop resulting in a correlation between the source signal and the loudspeaker (or receiver) signal of the hearing aid. For colored source signals the bias becomes very large and provokes distortion artifacts, which are called entrainment. To prevent this, the adaptation has to be controlled. Several control methods, which are mentioned in the following, were proposed in this thesis.

In order to minimize the bias of the adaptation, typically decorrelation methods are used. In this thesis a combination of the prediction error filter (PEF) and the frequency shift (FS) was proposed.
The prediction error filters and frequency shift are two well known and frequently used methods to reduce entrainment in the context of AFC. However, the combination of both was not presented in literature before. In this thesis, it is shown by extensive and objective studies that the combination of both methods is beneficial compared to an implementation of only one method. As performance criterion the bias of the adaptive filter is used, which is significantly reduced. Evaluations using the ECLG to measure the performance of the AFC system containing a constant step size (CSS), the PEF, and the FS as adaptation control confirm this as well. These simulations show that the adaptation stability is clearly increased while the tracking behavior is not influenced.

Even though the adaptation stability is increased, it is still not satisfying. Hence, the use of a variable step size (VSS) instead of a CSS was recommend in this thesis. Several approaches to calculate the VSS were theoretically analyzed. For several explained reasons the NPVSS by Benesty et al. [12] was chosen for further optimization in this thesis. The NPVSS performs excellently in theory even without the additional use of the mentioned decorrelation methods. However, the NPVSS is unsuitable for realistic systems as it requires a reliable estimate of the source signal power without feedback. Hence, a yet unknown estimate was theoretically derived, which is another important contribution of this thesis. In theory, with this estimate the AFC systems performs equally to an AFC system where the true source signal power, which is unknown in

reality, can be used. However, for real applications the estimate is only reliable with the previously proposed and already implemented decorrelation methods, PEF and FS. Hence, the decorrelation methods do not only decrease the adaptation bias but also ensure that the NPVSS performs as intended.
The AFC system with the combination of NPVSS, PEF, and FS shows a very good behavior in terms of adaptation stability and tracking behavior in realistic situations and clearly outperforms a system with a CSS, PEF, and FS. This is proven by several offline simulations and examinations with the real-time system.

To increase the adaptation performance of the AFC system even more some additional control methods were proposed.
The first one, is a correlation detection, which is able to distinguish between highly correlated source signal components and less correlated components. This is used to improve the trade-off between adaptation stability and tracking speed. Based on the correlation detection the VSS is additionally controlled. Hence, if no correlation occurs, the step size is increased in order to provide a faster adaptation. For high correlation the step size is kept lower to ensure the stability of the system. By using the correlation detection the trade-off between the adaptation stability and the tracking ability is significantly improved.
Two other additional control methods were derived to prepare the AFC systems for real world scenarios. Particularly, an impulse detection is implemented to handle wideband impulses, e.g., clapping and knocking, and a modified estimate of the source signal power is derived, which is needed in case of fast increases of the source signal power, e.g., jarring voices.

The complete AFC system was implemented on the real-time system Speedgoat to evaluate all proposed methods additional to the offline simulations. The real-time system allows a fast and realistic examination of the methods and the overall AFC system. Fast refers to the immediate results one gets by using the real-time system, in contrast to offline simulations, which require more time and computation capacity. Additionally, the methods can be evaluated with a large variety of realistic scenarios and naturally variable feedback paths. Theoretically, an unlimited number of different audio signals and feedback paths can be tested.
The evaluation on the real-time system is done by listening to the output signal of the hearing aid and by observing several informative signals, which can be analyzed in the AFC system. The signals are observed by an implemented graphical user interface, which can also be used to change parameter settings and to enable or disable certain

control methods.

All offline and real-time evaluations confirm the impressive performance in terms of adaptation stability and tracking behavior of the proposed AFC system in this thesis.

6.2 Future Directions

One major future research topic concerns the correlation detection. It is conceivable that the correlation detection can be used in other ways which could lead to an even better performance or decrease the computational cost of the AFC system.
One example would be that the correlation detection is used to distinguish between more than the two cases, correlated and uncorrelated. Hence, a finer quantization of the attenuation factor would be applied for a more precise adaptation control based on finer distinctions of the amount of correlation. This could increase the performance of the system.
Another possibility would be to apply the correlation detection directly to specify values of a CSS. This would decrease the computational costs as no VSS has to be calculated.

Other possible future investigations concern the final implementation in an actual hearing aid and could show a rise in performance of the overall system.
This could be a sub-band depending attenuation factor of the NPVSS. In particular, the attenuation factor of each sub-band can be defined depending on the general shape of the frequency response of feedback paths. For sub-bands, where the magnitude of the feedback paths is low, e.g., at high frequencies, or the probability that the paths is changing rapidly is low, the factor can be chosen low. For other sub-bands with a higher risk of howling and entrainment higher values can be assigned. In consequence, this could improve the adaptation stability and the tracking behavior simultaneously.
Other possibilities to increase the overall performance could be approaches which use additional information on the source signal obtained by a second (or third) microphone of the hearing aid or a combination of the left and right hearing aid. In the later case both hearing aids would communicate through a wireless connection. This could improve the differentiation between correlation based on the feedback path and correlation of the source signal. For example, a rapid feedback path change occurs typically only at one hearing aid while the source signal is received by all hearing aids.

List of Acronyms

AEC	Adaptive Echo Cancellation
AFC	Adaptive Feedback Cancellation
AGC	Automatic Gain Control
AIF	Adaptive Inverse Filtering
ASG	Added Stable Gain
CI	Confidence Interval
CSS	Constant Step Size
DFT	Discrete Fourier Transform
ECLG	Effective Closed Loop Gain
EM	Expectation-Maximization
FB	Filter Bank
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FM	Frequency Modulation
FS	Frequency Shift
GMM	Gaussian Mixture Model
GUI	Graphical User Interface
IDFT	Inverse Discrete Fourier Transform
IFB	Inverse Filter Bank
LBG	Linde-Buzo-Gray
LTI	Linear Time-Independent
MSE	Mean Squared Error
NLMS	Normalized Least Mean Squares
NPVSS	Non-Parametric Variable Step Size

OSS	Optimal Step Size
PC	Personal Computer
PDF	Probability Density Function
PEF	Prediction Error Filter
PF	Prediction Filter
PM	Phase Modulation
PSD	Power Spectral Density
SNR	Signal-to-Noise Ratio
SPL	Sound Pressure Level
VSS	Variable Step Size

List of Symbols

$^{\mathrm{T}}$	Transposition
$^{\mathrm{H}}$	Hermitian transposition
$^{-1}$	Matrix inversion
*	Complex conjugation
$*$	Convolution
$\mathrm{E}\{\cdot\}$	Expected value
n	Sampling time index
T_s	Sampling time
f_s	Sampling frequency
Ω	Normalized frequency
$x(n)$	Sample of a signal at time index n
$\mathbf{x}(\mathbf{n})$	Vector representation of a signal $x(n)$, contains the last N samples
$\|\|\mathbf{x}(\mathbf{n})\|\|^2$	Norm of a vector $\mathbf{x}(\mathbf{n})$
$\mathcal{F}\{x(n)\}$	Fourier transform of a signal $x(n)$
$\sigma_x^2(n)$	Power of a signal $x(n)$
$r_{xx}(n, n-l)$	Auto-correlation of a signal $x(n)$
$\mathbf{r}_{xx}(n)$	Auto-correlation vector of a signal $x(n)$
$\mathbf{R}_{xx}(n)$	Auto-correlation matrix of a signal $x(n)$
$r_{xy}(n, n-l)$	Cross-correlation of signals $x(n)$ and $y(n)$
$\mathbf{r}_{xy}(n)$	Auto-correlation vector of signals $x(n)$ and $y(n)$
$x_k(n)$	k-th sub-band signal of a signal $x(n)$
$u(n)$	Loudspeaker signal
$U(\Omega, n)$	Time depending spectrum of loudspeaker signal
$x(n)$	Source signal
$X(\Omega, n)$	Time depending spectrum source signal
$v(n)$	Feedback signal
$e(n)$	*A priori* error signal, respectively feedback-compensated microphone signal
$\epsilon(n)$	*A posteriori* error signal.
$y(n)$	Microphone signal

$\mathbf{g}(n)$	Impulse response of forward path
$G(\Omega, n)$	Short-term frequency response of forward path
$\mathbf{f}(n)$	Impulse response of feedback path
$F(\Omega, n)$	Short-term frequency response of feedback path
$C(\Omega, n)$	Short-term frequency response of closed loop system
$F_i(\Omega)$	Frequency response of i-th recorded feedback path
$\hat{f}(n)$	Impulse response of adaptive filter
$\hat{F}(\Omega, n)$	Short-term frequency response of adaptive filter
N	Number of adaptive filter coefficients
$\mu(n)$	Step size of NLMS algorithm
$\hat{v}(n)$	Estimated feedback signal
$\mathbf{b}(n)$	Bias vector of adaptation
$\mathrm{ECLG}(\Omega, n)$	Effective closed loop gain
$\mathrm{ECLG}_{\max}(n)$	Maximum of the $\mathrm{ECLG}(\Omega, n)$ over all frequency
$\mathrm{ASG}_{\mathrm{dB}}(n)$	Added stable gain
c_k	AGC knee point
c_r	AGC compression ratio
k	Sub-band index
K	Number of sub-bands
g_k	k-th bandpass filter
N_{FB}	Length of bandpass filters
r	Downsampling factor
$h(n)$	Impulse response of source signal model
$H(\Omega, n)$	Short-term frequency response of source signal model
$w(n)$	White signal
$\hat{\hat{h}}(n)$	Impulse response of prediction filter
$\hat{H}^{-1}(\Omega, n)$	Short-term frequency response of prediction error filter
M	Order of the prediction filter
$\tilde{y}(n)$, $\tilde{u}(n)$, $\tilde{e}(n)$	Prewhitened versions of $y(n)$, $u(n)$, $e(n)$

$d(n)$	Output of hearing aid processing
$d_{\mathrm{H}}(n)$	Hilbert transform of $d(n)$
$d_a(n)$	Analytical representation of $d(n)$
$h_{\mathrm{FS}}(n)$	Frequency shift modeled as a periodically time-varying filter
f_m	Amount of frequency shift
$\mu_{\mathrm{OSS}}(n)$	Optimal step size by Mader et al. [61]
$\mathbf{m}(n)$	System mismatch (distance) vector between true and estimated feedback path
$e_u(n)$	Undisturbed error signal
N_{D}	Number of artificial delay filter coefficients
b_{D}	Estimated system distance
$\mu_{\mathrm{NPVSS}}(n)$	Non-parametric variable step size by Benesty et al. [12]
$\epsilon(n)$	*A posteriori* error signal
$\hat{\sigma}_x^2(n)$	Proposed estimate of source signal power
$\hat{\sigma}_{\mathrm{opt},x}^2(n)$	Optimal estimate of source signal power
$\hat{\sigma}_{\mathrm{real},x}^2(n)$	Estimate of source signal power based on $x(n)$
$\hat{\sigma}_{\mathrm{Huang},x}^2(n)$	Estimate of source signal power by Huang and Lee [62]
α_1	Smoothing parameters to calculate variable step sizes
a	Attenuation parameter of variable step sizes
$\overline{y}_{\mathrm{fast}}(n)$	Fast smoothing of $y(n)$, follows the signal power fast
β_r	Fast smoothing parameter if signal power increases
β_f	Fast smoothing parameter if signal power decreases
$\overline{y}_{\mathrm{slow}}(n)$	Slow smoothing of $y(n)$, follows the signal power slowly
γ	Slow smoothing parameter
$I_k(n)$	Indicator function of impulse detection for sub-band k
$I_{\mathrm{comp}}(n)$	Complete indicator function of impulse detection
T_{ID}	Threshold of impulse detection
$\tilde{\mathbf{R}}_{uu}(n)$	Modified matrix which replaces $\mathbf{R}_{uu}(n)$ to estimate the source signal power

$\hat{\sigma}^2_{x,\mathrm{mod}}(n)$	Modified estimate of source signal power
α_2	Smoothing parameter used for modified estimate
$C_k(n)$	Indicator function of correlation detection
T_{CD}	Threshold of correlation detection
a_1	Attenuation parameters of variable step size if source signal is high correlated
a_2	Attenuation parameters of variable step size if source signal is less correlated

Bibliography

[1] T. van Waterschoot and M. Moonen, "Fifty years of acoustic feedback control: State of the art and future challenges," *Proceedings of the IEEE*, vol. 99, no. 2, pp. 288–327, 2011.

[2] E. Hänsler and G. Schmidt, *Acoustic echo and noise control: a practical approach*, vol. 40, John Wiley & Sons, 2005.

[3] S. Ibaraki, H. Furukawa, and H. Naono, "Howling canceller," May 24 1988, US Patent 4,747,132.

[4] A. Goertz, "An adaptive subtraction filter for feedback cancellation in public address sound systems," in *Proc. 15th Int. Congr. Acoust.(ICA95)*, 1995, pp. 69–72.

[5] H. Puder and B. Beimel, "Controlling the adaptation of feedback cancellation filters – problem analysis and solution approaches," in *Signal Processing Conference (EUSIPCO), 2004 Proceedings of the 12th European*, 2004, pp. 25 – 28.

[6] M.G. Siqueira and A. Alwan, "Steady-state analysis of continuous adaptation in acoustic feedback reduction systems for hearing-aids," *Speech and Audio Processing, IEEE Transactions on*, vol. 8, no. 4, pp. 443–453, 2000.

[7] M. Rotaru, F. Albu, and H. Coanda, "A variable step size modified decorrelated NLMS algorithm for adaptive feedback cancellation in hearing aids," in *Electronics and Telecommunications (ISETC), 2012 10th International Symposium on*, 2012, pp. 263 – 266.

[8] J.M. Gil-Cacho, T. van Watershoot, M. Moonen, and S.H. Jensen, "Transform domain prediction error method for improved acoustic echo and feedback cancellation," in *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, 2012, pp. 2422 – 2426.

[9] K. Ngo, T. Watershoot, M.G. Christensen, Moonen M., and S.H. Jensen, "Adaptive feedback cancellation in hearing aids using sinusoidal near-end signal model," in *Acoustics Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, 2010, pp. 181 – 184.

[10] Yi FanChiang, Cheng-Wen Wei, Yi-Le Meng, Yu-Wen Lin, Shyh-Jye Jou, and Tian-Sheuan Chang, "Low complexity formant estimation adaptive feedback cancellation for hearing aids using pitch based processing," *Audio, Speech, and Language Processing, IEEE/ACM Transactions on*, vol. 22, no. 8, Aug 2014.

[11] T. van Waterschoot and M. Moonen, "Adaptive feedback cancellation for audio applications," *Signal Processing*, vol. 89, no. 11, pp. 2185 – 2201, 2009.

[12] J. Benesty, H. Rey, L. Rey Vega, and S. Tressens, "A nonparametric VSS NLMS algorithm," *Signal Processing Letters, IEEE*, vol. 13, no. 10, pp. 581 – 584, 2006.

[13] H. Nyquist, "Regeneration theory," *Bell System Technical Journal*, vol. 11, no. 1, pp. 126–147, 1932.

[14] H.A.L. Joson, F. Asano, Y. Suzuki, and T. Sone, "Adaptive feedback cancellation with frequency compression for hearing aids," *The Journal of the Acoustical Society of America*, vol. 94, no. 6, pp. 3248 – 3254, 1993.

[15] M. Guo, S.H. Jensen, J. Jensen, and S.L. Grant, "On the use of a phase modulation method for decorrelation in acoustic feedback cancellation," in *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, 2012, pp. 2000 – 2004.

[16] M.R. Schroeder, "Improvement of acoustic feedback stability in public address systems," *The Journal of the Acoustical Society of America*, vol. 31, no. 6, pp. 851–852, 1959.

[17] M.R. Schroeder, "Improvement of acoustic-feedback stability by frequency shifting," *The Journal of the Acoustical Society of America*, vol. 36, no. 9, pp. 1718–1724, 1964.

[18] R.W. Guelke and A.D. Broadhurst, "Reverberation time control by direct feedback," *Acta Acustica united with Acustica*, vol. 24, no. 1, pp. 33–41, 1971.

[19] L.N. Mishin, "A method for increasing the stability of sound amplification systems," Jan./Mar. 1958.

[20] J.L. Nielsen and U.P. Svensson, "Performance of some linear time-varying systems in control of acoustic feedback," *The Journal of the Acoustical Society of America*, vol. 106, no. 1, pp. 240–254, 1999.

[21] G. Nishinomiya, "Improvement of acoustic feedback stability of public address system by warbling," in *Proc. 6th Int. Congr. Acoust*, 1968, pp. 93–96.

[22] D. Griesinger, "Improving room acoustics through time-variant synthetic reverberation," in *Audio Engineering Society Convention 90*. Audio Engineering Society, 1991.

[23] E.T. Patronis Jr, "Electronic detection of acoustic feedback and automatic sound system gain control," in *Audio Engineering Society Convention 57*. Audio Engineering Society, 1977.

[24] S. Ando, "Howling detection and prevention circuit and a loudspeaker system employing the same," Jun 2001, US Patent 6,252,969.

[25] Y. Nagata, S. Suzuki, M. Yamada, M. Yoshida, M. Kitano, K. Kuroiwa, and S. Kimura, "Howling remover composed of adjustable equalizers for attenuating complicated noise peaks," Mar 1998, US Patent 5,729,614.

[26] M. Hanajima, M. Yoneda, and T. Okuma, "Howling eliminating apparatus," Sep 2000, US Patent 6,125,187.

[27] Y. Terada and A. Murase, "Howling control device and howling control method," Mar 2007, US Patent 7,190,800.

[28] N. Osmanovic and V. Clarke, "Acoustic feedback cancellation system," Feb 2010, US Patent 7,664,275.

[29] S.M. Kuo and J. Chen, "New adaptive iir notch filter and its application to howling control in speakerphone system," *Electronics Letters*, vol. 28, no. 8, pp. 764–766, 1992.

[30] M. Tahernezhadi and L. Liu, "An adaptive notch filter for howling cancellation," *Acoustics letters*, vol. 18, no. 8, pp. 142–145, 1995.

[31] W. Leotwassana, R. Punchalard, and K. Silaphan, "Adaptive howling canceller using adaptive iir notch filter: Simulation and implementation," in *Neural Networks and Signal Processing, 2003. Proceedings of the 2003 International Conference on.* IEEE, 2003, vol. 1, pp. 848–851.

[32] G. Rombouts, T. van Waterschoot, and M. Moonen, "Proactive notch filtering for acoustic feedback cancellation," *Proc. 2nd Annual IEEE Benelux/DSP Valley Signal Process*, 2006.

[33] P. Gil-Cacho, T. Van Waterschoot, M. Moonen, and S.H. Jensen, "Regularized adaptive notch filters for acoustic howling suppression," in *Signal Processing Conference, 2009 17th European.* IEEE, 2009, pp. 2574–2578.

[34] Y. Takahashi, M. Tohyama, and Y. Yamasaki, "Cumulative spectral analysis for transient decaying signals in a transmission system including a feedback loop," *Journal of the Audio Engineering Society*, vol. 54, no. 7/8, pp. 620–629, 2006.

[35] C.P. Janse and H.J.W. Belt, "Sound reinforcement system having an echo suppressor and loudspeaker beamformer," May 2006, US Patent 7,054,451.

[36] K. Kobayashi, K. Furuya, and A. Kataoka, "A microphone array for howling cancellation," *J. Acoust. Soc. Jpn*, vol. 60, no. 3, pp. 115–125, 2004.

[37] G. Rombouts, A. Spriet, and M. Moonen, "Generalized sidelobe canceller based combined acoustic feedback-and noise cancellation," *Signal Processing*, vol. 88, no. 3, pp. 571–581, 2008.

[38] S. Ushiyama, T. Hirai, M. Tohyama, and Y. Shimizu, "Howling suppression by smoothing the open-loop transfer function," Tech. Rep., IEICE Tech. Rep, 1994.

[39] G. Schmidt and T. Haulick, "Signal processing for in-car communication systems," *Signal processing*, vol. 86, no. 6, pp. 1307–1326, 2006.

[40] S. Haykin, "Adaptive filter theory," *Prentice Hall*, vol. 2, 2002.

[41] J.M. Cioffi and T. Kailath, "Fast, recursive-least-squares transversal filters for adaptive filtering," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 32, no. 2, pp. 304–337, Apr 1984.

[42] D.T.M. Slock and T. Kailath, "Numerically stable fast transversal filters for recursive least squares adaptive filtering," *Signal Processing, IEEE Transactions on*, vol. 39, no. 1, pp. 92–114, Jan 1991.

[43] S.L. Gay and S. Tavathia, "The fast affine projection algorithm," in *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, May 1995, vol. 5, pp. 3023–3026.

[44] G. Rombouts and M. Moonen, "A sparse block exact affine projection algorithm," *Speech and Audio Processing, IEEE Transactions on*, vol. 10, no. 2, pp. 100–108, Feb 2002.

[45] B. Widrow and S.D. Stearns, "Adaptive signal processing," *Englewood Cliffs, NJ, Prentice-Hall, Inc.*, vol. 1, pp. 491, 1985.

[46] C.J. Clopper and E.S. Pearson, "The use of confidence or fiducial limits illustrated in the case of the binomial," *Biometrika*, vol. 26, no. 4, pp. 404–413, 1934.

[47] M. Guo, T.B. Elmedyb, S.H. Jensen, and J. Jensen, "Analysis of acoustic feedback/echo cancellation in multiple-microphone and single-loudspeaker systems using a power transfer function method," *Signal Processing, IEEE Transactions on*, vol. 59, no. 12, pp. 5774 – 5788, 2011.

[48] M. Guo, S.H. Jensen, and J. Jensen, "Evaluation of state-of-the-art acoustic feedback cancellation systems for hearing aids," *J. Audio Eng. Soc.*, vol. 61, no. 3, pp. 125 – 137, 2013.

[49] L.R. Rabiner, *Multirate digital signal processing*, Prentice Hall PTR, 1996.

[50] J.H. Stott and N.D. Wells, "Method and apparatus for reduction of unwanted feedback," July 31 2001, US Patent 6,269,165.

[51] A. Spriet, S. Doclo, M. Moonen, and J. Wouters, "Feedback control in hearing aids," in *Springer Handbook of Speech Processing*, J. Benesty, M.M. Sondhi, and Y. Huang, Eds., pp. 979 – 1000. Springer Berlin Heidelberg, 2008.

[52] G. Rombouts, T. van Watershoot, and M. Moonen, "Robust and efficient implementation of the PEM-AFROW algorithm for acoustic feedback cancellation," *J. Audio Eng. Soc.*, vol. 55, no. 11, pp. 955 – 966, 2007.

[53] J. Hellgren, "Analysis of feedback cancellation in hearing aids with filtered-x LMS and the direct method of closed loop identification," *Speech and Audio Processing, IEEE Transactions on*, vol. 10, no. 2, pp. 119 – 131, 2002.

[54] F. Strasser and H. Puder, "Sub-band feedback cancellation with variable step sizes for music signals in hearing aids," in *2014 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. IEEE, 2014, pp. 8207–8211.

[55] F. Strasser and H. Puder, "Adaptive feedback cancellation for realistic hearing aid applications," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 12, pp. 2322–2333, 2015.

[56] S.A. Samad, A. Hussain, and D. Isa, "Wave digital filters with minimum multiplier for discrete hilbert transformer realization," *Signal Processing*, vol. 86, no. 12, pp. 3761 – 3768, 2006.

[57] R.H. Kwong and E.W. Johnston, "A variable step size LMS algorithm," *Signal Processing, IEEE Transactions on*, vol. 40, no. 7, pp. 1633–1642, 1992.

[58] T. Aboulnasr and K. Mayyas, "A robust variable step-size LMS-type algorithm: analysis and simulations," *Signal Processing, IEEE Transactions on*, vol. 45, no. 3, pp. 631–639, 1997.

[59] H. Shin, A.H. Sayed, and W. Song, "Variable step-size NLMS and affine projection algorithms," *IEEE signal processing letters*, vol. 11, no. 2, pp. 132–135, 2004.

[60] J. Lee, H. Huang, and Y.N. Yang, "The generalized square-error-regularized LMS algorithm," *Proceedings of WCECS*, pp. 157–160, 2008.

[61] A. Mader, H. Puder, and G. U. Schmidt, "Step-size control for acoustic echo cancellation filters - an overview," *Signal Processing*, vol. 80, no. 9, pp. 1697 – 1719, 2000.

[62] H. Huang and J. Lee, "A new variable step-size NLMS algorithm and its performance analysis," *Signal Processing, IEEE Transactions on*, vol. 60, no. 4, pp. 2055 – 2060, 2012.

[63] C.M. Bishop, "Pattern recognition," *Machine Learning*, 2006.

[64] A.P. Dempster, N.M. Laird, and D.B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38, 1977.

[65] Y. Linde, A. Buzo, and R.M. Gray, "An algorithm for vector quantizer design," *Communications, IEEE Transactions on*, vol. 28, no. 1, pp. 84–95, 1980.

[66] F. Strasser and H. Puder, "Correlation detection for adaptive feedback cancellation in hearing aids," *IEEE Signal Processing Letters*, vol. 23, no. 7, pp. 979–983, 2016.

Curriculum Vitae

Name:	Falco Strasser
Date of birth:	17.11.1985
Place of birth:	Hanau (Germany)
Family status:	single

Education

03/2012 - 01/2017	PhD student at Graduate School of Excellence Computational Engineering (CE), Technische Universität Darmstadt
10/2005 - 12/2011	Technische Universität Darmstadt (Darmstadt, Germany) Electrical Engineering and Information Technology, Telecommunications Diplom–Ingenieur (Dipl.–Ing.)
12/2011	Diploma Thesis: "Motion Artifact Removal in ECG Signals"
10/2010	Student Thesis: "Robust Filtering of Autoregressive Processes"
06/2005	High school degree (Abitur) at Otto – Hahn – Schule, Kooperative Gesamtschule, Hanau, Germany

Work experience

03/2012 - 04/2015	Research Assistant at Signal Processing Group Technische Universität Darmstadt
10/2007 - 10/2011	Student Assistant at Several Departments Technische Universität Darmstadt
12/2010 - 05/2011	Specialized Internship at Siemens Audiologische Technik GmbH, Erlangen, Germany

Erklärung laut §9 der Promotionsordnung

Ich versichere hiermit, dass ich die vorliegende Dissertation allein und nur unter Verwendung der angegebenen Literatur verfasst habe. Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.