Rheinland-Pfälzische
Technische Universität
Kaiserslautern
Landau

Ein Beitrag zur akustischen Rückkopplungsunterdrückung mit Tisch- und Deckenarrays in Konferenzsystemen

Dissertation

Vom Fachbereich Elektrotechnik und Informationstechnik der Rheinland-Pfälzischen Technischen Universität Kaiserslautern-Landau zur Verleihung des akademischen Grades Doktor der Ingenieurwissenschaften (Dr.-Ing.) genehmigte Dissertation

von

Manuel Ruben Schalk
geboren in Kaiserslautern, Deutschland

Datum der wissenschaftlichen Aussprache	18.09.2026
Dekan	Prof. Dr.-Ing. Daniel Görges
Berichterstattende	Prof. habil. Dr.-Ing. Alexander Potchinkov
	Prof. Dr.-Ing. Christian Langen

D 386

Kurzfassung

Akustische Rückkopplung ist mindestens mit der Erfindung der elektroakustischen Beschallungssysteme und spätestens mit der fortschreitenden Entwicklung von Hörgeräten und Konferenzsystemen ein unumgängliches Thema geworden. Wird ein System durch die Postionierung von Lautsprecher(n) und Mikrofon(en) im gleichen akustischen Raum instabil, äußert sich das durch unangenehmes Pfeifen („Howling“) und will unter allen Umständen vermieden werden. Diese Instabilitäten treten erst ab bestimmten Verstärkungen auf, sodass das System in seiner maximalen Verstärkung deutlich beschränkt ist. In der Literatur vorgestellte Methoden versuchen entweder das Pfeifen als Symptom zu unterdrücken, oder die akustische Rückkopplung selbst aufzuheben und damit die maximal mögliche Verstärkung, bei der das System noch stabil ist, zu erhöhen. Übliche Methoden können dabei in Gruppen kategorisiert werden: Phasenmodulation, Verstärkungsreduktion, räumliche Filterung und Raummodellierung. Da die Methoden der adaptiven Rückkopplungsunterdrückung aus der Gruppe der Raummodellierung theoretisch in der Lage sind, die akustische Rückkopplung vollständig aufzuheben, haben sich diese mit der Zeit durchgesetzt. Hier wurden verschiedene Ansätze vorgeschlagen, die auf verschiedenen Optimierungsmethoden wie Least Squares, (Normalised) Least Mean Squares, oder auch Kalman-Filtern im Zeit-, Frequenz-, Teilband-, oder anderen Transformationsbereichen basieren. Die in der Literatur betrachteten Szenarien sind in der Regel einkanalig mit einem Lautsprecher und einem Mikrofon. Da moderne Konferenztechnikkomponenten meist aus einem Lautsprecher und einem Mikrofonarray, also mehreren Mikrofonkapseln bestehen, liegt es nahe, dies auch bei der akustischen Rückkopplungsunterdrückung auszunutzen. Vereinzelte Ansätze mit einem Lautsprecher und mehreren Mikrofonen (meist zwei) kombinieren die adaptive Rückkopplungsunterdrückung mit weiteren Methoden. Das Mikrofonarray wird dabei beispielsweise durch Nullsteering-Beamforming [1] oder mehrkanaliger Rauschunterdrückung [2, 3] ausgenutzt und im Anschluss eine einkanalige, adaptive Rückkopplungsunterdrückung durchgeführt. Diese Arbeit schlägt vor, die Mikrofonarraysignale direkt im adaptiven Filter auszunutzen, indem der mehrkanalige, akustische Rückkopplungspfad mit Kalman-Filtern im Frequenzbereich geschätzt wird.

Kapitel 1 gibt eine kurze Einführung in das Thema, Kapitel 2 erläutert die Problemstellung und das Signalmodell. In Kapitel 3 werden die in der Literatur vorgeschlagenen Ansätze zur akustischen Rückkopplungsunterdrückung vorgestellt. Zum objektiven Vergleich der verschiedenen Ansätze werden in Kapitel 4 die in der Literatur gängigen und in dieser Arbeit verwendeten Qualitätsmaße definiert. Kapitel 5 stellt die adaptive Rückkopplungsunterdrückung im Detail vor und geht auf eine bestehende Methode im Frequenzbereich mit

Kalmanfiltern ein, welche als Referenz für den in dieser Arbeit vorgestellten Algorithmus verwendet. Kapitel 6 stellt das mehrkanalige, adaptive Kalmanfilter im Frequenzbereich vor und diskutiert seine Eigenschaften im Vergleich zum einkanaligen Fall. Das siebte Kapitel fasst die Arbeit zusammen und diskutiert weitere Forschungsansätze.

Abstract

Acoustic feedback has become an unavoidable issue, at the very least since the invention of electroacoustic sound reinforcement systems and, at the latest, with the ongoing development of hearing aids and conference systems. If a system becomes unstable due to the placement of speakers and microphones in the same acoustic space, this manifests as an unpleasant whistling sound (known as "howling") and must be avoided at all costs. These instabilities do not occur until certain gain levels are reached, which means that the system is significantly limited in terms of its maximum gain. The methods described in the literature aim either to suppress the whistling as a symptom or to eliminate the acoustic feedback itself, thereby increasing the maximum possible gain at which the system remains stable. Common methods can be categorized into the following groups: phase modulation, gain reduction, spatial filtering, and room modeling. Since adaptive feedback cancellation methods based on room modeling are theoretically capable of completely eliminating acoustic feedback, they have become widely adopted over time. Various approaches have been proposed here, based on different optimization methods such as least squares, (normalized) least mean squares, or Kalman filtering in the time, frequency, subband, or other transform domains. The scenarios discussed in the literature are typically single-channel, involving a single loudspeaker and a single microphone. Since modern conference technology systems typically consist of a loudspeaker and a microphone array—that is, multiple microphone capsules—it makes sense to try to take advantage of this in the acoustic feedback control stage. A few approaches using a single loudspeaker and multiple microphones (usually two) combine adaptive feedback suppression with other methods. The microphone array is utilized, for example, through null-steering beamforming [1] or multichannel noise cancellation [2, 3], followed by single-channel adaptive feedback suppression. This thesis proposes utilizing the microphone array signals directly within the adaptive filter by estimating the multichannel acoustic feedback path using Kalman filters in the frequency domain.

Chapter 1 provides a brief introduction to the topic; Chapter 2 explains the problem and the signal model. Chapter 3 presents the approaches to acoustic feedback control proposed in the literature. To provide an objective comparison of the various approaches, Chapter 4 defines the quality metrics commonly used in the literature and employed in this thesis. Chapter 5 provides a detailed overview of adaptive feedback cancellation and discusses an existing frequency-domain method using Kalman filters that is used as the reference algorithm in this thesis. Chapter 6 introduces the multichannel adaptive Kalman filter in the frequency domain and discusses its properties in comparison to the single-

channel case. Chapter 7 summarizes the work and discusses further research directions.

Danksagung

Mein besonderer Dank gilt Herrn Prof. Dr.-Ing. habil. Alexander Potchinkov als mein Betreuer und Erstgutachter. Ihm möchte ich für sein entgegengebrachtes Vertrauen und seine Unterstützung in fachlichen und nicht-fachlichen Belangen danken.
Bei Herrn Prof. Dr.-Ing. Christian Langen möchte ich für die Übernahme des Zweitgutachtens bedanken. Dankbar bin ich besonders für seine Anregungen und Unterstützung während der Themenfindungsphase.
Meinem ehemaligen Kollegen Dr.-Ing. Stephan Herzog gilt besonderer Dank für seine uneingeschränkte fachliche und persönliche Unterstützung. Er war zu jeder Zeit offen für Fragen und konnte mir in allen Belangen weiterhelfen.
Bei meinem ehemaligen Kommilitonen und nun Kollegen Marvin Reski möchte ich mich für seine Anregungen und zahlreichen konstruktiven Diskussionen bedanken.
Weiterhin möchte ich mich bei Helga Mosbach-Theiss, Dr.-Ing. Christian De Schryver und Dr.-Ing. Bernd Gutheil bedanken, die mich besonders in der Endphase meiner Arbeit motiviert und unterstützt haben.
Besonderer Dank gilt zuletzt meiner Familie, die mich zu jedem Zeitpunkt uneingeschränkt unterstützt hat.

Inhaltsverzeichnis

Abbildungsverzeichnis **xi**

Abkürzungsverzeichnis **xiii**

Symbolverzeichnis **xv**

1. Einführung **1**
1.1. Beiträge 1
1.2. Gliederung 2

2. Akustische Rückkopplung **3**
2.1. Problemstellung und Signalmodell 3
2.2. Stabilitätsanalyse einer einkanaligen Rückkopplungsstrecke . . 5
2.3. Hörgeräte, Beschallungs- und Konferenzsystems 7

3. Ansätze der akustischen Rückkopplungsunterdrückung **9**
3.1. Einleitung 9
3.2. Manuelle Ansätze 9
3.3. Phasen-Modulation 10
3.4. Verstärkungsreduktion 11
3.5. Räumliche Filterung 13
3.6. Raummodellierung 14
3.6.1. Adaptive Inverse Filtering (AIF) 14
3.6.2. Adaptive Feedback Cancellation (AFC) 14

4. Qualitätsmaße **19**
4.1. Einleitung 19
4.2. Added Stable Gain 19
4.3. Misadjustment 20
4.4. Effective Closed-Loop Gain (ECLG) 20
4.5. Short-Time Objective Intelligibility (STOI) 20
4.6. Perceptual Evaluation of Speech Quality (PESQ) 21
4.7. Log-Spectral Distance (SD) 22

5. Adaptive Rückkopplungsunterdrückung **23**
5.1. Einleitung 23
5.2. Optimale Lösung und Bias 25
5.3. Dekorrelation zum Erhalt einer unbiased Lösung 26
5.4. PEM-Based Frequency-Domain Kalman Filter (PEM-FDKF) . 29

6. Mehrkanalige adaptive Rückkopplungsunterdrückung im Frequenzbereich mit Kalmanfiltern **33**
6.1. Einleitung 33
6.2. Signalmodell 34
6.3. Formulierung des Algorithmus 35
6.4. Simulation 41
6.5. Zusammenfassung des Kapitels 51

7. Zusammenfassung **53**
7.1. Ausblick 54

Literatur **55**

A. Lebenslauf **67**

Abbildungsverzeichnis

2.1. Modell einer akustischen Rückkopplungsstrecke mit S Quellen, M Mikrofonen und L Lautsprechern. 4
2.2. Einkanalige akustische Rückkopplungsstrecke mit dem Quellsignal $v(t)$ und dem Lautsprechersignal $u(t)$ 6

4.1. Blockschaltbild zur Berechnung des Short-Time Objective Intelligibility (STOI) Qualitätsmaß [126] 21
4.2. Struktur zur Berechnung des Perceptual Evaluation of Speech Quality (PESQ) Qualitätsmaß [127] 22

5.1. Einkanalige akustische Rückkopplungsstrecke mit dem Quellsignal $v(t)$ und dem Lautsprechersignal $u(t)$ 23
5.2. Einkanalige akustische Rückkopplungsstrecke mit adaptiver Rückkopplungsunterdrückung 24
5.3. Einkanalige akustische Rückkopplungsstrecke mit dem Prediction-Error-Method-basierten Ansatz für adaptive Rückkopplungsunterdrückung (vgl. [114]) 28

6.1. Modell einer akustischen Rückkopplungsstrecke mit einer Quelle, M Mikrofonen und einem Lautsprecher. 35
6.2. Mehrkanalige akustische Rückkopplungsstrecke mit einem Lautsprecher und M Mikrofonen und dem Prediction-Error-Method-basierten Ansatz für adaptive Rückkopplungsunterdrückung (vgl. [114]) 36
6.3. Zusammensetzungen $\mathcal{M}_1 - \mathcal{M}_4$ des verwendeten kreisförmigen Mikrofonarrays zur Generierung von akustischen Rückkopplungsstrecken mit unterschiedlicher Kanalanzahl 42
6.4. Achtkanaliges, ringförmiges Mikrofonarray und ein Genelec-Lautsprecher zur Generierung von Impulsantworten der mehrkanaligen, akustischen Rückkopplungsstrecke 43
6.5. Impulsantwort und einseitiges Betragsfrequenzspektrum der Rückkopplungsstrecke zwischen Lautsprecher und dem ersten Mikrofon 43
6.6. Zeitverlauf des verwendeten Quellsignals 44
6.7. Added Stable Gain (ASG) in dB für die Mikrofonzusammensetzungen $\mathcal{M}_1 - \mathcal{M}_4$ mit entsprechend 1, 2, 4 und 8 Mikrofonen über der Zeit für verschiedene breitbandige Verstärkungen K_i im elektroakustischen Vorwärtspfad 47

6.8. Misadjustment (Mis) in dB für die Mikrofonzusammensetzungen $\mathcal{M}_1 - \mathcal{M}_4$ mit entsprechend 1, 2, 4 und 8 Mikrofonen über der Zeit für verschiedene breitbandige Verstärkungen K_i im elektroakustischen Vorwärtspfad 48
6.9. Effective Closed Loop Gain (ECLG) in dB für die Mikrofonzusammensetzungen $\mathcal{M}_1 - \mathcal{M}_4$ mit entsprechend 1, 2, 4 und 8 Mikrofonen in Bezug auf die Verstärkung im Vorwärtspfad über dem Maximum Stable Gain (MSG) 49
6.10. Perceptual Evaluation of Speech Quality (PESQ) für die Mikrofonzusammensetzungen $\mathcal{M}_1 - \mathcal{M}_4$ mit entsprechend 1, 2, 4 und 8 Mikrofonen in Bezug auf die Verstärkung im Vorwärtspfad über dem Maximum Stable Gain (MSG) 49
6.11. Short-Time Objective Intelligibility (STOI) für die Mikrofonzusammensetzungen $\mathcal{M}_1 - \mathcal{M}_4$ mit entsprechend 1, 2, 4 und 8 Mikrofonen in Bezug auf die Verstärkung im Vorwärtspfad über dem Maximum Stable Gain (MSG) 50
6.12. Log-Spectral Distance (SD) in dB für die Mikrofonzusammensetzungen $\mathcal{M}_1 - \mathcal{M}_4$ mit entsprechend 1, 2, 4 und 8 Mikrofonen in Bezug auf die Verstärkung im Vorwärtspfad über dem Maximum Stable Gain (MSG) 50

Abkürzungsverzeichnis

AEQ	Automatic Equalization
AFC	Adaptive Feedback Cancellation
AGC	Automatic Gain Control
AIF	Adaptive Inverse Filtering
AR	Autoregressiver Prozess
ARMA	Autoregressiver gleitender Mittelwert-Prozess
ASG	Added Stable Gain
DFT	Diskrete Fourier-Tranfsormation
ECLG	Effective Closed-Loop Gain
ERB	Equivalent Rectangular Bandwidth
FIR-Filter	Finite Impulse Respone Filter - Filter mit endlich langer Impulsantwort
IIR-Filter	Infinite Impulse Respone Filter - Filter mit unendlich langer Impulsantwort
LS	Least Squares (Methode der kleinsten Quadrate)
Mis	Misadjustment
MSG	Maximum Stable Gain
NHS	Notch-Filter based Howling Suppression
NLMS	Normalized Least Mean Squares
PA	Public Address
PEM	Prediction-Error-Method
PEM-FDKF	PEM-based Frequency-Domain Kalman Filter
PESQ	Perceptual Evaluation of Speech Quality
SD	Frequency-weighted log-spectral Distance
STOI	Short-Time Objective Intelligibility
WLP	Frequenzverzerrte, lineare Prädiktion

Symbolverzeichnis

$A(z,t)$	Inverse polynomiale Übertragungsfunktion des Quellsignalmodells
A	Übergangsfaktor zur Modellierung der Zeitvariabilität des akustischen Rückkopplungspfades
$E\{\cdot\}$	Erwartungswert
$F(z,t)$	Polynomiale Übertragungsfunktion des elektroakustischen Rückkopplungspfads
$F(\omega,t)$	Übertragungsfunktion des elektroakustischen Rückkopplungspfads im Kurzzeit-Frequenzbereich
$G(z,t)$	Polynomiale Übertragungsfunktion des elektroakustischen Vorwärtspfades
$G(\omega,t)$	Übertragungsfunktion des elektroakustischen Vorwärtspfads im Kurzzeit-Frequenzbereich
$H(z,t)$	Polynomiale Übertragungsfunktion des Quellsignalmodells
$H_{sm}(z,t)$	Polynomische Übertragungsfunktion von Quelle s zu Mikrofon m zum Zeitpunkt t
$J(\cdot)$	Kostenfunktion
$J(z,t)$	Frequenzabhängiger, mittelwertfreier Verstärkungsfaktor des elektroakustischen Vorwärtspfades
$K(t)$	Frequenzunabhängiger, mittlerer Verstärkungsfaktor des elektroakustischen Vorwärtspfades
L	Anzahl der Lautsprecher
M	Anzahl der Mikrofone
N	Länge der diskreten Fouriertransformation
R	Verschiebungsintervall des Zeitfensters
S	Anzahl der Schallquellen
$U(\omega,t)$	Lautsprechersignal im Kurzzeit-Frequenzbereich
$V(\omega,t)$	Quellsignal im Kurzzeit-Frequenzbereich
$[\cdot]^H$	Hermitescher Transponierungsoperator
$[\cdot]^{-1}$	Matrix-Invertierungsoperator
$[\cdot]^T$	Transponierungsoperator
α	Vergessenfaktor zur Näherung der Kovarianzmatrizen

$\angle G(\omega,t)F(\omega,t)$	Rückkopplungsphase des rückgekoppelten Systems im Kurzzeit-Frequenzspektrum
$\mathbf{\Psi}_{\hat{f}\hat{f}}(k)$	Kovarianzmatrix des geschätzten akustischen Rückkopplungspfades
$\mathbf{\Psi}_{ee}(k)$	Kovarianzmatrix des Rauschterms zur Modellierung des Quellsignals als autoregressiver Prozess
$\mathbf{\Psi}_{nn}(k)$	Kovarianzmatrix des Rauschterms zur Modellierung der zufälligen, nicht erfassbaren Komponenten des akustischen Rückkopplungspfades
$G(\omega,t)F(\omega,t)$	Rückkopplungsantwort des rückgekoppelten Systems im Kurzzeit-Frequenzspektrum
$\hat{A}(z,t)$	Schätzung der inversen polynomialen Übertragungsfunktion des Quellsignalmodells
$\hat{F}(z,t)$	Schätzung der polynomialen Übertragungsfunktion des elektroakustischen Rückkopplungspfads
$\hat{F}(\omega,t)$	Übertragungsfunktion des geschätzten elektroakustischen Rückkopplungspfads im Kurzzeit-Frequenzbereich
$\hat{\mathbf{f}}(t)$	Koeffizienten des geschätzten einkanaligen akustischen Rückkopplungspfades zwischen Lautsprecher und Mikrofon zum Zeitpunkt t
$\hat{x}(t)$	Schätzung des rückgekoppelten Signals zum Zeitpunkt t
$\hat{x}_a(t)$	Vorgefiltertes geschätztes Rückkopplungssignal
$\|\cdot\|$	Euklidische Norm zur Berechnung der Länge von Vektoren
$\mathbf{C}_a(k)$	Messmatrix des Zustandsmodells
$\mathbf{E}(k)$	Signalfenster des weißen Rauschsignals zur Modellierung des Quellsignals als autoregressiver Prozess im Frequenzbereich zum Zeitfenster k
$\mathbf{F}(k)$	Zeitfenster-basierter Koeffizientenvektor des akustischen Rückkopplungspfades im Frequenzbereich
$\mathbf{F}_N$	Diskrete $N \times N$-Fourier-Transformationsmatrix
$\mathbf{F}$	Akustischer Rückkopplungspfad
$\mathbf{G}$	Elektroakustischer Vorwärtspfad
$\mathbf{H}_{SM}(z,t)$	Übertragungsmatrix mit Übertragungsfaktoren von den Quellen zu den Mikrofonen zum Zeitpunkt t
$\mathbf{I}$	Einheitsmatrix
$\mathbf{J}_{n_A}$	Spiegelnde $n_A \times n_A$ Einheitsmatrix
$\mathbf{K}(k)$	Kalman Gain
$\mathbf{N}(k)$	Normalverteilter, mittelwerter Rauschterm zur Modellierung der nicht erfassbaren Komponenten des akustischen Rückkopplungspfades

$\mathbf{P}(\cdot\|\cdot)$	Prädiktionsfehlermatrix
$\mathbf{Q}$	Projektionsmatrix zur korrekten linearen Faltung im Frequenzbereich mit der Overlap-Save Methode
$\mathbf{R}_{uu}(t)$	Autokorrelationsmatrix des Lautsprechersignals $\mathbf{u}(t)$ zum Zeitpunkt t
$\mathbf{U}_a(k)$	Vorgefilterte, diagonalisierte $N \times N$ Lautsprechersignalmatrix im Frequenzbereich zum Zeitfenster k
$\mathbf{Y}_a(k)$	Vorgefilterte Mikrofonsignalfenster im Frequenzbereich zum Zeitfenster k
$\hat{\mathbf{f}}(t)$	Koeffizientenvektor des optimalen, geschätzten, akustischen Rückkopplungspfades
$\hat{\mathbf{f}}_{opt}(t)$	Koeffizientenvektor des optimalen, geschätzten, akustischen Rückkopplungspfades
$\overline{\mathbf{E}}(k)$	Zeitfenster-basierte $R \times M$ Matrix der M Rauschterme zur Modellierung des Quellsignals als autoregressiver Prozess
$\overline{\mathbf{F}}(k)$	Zeitfenster-basierte $R \times M$ Matrix der M Rückkopplungspfade
$\overline{\mathbf{Y}}(k)$	Zeitfenster-basierte $R \times M$ Matrix der M Mikrofonsignale
$\overline{\mathbf{Y}}_a(k)$	Zeitfenster-basierte $R \times M$ Matrix der M vorgefilterten Mikrofonsignale
$\mathbf{a}(t)$	Koeffizientenvektor des autoregressiven Quellmodells zum Zeitpunkt t
$\mathbf{e}(k)$	Signalfenster des weißen Rauschsignals zur Modellierung des Quellsignals als autoregressiver Prozess der Länge R zum Zeitfenster k
$\mathbf{f}(t)$	Koeffizienten des einkanaligen akustischen Rückkopplungspfades zwischen Lautsprecher und Mikrofon zum Zeitpunkt t
$\mathbf{f}(k)$	Blockbasierter Koeffizientenvektor des akustischen Rückkopplungspfades
$\mathbf{f}_{lm}(t)$	Koeffizienten des akustischen Rückkopplungspfades zwischen dem Lautsprecher-Mikrofon-Paar (l, m) zum Zeitpunkt t
$\mathbf{r}_{uv}(t)$	Kreuzkorrelationsvektor des Lautsprechersignals $\mathbf{u}(t)$ und dem Quellsignal $v(t)$ zum Zeitpunkt t
$\mathbf{r}_{uy}(t)$	Kreuzkorrelationsvektor des Lautsprechersignals $\mathbf{u}(t)$ und dem Mikrofonsignal $y(t)$ zum Zeitpunkt t
$\mathbf{u}(k)$	Lautsprechersignalfenster der Länge N zum Zeitfenster k
$\mathbf{u}(t)$	Vektor aller Lautsprechersignale zum Zeitpunkt t
$\mathbf{u}_a(k)$	Vorgefiltertes Lautsprechersignalfenster der Länge N zum Zeitfenster k
$\mathbf{v}(t)$	Vektor aller Quellsignale zum Zeitpunkt t
$\mathbf{y}(k)$	Mikrofonsignalfenster der Länge R zum Zeitfenster k
$\mathbf{y}(t)$	Vektor aller Mikrofonsignale zum Zeitpunkt t

$\mathbf{y}_a(k)$	Vorgefiltertes Mikrofonsignalfenster der Länge R zum Zeitfenster k
$\mathcal{M}_i$	i-Te Menge der aktiven Mikrofone des verwendeten Mikrofonarrays zur Generierung von mehrkanaligen Rückkopplungsstrecken mit unterschiedlicher Kanalanzahl
$\mathcal{P}$	Menge der Frequenzen, für die die Phasenbedingung des Nyqusist Stabilitätskriteriums erfüllt ist
ω	Kreisfrequenz
$\mathbf{MSG}_{AFC}$	Maximale Verstärkung des Systems mit Rückkopplungsunterdrückung in Dezibel (dB), bei der das System gerade noch stabil ist
$\mathbf{diag}\{\cdot\}$	Diagonalisierungsoperator
$\varepsilon(t)$	Vorgefiltertes Fehlersignal
$d(t)$	Fehlersignal zum Zeitpunkt t
$e(t)$	Weißes Rauschsignal zur Gewinnung des Quellsignals durch einen autoregressiven Prozess
$e_a(t)$	Vorgefiltertes weißes Rauschsignal zur Modellierung des Quellsignals durch einen autoregressiven Prozess
f_s	Abtastrate
k	Diskreter Zeitfenster-Index
n_F	Ordnung der Übertragungsfunktion eines Lautsprecher-Mikrofon-Paars im elektroakustischen Rückkopplungspfad
n_G	Ordnung der Übertragungsfunktion eines Mikrofon-Lautsprecher-Paars im elektroakustischen Vorwärtspfad
n_H	Ordnung der Impulsantwort von Quelle s zu Mikrofon m
t	Diskrete Zeitvariable
$u(t)$	Lautsprechersignal zum Zeitpunkt t
$u_a(t)$	Vorgefiltertes Lautsprechersignal
$u_l(t)$	l-tes Lautsprechersignal im diskreten Zeitbereich
$v(t)$	Quellsignal zum Zeitpunkt t
$v_s(t)$	s-tes Quellsignal im diskreten Zeitbereich
$x(t)$	Rückgekoppeltes Signal zum Zeitpunkt t
$y(t)$	Mikrofonsignal zum Zeitpunkt t
$y_a(t)$	Vorgefiltertes Mikrofonsignal
$y_m(t)$	m-tes Mikrofonsignal im diskreten Zeitbereich
z	Verzögerungsoperator im diskreten Zeitbereich
$\lvert G(\omega,t)F(\omega,t)\rvert$	Rückkopplungsverstärkung des rückgekoppelten Systems im Kurzzeit-Frequenzspektrum

1. Einführung

Mit Erfindung der ersten Mikrofone durch Philipp Reis (Kontaktmikrofon, 1861), David Edward Hughes (Kohlemikrofon, 1878) und Edward Christopher Wente (Kondensatormikrofon, 1917) und der ersten Lautsprecher durch Peter L. Jensen und Edwin Priham („Magnavox“, 1915) und Edward Kellog und Chester Rice bei Western Electric (elektrodynamischer Konuslautsprecher, 1925) [4, 5] wurden die ersten Beschallungsanlagen (PA) möglich. Der erste Einsatz von Mikrofonen und Lautsprechern zur Beschallung einer größeren Menschenmenge ist auf den 10. Dezember 1915 mit dem „Magnavox“ von Jensen und Priham dokumentiert [5]. Spätestens mit steigender Bekanntheit und Weiterentwicklung solcher Beschallungsanlagen, wurde die Rückkopplung des verstärkten Signals in das Mikrofon zu einem Problem, das es zu lösen galt. Auch die Erfindung elektrischer Hörgeräte durch Miller Reese Hutchison im Jahr 1898 und deren stetigen Weiterentwicklung [6] trägt bis heute zur Weiterentwicklung von Methoden zur akustischen Rückkopplungsunterdrückung bei. Mit der fortschreitenden Hybridisierung von Konferenzen und Meetings wurden in den letzten Jahren zusätzlich Konferenzsysteme entwickelt, die Mikrofonarray und Lautsprecher zugleich sind. Diese sind meist als Tisch- oder Deckenarray ausgeführt, wie beispielsweise das MXA920 Decken-Mikrofonarray [7] von Shure, oder die TeamConnect-Reihe von Sennheiser [8]. Viele davon sind in der Lage, die Sprecherinnen und Sprecher vor Ort über den integrierten Lautsprecher zu verstärken um die Verständlichkeit zu verbessern. In der Folge sind auch hier Algorithmen zur akustischen Rückkopplungsunterdrückung nötig. In der Literatur haben sich Algorithmen der sogenannten adaptiven Rückkopplungsunterdrückung durchgesetzt. Diese versuchen den akustischen Rückkopplungspfad über ein adaptives Filter zu schätzen. Durch Filterung des Lautsprechersignals mit dem geschätzten Rückkopplungspfad lässt sich dieses vom Mikrofonsignal abziehen. In der Folge erhält man ein Signal, das im Idealfall dem reinen Quellsignal entspricht. Diese Methoden sind also theoretisch in der Lage, die akustische Rückkopplung vollständig aufzuheben. Über die Jahre wurden verschiedene Ansätze vorgeschlagen, die auf verschiedenen Optimierungsmethoden wie Least Squares, (Normalised) Least Mean Squares, oder auch Kalman-Filtern im Zeit-, Frequenz-, Teilband-, oder anderen Transformationsbereichen basieren. Meist sind die dort betrachteten Szenarien einkanalig mit einem Lautsprecher und einem Mikrofon.

1.1. Beiträge

Der Beitrag dieser Arbeit ist die Ausnutzung der Mikrofonarraysignale im adaptiven Filter, indem der mehrkanalige, akustische Rückkopplungspfad mit

Kalman Filtern im Frequenzbereich geschätzt wird. Hierdurch kann ein Verstärkungsgewinn von bis zu 5 dB im Vergleich zum einkanaligen Rückkopplungspfad gewonnen werden. Bisherige mehrkanalige Methoden mit einem Lautsprecher und mehreren Mikrofonen nutzen die Mehrkanaligkeit beispielsweise durch Nullsteering-Beamformer [1] oder mehrkanalige Rauschunterdrückung [3] aus und kombinieren diese mit einer einkanaligen adaptiven Rückkopplungsunterdrückung.

1.2. Gliederung

Kapitel 1 gibt eine kurze Einführung in das Thema, Kapitel 2 erläutert die Problemstellung und das Signalmodell. In Kapitel 3 werden die in der Literatur vorgeschlagenen Ansätze zur akustischen Rückkopplungsunterdrückung vorgestellt. Zum objektiven Vergleich der verschiedenen Ansätze werden in Kapitel 4 die in der Literatur gängigen und in dieser Arbeit verwendeten Qualitätsmaße definiert. Kapitel 5 stellt die adaptive Rückkopplungsunterdrückung im Detail vor und geht auf eine bestehende Methode im Frequenzbereich mit Kalmanfiltern ein, die im Rahmen dieser Arbeit als Referenz-Algorithmus verwendet wird. Kapitel 6 stellt das mehrkanalige, adaptive Kalmanfilter im Frequenzbereich vor und diskutiert seine Eigenschaften im Vergleich zum einkanaligen Fall. Das siebte Kapitel fasst die Arbeit zusammen und diskutiert weitere Forschungsansätze.

2. Akustische Rückkopplung

Inhalt

2.1. Problemstellung und Signalmodell 3
2.2. Stabilitätsanalyse einer einkanaligen Rückkopplungsstrecke . . 5
2.3. Hörgeräte, Beschallungs- und Konferenzsystems 7

2.1. Problemstellung und Signalmodell

Akustische Rückkopplung tritt auf, wenn eine Quelle über ein Mikrofon aufgenommen und im gleichen akustischen Raum über einen Lautsprecher verstärkt wiedergegeben wird. Dies ist beispielsweise bei Hörgeräten, bei klassischen Beschallungssystemen, auch PA-Systeme (Public Address) genannt, oder modernen Konferenzsystem der Fall. Zusätzlich zum Quellsignal wir auch das verstärkte Signal vom Mikrofon aufgenommen und wiederum verstärkt. Dieser geschlossene Signalkreislauf kann zur Instabilität eines Systems führen, die sich als unangenehmes Pfeifen bei bestimmten Frequenzen bemerkbar macht. Um die akustische Rückkopplung zu vermindern, gibt es verschiedene Methoden der akustischen Rückkkopplungsunterdrückung. Das Ziel dieser ist entweder, das Pfeifen als Symptom zu unterdrücken (reaktiv), oder zu verhindern, dass die Bedingungen für die Instablität eines Systems erfüllt sind (proaktiv).

Um den allgemeinen Fall eines solchen Systems zu modellieren, werden S Quellen, M Mikrofone und L Lautsprecher betrachtet. Alle kontinuierlichen Signale werden dabei als bandbegrenzt angenommen, sodass sie mit gängigen Abtastraten abgetastet und wert- und zeitdiskret dargestellt werden können. Werden die Quellsignale zum diskreten Zeitpunkt t über $v_s(t),\ s = 1, \ldots, S$, die Mikrofonsignale über $y_m(t), m = 1, \ldots, M$ und die Lautsprechersignale mit $u_l(t), l = 1, \ldots, L$ definiert, kann das grundlegende System über das Blockdiagramm in Abbildung 2.1 modelliert werden [9].
In diesem Fall trifft der Schall jeder Quelle auf jedes Mikrofon, sodass eine $M \times S$ Übertragungsmatrix

$$\mathbf{H}_{SM}(z,t) = \begin{bmatrix} H_{11}(z,t) & H_{21}(z,t) & \cdots & H_{S1}(z,t) \\ H_{12}(z,t) & H_{22}(z,t) & \cdots & H_{S2}(z,t) \\ \vdots & \vdots & \ddots & \vdots \\ H_{1M}(z,t) & H_{2M}(z,t) & \cdots & H_{SM}(z,t) \end{bmatrix} \tag{2.1}$$

definiert werden kann, deren Elemente

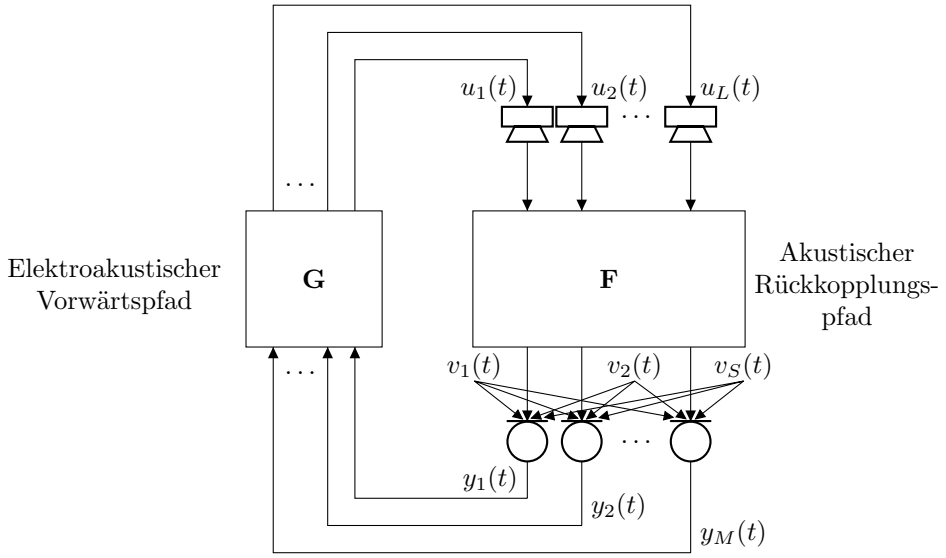


Abbildung 2.1.: *Modell einer akustischen Rückkopplungsstrecke mit S Quellen, M Mikrofonen und L Lautsprechern.*

$$H_{sm}(z,t) = h_{sm}^{(0)}(t) + h_{sm}^{(1)}(t)z^{-1} + \cdots + h_{sm}^{(n_H)}(t)z^{-n_H} \tag{2.2}$$

$$= \mathbf{z}^T \mathbf{h}_{sm}(t) \tag{2.3}$$

mit $\mathbf{h}_{sm} = \begin{bmatrix} h_{sm}^{(0)}(t) & h_{sm}^{(1)}(t) & \cdots & h_{sm}^{(n_H)}(t) \end{bmatrix}^T$ und $\mathbf{z} = \begin{bmatrix} 1 & z^{-1} & \cdots & z^{-n_H} \end{bmatrix}^T$ den polynomischen Übertragungsfunktionen der Ordnung n_H von Quelle s zu Mikrofon m entsprechen. $[\cdot]^T$ ist dabei der Transponierungsoperator und z stellt den Verzögerungsoperator im diskreten Zeitbereich dar. Der Ausdruck $z^{-1}h_{sm}^{(1)}(t)$ ist also äquivalent zu $h_{sm}^{(1)}(t-1)$.
Dieser Darstellung entsprechend kann auch die akustische Kopplung zwischen einem Lautsprecher-Mikrofon-Paar (l,m) mit $l = 1,\ldots,L$ und $m = 1.\ldots,M$ als zeitvariante, lineare Raumimpulsantwort der endlichen Ordnung n_F modelliert werden, sodass sie sich durch ein Filter mit endlich langer Impulsantwort (FIR-Filter) mit der polynomischen Übertragungsfunktion

$$F_{lm}(z,t) = f_{lm}^{(0)}(t) + f_{lm}^{(1)}(t)z^{-1} + \ldots + f_{lm}^{(n_F)}(t)z^{-n_F} \tag{2.4}$$

$$= \mathbf{z}^T \mathbf{f}_{lm}(t) \tag{2.5}$$

mit $\mathbf{f}_{lm}(t) = \begin{bmatrix} f_{lm}^{(0)}(t) & f_{lm}^{(1)}(t) & \cdots & f_{lm}^{(n_f)}(t) \end{bmatrix}^T$ darstellen lässt. Der mehrkanalige Rückkopplungspfad kann so über die Übertragungsfunktionen der Lautsprecher-Mikrofon-Paare als $M \times L$ Matrix

$$\mathbf{F}(z,t) = \begin{bmatrix} F_{11}(z,t) & \cdots & F_{1L}(z,t) \\ \vdots & \ddots & \vdots \\ F_{M1}(z,t) & \cdots & F_{ML}(z,t) \end{bmatrix} \tag{2.6}$$

dargestellt werden.
Der elektroakustische Vorwärtspfad $\mathbf{G}(z,t)$ müsste durch den üblichen Einsatz von nichtlinearer und zeitvarianter Signalverarbeitung wie zum Beispiel Kompression als nichtlineares und zeitvariantes Filter modelliert werden. Um jedoch eine Stabilitätsanalyse ausführen zu können, wird der elektroakustische Vorwärtspfad in der Regel wie der Rückkopplungspfad auch als lineare, zeitvariante Übertragungsfunktion endlicher Ordnung n_G zwischen den jeweiligen Mikrofon-Lautsprecher-Paaren (m,l) mit

$$G_{ml}(z,t) = g_{ml}^{(0)}(t) + g_{ml}^{(1)}(t)z^{-1} + \ldots + g_{ml}^{(n_G)}(t)z^{-n_G} \tag{2.7}$$

modelliert, sodass sich der mehrkanalige Vorwärtspfad zur $L \times M$-Matrix

$$\mathbf{G}(z,t) = \begin{bmatrix} G_{11}(z,t) & \cdots & G_{1M}(z,t) \\ \vdots & \ddots & \vdots \\ G_{L1}(z,t) & \cdots & G_{LM}(z,t) \end{bmatrix} \tag{2.8}$$

ergibt.
Werden abschließend die mehrkanaligen Quell-, Mikrofon- und Lautsprechersignale zum Zeitpunkt t zu den Vektoren

$$\mathbf{v}(t) = \begin{bmatrix} v_1(t) & v_2(t) & \cdots & v_S(t) \end{bmatrix}^T \tag{2.9}$$

$$\mathbf{y}(t) = \begin{bmatrix} y_1(t) & y_2(t) & \cdots & y_M(t) \end{bmatrix}^T \tag{2.10}$$

$$\mathbf{u}(t) = \begin{bmatrix} u_1(t) & u_2(t) & \cdots & u_L(t) \end{bmatrix}^T \tag{2.11}$$

zusammengefasst, kann das rückgekoppelte System über

$$\mathbf{y}(t) = \mathbf{F}(z,t)\mathbf{u}(t) + \mathbf{H}_{SM}(z,t)\mathbf{v}(t) \tag{2.12}$$

$$\mathbf{u}(t) = \mathbf{G}(z,t)\mathbf{y}(t) \tag{2.13}$$

beschrieben werden. In der Literatur wird in der Regel davon ausgegangen, dass Quellen so stark gerichet und so nah an einem Mikrofon positioniert sind, dass der Schall einer Quelle sich nur zu dem der Quelle zugehörigen Mikrofon ausbreitet. Somit vereinfacht sich die Übertragungsmatrix $\mathbf{H}_{SM}(z,t)$ unter Vernachlässigung der Übertragungsstrecke zur Einheitsmatrix $\mathbf{I}$ und damit Gleichung (2.12) zu

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{F}(z,t)\mathbf{u}(t) + \mathbf{I}\mathbf{v}(t) \\ &= \mathbf{F}(z,t)\mathbf{u}(t) + \mathbf{v}(t). \end{aligned} \tag{2.14}$$

2.2. Stabilitätsanalyse einer einkanaligen Rückkopplungsstrecke

Da in der Literatur überwiegend einkanalige Rückkopplungsstrecken betrachtet werden [9] und sich damit die Stabilitätsanalyse des Systems deutlich vereinfacht, wird hierzu das allgemeine auf das einkanalige System in Abbildung

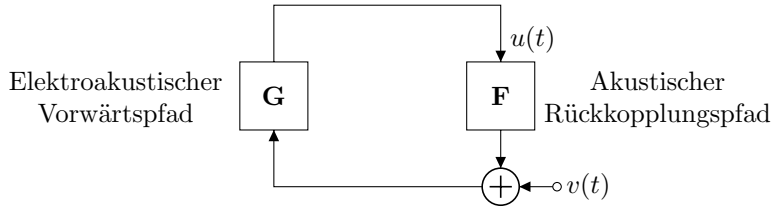


Abbildung 2.2.: *Einkanalige akustische Rückkopplungsstrecke mit dem Quellsignal $v(t)$ und dem Lautsprechersignal $u(t)$*

2.2 reduziert. Durch diese Vereinfachung ergibt sich das einkanalige System zum Zeitpunkt t zu

$$y(t) = F(z,t)u(t) + v(t). \tag{2.15}$$
$$u(t) = G(z,t)y(t) \tag{2.16}$$

Wird nun das Lautsprechersignal, das Quellsignal, der Vorwärtspfad und der Rückkopplungspfad im Kurzzeit-Frequenzspektrum mit der Kreisfrequenz $\omega \in [0, 2\pi]$ entsprechend durch $U(\omega,t)$, $V(\omega,t)$, $G(\omega,t)$ und $F(\omega,t)$ betrachtet, berechnet sich das Lautsprechersignal aus Abbildung 2.2 zu

$$U(\omega,t) = G(\omega,t)V(\omega,t) + G(\omega,t)F(\omega,t)U(\omega,t) \tag{2.17}$$

und somit die Übertragungsfunktion des geschlossenen Signalkreislaufs zu

$$\frac{U(\omega,t)}{V(\omega,t)} = \frac{G(\omega,t)}{1 - G(\omega,t)F(\omega,t)}. \tag{2.18}$$

Die Frequenzfunktion $G(\omega,t)F(\omega,t)$ im Nenner wird oft als Rückkopplungsantwort bezeichnet. Für den zugehörigen Betragsfrequenzgang $|G(\omega,t)F(\omega,t)|$ und den Phasenfrequenzgang $\angle G(\omega,t)F(\omega,t)$ werden die Begriffe Rückkopplungsverstärkung und Rückkopplungsphase verwendet. Mit der Annahme, dass sich der Vorwärtspfad und der Rückkopplungspfad nur langsam über der Zeit ändern, kann das Stabilitätskriterium von Nyquist [10] angewendet werden. Dieses besagt, dass ein rückgekoppeltes System unstabil ist, sobald für mindestens eine Kreisfrequenz ω die Bedingungen

$$\begin{cases} |G(\omega,t)F(\omega,t)| \geq 1 & (2.19) \\ \angle G(\omega,t)F(\omega,t) = n2\pi, \quad n \in \mathbb{Z} & (2.20) \end{cases}$$

erfüllt ist. Das Ziel der akustischen Rückkopplungsunterdrückung ist stets, mindestens eine der beiden Bedingungen (2.19) und (2.20) zu verhindern.
Mithilfe des Stabilitätskriteriums und einer Zerlegung des elektroakustischen Vorwärtspfades kann die maximal erreichbare Verstärkung des rückgekoppelten Systems bestimmt werden, bei dem das System gerade noch stabil ist. Hierzu wird der elektroakustische Vorwärtspfad in einen breitbandigen Verstärkungsfaktor

$$K(t) = \frac{1}{2\pi} \int_0^{2\pi} |G(\omega, t)| \, d\omega, \tag{2.21}$$

und den Faktor $J(z,t)$ aufgeteilt. $K(t)$ entspricht der frequenzunabhängigen, mittleren Verstärkung des Vorwärtspfads und $J(z,t)$ der frequenzabhängigen, mittelwertfreien Verstärkung des Vorwärtspfad, sodass

$$G(\omega, t) = K(t)J(z,t). \tag{2.22}$$

Mit der Menge der Frequenzen

$$\mathcal{P} = \{\omega | \measuredangle G(\omega,t)F(\omega,t) = n2\pi\}, \tag{2.23}$$

für die Bedingung (2.20) erfüllt ist und der Annahme, dass $J(z,t)$ bekannt und $K(t)$ veränderbar ist, kann die maximale Verstärkung, für die das System gerade noch stabil ist (Maximum Stable Gain (MSG)) über

$$\begin{aligned} \mathrm{MSG}(t) \overset{\Delta}{=} 20\log_{10} K(t) \quad & \text{unter der Bedingung} \\ & \max_{\omega\in\mathcal{P}} |G(\omega,t)F(\omega,t)| = 1 \qquad (2.24) \\ = -20\log_{10}\left[\max_{\omega\in\mathcal{P}} |J(z,t)F(\omega,t)|\right] & \qquad (2.25) \end{aligned}$$

definiert werden [9]. Mit der Vereinfachung, dass der Vorwärtspfad nur aus einer frequenzunabhängigen Verstärkung besteht, sodass $J(z,t) = \text{const.} = 1$, vereinfacht sich Gleichung (2.25) zu

$$\begin{aligned} \mathrm{MSG}(t) &= -20\log_{10}\left[\max_{\omega\in\mathcal{P}} |F(\omega,t)|\right] \qquad (2.26) \\ &= 10\log_{10}\left[\max_{\omega\in\mathcal{P}} \frac{1}{|F(\omega,t)|^2}\right]. \qquad (2.27) \end{aligned}$$

2.3. Hörgeräte, Beschallungs- und Konferenzsystems

Die Modelle des geschlossenen Signalkreislaufs und der Methoden der akustischen Rückkopplungsunterdrückung sind für Hörgeräte, klassische Beschallungs- und Konferenzsysteme grundlegend identisch. Es existieren dennoch ein paar Unterschiede, die es zu beachten gilt. Während in Hörgeräten und teilweise Konferenzsystemen die Positionen von Mikrofonen und Lautsprechern vor allem relativ zueinander im Voraus bekannt sind, ist das in klassischen Beschallungssystemen in der Regel nicht der Fall. Besonders im Livebetrieb sind die Anzahl und Positionen der Mikrofone stark variabel und verändern sich mitunter auch während des Betriebs. Vereinfachungen, die auf Basis von gemittelten Impulsanworten beruhen, sind dort also meist nicht anwendbar. Während in Hörgeräten kaum Nachhall im Rückkopplungspfad vorhanden ist, ist das in Konferenzsystemen und Beschallungsanlagen nicht der Fall. Dies

führt zu deutlich längeren Impulsantworten, die es zum Beispiel in der adaptiven Rückkopplungsunterdrückung zu schätzen gilt und bedeutet höheren Rechenaufwand und je nach Algorithmus unzumutbare Verzögerungen im Signalpfad.
Auch wenn viele der bekannten Ansätze vor allem in den letzten Jahren im Bezug auf Hörgeräte betrachtet wurden, sind die grundlegenden Überlegungen und Modelle bei allen Anwendungsfällen sehr ähnlich. Im Verlauf dieser Arbeit wird daher nur ausdrücklich unterschieden, wenn es nötig ist.

3. Ansätze der akustischen Rückkopplungsunterdrückung

Inhalt

3.1. Einleitung 9
3.2. Manuelle Ansätze 9
3.3. Phasen-Modulation 10
3.4. Verstärkungsreduktion 11
3.5. Räumliche Filterung 13
3.6. Raummodellierung 14
3.6.1. Adaptive Inverse Filtering (AIF) 14
3.6.2. Adaptive Feedback Cancellation (AFC) 14

3.1. Einleitung

Durch die akustische Rückkopplung des Lautsprechersignals in das Mikrofonsignal können unangenehme, pfeifende Geräusche („Howling“) entstehen. Diese sind besonders beim Einsatz von Hörgeräten kritisch, da hier der Lautsprecher direkt am Ohr positioniert ist und so im schlimmsten Fall das Gehör noch weiter geschädigt werden könnte. Des Weiteren wird durch die akustische Rückkopplung die maximal verfügbare Verstärkung begrenzt, da man das System in einem Bereich betreiben sollte, in dem diese pfeifenden Geräusche noch nicht auftreten. Es liegt also nahe, Ansätze und Methoden zu entwickeln, die entweder das „Howling“ als Symptom der Rückkopplung unterdrücken, oder die akustische Rückkopplung selbst so gut wie möglich aufheben. In den letzten 110 Jahren wurden verschiedene Ansätze vorgeschlagen, die nach [9] in vier Kategorien eingeteilt werden können: Phasen-Modulation, Verstärkungsreduktion, räumliche Filterung und Raummodellierung. Diese werden im Folgenden vorgestellt.

3.2. Manuelle Ansätze

Die ersten Ansätze zur Rückkopplungsunterdrückung waren manuelle Methoden. Dazu gehören beispielsweise die korrekte Positionierung und Nutzung von gerichteten Mikrofonen und Lautsprechern, das Herausfiltern einzelner Raummoden durch Notch-Filter, oder Filterung des gesamten Raumfreuenzspektrums mit Terzbandfiltern [11]. Diese manuellen Methoden werden auch

heute noch gerne von Tontechnikern im Live-Betrieb verwendet [9]. Da man hier jedoch nur eine Vergrößerung des MSG von $5 - 8$ dB erwarten kann [11] und das Ziel eine automatische Rückkopplungsunterdrückung ist, sind diese Methoden lediglich der Vollständigkeit halber erwähnt und werden im Folgenden nicht weiter betrachtet.

3.3. Phasen-Modulation

Die ersten Ansätze der automatischen Rückkopplungsunterdrückung gehören zur Gruppe der Phasen-Modulation. Das Ziel dieser ist es im Allgemeinen, die Phasenbedingung (2.20) des Nyquistkriteriums abzuschwächen.
Die Anfänge der Phasen-Modulation lassen sich hauptsächlich auf Schröder mit dem Ansatz der Frequenzverschiebung zurückführen [12–15]. Diese hat zum Ziel, das Betragsfrequenzspektrum der Systemübertragungsfunktion im elektroakustischen Vorwärtspfad zu glätten. Durch diese Glättung ist die maximal erreichbare stabile Verstärkung nicht mehr von den Maxima, sondern vom Mittel des Betragsfrequenzspektrums abhängig [14]. Die Frequenzverschiebung wird zwar der Phasen-Modulation zugeschrieben, da sie einen Spezialfall dieser darstellt [16, 17], zielt jedoch auf Grund der Glättung im Betragsfrequenzbereich darauf ab, die Amplitudenbedingung (2.19) des Nyquist-Kriteriums zu verhindern. Theoretisch kann der MSG in Abhängigkeit von den Eigenschaften des akustischen Rückkopplungspfades durch den Einsatz von Frequenzverschiebung um bis zu 12-14 dB erhöht werden [1, 12]. Jedoch wird das Signal durch den Einsatz des Algorithmus im Vorwärtspfad verzerrt, was hauptsächlich als schlagende Geräusche wahrgenommen werden kann. In der Praxis ist der Verstärkungsgewinn daher auf 6 dB unterhalb des theoretisch erreichbaren MSG beschränkt [14]. Die hörbaren Effekte hängen größtenteils von der Wahl der tatsächlichen Frequenzverschiebung ab. Schröder argumentiert, dass eine Verschiebung um 5 Hz optimal sei, da der mittlere Abstand zweier Spitzen im Betragsfrequenzgang von Raumimpulsantworten ungefähr 10 Hz beträgt [15]. In [18] zeigt Svensson jedoch, dass schon eine Verschiebung um $2,3 - 4,2$ Hz hörbare Artefakte erzeugt. Erste Implementierungen basierten auf analoger Einseitenbandmodulation [19], oder der analogen Hilberttransformation [20]. Später wurden auch digitale Ansätze entwickelt, die zum Beispiel auf der abgeschnittenen Hilberttransformation als FIR-Filter basieren [21, 22]. Ein großer Nachteil der Frequenzverschiebung ist, dass durch deren Einsatz die Harmonik des Signals verloren geht, was vor allem bei musikalischen Signalen hörbar ist. Alisobhani und Knorr [23] konnten jedoch zeigen, dass bei Einsatz einer Bandbreitenkompression die Harmonik erhalten bleibt und der MSG um bis zu 10 dB gesteigert werden kann.
Weitere Methoden der Phasenmodulation basieren auf der tatsächlichen Phasenmodulation [24, 25], Frequenzmodulation [26], Amplitudenmodulation [16, 17] und Delaymodulation [9, 25]. Auch diese werden im elektroakustischen Vorwärtspfad eingesetzt, weshalb auch hier der Verstärkungsgewinn vor allem bei musikalischen Signalen auf Grund schnell hörbarer Artefakte begrenzt ist [18]. Svensson [16] und Nielsen und Svensson [17] konnten den Ansatz der Phasen-

modulation zur Rückkopplungsunterdrückung verallgemeinern, indem sie die Modulatoren als lineare, periodisch zeitvariante Filter verstehen. Die Hauptvorteile der Phasenmodulation liegen in der Einfachheit des Konzepts dem geringen Rechenaufwand. Der größter Aufwand liegt dabei in der Berechnung des analytischen Mikrofonsignals für Phasenmodulation, Frequenzmodulation und Frequenzverschiebung und den Fractional Delay Inerpolationsfiltern für Delaymodulation. Wenige Überlegungen müssen auch für das Design durchgeführt werden, da nur die Art der Modulation, die Modulationsfrequenz und der Modulationsindex ausgewählt werden müssen. Die Wahl der Modulationsfrequenz hängt dabei hauptsächlich von der Wahl der Modulation ab. Bei der Frequenzverschiebung muss die Modulationsfrequenz höher sein als bei den anderen Modulationsarten. Eine höhere Modulationsfrqeuenz führt jedoch früher zu wahrnembaren Artefakten [17]. Die Wahl des Modulationsindex' stellt einen Kompromiss dar. Zum Einen führt ein Modulationsindex, der sich den Nullstellen der Besselfunktionen erster Ordnung [27] nähert, zu einer besseren Trägerunterdrückung und in der Folge zu einem größeren Verstärkungsgewinn. Auf der anderen Seite erzeugt ein größerer Modulationsindex mehr Seitenbänder, von denen erwartet wird, dass sie die Leistung der Rückkopplungsunterdrückung verbessern [17]. Die meisten Anwendungen der Phasenmodulation beziehen sich auf den einkanaligen Rückkopplungsfall. Untersuchungen für den mehrkanaligen Fall mit Frequenzverschiebung haben gezeigt, dass sich der Stabilitätsgewinn mit der Anzahl der Mikrofone verringert [28].

3.4. Verstärkungsreduktion

Ein zeitlich erst nach der Phasenmodulation aufkommender, aber deutlich intuiviverer Ansatz ist die Automatisierung der Maßnahmen, die beispielsweise ein Tontechniker im Falle von Instabilität durchführen würde. Wie der Name schon andeuetet, reduzieren die Ansätze der Verstärkungsreduktion die Verstärkung im elektroakustischen Vorwärtspfad. Die Rückkopplungsunterdrückung setzt daher bei der Verhinderung der Amplitudenbedingung (2.19) des Nyquist-Kriteriums an. Je nach Bandbreite, in der die Maßnahmen agieren, lassen sie sich in verschiedene Kategorien aufteilen [9]: Breitbandige Verstärkungsreduktion (Automatic Gain Control (AGC)), Teilbandige Verstärkungsreduktion (Automatic Equalization (AEQ)) und Notchfilter-basierte Verstärkungsreduktion (Notch-Filter based Howling Suppression (NHS)). Jede dieser Methoden muss aktiviert werden, sobald mindestens eine Tendenz zur Instabilität erkannt wird. Dies passiert in den meisten Fällen reaktiv. Eine eventuell sogar hörbare Instablität muss also schon vorliegen, bevor die Maßnahmen greifen können. Immerhin ein paar wenige Ansätze der Verstärkungsreduktion sind bekannt, die proaktiv arbeiten [9]. Diese basieren zum Beispiel auf einer Echtzeitmessung des Betragsfrequenzgangs des Rückkopplungspfades [29], oder der rechtzeitigen Erkennung des spektralen Akkumulierungseffektes, der bei kritischen Frequenzen im Falle einer Instabilität vorliegt [30, 31].
Zur Detektion von Instabilitäten wurden mit der Zeit verschiedene Kriterien vorgeschlagen, die auf zeitlichen oder spektralen Eigenschaften basieren

[9, 32]. Spektrale Eigenschaften versuchen anhand des gesamten Spektrums, Harmonischen oder den benachbarten Frequenzbändern, eine Instabilität zu erkennen und sie von tonalen Komponenten zu unterscheiden. Kriterien, die auf zeitlichen Eigenschaften basieren, versuchen dagegen, eine Instabilität auf Grund ihrer längeren Dauer und der exponentiell steigenden Energie im Vergleich zu tonalen Komponenten zu erkennen. Die Nutzung von lediglich einer Eigenschaft führt häufig dazu, dass tonale Komponenten fälschlicherweise als Instabilitäten erkannt werden. Eine Kombination der zeitlichen und spektralen Eigenschaften zur Detektion und Verifikation liegt daher nahe [1, 32, 33].

Wird eine Instabilität erkannt, wird eines der drei genannten Verfahren aktiv. Die breitbandige Verstärkungsreduktion (AGC) lässt sich größtenteils auf Patronis [34] zurückführen. Mit der Annahme, dass pfeifende Anteile im Signal länger klingen als tonale Quellkomponenten, werden Instabilitäten detektiert und anschließend die Verstärkung breitbandig reduziert. Nach einer vordefinierten Zeit wird die Verstärkung automatisch wieder auf den ursprünglichen Wert angehoben. Später wurden weitere Kriterien zur Erkennung herangezogen. Ando [35] zieht für die Detektion das Leistungsverhältnis zwischen benachbarten Teilbändern im Frequenzbereich heran, reduziert die Verstärkung jedoch breitbandig. Durch die breitbandige Verstärkungsreduktion wird eine vorliegende und erkannte Instabilität zwar temporär verhindert, der MSG aber nicht erhöht, da nicht die spektrale Form der Rückkopplungsverstärkung verändert wird [1, 9].

Anknüpfend an Andos Ansatz [35] wird bei der AEQ zusätzlich zur Detektion auch die Verstärkungsreduktion in Teilbändern durchgeführt. Osmanovic et al. [31] verwenden beispielsweise eine Kombination aus einer zeitlichen und einer spektralen Eigenschaft zur Detektion und 14 logarithmisch angeordnete Bandstop Filter mit unendlich langer Impulsantwort (IIR-Filter) 8. Ordnung zur Unterdrückung von Instabilitäten.

Die Methoden der NHS hingegen versuchen Instabilitäten so schmalbandig wie möglich herauszufiltern, um das Signal möglichst wenig hörbar zu verfärben. Je nachdem, ob die Detektion und Filterung gleichzeitig oder separat ausgeführt wird, spricht man von einstufiger oder zweistufiger Implementierung [9]. Die ersten NHS-Methoden waren einstufig und nutzten adaptive Notchfilter als FIR-Filter [36–38]. Da für eine ausreichende Schmalbandigkeit jedoch eine sehr hohe Ordnung für FIR-Filter nötig ist, findet man später lediglich Implementierungen, die auf IIR-Filtern basieren [39–41]. Diese können effizient in einer biquadratischen Struktur implementiert werden und benötigen deutlich weniger Koeffzienten und Rechenleistung [42]. Nachteil hierbei ist jedoch, dass für die Berechnung der adaptiven Notchfilter die Methode der kleinsten Quadrate (Least Squares (LS)) [43] angewendet wird und somit in der Regel eine nicht-konvexe Kostenfunktion optimiert werden muss [9]. Es ist daher nicht immer gegeben, dass das globale Minimum der Kostenfunktion gefunden wird. Verschiedene Ansätze versuchen daher, die Nicht-Konvexität zu umgehen, oder eine Lösung zu finden, die dem globalen Minimum so nahe wie möglich kommt [1, 39–41].

Die zweistufigen NHS-Methoden umgehen dieses Problem, indem die Detekti-

on und Verifikation von Instabilitäten separat ausgeführt wird. Die Frequenzen, bei denen eine Instabilität vorliegt, sind also zum Zeitpunkt des Entwurfs der Notchfilter schon bekannt. Unter den Ansätzen der Verstärkungsreduktion ist dieser daher der bekannteste und basiert vorrangig auf der Arbeit von Lewis et al. [44] und Er et al. [45].
Fast alle der bekannten NHS-Methoden beziehen sich auf den einkanaligen Fall. Jedoch wurde die NHS zum Beispiel von Abe [46] auch erfolgreich im mehrkanaligen Fall eingesetzt.
Der größte Vorteil der NHS liegt in ihrer Robustheit. Wird eine Instabilität korrekt erkannt und ausreichend gedämpft, kehrt ein System garantiert wieder in einen stabilen Zustand zurück [9]. Nachteil ist allerdings der Einsatz im elektroakustischen Vorwärtspfad, da dies das Signal verfärbt und ab einem gewissen Grad hörbar wird. Des Weiteren führt die Aktivierung der Notchfilter zu transienten Kompenenten im Lautsprechersignal. Diese können als kurzzeitige, klingelnde Artefakte wahrgenommen werden [45]. Werden die Filter zum Beispiel mit einem leckenden Integrator schrittweise (de-)aktiviert, können diese Artefakte verhindert werden [47]. Ein weiterer Nachteil ist der begrenzte MSG-Gewinn. Selbst wenn der Rückkopplungspfad durch den Einsatz geeigneter Notchfilter perfekt geglättet werden könnte, ist nach Schröders statistischer Analyse [15] ein maximaler Verstärkungsgewinn von 10 dB zu erwarten. In der Praxis würde dies jedoch zu einer zu starken Verfärbung des Signals führen, sodass der reelle Gewinn deutlich darunter liegt [9].

3.5. Räumliche Filterung

Wenn ein System aus mehreren Mikrofonen und/oder mehreren Lautsprechern besteht, können diese verwendet werden, um so genannte Null-Beamformer zu erzeugen. Mit den Mikrofonen kann somit eine „räumliche Null“ in Richtung des Lautsprechers oder alternativ mit einem Lautsprecher-Array eine Null in Richtung des Mikrofons erzeugt werden. Durch beide Ansätze wird im idealen Fall das rückgekoppelte Signal im Mikrofon ausgelöscht. Da ein Lautsprecher-Array im Gegensatz zu einem Mikrofon-Array keine Informationen über die akustischen Gegebenheiten erhalten kann [9], werden im Folgenden ledichglich Ansätze mit Mikrofon-Arrays betrachtet. Der erste Ansatz zur räumlichen Filterung im Bezug auf akustische Rückkopplungsunterdrückung lässt sich auf Duong et al. [48] zurückführen. Ihr Ansatz bezieht sich auf Freihand-Telefonie in einer einkanaligen Situation mit fester Mikrofon- und Lautsprecherposition. Da in den seltensten Fällen die Positionen fest oder im Voraus bekannt sind, sind Ansätze, die die Beamformer-Koeffizienten adaptiv berechnen, deutlich praktikabler. Sollen die Koeffizienten innerhalb des rückgekoppelten Systems berechnet werden, entsteht jedoch ein neues Problem. Durch die Rückkopplung sind Quell- und Lautsprechersignal miteinander korreliert, sodass konventionelle Beamformer-Algorithmen eine „biased“, also eine verzerrte, Lösung erhalten [9]. In der Folge wird ein Teil des Quellsignals herausgefiltert und Anteile des Rückkopplungssignals sind weiterhin im Mikrofonsignal vorhanden. Lösungsvorschläge beinhalten beispielsweise die Kombination mit Adaptive Feed-

back Cancellation (AFC) [49], die Berechnung der Beamformer-Koeffizienten im Voraus außerhalb der Rückkopplungskette [50], oder die Verwendung eines Prewhitening-Filters für das Quellsignal zur Dekorrelation von Mikrofon- und Lautsprechersignalen [51] und erreichen MSG-Gewinne zwischen 7 - 15 dB. Schepker [52–55] schlägt die Kombination eines Null-Beamformers mit zwei AFC-Algorithmen für den Einsatz in Hörgeräten mit mehrern Mikrofonen vor.

3.6. Raummodellierung

Die Methoden, die der Gruppe der Raummodellierung zugeschrieben werden können, versuchen, ein Modell des akustischen Rückkopplungspfades zu identifizieren. Dies kann entweder „offline“ während einer Initialisierungsphase des Systems stattfinden, oder aber „online“ im Echtzeitbetrieb. Je nachdem wie das Modell angewendet wird, können die Methoden in zwei Gruppen eingeteilt werden. [9]

3.6.1. Adaptive Inverse Filtering (AIF)

Bei den Ansätzen des sogenannten Adaptive Inverse Filtering (AIF) wird versucht, den inversen Rückkopplungspfad zu modellieren und zu identifizieren. Dieser wird im geschlossenen Regelkreis eingefügt und führt im Idealfall zu einer perfekten Glättung des Rückkopplungspfades. Nach Schröder [15] sind hier also bis zu 10 dB Verstärkungsgewinn zu erwarten. Es existieren vereinzelte Ansätze, die beispielsweise ein inverses Modell der minimalphasigen Anteile des Rückkopplungspfades im Voraus schätzen [56], basierend auf einer Offline-Messung des Rückkopplungspfades eine große Anzahl von Equalizern im elektroakustischen Vorwärtspfad automatisch anpassen [57], oder auch hybride AIF-AFC Ansätze wie in [49] oder [58]. Im Kontext der akustischen Rückkopplungsunterdrückung wurde das AIF jedoch im Allgemeinen vermutlich auf Grund des begrenzten Verstärkungsgewinns kaum beachtet [9].

3.6.2. Adaptive Feedback Cancellation (AFC)

Bei der Adaptive Feedback Cancellation (AFC) wird hingegen versucht, den tatsächlichen akustischen Rückkopplungspfad zu modellieren und seine Impulsantwort zu identifizieren. Damit wird die Rückkopplungskomponente im Mikrofonsignal geschätzt und vom Mikrofonsignal abgezogen. Das Ergebnis der Subtraktion stellt daher die Schätzung des ungestörten Quellsignals dar. Wird ein passendes Modell gefunden, kann die akustische Kopplung fast vollständig eliminiert werden. Theoretisch sind also sehr große Verstärkungsgewinne möglich [9]. Grundsätzlich ist der Ansatz sehr ähnlich zur akustischen Echounterdrückung (zum Beispiel [59]). Im Gegensatz dazu haben die Methoden der AFC das große Problem, dass Quell- und Lautsprechersignal korreliert sind [60]. Der Einsatz eines standardmäßigen adaptiven Filters würde daher zu einer „biased“, also verzerrten Schätzung der Impulsantwort des Rückkopplungspfades führen [61–63]. Das Quellsignal selbst würde dadurch teilweise untderdrückt und das rückgekoppelte Signal nicht vollständig entfernt werden

können. Die Dekorrelation zwischen dem Quell- und dem Lautsprechersignal ist bei der Anwendung von AFC-Methoden zur akustischen Rückkopplungsunterdrückung also unabdingbar. In der Literatur werden hierzu verschiedene Ansätze vorgeschlagen, die entweder im geschlossenen Signalkreislauf, oder im Pfad des adaptiven Filters zum Einsatz kommen [63]. Zur Anwendung im geschlossenen Signalkreislauf werden etwa Rauschinjektion, zeitvariable oder nichtlineare Signalverarbeitung, Verstärkungsreduktion, oder eine Verzögerung im elektroakusitschen Vorwärtspfad vorgeschlagen. Die Dekorrelation im Pfad des adaptiven Filters kann beispielsweise über das Hinzufügen von Verzögerungen, oder dekorrelierende Vorfilterung erreicht werden. Ein weiterer Ansatz schlägt den Einsatz eines Hilfsmikrofons vor. Im Folgenden wird auf die einzelnen Dekorrelationsmöglichkeiten eingegangen. Kapitel 5 disktutiert weiterhin die genaue Ursache der Korrelation und stellt die für diese Arbeit relevanten Algorithmen der AFC im Detail vor.

Dekorrelation im geschlossenen Signalkreislauf

Rauschinjektion Bei der Rauschinjektion wird dem Lautsprechersignal Rauschen hinzugefügt. Das Rauschen ist dabei so gestaltet, dass es mit dem eigentlichen Signal unkorreliert ist [64–67]. Das Injizieren von Rauschen kann zwar die Rückkopplungsunterdrückung verbessern und tatsächlich zu einer „unbiased“ Lösung führen [68], der Kompromiss zwischen Hörbarkeit und der Leistungsmaximierung des Rauschens zur Verbesserung der Dekorrelation stellt jedoch eine Herausforderung dar [1]. Zusätzlich verringert der Einsatz von Rauschen im adaptiven Filter-Pfad die Konvergenzeigenschaften des adaptiven Filters [66]. Der erste Einsatz von gezieltem Hinzufügen von Rauschen im Kontext von AFC lässt sich auf Ibaraki et al. [69] zurückführen und fügt Rauschen nicht-kontinuierlich in Pausen des Quellsignals ein, um die tieffrequenten Anteile der akustischen Rückkopplungspfads zu identifizieren. Patel und Panahi [70] nutzen die kurzzeitige Injektion von Rauschen, um den akustischen Rückkopplungspfad vollständig zu identifizieren. Liegt eine Änderung des Rückkopplungspfades und damit eine Instabilität vor, wird die Injektion über Instabilitäts-Detektoren erneut angestoßen. Ansätze mit kontinuierlicher Einspeisung des Rauschens [58, 71, 72] erreichen beispielsweise einen MSG-Gewinn von 5 dB in einem stark untermodellierten Szenarion [71]. Akhtar et al. [73] steuern die Rauschleistung so, dass sie während der Konvergenzphase der beiden adaptiven Filter größer und im stationären Betrieb geringer ist. Weiter wurde der Einsatz von A-gewichtetem [71], oder, auf Basis eines psychoakustischen Modells, zeitveränderlichem [74] Noise-Shaping zur Reduzierung der Wahrnehmung des injizierten Rauschens vorgeschlagen. Zur Reduzierung der Korrelation werden in [75, 76] die hochfrequenten Anteile des Eingangssignals durch ein „vocoded“ Signal ersetzt, das perzeptuell ähnlich zum orginalen Signal ist. Außerdem wurde die Verwendung von Langzeit-Prädiktionsfiltern zur Verbesserung der Identifikation [77, 78] und Umformungsfilter zur Verbesserung der Konvergenzeigenschaften im Zusammenhang mit Noise-Shaping [79]

vorgeschlagen.

Nichtlineare oder zeitvariante Signalverarbeitung Zur Dekorrelation von Quell- und Lautsprechersignal kann auch nichtlineare oder zeitvariante Signalverarbeitung im elektroakustischen Vorwärtspfad verwendet werden [1]. Dazu gehören beispielsweise Phasen-, Delaymodulation, oder Frequenzverschiebung [58, 80–83], Halbwellengleichrichtung [58, 84], oder zeitvariante Allpassfilter [85–87]. Theoretische Betrachtungen im Bezug auf stereophonische Echounterdrückung haben gezeigt, dass diese Methoden die Lautsprechersignale perfekt dekorrelieren können, sodass in der Theorie eine Verbesserung der Identifikation des Rückkopplungspfades zu erwarten ist [1].

Zeitverzögerung im elektroakustischen Vorwärtspfad Da die Kreuzkorrelation zwischen Lautsprecher- und Mikrofonsignal besonders für stimmlose Sprachsignale mit größerer Verzögerung abnimmt [9], wurde in [61, 88] für den Einsatz in Hörgeräten das Einfügen einer Verzögerung im elektroakustischen Vorwärtspfad vorgeschlagen.

Die verschiedenen Ansätze zur Dekorrelation im geschlossenen Signalkreislauf sind in ihrer Aufgabe sehr effektiv. Durch den Einsatz in eben diesem geschlossenen Signalkreislauf führen sie jedoch gezwungenermaßen zu ungewünschten Verzerrungen des Quellsignals. Um das zu verhindern, wurden verschiedene Ansätze vorgeschlagen, die sich in den Pfad des adaptiven Filters einfügen lassen und somit das Signal im geschlossnen Kreislauf möglichst unberührt lassen [9].

Dekorrelation und Biasreduktion im Pfad des adaptiven Filters

Beschränkung des adaptiven Filters In [89] wird vorgeschlagen, das adaptive Filter zu beschränken, sodass es nur in einem definierten Bereich um die im Voraus bestimmte „unbiased“ Schätzung abweichen darf. Dies führt zwar zu einer „Bias“-Reduktion und damit zu einer guten Schätzung des akustischen Rückkopplungspfades, wenn dieser nur wenig von der ursprünglichen Schätzung abweicht. Ist dies jedoch nicht gegeben, kann das adaptive Filter im schlimmsten Fall kein passendes Modell für den akustischen Rückkopplungspfad finden [90, 91]. Da akustische Rückkopplungen meist erst in einem Frequenzbereich ab 1,5 kHz auftreten, wird in [92] die Adaption auf diesen Frequenzbereich beschränkt. Generell kann der „Bias“ durch Beschränkung des adaptiven Filters jedoch nicht vollständig aufgehoben werden [1].

Vorverzögerung des adaptiven Filters Ähnlich zum Einfügen einer Zeitverzögerung im elektroakustischen Vorwärtspfad hat das Einfügen einer Vorverzögerung im adaptiven Filter [93, 94] auch einen positiven Einfluss auf die

Verringerung der Dekorrelation. Hier treten jedoch vor allem bei langen Verzögerungen keine störenden, hörbaren Effekte auf. Allerdings sollte die Vorverzögerung nicht länger sein als die Laufzeit des Schalls zwischen Lautsprecher und Mikrofon [9] und ist daher sehr stark von der relativen Positionierung abhängig.

Dekorrelierende Vorfilterung Die grundlegende Idee von dekorrelierenden Vorfiltern ist die „Entfärbung“ („Whitening“) des Quellsignals im Mikrofon- und Lautsprechersignal. Dieser Ansatz wurde als Erstes in Högeräten eingesetzt [62, 95, 96] und später unter anderem für Beschallungssysteme adaptiert [60, 63]. Das Problem, dass hier gleichzeitig sowohl das Modell des akustischen Rückkopplungspfades als auch der optimale Vorfilter geschätzt werden müssen, wurde durch Rombouts et al. [97] mit einem „Prediction-Error-Method (PEM)“-basierten Ansatz [98] gelöst. Hierbei wird das Quellsignal als autoregressiver (AR), linearer Prädiktionsprozess [99] modelliert. Es wird also angenommen, dass das Quellsignal aus einem mittelwertfreien Rauschprozess entsteht, dass mit einem Allpolfilter geformt wird. Das Inverse dieses Modells kann damit als zeitvariantes, entfärbendes FIR-Filter verwendet werden, um Mikrofon- und Lautsprechersignal vor der Filteradaption zu dekorrelieren („Prewhitening“). Ist das Allpolfilter exakt bekannt, kann das weiße Rauschsignal perfekt rekonstruiert und somit eine „unbiased“ Schätzung des akustischen Rückkopplungspfades identifiziert werden [100]. Für stimmlose Sprache ist die Modellierung über einen AR-Prozess eine gute Näherung, da der Vokaltrakt als langsam zeitvariantes Allpolfilter niedriger Ordnung und die glottale Lautbildung für stimmlose Laute als weißer Rauschprozess modelliert werden können [101]. Stimmhafte Sprache und Musik kann das Modell jedoch nur sehr grob nähern. Zum Einen müsste die Ordnung des AR-Prozesses mindestens dem Doppelten der im Signal enthaltenen Sinusterme entsprechen. Da jedoch reele Signale immer zu einem gewissen Teil verrauscht sind, reicht ein AR-Prozess nicht aus, um diese Signalanteile zu modellieren [101]. In der Praxis kann der „Bias“ daher nicht vollständig entfernt, sondern nur der Einfluss auf die Filteradaption verringert werden [1]. Für die gemeinsame Verwendung von AFC mit adaptiven, dekorrelierenden Vorfiltern wurden verschiedene Algorithmen im Zeitbereich [62, 100, 102–108], Frequenzbereich [109–115], Teilbandbereich [110, 116, 117] und anderen Transformationsbereichen [118] vorgeschlagen. Um auch stimmhafte Sprache oder Musik korrekt modellieren zu können, können komplexere Quellmodelle wie ein autoregressiver, gleitender Mittelwert-Prozess (Autoregressive-moving-average (ARMA)) als Nullstellen-Polfilter oder ein frequenzverzerrter, linearer Prädiktionsprozess (Warped linear prediction (WLP)) [101, 119] verwendet werden. Zuletzt existieren auch verschiedene PEM-AFC Ansätze in Kombination mit weiteren Rückkopplungsunterdrückungsmethoden [1] wie Notchfiltern [29, 110], Rauschinjektion [120], Rauschvocodern [75], einkanaliger Rauschunterdrückung [121, 122] oder inversen Verstärkungsfiltern [106].

Auf Grund des theoretisch sehr hohen Verstärkungsgewinns bei gleichzeitig niedriger Verzerrung des Quellsignals [123] hat sich in der Literatur hauptsächlich der PEM-basierte Ansatz in Kombination mit AFC durchgesetzt. Während die meisten Ansätze im einkanaligen Fall betrachtet werden, existieren auch mehrkanalige Ansätze der PEM-AFC. Diese nutzen beispielsweise Null-Beamformer [1] oder Rauschunterdrückung [2, 3] in Kombination mit einkanaliger AFC.

4. Qualitätsmaße

Inhalt

4.1. Einleitung 19
4.2. Added Stable Gain 19
4.3. Misadjustment 20
4.4. Effective Closed-Loop Gain (ECLG) 20
4.5. Short-Time Objective Intelligibility (STOI) 20
4.6. Perceptual Evaluation of Speech Quality (PESQ) 21
4.7. Log-Spectral Distance (SD) 22

4.1. Einleitung

Um die verschiedenen Methoden der akustischen Rückkopplungsunterdrückung vergleichen zu können, werden verschiedene Qualitätsmaße herangezogen. Die gängisten objektiven Maße sind dabei der Added Stable Gain (ASG) und das Misadjustment (Mis) bei adaptiven Filtern. Ein gutes Maß um Instabilität zu messen ist der Effective Closed-Loop Gain (ECLG). Wegen ihrer hohen Korrelation mit subjektiven Qualitätsmaßen [124] werden zusätzlich Short-Time Objective Intelligibility (STOI), Perceptual Evaluation of Speech Quality (PESQ) und Log-Spectral Distance (SD) betrachtet.

4.2. Added Stable Gain

Der Added Stable Gain (ASG) ist die durch eine Methode der akustischen Rückkopplungsunterdrückung hinzugewonnene maximale Verstärkung, bei der das System gerade noch nicht instabil ist. Rechnerisch ergibt sich der ASG aus der Subtraktion des MSG_{AFC} mit Rückkopplungsunterdrückung und dem MSG ohne Rückkopplungsunterdrückung zu

$$\mathrm{ASG}(t) = \mathrm{MSG}_{AFC}(t) - \mathrm{MSG}(t). \tag{4.1}$$

Wird AFC, also adaptive Filter zur Rückkopplungsunterdrückung verwendet, ergbit sich der MSG_{AFC} mit dem geschätzten Rückkopplungspfad $\hat{F}(\omega, t)$ zu

$$\mathrm{MSG}_{AFC}(t) = 10 \log_{10} \left[\max_{\omega \epsilon \mathcal{P}} \frac{1}{\left| F(\omega, t) - \hat{F}(\omega, t) \right|^2} \right] \tag{4.2}$$

[2] und (4.1) zusammen mit (2.27) [125] zu

$$\mathrm{ASG}(t) = 10\log_{10}\left[\max_{\omega\in\mathcal{P}}\frac{1}{\left|F(\omega,t)-\hat{F}(\omega,t)\right|^2}\right] - 10\log_{10}\left[\max_{\omega\in\mathcal{P}}\frac{1}{|F(\omega,t)|^2}\right]. \quad (4.3)$$

4.3. Misadjustment

Das Misadjustment (Mis) gibt die normalisierte Distanz in Dezibel zwischen dem echten und dem geschätzten Rückkopplungspfad mit den entsprechenden Übertragungskoeffizienten $\mathbf{f}(t)$ und $\hat{\mathbf{f}}(t)$ über

$$\mathrm{Mis}(t) = 20\log_{10}\left|\frac{\mathbf{f}(t)-\hat{\mathbf{f}}(t)}{\mathbf{f}(t)}\right| \quad \mathrm{dB} \quad (4.4)$$

an und ist daher ein gutes Maß, um die Schätzung der Impulsantwort des Rückkopplungspfades zu bewerten [124]. Da dieses Maß jedoch keine Information über die Qualtität der akustischen Rückkopplungsunterdrückung im Bezug auf Verstärkungsgewinn oder Klangqualität liefert [123], wird das Misadjustment in der Regel in Verbindung mit weiteren Maßen angegeben. Je negativer der Wert des Mis ist, desto besser ist die Schätzung des akustischen Rückkopplungspfades. 0 dB bedeuten dabei, dass keine Anpassung stattfindet, also $\hat{F}(\omega,t) = 0 \ \forall\omega,t$.

4.4. Effective Closed-Loop Gain (ECLG)

Der Effective Closed-Loop Gain (ECLG) ist ein Maß für den vorhandenen Verstärkungsspielraum im geschlossenen Signalkreislauf, indem das Maß zusätzlich den Vorwärtspfad einbezieht [1, 117]. Zusammen mit dem Vorwärtspfad $G(\omega,t)$, dem tatsächlichen Rückkopplungspfad $F(\omega,t)$ und dem geschätzten Rückkopplungspfad $\hat{F}(\omega,t)$ ergibt sich der ECLG zu

$$ECLG(t) = \max_{\omega\in\mathcal{P}}|G(\omega,t)(F(\omega,t)-\hat{F}(\omega,t))|. \quad (4.5)$$

Das System ist stabil, solange $ECLG(t) < 1$. Alternativ kann der ECLG mit

$$ECLG_{\mathrm{dB}}(t) = 20\log_{10}\left[\max_{\omega\in\mathcal{P}}|G(\omega,t)(F(\omega,t)-\hat{F}(\omega,t))|\right]\mathrm{dB} \quad (4.6)$$

auch in Dezibel angegeben werden, wobei hier $ECLG_{\mathrm{dB}}(t) < 0$ Stabilität bedeutet.

4.5. Short-Time Objective Intelligibility (STOI)

Das STOI Qualitätsmaß beschreibt die Verständlichkeit von Sprache über die Korrelation von zeitlichen Hüllkurven des ungestörten Quellsignals und des degradierten Signals in kurzen Zeitabschnitten von 382 Millisekunden [126].

Abbildung 4.1 zeigt den groben Ablauf zur Bestimmung des STOI. Beide Signale werden im ersten Schritt in Terzbänder und anschließend in Zeitfenster der Länge 382 ms aufgeteilt. Vor der Berechnung der Korrelationskoeffizienten zwischen den zeitlichen Hüllkurven der beiden Signale wird das degradierte Signal normalisiert und der Wertebereich begrenzt, um globale Pegelunterschiede zu kompensieren. Das Qualitätsmaß ergibt sich im letzten Schritt aus dem Durchschnitt der Koeffizienten in allen Frequenzbändern und Zeitfenstern. Das Maß gibt Werte zwischen 0 und 1 beziehungsweise 0 und 100 Prozent aus. Bei einem Wert von 0 ist das degradierte Signal völlig unverständlich. Perfekte Sprachverständlichkeit wird durch den Wert 1 signalisiert. Das Qualitätsmaß zeigt sowohl für Algorithmen der Rauschunterdrückung [126] als auch der akustischen Rückkopplungsunterdrückung [124] eine sehr hohe Korrelation zu subjektiven Qualitätsmaßen, die durch Hörtests gewonnen werden.

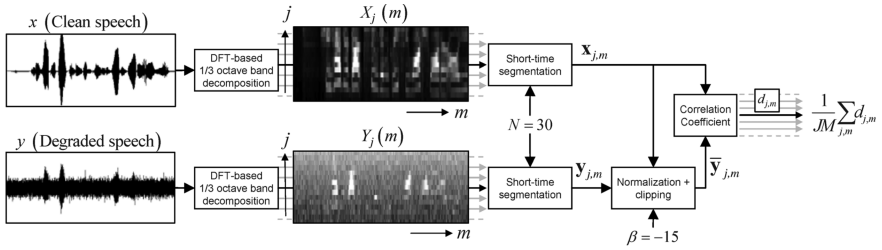


Abbildung 4.1.: *Blockschaltbild zur Berechnung des STOI Qualitätsmaß [126]*

4.6. Perceptual Evaluation of Speech Quality (PESQ)

PESQ versucht die wahrgenommene Qualität eines Signals vorherzusagen, die Menschen in einem subjektiven Hörtest angeben würden. Hierfür vergleicht der Algorithmus das ungestörte Signal mit dem degradierten Signal über das Wahrnehmungsmodell in Abbildung 4.2. Beide Signale werden so repräsentiert, dass sie der psychophysikalischen Repräsentiation von Audiosignalen im menschlichen Gehör entsprechen. Im kognitiven Modell werden anschließend zwei Fehlerparameter berechnet, deren Kombination das objektive Qualitätsmaß PESQ ergibt [127, 128]. Der Wertebereich erstreckt sich von -0.5 bis 4.5. Je höher der Wert, desto besser ist die Signalqualität. Bei einem Wert von 2.5 bis 3 sind schon deutliche Artefakte im Signal zu hören, die Sprachverständlichkeit ist jedoch noch gegeben.

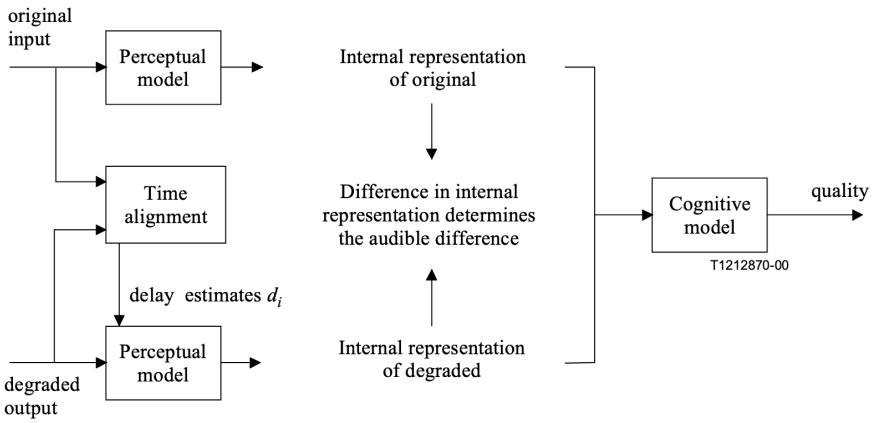


Abbildung 4.2.: *Struktur zur Berechnung des PESQ Qualitätsmaß [127]*

4.7. Log-Spectral Distance (SD)

Um die Verzerrung des Lautsprechersignals im Vergleich zum Quellsignal zu quantifizieren, kann die frequenzgewichtete, logarithmisch-spektrale Signalverzerrung SD(t) (Frequency-Weighted Log-Spectral Signal Distortion) [129, 130] herangezogen werden. Mit den Kurzzeit-Leistungsdichtespektren $P_u(f,t)$ und $P_v(f,t)$ des Lautsprecher- und Quellsignals ergibt sich die SD(t) zu

$$\mathrm{SD}(t) = \sqrt{\int_{f_l}^{f_h} w(f) \left(10 \log_{10} \frac{P_u(f,t)}{P_v(f,t)}\right)^2 df} \quad \mathrm{dB}. \tag{4.7}$$

Mit der Gewichtungsfunktion $w(f)$ wird jedes kritische Band zwischen f_l = 300 Hz und f_h = 6400 Hz gleich gewichtet. Die Frequenzbänder können über die äquivalente, rechteckige Bandbreite (Equivalent Rectangular Bandwidth (ERB)) für auditorische Filter bestimmt werden [131]. Da die Bandbreiten für steigende Frequenzen größer werden, erhalten höhere Frequenzbänder als Ausgleich ein kleineres Gewicht. Die Gewichtsfunktion wird so normalisiert, dass

$$\int_{f_l}^{f_h} w(f) df = 1. \tag{4.8}$$

Bei einem Wert von $SD(t) < 1$ dB ist die Verzerrung für das menschliche Gehör nahezu unbemerkbar. Ab einem $SD(t) > 3$ dB können Verzerrungen des Signals deutlich wahrgenommen werden.

5. Adaptive Rückkopplungsunterdrückung

Inhalt

5.1. Einleitung 23
5.2. Optimale Lösung und Bias 25
5.3. Dekorrelation zum Erhalt einer unbiased Lösung 26
5.4. PEM-Based Frequency-Domain Kalman Filter (PEM-FDKF) . 29

5.1. Einleitung

Dieses Kapitel beleuchtet die Methodik der adaptiven Rückkopplungsunterdrückung (AFC) im Detail. Besonders liegt der Fokus auf der Verwendung von Kalman-Filtern zur adaptiven Schätzung des Modells der akustischen Rückkopplungsstrecke. Die grundlegende Funktion der AFC lässt sich anhand des einkanaligen Systemmodells ohne AFC aus Kapitel 2.2, wie hier nochmal in Abbildung 5.1 zu sehen, anschaulich zeigen.
Das Mikrofonsignal $y(t)$ zum diskreten Zeitpunkt t setzt sich aus dem Quellsignal $v(t)$ und dem rückgeppelten Signal $x(t)$ zusammen, sodass

$$y(t) = v(t) + x(t). \tag{5.1}$$

Das Lautsprechersignal bildet sich aus dem vom elektroakustischen Vorwärtspfad $G(z,t)$ verarbeiteten Mikrofonsignal über

$$u(t) = G(z,t)y(t). \tag{5.2}$$

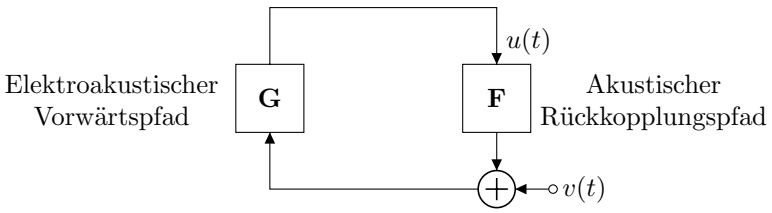


Abbildung 5.1.: *Einkanalige akustische Rückkopplungsstrecke mit dem Quellsignal $v(t)$ und dem Lautsprechersignal $u(t)$*

Über den akustischen Rückkopplungspfad $F(z,t)$ der Ordnung n_F wird das Lautsprechersignal über

$$x(t) = F(z,t)u(t) \tag{5.3}$$

in das Mikrofonsignal rückgekoppelt. Für die Übertragungsfunktion ergibt sich somit

$$\frac{U(\omega,t)}{V(\omega,t)} = \frac{G(\omega,t)}{1 - G(\omega,t)F(\omega,t)} \tag{5.4}$$

mit den entsprechenden Stabilitätskriterien nach Nyquist [10] aus Gleichungen (2.19) und (2.20).
Die grundlegende Idee der AFC ist die Schätzung des Rückkopplungspfades, sodass über Faltung des Lautsprechersignals mit dem geschätzten Rückkopplungspfad das geschätzte Rückkopplungssignal

$$\hat{x}(t) = \hat{F}(z,t)u(t) \tag{5.5}$$

gewonnen werden kann. Dieses wird vom Mikrofonsignal abgezogen, sodass im idealen Fall das Fehlersignal

$$d(t) = y(t) - \hat{x}(t) \tag{5.6}$$

dem reinen Quellsignal entspricht. Das Lautsprechersignal $u(t)$ wird anschließend über Signalverarbeitung des Fehlersignals im elektroakustischen Vorwärtspfad $G(z,t)$ über

$$u(t) = G(z,t)d(t) \tag{5.7}$$

gewonnen. Das Modell der AFC ist in Abbildung 5.2 zu finden. Die Übertra-

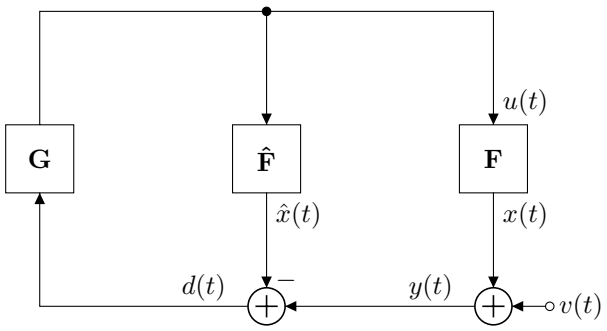


Abbildung 5.2.: *Einkanalige akustische Rückkopplungsstrecke mit adaptiver Rückkopplungsunterdrückung*

gungsfunktion ergibt sich hier zu

$$\frac{U(\omega,t)}{V(\omega,t)} = \frac{G(\omega,t)}{1 - G(\omega,t)(F(\omega,t) - \hat{F}(\omega,t))}. \tag{5.8}$$

Ist das Filter $\hat{F}(z,t)$ in der Lage, den tatsächlichen akustischen Rückkopplungspfad perfekt zu schätzen, sodass $\hat{F}(z,t) = F(z,t)$, ergibt sich das Lautsprechersignal zu

$$u(t) = G(z,t)v(t) \tag{5.9}$$

und das System ist für alle möglichen Verstärkungen im elektroakustischen Vorwärtspfad stabil.

5.2. Optimale Lösung und Bias

Der optimale Filterkoeffizientenvektor $\hat{\mathbf{f}}_{opt}(t)$ des adaptiven Filters $\hat{F}(z,t)$ kann durch Minimierung des quadratischen Fehlersignals und somit der Kostenfunktion

$$
\begin{aligned}
J(\hat{\mathbf{f}}(t)) &= E\left\{d^2(t)\right\} && (5.10)\\
&= E\left\{\left[y(t)-\hat{x}(t)\right]^2\right\} && (5.11)\\
&= E\left\{\left[y(t)-\hat{\mathbf{f}}(t)\mathbf{u}(t)\right]^2\right\} && (5.12)\\
&= E\left\{\hat{\mathbf{f}}^T(t)\mathbf{u}(t)\mathbf{u}^T(t)\hat{\mathbf{f}}(t)\right\} - E\left\{2\hat{\mathbf{f}}(t)\mathbf{u}(t)y(t)\right\} + E\left\{y^2(t)\right\} && (5.13)\\
&= \hat{\mathbf{f}}^T(t)\mathbf{R}_{uu}(t)\hat{\mathbf{f}}(t) - 2\hat{\mathbf{f}}^T(t)\mathbf{r}_{uy}(t) + y^2(t) && (5.14)
\end{aligned}
$$

gewonnen werden [132]. $E\{\cdot\}$ stellt dabei den mathematischen Erwartungswert dar, $\mathbf{u}(t) = \begin{bmatrix} u(t) & u(t-1) & \cdots & u(t-n_F)\end{bmatrix}^T$ ist der Lautsprechersignalvektor mit den bis n_F verzögerten Werten,

$$\mathbf{R}_{uu}(t) = E\left\{\mathbf{u}(t)\mathbf{u}^T(t)\right\} \tag{5.15}$$

ist die Autokorrelationsmatrix des Lautsprechersignals und

$$\mathbf{r}_{uy}(t) = E\left\{\mathbf{u}(t)y(t)\right\} \tag{5.16}$$

ist der Kreukorrelationsvektor zwischen Lautsprecher- und Mikrofonsignal. Über die Wiener-Hopf Gleichung lässt sich $\hat{\mathbf{f}}_{opt}(t)$ als geschlossene Lösung über

$$\hat{\mathbf{f}}_{opt}(t) = {\mathbf{R}_{uu}}^{-1}(t)\mathbf{r}_{uy}(t) \tag{5.17}$$

formulieren [132]. Unter der Annahme, dass der akustische Rückkopplungspfad mit einer ausreichenden Ordnung geschätzt wird, sodass $n_{\hat{F}} = n_F$, kann der Kreuzkorrelationsvektor von Lautsprecher- und Mikrofonsignal über das Quellsignal $v(t)$ und den rückgekoppelten Signalanteil $x(t)$ ausgedrückt werden, sodass sich (5.17) zu

$$
\begin{aligned}
\hat{\mathbf{f}}_{opt}(t) &= \mathbf{R}_{uu}^{-1}(t)E\left\{\mathbf{u}(t)\left[\mathbf{f}^T(t)\mathbf{u}(t)+v(t)\right]^T\right\} && (5.18)\\
&= \mathbf{R}_{uu}^{-1}(t)E\left\{\mathbf{u}(t)\mathbf{u}^T(t)\mathbf{f}(t)+\mathbf{u}(t)v(t)\right\} && (5.19)\\
&= \mathbf{R}_{uu}^{-1}(t)\left[\mathbf{R}_{uu}(t)\mathbf{f}(t)+\mathbf{r}_{uv}(t)\right] && (5.20)\\
&= \mathbf{R}_{uu}^{-1}(t)\mathbf{R}_{uu}(t)\mathbf{f}(t)+\mathbf{R}_{uu}^{-1}(t)\mathbf{r}_{uv}(t) && (5.21)\\
&= \mathbf{f}(t)+\mathbf{R}_{uu}^{-1}(t)\mathbf{r}_{uv}(t) && (5.22)
\end{aligned}
$$

ergibt. Aus (5.22) lässt sich erkennen, dass $\hat{\mathbf{f}}_{opt}(t) = \mathbf{f}(t)$ nur dann erfüllt sein und damit eine „unbiased“ Lösung vorliegen kann, wenn Lautsprecher- und Quellsignal unkorreliert sind, also $\mathbf{r}_{uv}(t) = \mathbf{0}$.

Unter der weiteren Annahme, dass der elektroakustische Vorwärtspfad lediglich aus einer breitbandigen Verstärkung $|G|$ und einer Verzögerung d_G besteht, sodass $G(z,t) = |G|z^{-d_G}$, lässt sich $\mathbf{r}_{uv}(t)$ unter Zuhilfenahme von (5.8) und Zusammenfassung dessen Nenners zu $C(z,t)$ durch

$$\mathbf{r}_{uv}(t) = C(z,t)|G|z^{-d_G} E\left\{\begin{bmatrix} v(t-d_G) \\ v(t-d_G-1) \\ \dots \\ v(t-d_G-n_{\hat{F}}+1) \end{bmatrix} v(t)\right\} \tag{5.23}$$

ausdrücken [1]. Eine Möglichkeit der Dekorrelation stellt also das Hinzufügen einer ausreichend großen Verzögerung d_G im elektroakustischen Vorwärtspfad dar, sodass diese länger ist als die Autokorrelation des Quellsignals. Bei weißem Rauschen als Quellsignal würde also beispielsweise eine Verzögerung von $d_G = 1$ ausreichen [1]. Für Sprach- und Audiosignale ist die Autokorrelation jedoch deutlich länger. Die Autokorrelationszeit von Sprache liegt beispielsweise bei um die 20 ms und für anhaltende Töne in Musik bei bis zu mehreren Sekunden [1]. Da aber Verzögerungen im elektroakustischen Vorwärtspfad schon ab 10 ms [133] wahrnehmbar sind, ist eine beliebig große Verzögerung zur Dekorrelation im Allgemeinen nicht möglich.

5.3. Dekorrelation zum Erhalt einer unbiased Lösung

In Kapitel 3.6.2 wurden verschiedene Methoden zur Dekorrelation der Signale und damit der „Bias“-Reduktion im Zusammenhang mit der AFC vorgestellt. Dazu gehören Rauschinjektion, nichtlineare oder zeitvariante Signalverarbeitung, das Hinzufügen einer Verzögerung im elektroakustischen Vorwärtspfad oder im adaptiven Filter, die Beschränkung des adaptiven Filters, oder die Prediction-Error-Method (PEM) mit entfärbenden Vorfiltern. Verfahren im elektroakustischen Vorwärtspfad haben den Nachteil, dass sie das Lautsprechersignal verändern und damit die Qualität des Sprachsignals verschlechtern. Es werden daher Methoden bevorzugt, die im Pfad des adaptiven Filters angewendet werden. In der Literatur hat sich der PEM-basierte Ansatz durchgesetzt, da hier ein guter Verstärkungsgewinn möglich ist ohne das Lautsprechersignal zu degradieren [123]. Im Folgenden wird daher der PEM-basierte Ansatz im Detail und anschließend ein Verfahren vorgestellt, das zur adaptiven Schätzung des autoregressiven Modells des Sprachsignals verwendet werden kann.

Das Ziel der PEM ist die Reduzierung des Bias' im adaptiven Filter, indem das Quellsignal $v(t)$ und das Lautsprechersignal $u(t)$ über entfärbende Vorfilter dekorreliert werden. Sind die Korrelationseigenschaften des Quellsignals bekannt, können statische Vorfilter verwendet werden. In der Regel sind diese jedoch nicht bekannt, sodass meist adaptive Vorfilter genutzt werden, die versuchen, die Korrelationseigenschaften des Quellsignals zu schätzen. Hierzu wird angenommen, dass das Quellsignal über ein autoregressives Modell [99] modelliert werden kann. Für Sprachsignale ist dieses Modell ausreichend, für tonalere Signale und Musik sind komplexere Modelle wie wie ein autoregressiver, gleitender Mittelwert-Prozess (Autoregressive-moving-average (ARMA))

als Nullstellen-Polfilter oder ein frequenzverzerrter, linearer Prädiktionsprozess (Warped linear prediction (WLP)) [101, 119] nötig. Abbildung 5.3 zeigt das Blockschaltbild einer einkanaligen Rückkopplungsstrecke mit dem PEM-basierten Ansatz zur adaptiven Rückkopplungsunterdrückung. Das Quellsignal wird so modelliert, dass

$$v(t) = H(z,t)e(t) = \frac{1}{A(z,t)}e(t). \tag{5.24}$$

$H(z,t)$ wird dabei als monisches, invers-stabiles, zeitvariantes Modell des Quellsignals $v(t)$ angenommen, das durch das weiße Rauschsignal $e(t)$ angeregt wird. Durch

$$H(z,t) = \frac{1}{A(z,t)} \tag{5.25}$$

und der Annahme, dass $H(z,t)$ und damit auch $A(z,t)$ inversstabil sind, kann das Modell als autoregressiver Prozess der Ordnung n_A verstanden werden. Der Koeffizientenvektor des autoregressiven Modells kann über

$$\mathbf{a}(t) = \begin{bmatrix} 1 & a_1(t) & \cdots & a_{n_A}(t) \end{bmatrix}^T \tag{5.26}$$

definiert werden. Die Schätzung $\hat{A}(z,t)$ des Modells kann dann als entfärbendes Vorfilter verwendet werden. Anstelle der Minimierung der Kostenfunktion $J(\hat{\mathbf{f}}(t))$ in (5.10) wird bei der PEM das entfärbte Fehlersignal

$$e_a(t) = \hat{A}(z,t)e(t) \tag{5.27}$$

über die Kostenfunktion

$$\begin{aligned} J(\mathbf{a}(t), \hat{\mathbf{f}}(t)) &= E\left\{e_a^2(t)\right\} && (5.28)\\ &= E\left\{\left[\hat{A}(z,t)e(t)\right]^2\right\} && (5.29)\\ &= E\left\{\left[\hat{A}(z,t)\left(y(t) - \hat{\mathbf{f}}^T(t)\mathbf{u}(t)\right)\right]^2\right\} && (5.30) \end{aligned}$$

minimiert. In der Praxis werden hierzu das Mikrofonsignal $y(t)$ und das Lautsprechersignal $u(t)$ zu

$$y_a(t) = \hat{A}(z,t)y(t) \tag{5.31}$$

und

$$u_a(t) = \hat{A}(z,t)u(t) \tag{5.32}$$

vorgefiltert. Durch Filterung des vorgefilterten Lautsprechersignals $u_a(t)$ mit dem geschätzten akustischen Rückkopplungspfad $\hat{F}(z,t)$ wird das vorgefilterte, geschätzte Rückkolpplungssignal

$$\hat{x}_a(t) = \hat{A}(z,t)\hat{x}(t) \tag{5.33}$$

gewonnen. Aus der Subtraktion des vorgefilterten Mikrofonsignals $y_a(t)$ und dem vorgefilterten, geschätzten Rückkopplungssignal $\hat{x}_a(t)$ erhält man das Fehlersignal

$$\varepsilon(t) = y_a(t) - \hat{x}_a(t), \tag{5.34}$$

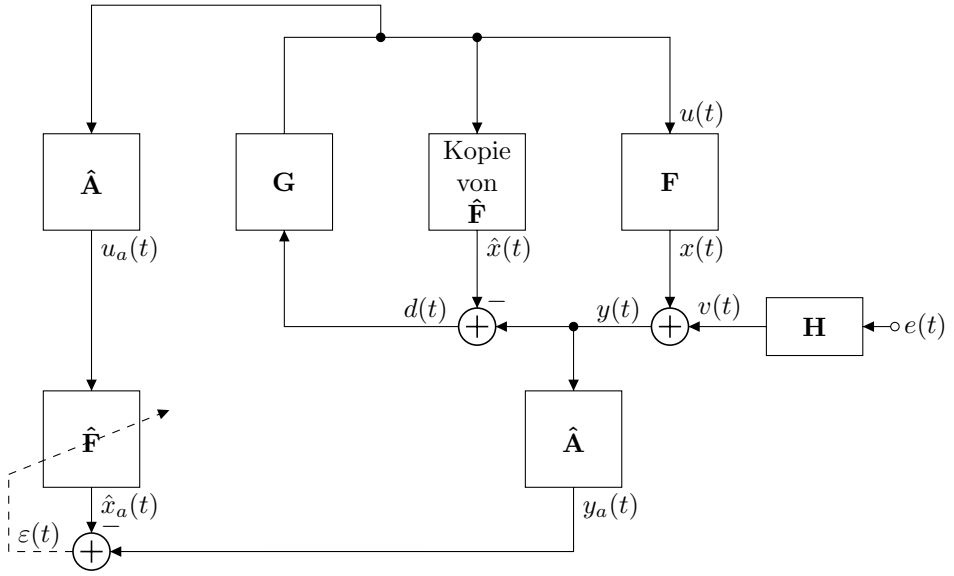


Abbildung 5.3.: *Einkanalige akustische Rückkopplungsstrecke mit dem Prediction-Error-Method-basierten Ansatz für adaptive Rückkopplungsunterdrückung (vgl. [114])*

das zur Adaption des adaptiven Filters verwendet wird. Die hieraus gewonnenen Filterkoeffizienten $\hat{\mathbf{f}}(t)$ werden anschließend verwendet, um das eigentliche, geschätzte Rückkopplungssignal $\hat{x}(t)$ über (5.5) zu gewinnen. Der PEM-basierte Ansatz ist also in der Lage, eine „unbiased“ Lösung zu generieren, falls die Korrelationseigenschaften des Quellsignals bekannt sind oder korrekt geschätzt werden können. Die größte Herausforderung ist hierbei die simultane Schätzung des adaptiven Filters und des autoregressiven Modells des Quellsignals [9].
Da angenommen wird, dass das autoregressive Modell inversstabil ist, wird in der Regel versucht, das Modell über ein Modell des Fehlersignals $d(t)$ zu schätzen. Es können daher Methoden zur Schätzung von inversstabilen Filtern herangezogen werden [100]. Hier wird die blockbasierte, rekursive Levinson-Durbin Methode [134] verwendet wie sie in Algorithmus 1 zu finden ist.

$$\mathbf{J}_{n_A} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{5.35}$$

stellt dabei die $n_A \times n_A$ spiegelnde Einheitsmatrix dar, $E_n(t)$ ist die Leistung des Prädiktionsfehlers, $\kappa_n(t)$ sind die Reflexionskoeffizienten und $r_{dd,n}(t)$ ist das n-te Element des Autokorrealtionsvektors

$$\mathbf{r}_{dd}(t) = E\{\mathbf{d}(t-1)d(t)\}, \tag{5.36}$$

der zum Zeitpunkt $(t-1)$ n_A vergangenen Werte und dem aktuellen Wert des Fehlersignals $d(t)$ mit $\mathbf{d}(t-1) = \begin{bmatrix} d(t-1) & d(t-2) & \cdots & d(t-n_A) \end{bmatrix}^T$.

Algorithmus 1: Levinson-Durbin Algorithmus zur Schätzung der entfärbenden Vorfiltern [1, 134]

Eingabe: Prädiktionsordnung n_A
Ausgabe: Koeffizientenvektor des autoregressiven Modells $\mathbf{a}(t)$

Initialisiere $n_A \times n_A$-dimensionale Gegeneinhatsmatrix $\mathbf{J}_{n_A}$
Initialisiere $E_0(t) \leftarrow r_{dd,0}(t)$
for $n \leftarrow 1$ **to** n_A **do**
    Berechne Reflexionskoeffizienten
    $\kappa_n(t) \leftarrow \frac{1}{E_{n-1}(t)} \left[r_{dd,n}(t) - \mathbf{a}_{n-1}^T(t) \mathbf{J}_{n-1} \mathbf{r}_{dd,n-1}(t) \right]$
    Aktualisiere Prädiktionsfilterkoeffizienten
    $\mathbf{a}_n(t) = \begin{pmatrix} \mathbf{a}_{n-1}(t) \\ 0 \end{pmatrix} - \kappa_n(t) \mathbf{J}_n \begin{pmatrix} -1 \\ \mathbf{a}_{n-1}(t) \end{pmatrix}$
    $E_n(t) = E_{n-1}(t)(1 - \kappa_n^2(t))$
end

5.4. PEM-Based Frequency-Domain Kalman Filter (PEM-FDKF)

Wie in Kapitel 3.6.2 zusammengefasst, wurden für die Schätzung der akustischen Rückkopplungsstrecke mit adaptiven Filtern verschiedene Algorithmen in Zeit-, Teilband-, Frequenz-, oder anderen Transformationsbereichen vorgeschlagen. Durchgesetzt hat sich zum Einen die Verarbeitung im Frequenzbereich. Hier können die Dekorrelationseigenschaften der diskreten Fouriertransformation (DFT) ausgenutzt werden, um der „unbiased“ Lösung näher zu kommen [107]. Im Vergleich zur Verarbeitung im Zeitbereich macht die Berechnung im Frequenzbereich einen deutlich reduzierten Rechenaufwand möglich [115]. Bei der Wahl der Optimierungsmethoden haben sich vor allem Normalized Least Mean Squares (NLMS) [1, 125] und Kalman Filter [2, 3, 114, 115] durchgesetzt. Während bei den „klassischen“ Methoden wie NLMS zusätzlich zur Schätzung des Rückkopplungspfades und dem Sprachmodell auch die optimale Schrittweite des Algorithmus' gefunden werden muss, erlaubt die Formulierung über Zustandsvariablen bei Kalman Filtern die automatische Schätzung der optimalen Schrittweite als Teil der Berechnungsvorschriften [59, 135]. Dies führt zusätzlich zu verbesserten Konvergenzeigenschaften von Kalman Filtern [114].
Aufgrund dieser Vorteile wird im Folgenden der PEM-basierte Kalman Filter im Frequenzbereich (PEM-FDKF) [114] vorgestellt und im Rahmen dieser Arbeit für den mehrkanaligen Anwendungsfall erweitert. Zur Problemformulierung im Frequenzbereich werden die Signale anstatt über die diskrete Zeit t über den diskreten Zeitfenster-Index k dargestellt. Mit der Länge des Fourier-

transformation N und dem Verschiebungsintervall R lässt sich das N-Werte blockbasierte Lautsprechersignal zum Zeitpunkt k über

$$\mathbf{u}(k) = \begin{bmatrix} u(kR-N+1) & \dots & u(kR-1) & u(kR) \end{bmatrix}^T \tag{5.37}$$

darstellen. Mit der Schätzung des autoregressiven Sprachmodells des letzten Zeitfensters $\hat{A}(z,t-1)$ kann das vorgefilterte Lautsprechersignalfenster mit

$$\mathbf{u}_a(k) = \begin{bmatrix} \hat{A}(z,t-1)u(kR) \\ \vdots \\ \hat{A}(z,t-1)u(kR-N+1) \end{bmatrix} \tag{5.38}$$

$$= \begin{bmatrix} u_a(kR-N+1) & \dots & u_a(kR-1) & u_a(kR) \end{bmatrix}^T \tag{5.39}$$

definiert werden. Über die $N \times N$ Fourier-Transformationsmatrix [136]

$$\mathbf{F}_N = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_M & W_M^2 & \cdots & W_M^{M-1} \\ 1 & W_M^2 & W_M^4 & \cdots & W_M^{2(M-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_M^{M-1} & W_M^{2(M-1)} & \cdots & W_M^{(M-1)(M-1)} \end{bmatrix} \tag{5.40}$$

mit

$$W_M = e^{-j2\pi/M} \tag{5.41}$$

und dem Diagonalisierungsoperator $\mathrm{diag}\{\cdot\}$ kann das diagonalisierte Lautsprechersignalfenster im Frequenzbereich

$$\mathbf{U}_a(k) = \mathrm{diag}\{\mathbf{F}_N \mathbf{u}_a(k)\} \tag{5.42}$$

gewonnen werden. Auch der akustische Rükkopplungspfad $\mathbf{f}(t)$ benötigt zur Darstellung im Frequenzbereich eine blockbasierte Darstellung

$$\mathbf{f}(k) = \begin{bmatrix} f(0,k) & f(1,k) & \cdots & f(N-R-1,k) \end{bmatrix}^T, \tag{5.43}$$

sodass

$$\mathbf{F}(k) = \mathbf{F}_N \begin{pmatrix} \mathbf{f}(k) \\ \mathbf{0} \end{pmatrix}. \tag{5.44}$$

Bei der Wahl der Fenstergröße N und dem Verschiebungsintervall R sollte die Länge des tatsächlichen, akustischen Rückkopplungspfades beachtet werden. Gute Werte sind beispielsweise mit $R = n_F$ und $N = 2R$ gegeben [112, 114, 137]. Das Mikrofonsignalfenster $\mathbf{y}(k)$ der Länge R zum Zeitfenster k ist über

$$\mathbf{y}(k) = \begin{bmatrix} y(kR-R+1) & \dots & y(kR-1) & y(kR) \end{bmatrix}^T \tag{5.45}$$

und das vorgefilterte Mikrofonsignalfenster $\mathbf{y}_a(k)$ ähnlich zu (5.38) über

$$\mathbf{y}_a(k) = \begin{bmatrix} \hat{A}(z,t-1)y(kR) \\ \vdots \\ \hat{A}(z,t-1)y(kR-R+1) \end{bmatrix} \tag{5.46}$$

$$= \begin{bmatrix} y_a(kR-R+1) & \dots & y_a(kR-1) & y_a(kR) \end{bmatrix}^T \tag{5.47}$$

definiert. Auch das weiße Rauschsignal $e(t)$, dass in der Modellvorstellung das Sprachsignal über den autoregressiven Prozess erzeugt, wird mit

$$\mathbf{e}(k) = \begin{bmatrix} e(kR-R+1) & \dots & e(kR-1) & e(kR) \end{bmatrix}^T \tag{5.48}$$

in eine blockbasierte Darstellung überführt. Mit der $R \times R$ Einheitsmatrix $\mathbf{I}_R$ und der $N \times R$ Projektionsmatrix

$$\mathbf{Q} = \begin{pmatrix} \mathbf{0}_{N-R\times R} \\ \mathbf{I}_R \end{pmatrix} \tag{5.49}$$

können das vorgefilterte Mikrofon- und Quellsignalfenster über

$$\mathbf{Y}_a(k) = \mathbf{F}_N \mathbf{Q} \mathbf{y}_a(k) \tag{5.50}$$

$$\mathbf{E}(k) = \mathbf{F}_N \mathbf{Q} \mathbf{e}(k) \tag{5.51}$$

dargestellt werden. Über $\mathbf{Q}$ und ihrer Transformierten $\mathbf{Q}^T$ ist lineare Faltung im Frequenzbereich über die Overlap-Save Methode [136] möglich.
Anstelle der Minimierung der Kostenfunktion in (5.10) erhält man bei der Verwendung von Kalman Filtern die lineare, minimale, mittlere quadratische Abweichung (Minimum Mean Square Error (MMSE)) durch Minimierung der Kostenfunktion [138]

$$J(\hat{\mathbf{F}}(k)) = E\left\{ \left\| \mathbf{F}(k) - \hat{\mathbf{F}}(k) \right\|^2 \right\}. \tag{5.52}$$

$\|\cdot\|$ entspricht dabei der euklidischen Norm. Da (5.52) von den unbekannten Größen $v(t)$ und $\mathbf{F}(k)$ abhängig ist, muss das Problem aus einer stochastischen Sicht betrachtet werden. Für das Quellsignal $v(t)$ ist über den autoregressiven Prozess schon ein geeignetes Modell eingeführt. Da das Änderungsverhalten des akustischen Rückkopplungspfades als langsam angenommen werden kann [59], kann die zeitliche Änderung der Rückkopplungsstrecke als Gauß-Markov Prozess erster Ordnung [138] modelliert werden. Zusammen mit einem linearen Modell des vorgefilterten Mikrofonsignals $\mathbf{Y}_a(k)$ kann das System über das Zustandsmodell mit

$$\mathbf{F}(k) = A \cdot \mathbf{F}(k-1) + \mathbf{N}(k) \tag{5.53}$$

$$\mathbf{Y}_a(k) = \mathbf{C}_a(k)\mathbf{F}(k) + \mathbf{E}(k) \tag{5.54}$$

beschrieben werden. $\mathbf{C}_a(k) = \mathbf{F}_N \mathbf{Q}\mathbf{Q}^T {\mathbf{F}_N}^{-1} \mathbf{U}_a(k)$ stellt dabei die Messmatrix dar und erlaubt die korrekte lineare Faltung zwischen $\mathbf{u}_a(k)$ und $\mathbf{f}(k)$ im Frequenzbereich durch lineare Transformation von $\mathbf{U}_a(k)$. $\mathbf{N}(k)$ ist ein normalverteileter, mittelwertfreier Rauschterm, der die zufälligen, nicht erfassbaren Komponenten des akustischen Rückkopplungspfades modelliert und A modelliert als Übergangsfaktor die Zeitvariablität des akustischen Rückkopplungspfades [59]. Für die Schätzung von $\mathbf{N}(k)$ wird angenommen, dass sich die statistischen Eigenschaften des akustischen Rückkopplungspfades deutlich langsamer ändern als die des Quellsignals [114] und A sollte im Bereich von $0 \ll A < 1$ liegen [59, 114, 139]. Mit den Näherungen von Enzner und Vary

[59], die durch Diagonaloperationen den Rechenaufwand verringern und die Problemstellung verbessern, ergeben sich folgende Gleichungen zur Aktualisierung des adaptiven Filters [114, 138].

Schätzung:

$$\hat{\mathbf{F}}(k|k-1) = A \cdot \hat{\mathbf{F}}(k-1|k-1) \tag{5.55}$$

Minimale Prädiktions-MSE-Matrix:

$$\mathbf{P}(k|k-1) = A^2 \cdot \mathbf{P}(k-1|k-1) + \mathbf{\Psi}_{nn}(k) \tag{5.56}$$

Kalman Gain:

$$\mathbf{K}(k) = \mathbf{P}(k|k-1)\mathbf{U}_a^H(k)\left[\mathbf{U}_a(k)\mathbf{P}(k|k-1)\mathbf{U}_a^H(k) + \mathbf{\Psi}_{ee}(k)\right]^{-1} \tag{5.57}$$

Korrektur:

$$\begin{aligned}\hat{\mathbf{F}}(k|k) = \hat{\mathbf{F}}(k|k-1) &+ \mathbf{F}_N\left(\mathbf{I}_N - \mathbf{Q}\mathbf{Q}^T\right)\mathbf{F}_N{}^{-1}\mathbf{K}(k) \\ &\times \underbrace{\mathbf{F}_N\mathbf{Q}\left[\mathbf{y}_a(k) - \mathbf{Q}^T\mathbf{F}_N{}^{-1}\mathbf{U}_a(k)\hat{\mathbf{F}}(k|k-1)\right]}_{\varepsilon(k)}\end{aligned} \tag{5.58}$$

Minimale MSE-Matrix:

$$\mathbf{P}(k|k) = \left[\mathbf{I}_N - \frac{R}{N}\mathbf{K}(k)\mathbf{U}_a(k)\right]\mathbf{P}(k|k-1) \tag{5.59}$$

Der Index $(k|k-1)$ stellt eine Größe zum Zeitfenster k unter Bezugnahme von allen Messwerten bis zum Zeitpunkt $k-1$, also ohne der aktuellen Messwerte, dar. Diese werden daher „a priori“ Werte genannt. Größen mit Index $(k|k)$ sind „a posteriori“ Werte, die die aktuellen Messwerte mit einbeziehen. $\mathbf{P}(\cdot|\cdot)$ ist die Prädiktionsfehler-Kovarianzmatrix im Frequenzbereich, $\mathbf{K}(k)$ der Kalman Gain und $\mathbf{\Psi}_{nn}(k)$ und $\mathbf{\Psi}_{ee}(k)$ sind die Kovarianzmatrizen von $\mathbf{N}(k)$ und $\mathbf{E}(k)$. Diese sind nicht bekannt und können unter der Annahme eines sich nur langsam ändernden akustischen Rückkopplungsfpad geschätzt werden. $\mathbf{\Psi}_{ee}(k) = E\left\{\mathbf{E}(k)\mathbf{E}^H(k)\right\}$ kann direkt über

$$\mathbf{\Psi}_{ee}(k) \approx \alpha\mathbf{\Psi}_{ee}(k-1) + (1-\alpha)\mathbf{E}(k)\mathbf{E}^H(k) \tag{5.60}$$

genähert werden [114], wobei $0 \ll \alpha < 1$ einen Vergessensfaktor darstellt. Die Kovarianz $\Psi_{nn}(k)$ der zeitlichen Veränderung des akustischen Rückkopplungspfades kann mit

$$\mathbf{\Psi}_{nn}(k) = (1-A^2)\mathbf{\Psi}_{\hat{f}\hat{f}}(k) \tag{5.61}$$

über die Kovarianz des geschätzten akustischen Rückkopplungspfades

$$\begin{aligned}\mathbf{\Psi}_{\hat{f}\hat{f}}(k) &= E\left\{\hat{\mathbf{F}}(k)\hat{\mathbf{F}}^H(k)\right\} \\ &\approx \alpha\mathbf{\Psi}_{\hat{f}\hat{f}}(k-1) + (1-\alpha)\hat{\mathbf{F}}(k)\hat{\mathbf{F}}^H(k)\end{aligned} \tag{5.62}$$

geschätzt werden [59].

6. Mehrkanalige adaptive Rückkopplungsunterdrückung im Frequenzbereich mit Kalmanfiltern

Inhalt

6.1. Einleitung 33
6.2. Signalmodell 34
6.3. Formulierung des Algorithmus 35
6.4. Simulation 41
6.5. Zusammenfassung des Kapitels 51

6.1. Einleitung

Mit der Entdeckung, dass man durch Signalverarbeitung von mehreren Mikrofonkapseln ein Signal nicht nur zeitlich, sondern auch räumlich abtasten kann, wurde den sogenannten Mikrofonarrays immer mehr Beachtung geschenkt. Durch Delay and Sum Beamformer [140] war es erstmals möglich, elektronisch gesteuert Nutzschall aus einer bestimmten Richtung aufzunehmen und Störschall aus anderen Richtungen zu unterdrücken. Die räumliche Abtastung ermöglicht weiterhin, eintreffenden Schall zu lokalisieren und die Richtcharakteristik adaptiv darauf auszurichten [141], oder über Kopplung mit einer Videokamera Schalldrücke sichtbar zu machen (akustische Kamera) [142]. Stochastische Signalverarbeitung in Zusammenhang mit Mikrofonarrays ermöglicht beispielsweise auch eine verbesserte Nachhall- und Rauschunterdrückung, oder die Separierung von simultan aktiven Quellen [140]. Durch ihre Vorteile haben sich Mikrofonarrays auch in Hörgeräten und Konferenzsystemen etabliert. Im Bezug auf adaptive Rückkopplungsunterdrückung mit Kalman Filtern wird die Mehrkanaligkeit nach Stand des Autors nicht im adaptiven Filter selbst, sondern im Zusammenhang mit weiteren Methoden, wie Nullsteering-Beamforming [1] oder mehrkanaliger Rauschunterdrückung [2, 3] ausgenutzt und die adaptive Rückkopplungsunterdrückung einkanalig durchgeführt. Die Überlegung liegt also nahe, die Mehrkanaligkeit in Systemen mit Mikrofonarrays und einem Lautsprecher, wie es in modernen Konferenzsystemen üblich ist [7, 8], auch direkt im adaptiven Filter auszunutzen mit dem Ziel eines größeren Verstärkungsgewinns durch verbesserte Schätzung des Rückkopplungspfades in einem Referenzkanal.

6.2. Signalmodell

Das Modell des Systems mit einem Lautpsrecher und M Mikrofonen ist in Abbildung 6.1 zu finden. Der mehrkanalige akustische Rückkopplungspfad $\mathbf{F}(t)$ kann auch hier als zeivariante, lineare Raumimpulsantwort der Ordnung n_F modelliert werden, sodass sich für die Rückkopplung zwischen dem Lautsprecher und einem Mikrofon m die polynomische Übertragungsfunktion

$$\begin{aligned} F_m(z,t) &= f_m^{(0)}(t) + f_m^{(1)}(t)z^{-1} + \cdots + f_m^{(n_F)}(t)z^{-n_F} && (6.1)\\ &= \mathbf{z}^T\mathbf{f}_m(t) && (6.2)\end{aligned}$$

mit $\mathbf{f}_m(t) = \begin{bmatrix} f_m^{(0)}(t) & f_m^{(1)}(t) & \cdots & f_m^{(n_F)}(t)\end{bmatrix}^T$ und $\mathbf{z} = \begin{bmatrix}1 & z^{-1} & \cdots & z^{-n_F}\end{bmatrix}^T$ ergibt. Der mehrkanalige akustische Rückkopplungspfad lässt sich damit zu

$$\mathbf{F}(t) = \begin{bmatrix} F_1(z,t) \\ F_2(z,t) \\ \vdots \\ F_M(z,t)\end{bmatrix} \tag{6.3}$$

zusammenfassen. Ähnlich lässt sich unter der Vereinfachung von reiner linearer, zeitvarianter Signalverarbeitung der elektroakustische Vorwärtspfad mit Ordnung n_G

$$\mathbf{G}(t) = \begin{bmatrix} G_1(z,t) \\ G_2(z,t) \\ \vdots \\ G_M(z,t)\end{bmatrix} \tag{6.4}$$

formulieren mit der polynomischen Übertragungsfunktion von Mikrofon m zum Ausgangssignal $u(t)$

$$\begin{aligned} G_m(z,t) &= g_m^{(0)}(t) + g_m^{(1)}(t)z^{-1} + \cdots + g_m^{(n_G)}(t)z^{-n_G} && (6.5)\\ &= \mathbf{z}^T\mathbf{g}_m(t) && (6.6)\end{aligned}$$

im elektroakustischen Vorwärtspfad mit $\mathbf{g}_m(t) = \begin{bmatrix} g_m^{(0)}(t) & g_m^{(1)}(t) & \cdots & g_m^{(n_G)}(t)\end{bmatrix}^T$ und $\mathbf{z} = \begin{bmatrix}1 & z^{-1} & \cdots & z^{-n_G}\end{bmatrix}^T$. Die Mikrofonsignale $y_m(t)$ mit $m = 1,\ldots,M$ setzen sich aus dem rückgekoppelten Signal $x_m(t)$ mit $m = 1,\ldots,M$ und dem Quellsignal über

$$\begin{aligned} y_m(t) &= x_m(t) + H_{Sm}(z,t)v(t) \\ &= F_m(z,t)u(t) + H_{Sm}(z,t)v(t)\end{aligned} \tag{6.7}$$

zusammen. $H_{Sm}(z,t)$ ist dabei die polynomiale Übertragunsfunktion von der Quelle zum Mikrofon m. Werden die Übertragungsstrecken von der Quelle zu den Mikrofonen über

$$\mathbf{H}_{SM}(t) = \begin{bmatrix} H_{S1}(z,t) \\ H_{S2}(z,t) \\ \vdots \\ H_{SM}(z,t)\end{bmatrix} \tag{6.8}$$

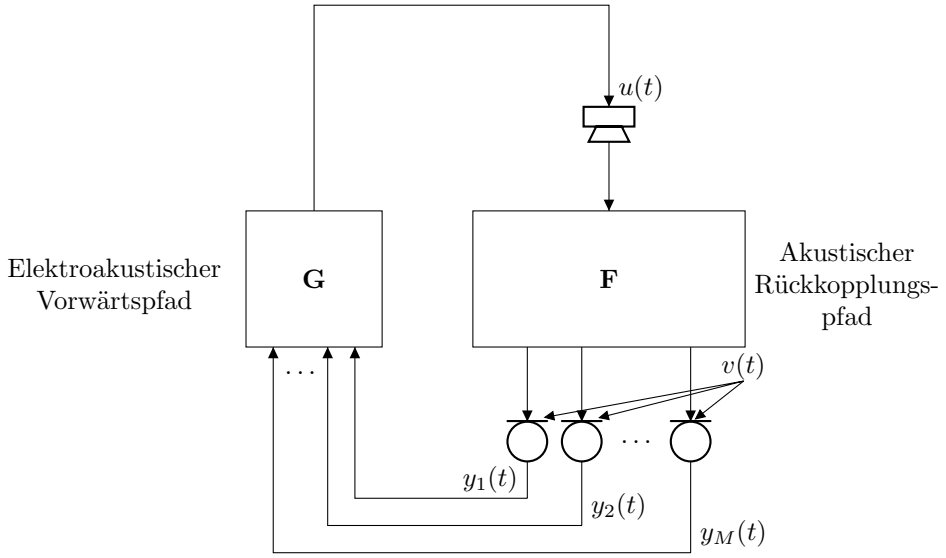


Abbildung 6.1.: *Modell einer akustischen Rückkopplungsstrecke mit einer Quelle, M Mikrofonen und einem Lautsprecher.*

zusammengefasst, kann (6.7) in Vektorenschreibweise über

$$\begin{aligned}\mathbf{y}(t) &= \mathbf{x}(t) + \mathbf{H}_{SM}(t)v(t) \\ &= \mathbf{F}(t)u(t) + \mathbf{H}_{SM}(t)v(t)\end{aligned} \tag{6.9}$$

mit

$$\mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_M(t) \end{bmatrix} \tag{6.10}$$

und

$$\mathbf{x}(t) = \mathbf{F}(t)u(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_M(t) \end{bmatrix} \tag{6.11}$$

definiert werden. Die M Mikrofonsignale $y_m(t)$ mit $m = 1, \ldots, M$ können beispielsweise im elektroakustischen Vorwärtspfad $\mathbf{G}(t)$ zu einem Ausgangssignal

$$\begin{aligned}u(t) &= \sum_{m=1}^{M} G_m(z,t)y_m(t) \\ &= \mathbf{G}^T(t)\mathbf{y}(t)\end{aligned} \tag{6.12}$$

verarbeitet werden. Das Lautsprechersignal wird dann über den mehrkanaligen, akustischen Rückkopplungspfad $\mathbf{F}(t)$ in die Mikrofone rückgekoppelt.

6.3. Formulierung des Algorithmus

Auf dieses mehrkanalige System mit einem Lautsprecher und M Mikrofonen soll nun der PEM-basierte Kalman Filter Ansatz im Frequenzbereich zur ad-

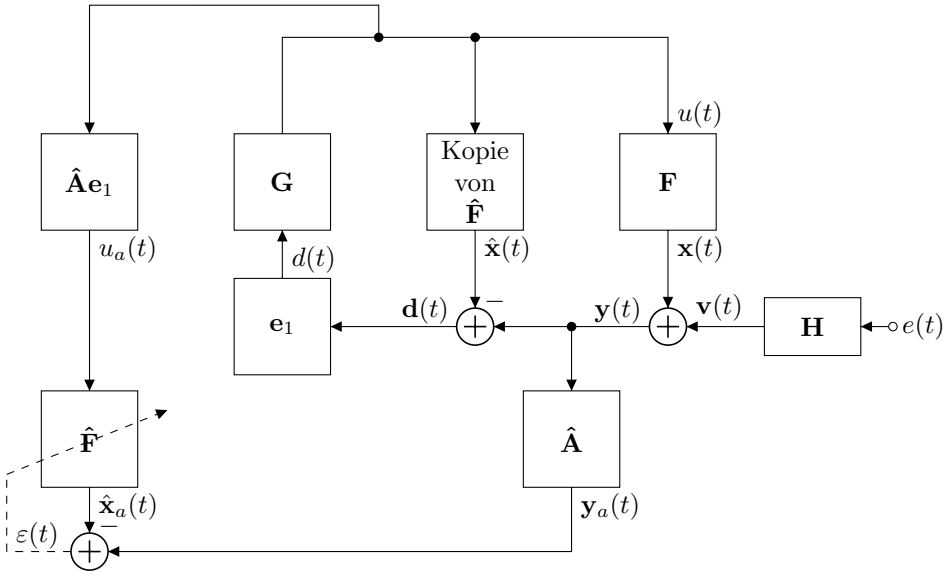


Abbildung 6.2.: *Mehrkanalige akustische Rückkopplungsstrecke mit einem Lautsprecher und M Mikrofonen und dem Prediction-Error-Method-basierten Ansatz für adaptive Rückkopplungsunterdrückung (vgl. [114])*

aptiven Rückkopplungsunterdrückung angewendet werden. Die Mehrkanaligkeit wird hierfür bei der Schätzung der adaptiven Filter ausgenutzt, die den mehrkanaligen Rückkopplungspfad modellieren sollen. Das Systemmodell mit mehrkanaliger, adaptiver Rückkopplungsunterdrückung ist in Abbildung 6.2 zu finden. Wie im einkanaligen Fall wird das Quellsignal als autoregressiver Prozess mit dem weißen Anregungssignal $e(t)$ und dem monischen, inversstabilen Filter $H(z,t)$ modelliert. Die räumliche Verteilung der Mikrofone hat zur Folge, dass zwischen der Quelle und jedem Mikrofon eine unterschiedliche Übertragungsstrecke liegt. Das Quellsignal am m-ten Mikrofon setzt sich daher aus der Übertragunsstrecke von der Quelle zum m-ten Mikrofon $H_{Sm}(z,t)$ und dem Quellsignal $v(t)$ über

$$\begin{aligned} v_m(t) &= H_{Sm}(z,t)v(t) \\ &= H_{Sm}(z,t)H(z,t)e(t) \end{aligned} \tag{6.13}$$

zusammen. Mit der Zusammenfassung von $H_{Sm}(z,t)$ und dem Quellenmodell $H(z,t)$ zu

$$H_m(z,t) = H_{Sm}(z,t)H(z,t) \tag{6.14}$$

kann (6.13) durch

$$v_m(t) = H_m(z,t)e(t) \tag{6.15}$$

$$= \frac{1}{A_m(z,t)}e(t) \tag{6.16}$$

oder

$$\mathbf{v}(t) = \begin{bmatrix} v_1(t) \\ v_2(t) \\ \vdots \\ v_M(t) \end{bmatrix} \tag{6.17}$$

in Vektorschreibweise ausgedrückt werden. Unter der Annahme, dass $H_m(z,t)$ inversstabil ist, kann der Quellsignalanteil an jedem Mikrofon durch einen eigenen autoregressiven Prozess $A_m(z,t)$ der Ordnung n_A modelliert werden. Mit dem Koeffizientenvektor des autoregressiven Modells für das Quellsignal am m-ten Mikrofon

$$\mathbf{a}_m(t) = \begin{bmatrix} 1 & a_{m,1}(t) & \cdots & a_{m,n_A}(t) \end{bmatrix}^T \tag{6.18}$$

lassen sich die M Quellmodelle zu

$$\mathbf{A}(t) = \begin{bmatrix} \mathbf{z}^T\mathbf{a}_1(t) & \mathbf{z}^T\mathbf{a}_2(t) & \cdots & \mathbf{z}^T\mathbf{a}_M(t) \end{bmatrix}^T \tag{6.19}$$

$$= \begin{bmatrix} A_1(z,t) & A_2(z,t) & \cdots & A_M(z,t) \end{bmatrix}^T \tag{6.20}$$

mit $\mathbf{z} = \begin{bmatrix} 1 & z^{-1} & \cdots & z^{-n_A} \end{bmatrix}^T$ zusammenfassen. Die Quellmodelle werden wie in Kapitel 5 über den Levinson-Durbin-Algorithmus [1, 134] zu $\hat{\mathbf{a}}_m(t)$ geschätzt, sodass sich eine geschätze AR-Matrix

$$\hat{\mathbf{A}}(t) = \begin{bmatrix} \mathbf{z}^T\hat{\mathbf{a}}_1(t) & \mathbf{z}^T\hat{\mathbf{a}}_2(t) & \cdots & \mathbf{z}^T\hat{\mathbf{a}}_M(t) \end{bmatrix}^T \tag{6.21}$$

$$= \begin{bmatrix} \hat{A}_1(z,t) & \hat{A}_2(z,t) & \cdots & \hat{A}_M(z,t) \end{bmatrix}^T \tag{6.22}$$

ergibt. Die geschätzten, rückgekoppelten Signale

$$\hat{\mathbf{x}}(t) = \hat{\mathbf{F}}(t)u(t) \tag{6.23}$$

werden von den Mikrofonsignalen $\mathbf{y}(t)$ abgezogen, sodass sich das mehrkanalige Fehlersignal

$$\begin{aligned} \mathbf{d}(t) &= \mathbf{y}(t) - \hat{\mathbf{x}}(t) \\ &= \begin{bmatrix} d_1(t) & d_2(t) & \cdots & d_M(t) \end{bmatrix}^T \end{aligned} \tag{6.24}$$

ergibt. Wird der mehrkanalige, akustische Rückkopplungspfad fehlerfrei geschätzt entspricht $\mathbf{d}(t)$ den reinen Quellsignalen $\mathbf{v}(t)$ an den Mikrofonen. Zur Bildung des einkanaligen Lautsprechersignals wird im Rahmen dieser Arbeit das Fehlersignal

$$d_1(t) = \mathbf{d}(t)^T\mathbf{e}_1 \tag{6.25}$$

des ersten Mikrofons verwendet. Der M-Werte Vektor $\mathbf{e}_1 = \begin{bmatrix} 1 & \mathbf{0} \end{bmatrix}^T$ wählt dabei die erste Zeile von $\mathbf{d}(t)$ aus. Das Lautsprechersignal bildet sich anschließend über

$$u(t) = G(z,t)d_1(t). \tag{6.26}$$

Zum Erhalt einer „unbiased“ Schätzung des mehrkanaligen Rückkopplungsfpads werden die Mikrofonsignale und das Lautsprechersignal mit

$$\mathbf{y}_a(t) = \hat{\mathbf{A}}(t)^T \mathbf{y}(t) \tag{6.27}$$

$$\begin{aligned} u_a(t) &= \hat{\mathbf{A}}(t)^T \mathbf{e}_1 u(t) \\ &= A_1(z,t) u(t) \end{aligned} \tag{6.28}$$

vorgefiltert und das vorgefilterte Lautsprechersignal mit dem geschätzten, mehrkanaligen, akustischen Rückkopplungspfad zu den geschätzten, vorgefilterten Rückkopplungssignalen

$$\hat{\mathbf{x}}_a(t) = \hat{\mathbf{F}}(t)^T u_a(t) \tag{6.29}$$

mit

$$\hat{\mathbf{F}}(t) = \begin{bmatrix} \hat{F}_1(z,t) \\ \hat{F}_2(z,t) \\ \vdots \\ \hat{F}_M(z,t) \end{bmatrix} \tag{6.30}$$

gefiltert. Das Fehlersignal

$$\boldsymbol{\varepsilon}(t) = \mathbf{y}_a(t) - \hat{\mathbf{x}}_a(t) \tag{6.31}$$

wird zur Adaption der geschätzten akustischen Rückkopplungspfade verwendet.
Zur Formulierung des Algorithmus' im Frequequenzbereich muss das Modell auch hier in eine blockbasierte Darstellung mit Zeitfenster-Index k überführt werden. Da nur ein Lautsprecher vorhanden ist, wird das einkanalige Lautsprechersignalfenster im Zeitbereich zum Zeitpunkt k wie in (5.37) als

$$\mathbf{u}(k) = \begin{bmatrix} u(kR-N+1) & \dots & u(kR-1) & u(kR) \end{bmatrix}^T \tag{6.32}$$

der Länge N definiert. Das vorgefilterte Lautsprechersignalfenster ergibt sich damit zu

$$\mathbf{u}_a(k) = \begin{bmatrix} \hat{\mathbf{A}}^T(z,t-1)\mathbf{e}_1 u(kR) \\ \vdots \\ \hat{\mathbf{A}}^T(z,t-1)\mathbf{e}_1 u(kR-N+1) \end{bmatrix} = \begin{bmatrix} \hat{A}_1(z,t-1) u(kR) \\ \vdots \\ \hat{A}_1(z,t-1) u(kR-N+1) \end{bmatrix} \tag{6.33}$$

$$= \begin{bmatrix} u_a(kR-N+1) & \cdots & u_a(kR-1) & u_a(kR) \end{bmatrix}. \tag{6.34}$$

Genau wie in (5.42) wird das Lautsprechersignalfenster im Frequenzbereich als $N \times N$-Diagonalmatrix mit der $N \times N$-DFT-Matrix $\mathbf{F}_N$ und dem Diagonalisierungsoperator $\mathrm{diag}\{\cdot\}$ über

$$\mathbf{U}_a(k) = \mathrm{diag}\left\{\mathbf{F}_N \mathbf{u}_a(k)\right\} \tag{6.35}$$

gebildet. Der Rückkopplungspfad von Lautsprecher zum m-ten Mikrofon der Ordnung $n_F = R-1$ kann im Zeitbereich als Länge R-Vektor

$$\mathbf{f}_m(k) = \begin{bmatrix} f_1(0,k) & f_1(1,k) & \cdots & f_1(N-R-1,k) \end{bmatrix}^T \tag{6.36}$$

in blockbasierter Darstellung angegeben und damit alle M Rückkopplungspfade zur $R \times M$-Matrix

$$\begin{aligned}\overline{\mathbf{F}}(k) &= \begin{bmatrix}\mathbf{f}_1(k) & \mathbf{f}_2(k) & \cdots & \mathbf{f}_M(k)\end{bmatrix} \\ &= \begin{bmatrix} f_1(0,k) & f_2(0,k) & \cdots & f_M(0,k) \\ f_1(1,k) & f_2(1,k) & \cdots & f_M(1,k) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(N-R-1,k) & f_2(N-R-1,k) & \cdots & f_M(N-R-1,k)\end{bmatrix}\end{aligned} \tag{6.37}$$

zusammengefasst werden. Zur Überführung in den Frequenzbereich werden die Koeffizientenvektoren mit $N-R$ Nullen auf eine Länge von N verlängert und mit der DFT-Matrix über

$$\begin{aligned}\mathbf{F}(k) &= \mathbf{F}_N \begin{bmatrix}\overline{\mathbf{F}}(k) \\ \mathbf{0}_{N-R\times M}\end{bmatrix} \\ &= \mathbf{F}_N \begin{bmatrix}\mathbf{f}_1(k) & \mathbf{f}_2(k) & \cdots & \mathbf{f}_M(k) \\ \mathbf{0}_{N-R\times 1} & \mathbf{0}_{N-R\times 1} & \cdots & \mathbf{0}_{N-R\times 1}\end{bmatrix} \\ &= \begin{bmatrix}\mathbf{F}_1(k) & \mathbf{F}_2(k) & \cdots & \mathbf{F}_M(k)\end{bmatrix}\end{aligned} \tag{6.38}$$

zur $N \times M$-Matrix im Frequenzbereich multipliziert. Auch das Zeitfenster der Mikrofonsignale hat für eine korrekte lineare Faltung im Frequenzbereich die Länge R. Das Mikrofonsignalfenster des m-ten Mikrofons zum Zeitpunkt k ergibt sich also zu

$$\mathbf{y}_m(k) = \begin{bmatrix}y_m(kR-R+1) & \cdots & y_m(kR-1) & y_m(kR)\end{bmatrix}^T \tag{6.39}$$

und die M Mikrofonsignale lassen sich zu einer $R \times M$-Matrix

$$\begin{aligned}\overline{\mathbf{Y}}(k) &= \begin{bmatrix}\mathbf{y}_1(k) & \mathbf{y}_2(k) & \cdots & \mathbf{y}_M(k)\end{bmatrix} \\ &= \begin{bmatrix} y_1(kR-R+1) & y_2(kR-R+1) & \cdots & y_M(kR-R+1) \\ \vdots & \vdots & \ddots & \vdots \\ y_1(kR-1) & y_2(kR-1) & \cdots & y_M(kR-1) \\ y_1(kR) & y_2(kR) & \cdots & y_M(kR)\end{bmatrix}\end{aligned} \tag{6.40}$$

zusammenfassen. Die vorgefilterten Mikrofonsignale $\overline{\mathbf{Y}}_a(k)$ ergeben sich anschließend zu

$$\begin{aligned}\overline{\mathbf{Y}}_a(k) &= \begin{bmatrix} \hat{A}_1(z,t-1)y_1(kR) & \cdots & \hat{A}_M(z,t-1)y_M(kR) \\ \vdots & \ddots & \vdots \\ \hat{A}_1(z,t-1)y_1(kR-R+1) & \cdots & \hat{A}_M(z,t-1)y_M(kR-R+1)\end{bmatrix} \\ &= \begin{bmatrix} y_{1,a}(kR-R+1) & \cdots & y_{M,a}(kR-R+1) \\ \vdots & \ddots & \vdots \\ y_{1,a}(kR-1) & \cdots & y_{M,a}(kR-1) \\ y_{1,a}(kR) & \cdots & y_{M,a}(kR)\end{bmatrix}.\end{aligned} \tag{6.41}$$

Zuletzt wird das weiße Rauschsignal $e(t)$, das in der autoregressiven Modellvorstellung das Quellsignal erzeugt, in eine blockbasierte Darstellung mit

$$\mathbf{e}(k) = \begin{bmatrix}e(kR-R+1) & \cdots & e(kR-1) & e(kR)\end{bmatrix}^T \tag{6.42}$$

überführt. Da für jedes Mikrofonsignal ein eigenes autoregressives Modell geschätzt wird, ergibt sich für jedes Mikrofonsignal ein eigenes Rauschsignal $\mathbf{e}_m(k)$ mit $m = 1, \ldots, M$. Für eine mehrkanalige Matrixdarstellung ergibt sich daher eine $R \times M$ Matrix

$$\overline{\mathbf{E}}(k) = \begin{bmatrix} \mathbf{e}_1(k) & \mathbf{e}_2(k) & \cdots & \mathbf{e}_M(k) \end{bmatrix}. \tag{6.43}$$

Im idealen Fall, in dem die Quellsignale perfekt geschätzt werden, gilt $\mathbf{e}_1(k) = \mathbf{e}_1(k) = \cdots = \mathbf{e}_M(k) = \mathbf{e}(k)$. Mit der Projektionsmatrix $\mathbf{Q}$ aus (5.49) lassen sich anschließend die vorgefilterten Mikrofon- und Quellsignalfenster über

$$\mathbf{Y}_a(k) = \mathbf{F}_N \mathbf{Q} \overline{\mathbf{Y}}_a(k) \tag{6.44}$$

$$\mathbf{E}(k) = \mathbf{F}_N \mathbf{Q} \overline{\mathbf{E}}(k) \tag{6.45}$$

in den Frequenzbereich transformieren. Analog zum einkanaligen Fall erhält man im mehrkanaligen Fall durch die Minimierung der Fehlerfunktion

$$J(\hat{\mathbf{F}}) = E\left\{ \left\| \mathbf{F}(k) - \hat{\mathbf{F}}(k) \right\|_F^2 \right\} \tag{6.46}$$

den MMSE. Da es sich hier jedoch nicht um Vektoren, sondern Matrizen handelt, wird statt der L2-Norm die Frobenius-Norm $\|\cdot\|_F$ angewendet. Mit der Modellierung des zeitlichen Änderungsverhaltens der akustischen Rückkopplungspfade als Gauß-Markov Prozess erster Ordnung, kann das mehrkanalige System als Zustandsmodell über

$$\mathbf{F}(k) = A \cdot \mathbf{F}(k-1) + \mathbf{N}(k) \tag{6.47}$$

$$\mathbf{Y}_a(k) = \mathbf{C}_a(k)\mathbf{F}(k) + \mathbf{E}(k) \tag{6.48}$$

mit der Messmatrix $\mathbf{C}_a(k) = \mathbf{F}_N \mathbf{Q}\mathbf{Q}^T {\mathbf{F}_N}^{-1} \mathbf{U}_a(k)$ zur korrekten, linearen Faltung zwischen $\mathbf{u}_a(k)$ und $\overline{\mathbf{F}}(k)$ im Frequenzbereich beschrieben werden. Auch wenn die Zustandsgleichungen (6.47) und (6.48) denen aus (5.53) und (5.54) ähnlich sehen, unterscheiden sie sich in ihrer Mehrkanaligkeit. So stellen die Größen keine einkanaligen Vektoren, sondern mehrkanalige $N \times M$ Matrizen dar.

$$\mathbf{N}(k) = \begin{bmatrix} \mathbf{N}_1(k) & \mathbf{N}_2(k) & \cdots & \mathbf{N}_M(k) \end{bmatrix} \tag{6.49}$$

mit

$$\mathbf{N}_m(k) = \begin{bmatrix} N_m(0,k) & N_m(1,k) & N_m(N-1,k) \end{bmatrix}^T \quad \text{mit } m = 1, \ldots, M \tag{6.50}$$

modelliert die zufälligen, nicht erfassbaren Komponenten und A die Zeitvariabilität als Übergangsfaktor der akustischen Rückkopplungspfade. Zur Aktualisierung der adaptiven Filter ergeben sich nahezu die gleichen Gleichungen (6.52) - (6.55) wie im einkanaligen Fall. Der Unterschied liegt vor allem in der Skalierung der Kovarianzmatrizen $\mathbf{\Psi}_{\hat{f}\hat{f}}(k)$ in (6.58) und $\mathbf{\Psi}_{ee}(k)$ in (6.56) mit der Anzahl der Mikrofone M und in den mehrkanaligen Matrizen statt der einkanaligen Vektoren.

Schätzung:

$$\hat{\mathbf{F}}(k|k-1) = A \cdot \hat{\mathbf{F}}(k-1|k-1) \tag{6.51}$$

Minimale Prädiktions-MSE-Matrix:

$$\mathbf{P}(k|k-1) = A^2 \cdot \mathbf{P}(k-1|k-1) + \mathbf{\Psi}_{nn}(k) \tag{6.52}$$

Kalman Gain:

$$\mathbf{K}(k) = \mathbf{P}(k|k-1)\mathbf{U}_a^H(k)\left[\mathbf{U}_a(k)\mathbf{P}(k|k-1)\mathbf{U}_a^H(k) + \mathbf{\Psi}_{ee}(k)\right]^{-1} \tag{6.53}$$

Korrektur:

$$\begin{aligned}\hat{\mathbf{F}}(k|k) = \hat{\mathbf{F}}(k|k-1) &+ \mathbf{F}_N\left(\mathbf{I}_N - \mathbf{Q}\mathbf{Q}^T\right)\mathbf{F}_N{}^{-1}\mathbf{K}(k) \\ &\times \underbrace{\mathbf{F}_N\mathbf{Q}\left[\mathbf{y}_a(k) - \mathbf{Q}^T\mathbf{F}_N{}^{-1}\mathbf{U}_a(k)\hat{\mathbf{F}}(k|k-1)\right]}_{\varepsilon(k)}\end{aligned} \tag{6.54}$$

Minimale MSE-Matrix:

$$\mathbf{P}(k|k) = \left[\mathbf{I}_N - \frac{R}{N}\mathbf{K}(k)\mathbf{U}_a(k)\right]\mathbf{P}(k|k-1) \tag{6.55}$$

Die Indizes $(k|k-1)$ und $(k|k)$ kennzeichnen auch hier „a priori„ beziehungsweise „a posteriori„ Werte. Die Koviarianzmatrizen $\mathbf{\Psi}_{nn}(k)$ und $\mathbf{\Psi}_{ee}(k)$ von $\mathbf{N}(k)$ und $\mathbf{E}(k)$ sind hier ebenfalls nicht bekannt und müssen ähnlich zu (5.60 - 5.62) über

$$\begin{aligned}\mathbf{\Psi}_{ee}(k) &= E\left\{\mathbf{E}(k)\mathbf{E}^H(k)\right\} \\ &\approx \alpha\mathbf{\Psi}_{ee}(k-1) + (1-\alpha)\frac{1}{M}\mathbf{E}(k)\mathbf{E}^H(k)\end{aligned} \tag{6.56}$$

und

$$\mathbf{\Psi}_{nn}(k) = (1-A^2)\mathbf{\Psi}_{\hat{f}\hat{f}}(k) \tag{6.57}$$

mit

$$\begin{aligned}\mathbf{\Psi}_{\hat{f}\hat{f}}(k) &= E\left\{\hat{\mathbf{F}}(k)\hat{\mathbf{F}}^H(k)\right\} \\ &\approx \alpha\mathbf{\Psi}_{\hat{f}\hat{f}}(k-1) + (1-\alpha)\frac{1}{M}\hat{\mathbf{F}}(k)\hat{\mathbf{F}}^H(k)\end{aligned} \tag{6.58}$$

und dem Vergessenfaktor $0 \ll \alpha < 1$ geschätzt werden. Für $M = 1$ entspricht der vorgeschlagene Algorithmus dem einkanaligen PEM-FDKF aus Kapitel 5.4.

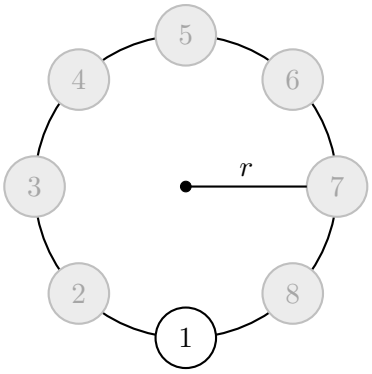
6.4. Simulation

Zur Evaluierung der vorgeschlagenen PEM-basierten mehrkanaligen akustischen Rückkopplungsunterdrückung im Frequenzbereich mit Kalman Filtern wurden in einem reflexionsarmen Raum Impulsantworten mit einem achtkanaligen, kreisförmigen Mikrofonarray aus Abbildung 6.3 aufgenommen, die die

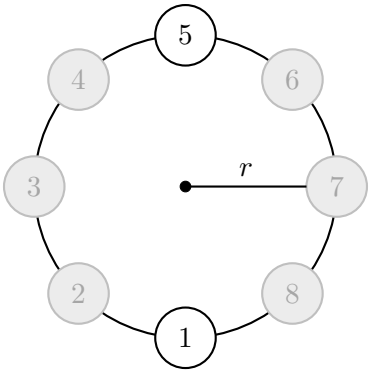
mehrkanalige akustische Rückkopplungsstrecke simulieren sollen. Der Radius des Mikrofonarrays beträgt $r = 3.5$ cm. Aus diesem Array wurden 4 verschiedene Zusammensetzungen

$$\mathcal{M}_1 = \{1\} \tag{6.59}$$
$$\mathcal{M}_2 = \{1, 5\} \tag{6.60}$$
$$\mathcal{M}_3 = \{1, 3, 5, 7\} \tag{6.61}$$
$$\mathcal{M}_4 = \{1, 2, 3, 4, 5, 6, 7, 8\} \tag{6.62}$$

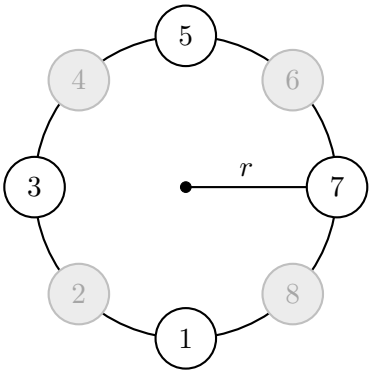
gewählt, um die Leistung des vorgeschlagenen Algorithmus bei unterschiedlicher Anzahl der Mikrofonkanäle zu bestimmen. Die Zahlen innerhalb der Mengendefinition entsprechen dabei den Mikrofonen aus Abbildung 6.3. $\mathcal{M}_1$ stellt



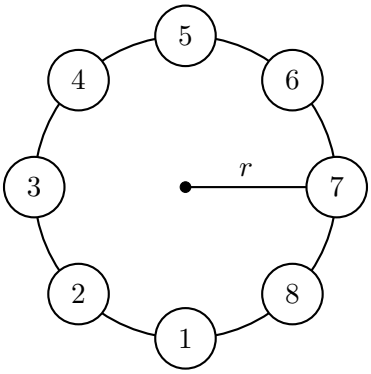
(a) *Zusammensetzung $\mathcal{M}_1$ des Mikrofonarrays*



(b) *Zusammensetzung $\mathcal{M}_2$ des Mikrofonarrays*



(c) *Zusammensetzung $\mathcal{M}_3$ des Mikrofonarrays*



(d) *Zusammensetzung $\mathcal{M}_4$ des Mikrofonarrays*

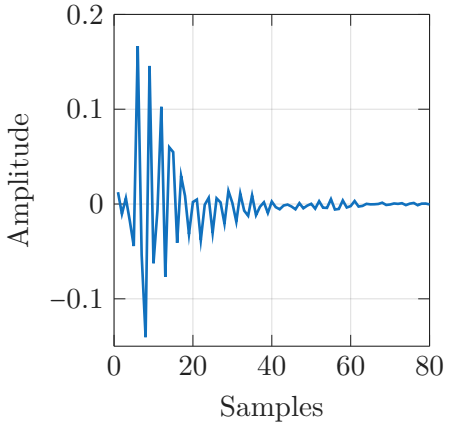
Abbildung 6.3.: *Zusammensetzungen $\mathcal{M}_1$ – $\mathcal{M}_4$ des verwendeten kreisförmigen Mikrofonarrays zur Generierung von akustischen Rückkopplungsstrecken mit unterschiedlicher Kanalanzahl*

also den einkanaligen Fall mit $M = |\mathcal{M}_1| = 1$ Mikrofonen dar, mit der Zusammensetzung $\mathcal{M}_2$ wird mehrkanalige akustische Rückkopplungsunterdrückung mit einem linearen Array mit $M = |\mathcal{M}_2| = 2$ Mikrofonen durchgeführt und $\mathcal{M}_3$ und $\mathcal{M}_4$ unterdrücken akustische Rückkopplungen mit einem kreisförmigen Mikrofonarray mit $M = |\mathcal{M}_3| = 4$ beziehungsweise $M = |\mathcal{M}_4| = 8$ Mikrofonen.

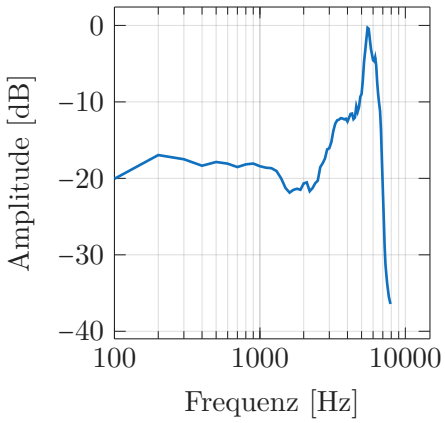
Der Versuchsaufbau zur Gewinnung der Impulsantworten der mehrkanaligen akustischen Rückkopplungsstrecke ist in Abbildung 6.4 zu finden. Der Abstand zwischen Lautsprecher und dem ersten Mikrofon beträgt 11 cm. Mit einem

Abbildung 6.4.: *Achtkanaliges, ringförmiges Mikrofonarray und ein Genelec-Lautsprecher zur Generierung von Impulsantworten der mehrkanaligen, akustischen Rückkopplungsstrecke*

Verschiebungsintervall $R = 80$ und der Größe des Zeitfensters $N = 2R = 160$ bei einer Abtastrate von $f_s = 16$ kHz ergibt sich beispielsweise die Impulsantwort $\mathbf{f}_1(t)$ zwischen dem Lautsprecher und Mikrofon 1 und dessen einseitiger Betragsfrequenzgang $|\mathbf{F}_1(t)|$ aus den Abbildungen 6.5a und 6.5b.



(a) *Impulsantwort $\mathbf{f}_1(t)$ der Rückkopplungsstrecke von Lautsprecher zum ersten Mikrofon der Länge $R = 80$ Samples*



(b) *Einseitiger Betragsfrequenzgang $|\mathbf{F}_1(t)|$ der Rückkopplungsstrecke zwischen Lautsprecher und erstem Mikrofon*

Abbildung 6.5.: *Impulsantwort und einseitiges Betragsfrequenzspektrum der Rückkopplungsstrecke zwischen Lautsprecher und dem ersten Mikrofon*

Zusätzlich wurden aus einer anderen Lautsprecherposition weitere Impulsantworten aufgenommen, um das Quellsignal $v(t)$ aus Abbildung 6.6 mit den korrekten Laufzeit- und Lautstärkeunterschieden an den Mikrofonen darzu-

stellen. Als Quellsignal wurde eine 8-sekündige Sprachsequenz aus der ITU-T P.501 Datenbank [143] (C2.3 British English - Female 2) herangezogen.

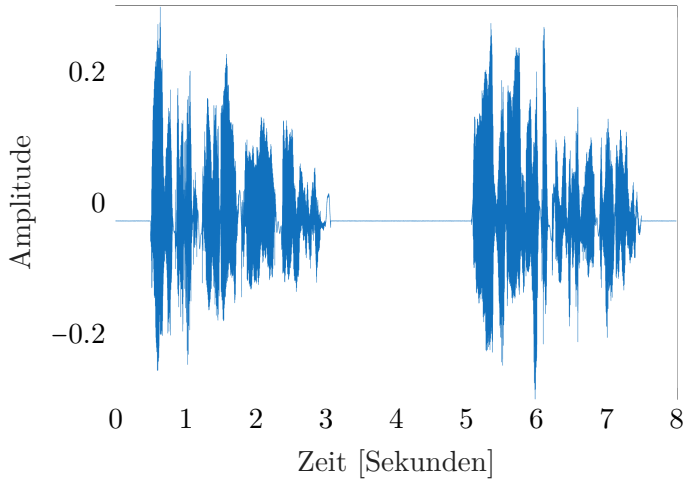


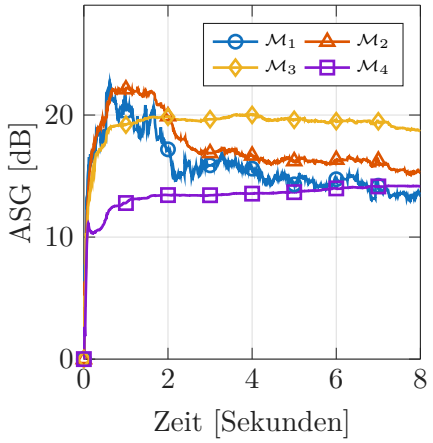
Abbildung 6.6.: *Zeitverlauf des verwendeten Quellsignals*

Das Verschiebungsintervall und die DFT-Größe wurden für die Simulation auf $R = 80$ und $N = 2R = 160$ festgelegt. Die Länge der autoregressiven Sprachmodelle wurde mit $n_A = R = 80$ definiert. Zur Evaluierung des vorgeschlagenen Algorithmus' wurde für jede der vier Mikrofonzusammensetzungen das Rückkopplungsszenario mit einer Verzögerung von 80 Samples ($\geq R$) und breitbandigen Verstärkungsfaktoren $K(t)$ im elektroakustischen Vorwärtspfad von 0 - 20 dB über dem MSG in 1 dB Schritten simuliert. Der MSG wurde experimentell bestimmt. Die Rückkopplungspfade blieben während der Simulation unverändert, auch wenn der Gauß-Markov Prozess in Gleichung (6.47) Veränderungen über der Zeit annimmt. Dies wurde wie in [114] so gewählt, um zu zeigen, dass die Wahl des Übergangsfaktors A vom letzten zum aktuellen Zustand auch im mehrkanaligen Szenario nicht extrem kritisch ist und wurde auf $A = 0.99999$ festgelegt. Weiter wurde der Vergessensfaktor α zur Näherung der Kovarianzmatrizen mit $\alpha = 0.85$ definiert. Die Prädiktionsfehlermatrix $\mathbf{P}(k)$ wurde mit Vorwissen zu den stochastischen Eigenschaften der Rückkopplungspfade und in Abhänigkeit der Anzahl der Mikrofone M mit $\mathbf{P}(0) = M^2\sigma_{f1}^2\mathbf{I}_N$ initialisiert. $\sigma_{f1}^2 = E\left\{(\mathbf{f}_1(t) - \boldsymbol{\mu}_{f_1})^2\right\}$ mit $\boldsymbol{\mu}_{f_1} = E\{\mathbf{f}_1(t)\}$ entspricht dabei der Varianz der Impulsantwort der Rückkopplungsstrecke zwischen Lautsprecher und dem ersten Mikrofon. Der Kalman Gain wurde mit $\mathbf{K}(0) = \mathbf{I}_N$, die geschätzten autoregressiven Sprachmodelle zu $\hat{\mathbf{a}}_m(0) = \begin{bmatrix}1 & \mathbf{0}\end{bmatrix}^T$ und die Impulsantworten der geschätzten Rückkopplungspfade mit $\hat{\mathbf{f}}_m(0) = \mathbf{0}$ mit $m = 1, \ldots, M$ initalisiert.

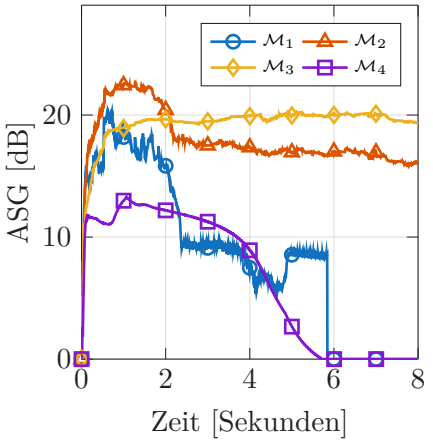
Abbildung 6.7 zeigt für ausgewählte breitbandige Verstärkungsfaktoren K_i des elektroakustischen Vorwärtspfades den Added Stable Gain (ASG) für alle Mikrofonzusammenstellungen $\mathcal{M}_1 - \mathcal{M}_4$. Da die Verstärkungsfaktoren während einem Durchlauf der Simulation nicht verändert werden, wurde die Zeitabhängigkeit (t) in der Darstellung entfernt. Zum einen ist zu erkennen, dass der

Algorithmus unabhängig von der Anzahl der Mikrofone nach ungefähr 0.5 Sekunden Stabilität erreicht. Sowohl die einkanalige Rückkopplungsstrecke mit $\mathcal{M}_1$ als Referenz-Algorithmus als auch die mehrkanaligen Rückkopplungsstrecken $\mathcal{M}_2 - \mathcal{M}_4$ unabhängig der Anzahl der Mikrofone erreichen damit Stabilität in der gleichen Zeit wie in [114]. Abbildung 6.7a zeigt, dass bei einer Verstärkung von $K_1 = 13$ dB über dem MSG alle simulierten Mikrofonzusammensetzungen nach einer kurzen Konvergenzzeit Stabilität erreichen. Bei einer Verstärkung von $K_2 = MSG + 14$ dB in Abbildung 6.7b können sowohl das einkanalige als auch das achtkanalige System $\mathcal{M}_1$ und $\mathcal{M}_4$ keinen ausreichenden zusätzätzlichen Verstärkungsgewinn mehr erzielen. Der ASG fällt daher unter die Verstärkung K_2 und das rückgekoppelte System wird in diesen beiden Fällen instabil und es treten entsprechende Artefakte im Lautsprechersignal auf. Die zweikanaligen und vierkanaligen Mikrofonarrays $\mathcal{M}_2$ und $\mathcal{M}_3$ können jedoch weiterhin ausreichende zusätzliche Verstärkung erzielen und garantieren weiterhin Stabilität. Wie in Abbildung 6.7c zu erkennen ist, ist dies auch bei einer Verstärkung von $K_3 = 15$ dB über dem MSG der Fall. Erst bei einer Verstärkung von $K_4 = MSG + 16$ dB in Abbildung 6.7d wird das zweikanalige Rückkopplungssystem $\mathcal{M}_2$ nach kurzer Zeit instabil. Das vierkanalige System $\mathcal{M}_3$ erreicht auch hier eine ausreichende zusätzliche Verstärkung, sodass dieses System weiterhin stabil bleibt. Selbst eine Verstärkung von $K_5 = 18$ dB über dem MSG in Abbildung 6.7e lässt das vierkanalige System nicht instabil werden. Ab einer Verstärkung von $K_6 = 19$ dB über dem MSG ist jedoch auch das vierkanalige System nicht mehr in der Lage, einen ausreichenden Verstärkungsgewinn zur erreichen und wird daher auch instabil. Im Bezug auf den Verstärkungsgewinn (ASG) durch den vorgeschlagenen Algorithmus in Zusammenhang mit Mikrofonarrays ist im zweikanaligen System $\mathcal{M}_2$ eine 2 dB und im vierkanaligen System $\mathcal{M}_3$ eine 5 dB höhere Verstärkung zu erreichen als im einkanaligen Fall. Lediglich das achtkanalige System $\mathcal{M}_4$ scheint keinen Leistungsgewinn zu erbringen. Diese Ergebnisse werden durch die Auswertung des Misadjustment (Mis) in den Abbildungen 6.8a - 6.8f bestätigt. Da das Lautsprechersignal dem rückkopplungsunterdrückten Signal des ersten Mikrofons entspricht ist der Mis lediglich für die Schätzung des Rückkopplungspfades zwischen Lautsprecher und dem ersten Mikrofon aufgetragen. Bei den gleichen Verstärkungen K_i im elektroakustischen Vorwärtspfad wie in den Grafiken zum ASG ist eine gute beziehungsweise schlechte Schätzung der Rückkopplungsstrecke zu erkennen. Auffällig ist auch hier, dass das achtkanalige System $\mathcal{M}_4$ deutlich schlechtere Ergebnisse erzielt als die anderen betrachteten Systeme. Die Betrachtung des Effective Closed-Loop Gain (ECLG) der Mikrofonzusammensetzungen $\mathcal{M}_1 - \mathcal{M}_4$ über der Verstärkung über dem MSG in Abbildung 6.9 ergibt ähnliche Ergebnisse. Hier erkennt man jedoch, dass das achtkanalige System $\mathcal{M}_4$ schon für $K = 12$ dB über dem MSG instabil war, da hierfür kein Datenpunkt im Diagramm existiert. Durch die Instabilität ($ECLG > 0$ dB) werden die Werte duch die Matrixinversion in (6.53) zu klein, sodass sie digital nicht mehr darstellbar sind und durch ein „nan (not a number)“ ersetzt. Des Weiteren zeigt der ECLG für die ein- und zweikanaligen Systeme $\mathcal{M}_1$ und $\mathcal{M}_2$ enstprechend eine Stabilität bei $K = 14$ dB und $K = 16$

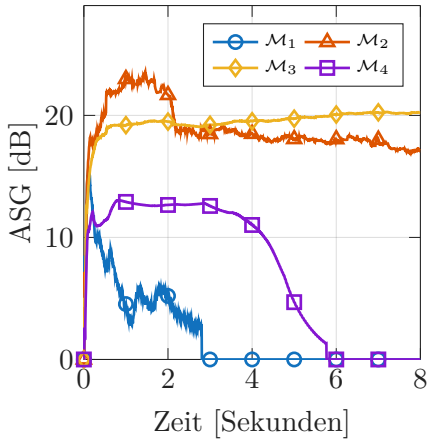
dB über dem MSG an. ASG und Mis zeigen hier jedoch schon Instabilitäten auf. Dies liegt daran, dass für die Berechnung des ECLG der Median über der Zeit herangezogen wird. In den simulierten Fällen, in denen ein System ab einer bestimmten Verstärkung instabil wird (wie $K_2 = MSG + 14$ dB für $\mathcal{M}_1$) ist das System zu anfang noch stabil und wird erst nach einiger Zeit instabil. Der Median entspricht in bestimmten Fällen dann einem Wert, der zu einem Zeitpunkt aufgetreten ist, in dem das System gerade noch stabil war.
Die bisher betrachteten Metriken lassen ein Urteil zu, wie lange ein System stabil ist und ab wann Artefakte auf Grund von Instabilitäten auftreten können. Jedoch lassen sich hiermit noch keine Aussagen darüber treffen, wie das eigentliche Lautsprechersignal durch die akustische Rückkopplungsunterdrückung verändert und verzerrt wird. Hierfür werden drei objektive Metriken verwendet, die eine starke Korrelation zu subjektiven Hörtests aufweisen [124]. Die Betrachtung des PESQ in Abbildung 6.10, des STOI in Abbildung 6.11 und des SD in Abbildung 6.12 zeigt auch hier deutlich, dass das System $\mathcal{M}_4$ mit acht Mikrofonen die schlechtesten Ergebnisse erzielt. Sind die Werte für eine Verstärkung von $K = 8$ dB über dem MSG noch vergleichbar mit den anderen Systemen, verschlechtern sich die Metriken darüber hinaus stark. Die Werte von PESQ und STOI zeigen, dass die zwei- und vierkanaligen Systeme $\mathcal{M}_2$ und $\mathcal{M}_3$ sowohl im Bereich, in dem der Referenz-Algorithmus mit $\mathcal{M}_1$ stabil ist, als auch darüber hinaus in der Signalwahrnehmung bessere und qualitativ sehr hochwerige Ergebnisse erzielen. Die Signalverzerrung SD des zweikanaligen Systems $\mathcal{M}_2$ liegt im Bereich des Referenz-Algorithmus' und ist auch über den Stabilitätsbereich von $\mathcal{M}_1$ hinaus stabil. Das vierkanalige System $\mathcal{M}_3$ erreicht eine leicht höhere Signalverzerrung SD, die aber noch in einem Bereich liegt, der für den Menschen kaum wahrnehmbar ist. Dies wird durch die Werte von PESQ und STOI bestätigt.



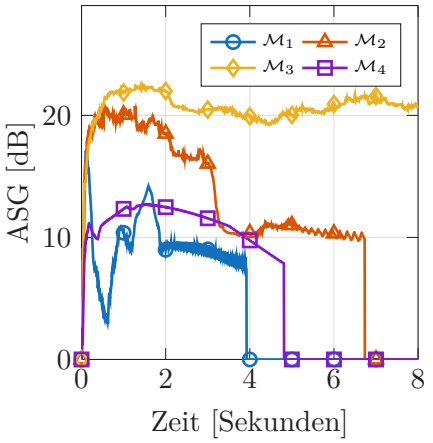
(a) *ASG in dB für $K_1 = MSG + 13$ dB*



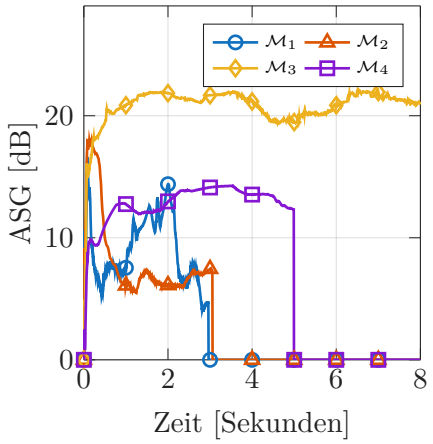
(b) *ASG in dB für $K_2 = MSG + 14$ dB*



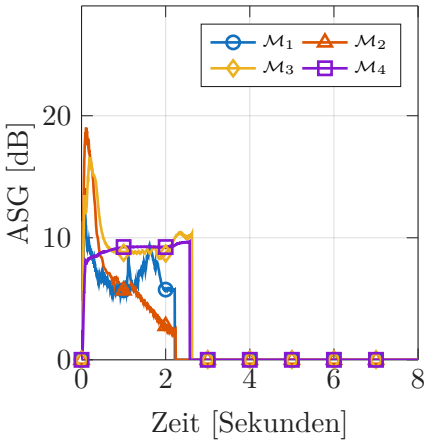
(c) *ASG in dB für $K_3 = MSG + 15$ dB*



(d) *ASG in dB für $K_4 = MSG + 16$ dB*

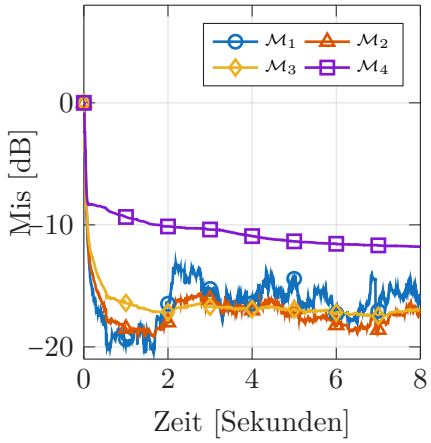


(e) *ASG in dB für $K_5 = MSG + 18$ dB*

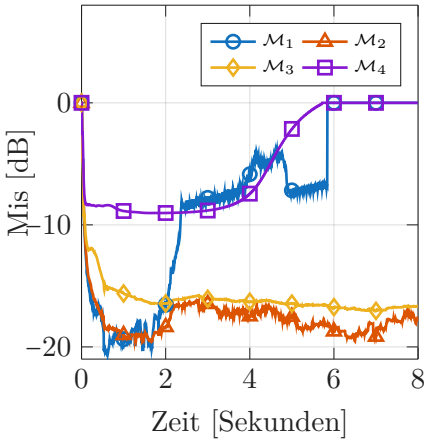


(f) *ASG in dB für $K_6 = MSG + 19$ dB*

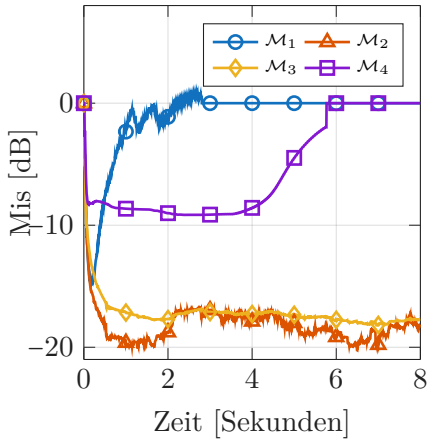
Abbildung 6.7.: *Added Stable Gain (ASG) in dB für die Mikrofonzusammensetzungen $\mathcal{M}_1$ – $\mathcal{M}_4$ mit entsprechend 1, 2, 4 und 8 Mikrofonen über der Zeit für verschiedene breitbandige Verstärkungen K_i im elektroakustischen Vorwärtspfad*



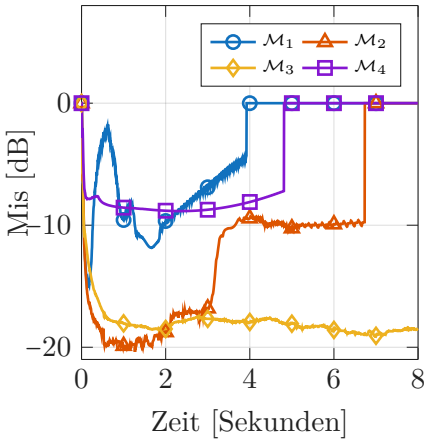
(a) *Mis in dB für $K_1 = MSG + 13$ dB*



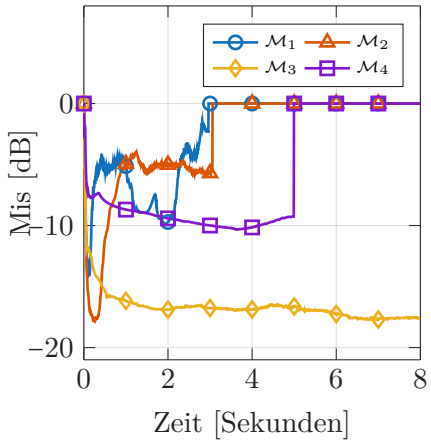
(b) *Mis in dB für $K_2 = MSG + 14$ dB*



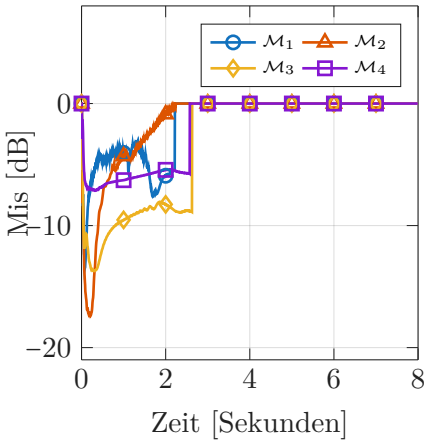
(c) *Mis in dB für $K_3 = MSG + 15$ dB*



(d) *Mis in dB für $K_4 = MSG + 16$ dB*



(e) *Mis in dB für $K_5 = MSG + 18$ dB*



(f) *Mis in dB für $K_6 = MSG + 19$ dB*

Abbildung 6.8.: *Misadjustment (Mis) in dB für die Mikrofonzusammensetzungen $\mathcal{M}_1 - \mathcal{M}_4$ mit entsprechend 1, 2, 4 und 8 Mikrofonen über der Zeit für verschiedene breitbandige Verstärkungen K_i im elektroakustischen Vorwärtspfad*

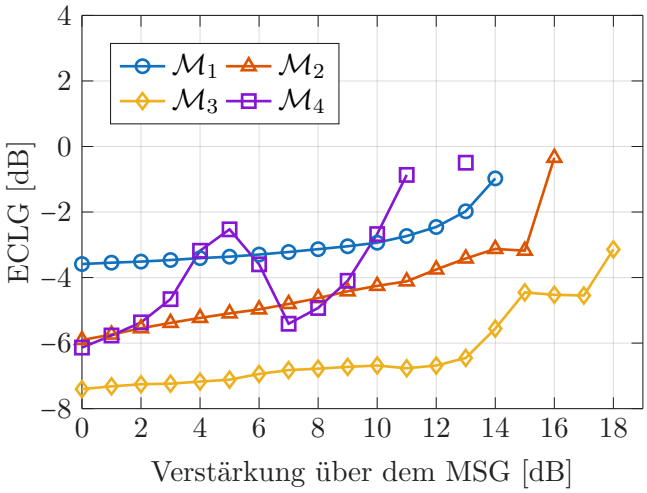


Abbildung 6.9.: *Effective Closed Loop Gain (ECLG) in dB für die Mikrofonzusammensetzungen $\mathcal{M}_1 - \mathcal{M}_4$ mit entsprechend 1, 2, 4 und 8 Mikrofonen in Bezug auf die Verstärkung im Vorwärtspfad über dem Maximum Stable Gain (MSG)*

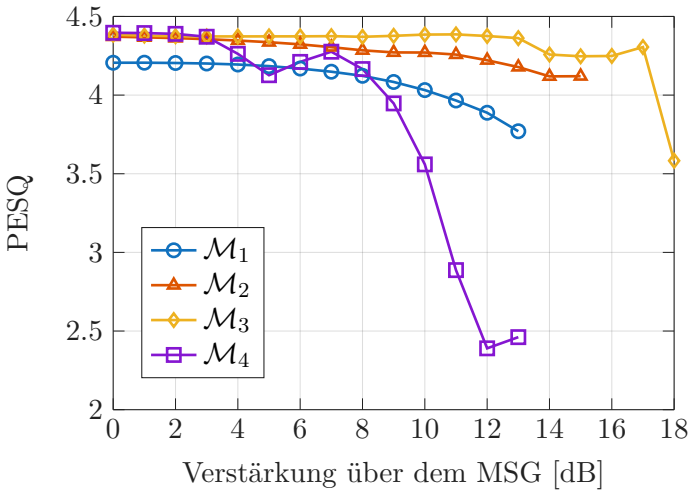


Abbildung 6.10.: *Perceptual Evaluation of Speech Quality (PESQ) für die Mikrofonzusammensetzungen $\mathcal{M}_1 - \mathcal{M}_4$ mit entsprechend 1, 2, 4 und 8 Mikrofonen in Bezug auf die Verstärkung im Vorwärtspfad über dem Maximum Stable Gain (MSG)*

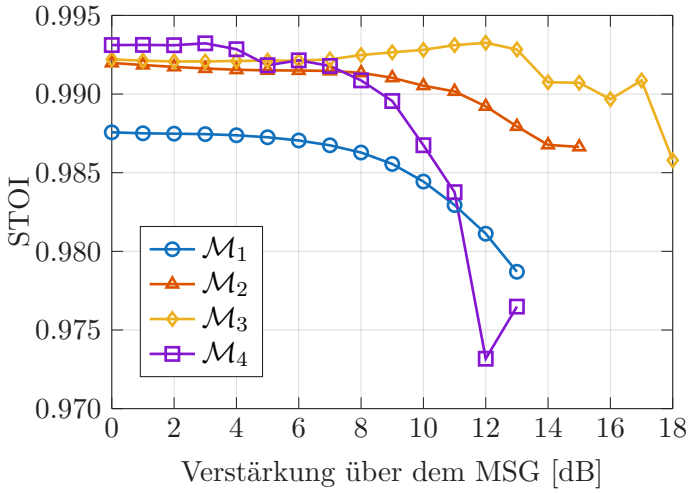


Abbildung 6.11.: *Short-Time Objective Intelligibility (STOI) für die Mikrofonzusammensetzungen $\mathcal{M}_1 - \mathcal{M}_4$ mit entsprechend 1, 2, 4 und 8 Mikrofonen in Bezug auf die Verstärkung im Vorwärtspfad über dem Maximum Stable Gain (MSG)*

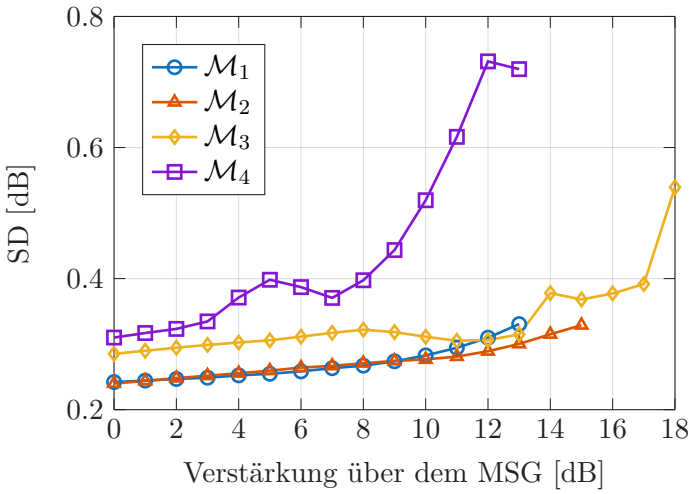


Abbildung 6.12.: *Log-Spectral Distance (SD) in dB für die Mikrofonzusammensetzungen $\mathcal{M}_1 - \mathcal{M}_4$ mit entsprechend 1, 2, 4 und 8 Mikrofonen in Bezug auf die Verstärkung im Vorwärtspfad über dem Maximum Stable Gain (MSG)*

6.5. Zusammenfassung des Kapitels

In diesem Kapitel wurde der vorgeschlagene Algorithmus zur mehrkanaligen akustischen Rückkopplungsunterdrückung im Frequenzbereich mit Kalman Filtern unter Ausnutzung der Mehrkanaligkeit von modernen Tisch- und Deckenarrays in Konferenzsystemen direkt im adaptiven Filter vorgestellt. Das entsprechende Signalmodell wurde definiert und die Berechnungsvorschriften zur Aktualisierung der mehrkanaligen adaptiven Filter beschrieben. Anschließend wurde der vorgeschlagene Algorithmus mit einem achtkanaligen Mikrofonarray evaluiert. Hierzu wurden die Mikrofone des Mikrofonarrays zu vier verschiedenen Zusammensetzungen $\mathcal{M}_1 - \mathcal{M}_4$ mit einem, zwei, vier und acht Mikrofonen aufgeteilt, um die Leistung des Algorithmus in Abhängigkeit der Anzahl der Kanäle in der akustischen Rückkopplungsstrecke zu messen. Die Zusammensetzung $\mathcal{M}_1$ mit einem Mikrofon entspricht dabei dem Referenz-Algorithmus aus Kapitel 5.4. Anhand von Simulationen mit gemessenen Rückkopplungsstrecken konnte gezeigt werden, dass der vorgeschlagene Algorithmus mit zwei- und vierkanaliger Rückkopplungsstrecke ($\mathcal{M}_2$ und $\mathcal{M}_3$) einen Verstärkungsgewinn von 2 dB beziehungsweise 5 dB bei ähnlicher oder besserer Leistung der wahrnehmungskorrelierten Metriken im Vergleich zum Referenz-Algoritmus erreicht. Lediglich die achtkanalige Rückkopplungsstrecke mit $\mathcal{M}_4$ erreicht schlechtere Ergebnisse.

7. Zusammenfassung

Diese Arbeit befasst mit dem Problem der akustischen Rückkopplung im Bezug auf Decken- und Tischarrys in modernen Konferenzsystemen und den Methoden zu dessen Unterdrückung. Hierzu wurde das Systemmodell aufgestellt und die Problemstellung diskutiert. Anschließend wurden bestehende Methoden zur akustischen Rückkopplungsunterdrückung vorgestellt. Diese beziehen sich größtenteils auf Hörgeräte und Beschallungssysteme. Das Systemmodell ist aber zu dem mit Tisch- und Deckenarrays in modernen Konferenzsystemen identisch, sodass diese direkt auch auf diesen Fall angewendet werden können. Das Ziel der Arbeit ist die Ausnutzung der Mehrkanaligkeit der Tisch- und Deckenarrays im adaptiven Filter bei der Verwendung von adaptiver Rückkopplungsunterdrückung. Als Baseline-Algorithmus wurde der Prediction Error Method Frequency Domain Kalman Filter (PEM-FDKF) identifiziert. Durch die Berechnung im Frequenzbereich können Dekorrelationseigenschaften der diskreten Fouriertransformation ausgenutzt werden, um den „Bias“ der geschätzten Lösung zu reduzieren. Die Anwendung des Kalman Filters erlaubt darüber hinaus die automatische Schätzung der optimalten Schrittweite des adaptiven Filters, sodass im Vergleich zu anderen Optimierungsmethoden verbesserte Konvergenzeigenschaften erreicht werden. Zur Ausnutzung der Mehrkanaligkeit im adaptiven Filter wurde der Baseline-Algorithmus erweitert, sodass die Rückkopplungsfpade aller Mikrofone gemeinsam geschätzt werden, um höhere Verstärkungsgewinne zu erzielen. Hierzu wurde das entsprechende Systemmodell definiert und über Zustandsgleichungen beschrieben. Anschließend wurden die Berechnungsvorschriften zur Aktualisierung der adaptiven Filter im mehrkanaligen System aufgestellt. Zur Evaluierung des vorgestellten Algorithmus' wurden mit einem Lautsprecher und einem achtkanaligen, kreisförmigen Mikrofonarray Rückkopplungsstrecken in einem reflexionsarmen Raum gemessen. Mit diesen wurde ein rückgekoppeltes System für vier Mikrofonzusammenstellungen mit einem, zwei, vier und acht Mikrofonen für verschiedene breitbandige Verstärkungen im elektroakustischen Vorwärtspfad simuliert. Über verschiedene Metriken wurde gezeigt, dass der vorgeschlagene Algorithmus mit zwei- und vierkanaligen Rückkopplungsstrecken einen Verstärkungsgewinn von 2 dB beziehungsweise 5 dB im Vergleich zum Baseline-Algorihmus bei ähnlicher oder besserer wahrnehmungskorrelierter Metriken erreicht. Das Ziel der Ausnutzung der Mehrkanaligkeit von Tisch- und Deckenarrays in modernen Konferenzsystemen im adaptiven Filter bei adaptiver Rückkopplungsunterdrückung konnte daher erfolgreich umgesetzt werden.

7.1. Ausblick

Zur weiteren Verbesserung der akustischen Rückkopplungsunterdrückung ähnlich zu [1] soll in Zukunft die gemeinsame Verwendung des vorgeschlagenen Algorithmus mit einem Nullsteering-Beamformers untersucht werden. In diesem Fall soll der Beamformer jedoch mit den vorgefilterten Mikrofonsignalen hinter der AFC angeordnet werden. Nach [1] wäre eine Verdopplung des Verstärkungsgewinns zu erwarten. Des Weiteren wurden hier Signale betrachtet, die in einem reflexionsarmen Raum erzeugt wurden. Diese bestehen daher fast nur aus dem Direktsignal und haben eine sehr kurze Länge. In einem realen Raum wird dem Direksignal Nachhall hinzugefügt, sodass die Längen der Impulsantworten deutlich größer sind. Dies würde im vorgestellten Algorithmus zu Signalverzögerungen führen, die am Lautsprechersignal deutlich zu hören wären. Um dieses Problem zu umgehen, soll in Zukunft auch der Prediction Error Method-basiere Partioned Block Frequency Domain Kalman Filter (PEM-PBFDKF) [115] auf den mehrkanaligen Fall erweitert werden. Durch die Aufteilung der Signale in kleinere Blöcke kann die Signalverzögerung bei langen Impulsantworten deutlich reduziert werden. Zuletzt soll untersucht werden, ob die Anordnung der Mikrofone einen Einfluss auf den Verstärkungsgewinn hat. So wäre zu untersuchen, ob durch eine andere Anordnung der acht Mikrofone in der hier vorgestellten Simulation, die Leistung der adaptiven Rückkopplungsunterdrückung mit acht Mikrofonen verbessert werden kann. Falls ja, wäre zu untersuchen, ob eine optimale relative Positionierung der Mikrofone exisitert. In diesem Fall könnte der vorgestellte Algorithmus anschließend auch mit mehr als acht Mikrofonen evaluiert werden. In Tischarrays sind zwar vier verbaute Mikrofone die Regel, Deckenarrays verwenden jedoch üblicherweise deutlich mehr. Im Deckenarray aus [8] sind beispielsweise 28 Mikrofonkapseln verbaut.

Literatur

[1] Henning Schepker. „Robust feedback suppression algorithms for single- and multi-microphone hearing aids“. Diss. 19. Dez. 2017.

[2] Santiago Ruiz, Toon Van Waterschoot und Marc Moonen. „Cascade Multi-Channel Noise Reduction and Acoustic Feedback Cancellation“. In: *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Singapur, Singapur: IEEE, 2022, S. 676–680.

[3] S. Ruiz, T. van Waterschoot und M. Moonen. „Centralized Cascade Multi-Channel Noise Reduction and Acoustic Feedback Cancellation in a Wireless Acoustic Sensor and Actuator Network“. In: *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Rhodes Island, Griechenland: IEEE, Juni 2023, S. 1–5.

[4] Stefan Weinzierl. „Grundlagen“. In: *Handbuch der Audiotechnik*. Hrsg. von Stefan Weinzierl. Springer Berlin, Heidelberg, 2009, S. 1–39.

[5] The Vintage Radio & Phonograph Society. „Sound Waves VRPS Fall 2017“. In: *Sound Waves* (2017).

[6] amplifon. *Die Geschichte der Hörgeräte*. 2026. URL: https://www.amplifon.com/de/hoergeraete/technologie/wie-funktioniert-ein-hoergeraet/die-geschichte-der-hoergeraete.

[7] Shure. *MXA920 Decken-Mikrofonarray*. 2026. URL: https://www.shure.com/de-DE/produkte/mikrofone/mxa920?variant=MXA920W-S-60CM.

[8] Sennheiser. *TeamConnect Ceiling 2 Black*. 2026. URL: https://www.sennheiser.com/de-de/catalog/products/meeting-und-konferenzsysteme/teamconnect-ceiling-2/teamconnect-ceiling-2-black-509161.

[9] Toon Van Waterschoot und Marc Moonen. „Fifty Years of Acoustic Feedback Control: State of the Art and Future Challenges“. In: *Proceedings of the IEEE* 99.2 (2011), S. 288–327.

[10] H. Nyquist. „Regeneration Theory“. In: *Bell System Technical Journal* 11.1 (Jan. 1932), S. 126–147.

[11] Peter Mapp und Christien Ellis. „Improvements in Acoustic Feedback Margin in Sound Reinforcement Systems“. In: *Journal of the Audio Engineering Society*. 4978. AES Convention 106, Mai 1999.

[12] M. R. Schroeder. „Improvement of acoustic feedback stability in public address systems“. In: *Proc. 3rd Int. Congr. Acoust.* Stuttgart, Deutschland, 1959, S. 771–775.

[13] M. R. Schroeder. „Stop feedback in public address systems“. In: *Radio Electron.* 31 (Feb. 1960), S. 40–42.

[14] M. R. Schroeder. „Improvement of feedback stability of public address systems by frequency shifting“. In: *J. Audio Eng. Soc.* 10.2 (Apr. 1962), S. 108–109.

[15] M. R. Schroeder. „Improvement of acoustic-feedback stability by frequency shifting“. In: *J. Acoust. Soc. Amer.* 36.9 (Sep. 1964), S. 1718–1724.

[16] P. U. Svensson. „Computer simulations of periodically time-varying filters for acoustic feedback control“. In: *J. Audio Eng. Soc.* 43.9 (Sep. 1995), S. 667.

[17] J. L. Nielsen und P. U. Svensson. „Performance of some linear time-varying systems in control of acoustic feedback“. In: *J. Acoust. Soc. Amer.* 106.1 (Juli 1999), S. 240–254.

[18] P. U. Svensson. „On reverberation enhancement in auditoria“. Diss. Chalmers University of Technology, Göteborg, Sweden, Nov. 1994.

[19] A. J. Prestigiacomo und D. J. MacLean. „A frequency shifter for improving acoustic feedback stability“. In: *J. Audio Eng. Soc.* 10.2 (Apr. 1962), S. 110–113.

[20] M. D. Burkhard. „A simplified frequency shifter for improving acoustic feedback stability“. In: *J. Audio Eng. Soc.* 11.3 (Juli 1963), S. 234–237.

[21] S. Wardle. „A hilbert-transform frequency shifter for audio“. In: *Proc. Workshop on Dig. Audio Eff. at DAFx.* Barcelona, Spanien, 1998, S. 25–29.

[22] C. V. Deutschbein. „Digital frequency shifting for electroacoustic feedback suppression“. In: *Audio Engineering Society Convention 118.* Barcelona, Spanien, Mai 2005.

[23] J. Alisobhani und S. G. Knorr. „Improvement of acoustic-feedback stability by bandwidth compression“. In: *IEEE Trans. Acoust., Speech, Signal Process.* ASSP-28.6 (1980), S. 636–644.

[24] R. W. Guelke und A. D. Broadhurst. „Reverberation time control by direct feedback“. In: *Acustica* 24 (1971), S. 33–41.

[25] L. N. Mishin. „A method for increasing the stability of sound amplification systems“. In: *Akust. Z.* 4 (Jan. 1971), S. 64–72.

[26] G. Nishinomiya. „Improvement of acoustic feedback stability of public address system by warbling“. In: *Proc. 6th Int. Congr. Acoust.* Bd. E. Tokyo, Japan, 1968, S. 93–96.

[27] Guido Walz. *Lexikon der Mathematik: Band 1.* Springer Spektrum Berlin, Heidelberg, 2017, S. 200–201.

[28] M. A. Poletti. „The stability of multichannel sound systems with frequency shifting“. In: *J. Acoust. Soc. Amer.* 116.2 (Aug. 2004), S. 853–871.

[29] G. Rombouts, T. van Waterschoot und M. Moonen. „Proactive notch filtering for acoustic feedback cancellation“. In: *Proc. 2nd Annu. IEEE Benelux/DSP Valley Signal Process. Symp.* Antwerp, Belgien, März 2006, S. 169–172.

[30] S. Ibaraki, H. Furukawa und H. Naono. „Pre-howling howlback detection method“. In: *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.* Tokyo, Japan, Apr. 1986, S. 941–944.

[31] N. Osmanovic, V. E. Clarke und E. Velandia. „An in-flight low latency acoustic feedback cancellation algorithm“. In: *Preprints AES 123rd Conv.* New York, Okt. 2007.

[32] T. van Waterschoot und M. Moonen. „Comparative evaluation of howling detection criteria in notch-filter-based howling suppression“. In: *J. Audio Eng. Soc.* 58.11 (Nov. 2010), S. 923–940.

[33] J. Flocon-Cholet, J. Faure, A. Guerin und P. Scalart. „A robust howling detection algorithm based on a statistical approach“. In: *Proc. Int. Workshop Acoust. Signal Enhance. IWAENC*. Juan-les-Pins, Frankreich, Sep. 2014, S. 65–69.

[34] E. T. Patronis, Jr. „Electronic detection of acoustic feedback and automatic sound system gain control“. In: *J. Audio Eng. Soc.* 26.5 (Mai 1978), S. 323–326.

[35] S. Ando. „Howling detection and prevention circuit and a loudspeaker system employing the same“. US-Pat. 6 252 969. Juni 2001.

[36] J. B. Foley. „Adaptive periodic noise cancellation for the control of acoustic howling“. In: *Proc. IEE Colloq. Adaptive Filters*. London, U.K., März 1989, S. 7/1–7/4.

[37] W. Staudacher. „Acoustic feedback cancellation for equalized amplifying systems“. US-Pat. 5 533 120. Juli 1996.

[38] D. K. Bustamante, T. L. Worrall und M. J. Williamson. „Measurement and adaptive suppression of acoustic feedback in hearing aids“. In: *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP*. Glasgow, Schottland, Mai 1989.

[39] S. M. Kuo und J. Chen. „New adaptive IIR notch filter and its application to howling control in speakerphone system“. In: *IEE Electron. Lett.* 28.8 (Apr. 1992), S. 764–766.

[40] W. Loetwassana, R. Punchalard, A. Lorsawatsiri, J. Koseeyaporn und P. Wardkein. „Adaptive howling suppressor in an audio amplifier system“. In: *Proc. Asia-Pacific Conf. Commun.* Bangkok, Thailand, Okt. 2007, S. 445–448.

[41] P. Gil-Cacho, T. van Waterschoot, M. Moonen und S. H. Jensen. „Regularized adaptive notch filters for acoustic howling suppression“. In: *Proc. 17th Eur. Signal Process. Conf.* Glasgow, Schottland, Aug. 2009, S. 2574–2578.

[42] J. G. Proakis und D. G. Manolakis. „Structures for IIR Systems“. In: *Third Edition Digital Signal Processing. Principles, Algorithms and Applications*. Prentice Hall, New Jersey, 1996, S. 519–539.

[43] H.R. Schwarz und N. Köckler. „Ausgleichsprobleme, Methode der kleinsten Quadrate“. In: *Numerische Mathematik*. 7. Aufl. Vieweg+Teubner Verlag Wiesbaden, 2009, S. 274–306.

[44] M. P. Lewis, T. J. Tucker und D. M. Oster. „Method and apparatus for adaptive audio resonant frequency filtering“. US-Pat. 5 245 665. Sep. 1993.

[45] M. H. Er, T. H. Ooi, L. S. Li und C. J. Liew. „A DSP-based acoustic feedback canceller for public address systems“. In: *Proc. Int. Conf. Signal Process.* Beijing, China, Okt. 1993, S. 1251–1254.

[46] R. Abe. „Howling suppression device and howling suppression method“. US-Pat. 7 295 670. Nov. 2007.

[47] D. Somasundaram. „Feedback cancellation in a sound system“. Europ. Pat. 1 903 833 A1. März 2008.

[48] T. K. Duong, E. Lefort und M. G. Bellanger. „Acoustic feedback cancelling electro-acoustic transducer network“. US-Pat. 4 485 272. Nov. 1984.

[49] C. P. Janse und H. J. W. Belt. „Sound reinforcement system having an echo suppressor and loudspeaker beamformer“. WIPO-Pat. WO/2003/010 996. Feb. 2003.

[50] K. Kobayashi, K. Furuya und A. Kataoka. „An adaptive microphone array for howling cancellation“. In: *Acoust. Sci. Technol.* 24.1 (Jan. 2003), S. 45–47.

[51] G. Rombouts und M. Spriet A. und Moonen. „Generalized sidelobe canceller based combined acoustic feedback- and noise cancellation“. In: *Signal Process.* 88.3 (März 2008), S. 571–581.

[52] H. Schepker, L. T. T. Tran, S. Nordholm und S. Doclo. „Acoustic feedback cancellation for a multi-microphone earpiece based on a null-steering beamformer“. In: *2016 IEEE International Workshop on Acoustic Signal Enhancement (IWAENC)*. 2016.

[53] H. Schepker, L. T. T. Tran, S. Nordholm und S. Doclo. „A Robust Null-Steering Beamformer for Acoustic Feedback Cancellation for a Multi-Microphone Earpiece“. In: *Speech Communication; 12. ITG Symposium*. Paderborn, Deutschland, Okt. 2016, S. 165–169.

[54] H. Schepker, L. T. T. Tran, S. Nordholm und S. Doclo. „Null-steering beamformer for acoustic feedback cancellation in a multi-microphone earpiece optimizing the maximum stable gain“. In: *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP*. New Orleans, USA, 2017, S. 341–345.

[55] H. Schepker, L. T. T. Tran, S. Nordholm und S. Doclo. „Combining null-steering and adaptive filtering for acoustic feedback cancellation in a multi-microphone earpiece“. In: *Proc. European Signal Processing Conference EUSIPCO*. Kos Island, Griechenland, Aug. 2017, S. 241–245.

[56] S. Ushiyama, T. Hirai, M. Tohyama und Y. Shimizu. „Howling suppression by smoothing the open-loop transfer function“. In: *IEICE Tech. Rep.* 94.20 (Apr. 1994), S. 23–28.

[57] Y. Nagata, S. Suzuki, M. Yamada, M. Yoshida, M. Kitano, K. Kuroiwa und S. Kimura. „Howling remover composed of adjustable equalizers for attenuating complicated noise peaks“. US-Pat. 5 729 614. März 1998.

[58] G. Schmidt und T. Haulick. „Signal processing for in-car communication systems“. In: *Signal Process.* 86, Special Issue on Applied Speech and Audio Processing.6 (Juni 2006), S. 1307–1326.

[59] G. Enzner und P. Vary. „Frequency-domain adaptive Kalman filter for acoustic echo control in hands-free telephones“. In: *Signal Processing* 86.6 (Juni 2006), S. 1140–1156.

[60] T. van Waterschoot, K. Eneman und M. Moonen. *Instrumental Variable Methods for Acoustic Feedback Cancellation*. Techn. Ber. ESAT-SISTA TR 05-14. Katholieke Universiteit Leuven, Belgium, Okt. 2004. URL: `ftp://ftp.esat.kuleuven.be/pub/sista/vanwaterschoot/abstracts/05-14.html`.

[61] M. G. Siqueira und A. Alwan. „Steady-state analysis of continuous adaptation in acoustic feedback reduction systems for hearing-aids“. In: *IEEE Trans. Speech Audio Process.* 8.4 (Juli 2000), S. 443–453.

[62] J. Hellgren und U. Forssell. „Bias of feedback cancellation algorithms in hearing aids based on direct closed loop identification“. In: *IEEE Trans. Speech Audio Process.* 9.7 (Nov. 2001), S. 906–913.

[63] T. van Waterschoot, G. Rombouts und M. Moonen. „On the performance of decorrelation by prefiltering for adaptive feedback cancellation in public address systems“. In: *Proc. 4th IEEE Benelux Signal Process. Symp.* Hilvarenbeek, Niederlande, Apr. 2004, S. 167–170.

[64] J. M. Kates. „Feedback cancellation in hearing aids: Results from a computer simulation“. In: *IEEE Trans. Signal Process.* 39.3 (März 1991), S. 553–562.

[65] Y.-C. Park, D.-W. Kim und Y.-I. Kim. „An efficient feedback cancellation for multiband compression in hearing aids“. In: *Proc. Ann. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBS)*. Hong Kong, China, Nov. 1998, S. 2706–2709.

[66] M. Guo, T. B. Elmedyb, S. H. Jensen und J. Jensen. „On acoustic feedback cancellation using probe noise in multiple-microphone and single-loudspeaker systems“. In: *IEEE Signal Process. Lett.* 19.5 (Mai 2012), S. 283–286.

[67] M. T. Akhtar und A. Nishihara. „Automatic tuning of probe noise for continuous acoustic feedback cancellation“. In: *Proc. Europ. Signal Process. Conf. (EUSIPCO)*. Budapest, Ungarn, Sep. 2016, S. 888–892.

[68] C. R. C. Nakagawa, S. Nordholm und W.-Y. Yan. „New insights into optimal acoustic feedback cancellation“. In: *IEEE Signal Process. Lett.* 20.9 (Sep. 2013), S. 869–872.

[69] S. Ibaraki, H. Furukawa und H. Naono. „Howling canceller“. US-Pat. 4 747 132. Mai 1988.

[70] Kashyap Patel und Issa M. S. Panahi. „Efficient Real-Time Acoustic Feedback Cancellation using Adaptive Noise Injection Algorithm“. In: *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. Montreal, QC, Kanada: IEEE, Juli 2020, S. 972–975.

[71] A. Goertz. „An adaptive subtraction filter for feedback cancellation in public address sound systems“. In: *Proc. 15th Int. Congr. Acoust.* Trondheim, Norwegen, Juni 1995, S. 69–72.

[72] J. H. Stott und N. Wells. „Method and apparatus for reduction of unwanted feedback“. US-Pat. 6 269 165. Juli 2001.

[73] Muhammad Tahir Akhtar, Felix Albu und Akinori Nishihara. „Acoustic feedback cancellation in hearing aids using dual adaptive filtering and gain-controlled probe signal“. In: *Biomedical Signal Processing and Control* 52 (2019), S. 1–13.

[74] C. P. Janse und C. C. Tchang. „Acoustic feedback suppression“. WIPO-Pat. WO/2005/079 109. Aug. 2005.

[75] G. Ma, F. Gran, F. Jacobsen und F. T. Agerkvist. „Adaptive feedback cancellation with band-limited lpc vocoder in digital hearing aids“. In: *IEEE Trans. Audio, Speech, Lang. Process.* 19.4 (Mai 2011), S. 677–687.

[76] A. Anand und R. Bhatia. „Performance evaluation of band-limited lpc vocoder and band-limited relp vocoder in adaptive feedback cancellation“. In: *Proc. Int. Conf. Adv. Comput. Comm. Inf. (ICACCI)*. Kochi, Indien, Aug. 2015, S. 126–132.

[77] M. Guo, S. H. Jensen und J. Jensen. „An Improved Probe Noise Approach for Acoustic Feedback Cancellation“. In: *Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*. Hoboken, USA, 2012.

[78] Meng Guo, Søren Holdt Jensen und Jesper Jensen. „Novel Acoustic Feedback Cancellation Approaches in Hearing Aid Applications Using Probe Noise and Probe Noise Enhancement“. In: *IEEE Transactions on Audio, Speech, and Language Processing* 20.9 (Nov. 2012), S. 2549–2563.

[79] C. R. C. Nakagawa, S. Nordholm und W.-Y. Yan. „Feedback Cancellation with Probe Shaping Compensation“. In: *IEEE Signal Processing Letters* 21.3 (März 2014), S. 365–369.

[80] C. P. Janse und P. A. A. Timmermans. *Signal Amplifier System with Improved Echo Cancellation*. U.S. Patent 5,748,751. Mai 1998.

[81] S. Kamerling, K. Janse und F. van der Meulen. „A New Way of Acoustic Feedback Suppression“. In: *Preprints of the 104th AES Convention*. AES Preprint 4735. Amsterdam, Niederlande, Mai 1998.

[82] M. Guo, S. H. Jensen, J. Jensen und S. L. Grant. „On the Use of Phase Modulation Method for Decorrelation in Acoustic Feedback Cancellation“. In: *Proceedings of the European Signal Processing Conference (EUSIPCO)*. Bucharest, Rumänien, Aug. 2012.

[83] M. Guo und B. Kuenzle. „On the Periodically Time-Varying Bias in Adaptive Feedback Cancellation Systems with Frequency Shifting“. In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. Shanghai, China, März 2016, S. 539–543.

[84] T. van Waterschoot und M. Moonen. „Assessing the Acoustic Feedback Control Performance of Adaptive Feedback Cancellation in Sound Reinforcement Systems“. In: *Proceedings of the European Signal Processing Conference (EUSIPCO)*. Glasgow, Schottland, Aug. 2009, S. 1997–2001.

[85] A. Chankawee und N. Tansangiumvisai. „On the Improvement of Acoustic Feedback Cancellation in Hearing-Aid Devices“. In: *Proceedings of the IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*. Hiroshima, Japan, Juli 2004, S. II-17–II-20.

[86] C. Boukis, D. P. Mandic und A. G. Constantinides. „Bias Reduction in Acoustic Feedback Cancellation Systems with Varying All-Pass Filters“. In: *Electronics Letters* 42.9 (Apr. 2006).

[87] C. Boukis, D. P. Mandic und A. G. Constantinides. „Towards Bias Minimization in Acoustic Feedback Cancellation Systems“. In: *Journal of the Acoustical Society of America* 121.3 (März 2007), S. 1529–1537.

[88] P. Estermann und A. Kaelin. „Feedback Cancellation in Hearing Aids: Results from Using Frequency-Domain Adaptive Filters“. In: *Proceedings of the IEEE International Symposium on Circuits and Systems*. Bd. 2. London, U.K., Mai 1994, S. 257–260.

[89] J. M. Kates. „Feedback Cancellation in Hearing Aids Using Constrained Adaptation“. In: *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*. New Paltz, USA, Okt. 1999, S. 231–234.

[90] J. Hellgren, T. Lunner und S. Arlinger. „Variations in the Feedback of Hearing Aids“. In: *Journal of the Acoustical Society of America* 106.5 (Nov. 1999), S. 2821–2833.

[91] M. R. Stinson und G. A. Daigle. „Effect of Handset Proximity on Hearing Aid Feedback“. In: *Journal of the Acoustical Society of America* 115.3 (März 2004), S. 1147–1156.

[92] H.-F. Chi, S. X. Gao, S. D. Soli und A. Alwan. „Band-Limited Feedback Cancellation with a Modified Filtered-X LMS Algorithm for Hearing Aids“. In: *Speech Communication* 39.1–2 (Jan. 2003), S. 147–161.

[93] F. Gallego, E. Lleida, E. Masgrau und A. Ortega. „Method and System for Suppressing Echoes and Noises in Environments under Variable Acoustic and Highly Feedback Conditions“. WIPO-Pat. WO/2002/101 728. Dez. 2002.

[94] A. Ortega, E. Lleida und E. Masgrau. „Speech Reinforcement System for Car Cabin Communications“. In: *IEEE Transactions on Speech and Audio Processing* 13.5 (Sep. 2005), S. 917–929.

[95] J. Hellgren. „Analysis of Feedback Cancellation in Hearing Aids with Filtered-X LMS and the Direct Method of Closed-Loop Identification“. In: *IEEE Transactions on Speech and Audio Processing* 10.2 (Feb. 2002), S. 119–131.

[96] A. Spriet, M. Moonen und I. Proudler. „Feedback Cancellation in Hearing Aids: An Unbiased Modelling Approach“. In: *Proceedings of the 11th European Signal Processing Conference (EUSIPCO)*. Toulouse, Frankreich, Sep. 2002, S. 531–534.

[97] G. Rombouts, T. van Waterschoot, K. Struyve und M. Moonen. „Acoustic Feedback Suppression for Long Acoustic Paths Using a Nonstationary Source Model“. In: *Proceedings of the 13th European Signal Processing Conference (EUSIPCO)*. Antalya, Türkei, Sep. 2005.

[98] S. Haykin. „Linear Prediction“. In: *Adaptive Filter Theory*. 4. Aufl. Upper Saddle River: Prentice Hall, 2002. Kap. 3.

[99] S. Haykin. „Stochastic Models“. In: *Adaptive Filter Theory*. 4. Aufl. Upper Saddle River: Prentice Hall, 2002. Kap. 1.5.

[100] A. Spriet, I. Proudler, M. Moonen und J. Wouters. „Adaptive Feedback Cancellation in Hearing Aids with Linear Prediction of the Desired Signal“. In: *IEEE Transactions on Signal Processing* 53.10 (2005), S. 3749–3763.

[101] Toon Van Waterschoot und Marc Moonen. „Comparison of Linear Prediction Models for Audio Signals“. In: *EURASIP Journal on Audio, Speech, and Music Processing* 2008 (2008), S. 1–24.

[102] G. Rombouts, T. van Waterschoot, K. Struyve und M. Moonen. „Acoustic Feedback Cancellation for Long Acoustic Paths Using a Nonstationary Source Model“. In: *IEEE Transactions on Signal Processing* 54.9 (Sep. 2006), S. 3426–3434.

[103] T. van Waterschoot, G. Rombouts und M. Moonen. „MSE Optimal Regularization of APA and NLMS Algorithms in Room Acoustic Applications“. In: *Proceedings of the International Workshop on Acoustic Echo and Noise Control (IWAENC)*. Paris, Frankreich, Sep. 2006.

[104] T. van Waterschoot, G. Rombouts und M. Moonen. „Dually Regularized Recursive Prediction Error Identification for Acoustic Feedback and Echo Cancellation“. In: *Proceedings of the 15th European Signal Processing Conference (EUSIPCO)*. Poznan, Polen, Sep. 2007, S. 1610–1614.

[105] T. van Waterschoot, G. Rombouts und M. Moonen. „Optimally Regularized Adaptive Filtering Algorithms for Room Acoustic Signal Enhancement“. In: *Signal Processing* 88.3 (März 2008), S. 594–611.

[106] K. Lee, Y.-H. Baik, Y. Park, D. Kim und J. Sohn. „Robust Adaptive Feedback Canceller Based on Modified Pseudo Affine Projection Algorithm“. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*. Boston, USA, Aug. 2011, S. 3760–3763.

[107] J. M. Gil-Cacho, T. van Waterschoot und M. Moonen. „Wiener Variable Step Size and Gradient Spectral Variance Smoothing for Double-Talk-Robust Acoustic Echo Cancellation and Acoustic Feedback Cancellation“. In: *Signal Processing* 104 (Nov. 2014), S. 1–14.

[108] L. T. T. Tran, H. Schepker, S. Doclo, H. H. Dam und S. Nordholm. „Proportionate NLMS for Adaptive Feedback Cancellation in Hearing Aids“. In: *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. New Orleans, USA, 2017, S. 211–215.

[109] A. Spriet, G. Rombouts, M. Moonen und J. Wouters. „Adaptive Feedback Cancellation in Hearing Aids“. In: *Journal of the Franklin Institute* 343.6 (Aug. 2006), S. 545–573.

[110] G. Rombouts, T. van Waterschoot und M. Moonen. „Robust and Efficient Implementation of the PEM-AFROW Algorithm for Acoustic Feedback Cancellation“. In: *Journal of the Audio Engineering Society* 55.11 (Nov. 2007), S. 955–966.

[111] A. Spriet, S. Doclo, M. Moonen und J. Wouters. „Feedback Control in Hearing Aids“. In: *Springer Handbook of Speech Processing*. Berlin, Heidelberg: Springer, 2008, S. 979–1000.

[112] J. M. Gil-Cacho, T. van Waterschoot, M. Moonen und S. H. Jensen. „A Frequency-Domain Adaptive Filter (FDAF) Prediction Error Method (PEM) Framework for Double-Talk-Robust Acoustic Echo Cancellation“. In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 22.12 (Dez. 2014), S. 2074–2086.

[113] Giuliano Bernardi, Toon Van Waterschoot, Jan Wouters und Marc Moonen. „An all-frequency-domain adaptive filter with PEM-based decorrelation for acoustic feedback control“. In: *2015 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*. New Paltz, NY, USA: IEEE, Okt. 2015, S. 1–5.

[114] G. Bernardi, T. van Waterschoot, J. Wouters, M. Hillbmtt und M. Moonen. „A PEM-Based Frequency-Domain Kalman Filter for Adaptive Feedback Cancellation“. In: *Proceedings of the 23rd European Signal Processing Conference (EUSIPCO)*. IEEE, Aug. 2015, S. 270–274.

[115] G. Bernardi, T. van Waterschoot, J. Wouters und M. Moonen. „Adaptive Feedback Cancellation Using a Partitioned-Block Frequency-Domain Kalman Filter Approach with PEM-Based Signal Prewhitening“. In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 25.9 (Sep. 2017), S. 1784–1798.

[116] P. Chaisakul, N. Tangsangiumvisai, P. Luangpitakchumpon und A. Nishihara. „On the Use of Linear Prediction for Acoustic Feedback Cancellation in Multi-Band Hearing Aids“. In: *Proceedings of the European Signal Processing Conference (EUSIPCO)*. Florence, Italien, Sep. 2006.

[117] Falco Strasser und Henning Puder. „Adaptive Feedback Cancellation for Realistic Hearing Aid Applications“. In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 23.12 (Dez. 2015), S. 2322–2333.

[118] J. Gil-Cacho, T. van Waterschoot, M. Moonen und S. H. Jensen. „Transform Domain Prediction Error Method for Improved Acoustic Echo and Feedback Cancellation“. In: *Proceedings of the European Signal Processing Conference (EUSIPCO)*. Bucharest, Rumänien, Aug. 2012, S. 2422–2426.

[119] T. van Waterschoot und M. Moonen. „Adaptive Feedback Cancellation for Audio Applications“. In: *Signal Processing* 89.11 (Nov. 2009), S. 2185–2201.

[120] G. Ma, F. Gran, F. Jacobsen und F. T. Agerkvist. „Noise Injection for Feedback Cancellation with Linear Prediction“. In: *Proceedings of the International Congress on Acoustics (ICA)*. Sydney, Australien, Aug. 2010.

[121] A. Bastari, S. Squartini und F. Piazza. „Joint Acoustic Feedback Cancellation and Noise Reduction Within the Prediction Error Method Framework“. In: *Proceedings of the International Workshop on Hands-Free Communication and Microphone Arrays (HSCMA)*. Trento, Italien, Mai 2008, S. 228–231.

[122] S. Cifani, R. Rotili, E. Principi, S. Squartini und F. Piazza. „Real-Time Implementation of Robust PEM-AFROW Based Solutions for Acoustic Feedback Control“. In: *Audio Engineering Society Convention 127*. New York, USA, Okt. 2009.

[123] T. van Waterschoot und M. Moonen. „Assessing the acoustic feedback control performance of adaptive feedback cancellation in sound reinforcement systems“. In: *Proc. 17th Eur. Signal Process. Conf.* Glasgow, Schottland, U.K., Aug. 2009, S. 1997–2001.

[124] Giuliano Bernardi, Toon Van Waterschoot, Jan Wouters und Marc Moonen. „Subjective and Objective Sound-Quality Evaluation of Adaptive Feedback Cancellation Algorithms“. In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 26.5 (2018), S. 1010–1024.

[125] Sven Nordholm, Henning Schepker, Linh T. T. Tran und Simon Doclo. „Stability-controlled hybrid adaptive feedback cancellation scheme for hearing aids“. In: *The Journal of the Acoustical Society of America* 143.1 (2018), S. 150–166.

[126] Cees H. Taal, Richard C. Hendriks, Richard Heusdens und Jesper Jensen. „An Algorithm for Intelligibility Prediction of Time–Frequency Weighted Noisy Speech“. In: *IEEE Transactions on Audio, Speech, and Language Processing* 19.7 (2011), S. 2125–2136.

[127] ITU-T Rec. P.862. *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*. Feb. 2001.

[128] A.W. Rix, J.G. Beerends, M.P. Hollier und A.P. Hekstra. „Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs“. In: *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.01CH37221)*. Bd. 2. Salt Lake City, UT, USA: IEEE, 2001, S. 749–752.

[129] Arne Leijon, Koen Eneman, Ann Spriet, Robert Aichner u. a. „D-5-1: Sub-set of signal enhancement techniques operational on PC system“. In: *Deliverable D.5.1 of the European project FP6004171 HEARCOM Hearing in the Communication Society* (2005).

[130] Ann Spriet, Marc Moonen und Jan Wouters. „Evaluation of feedback reduction techniques in hearing aids based on physical performance measures“. In: *The Journal of the Acoustical Society of America* 128.3 (1. Sep. 2010), S. 1245–1261.

[131] B. Moore. *An Introduction to the Psychology of Hearing*. 5. Aufl. London: Academic Press, 2003.

[132] S. Haykin. „Wiener Filters“. In: *Adaptive Filter Theory*. 4. Aufl. Upper Saddle River: Prentice Hall, 2002. Kap. 2.

[133] J. Agnew und J. M. Thornton. „Just Noticeable and Objectionable Group Delays in Digital Hearing Aids“. In: *Journal of the American Academy of Audiology* 11.6 (Juni 2000), S. 330–336.

[134] J. Benesty, J. Chen und Y. Huang. „Linear Prediction“. In: *Springer Handbook of Speech Processing*. Springer, 2007, S. 121–134.

[135] B. H. Nitsch. „A Frequency-Selective Stepfactor Control for an Adaptive Filter Algorithm Working in the Frequency Domain“. In: *Signal Processing* 80.9 (Sep. 2000), S. 1733–1745.

[136] J. G. Proakis und D. G. Manolakis. „The Discrete Fourier Transform: Its Properties and Applications“. In: *Third Edition Digital Signal Processing. Principles, Algorithms and Applications*. Prentice Hall, New Jersey, 1996, S. 394–447.

[137] J. J. Shynk. „Frequency-Domain and Multirate Adaptive Filtering“. In: *IEEE Signal Processing Magazine* 9.1 (Jan. 1992), S. 14–37.

[138] S. Kay. „Kalman Filters“. In: *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall PTR, 1993.

[139] G. Enzner. „Bayesian Inference Model for Applications of Time-Varying Acoustic System Identification“. In: *Proceedings of the 18th European Signal Processing Conference (EUSIPCO)*. Aug. 2010.

[140] J. Benesty, J. Chen und Y. Huang. *Microphone Array Signal Processing*. Bd. 1. Springer Topics in Signal Processing. Berlin, Heidelberg: Springer, 2008.

[141] M. Brandstein und D. Ward. *Microphone Arrays: Signal Processing Techniques and Applications*. Digital Signal Processing. Berlin, Heidelberg: Springer, 2001.

[142] D. Döbler und G. Heilmann. „Perspectives of the Acoustic Camera“. In: *Proceedings of Internoise 2005*. Paper No. 83. Rio de Janeiro, Brasilien, 2005.

[143] ITU-T. *Test Signals for Use in Telephony and Other Speech-Based Applications*. Recommendation P.501. International Telecommunication Union (ITU-T), Apr. 2025. URL: `https://www.itu.int/rec/T-REC-P.501-202504-I/en`.

A. Lebenslauf

Persönliche Daten

Name:	Manuel Ruben Schalk
Geburtsort:	Kaiserslautern

Berufserfahrung

2022–heute	**Wissenschaftlicher Mitarbeiter** RPTU Kaiserslautern-Landau Fachbereich Elektrotechnik und Informationstechnik, Lehrstuhl Digitale Signalverarbeitung
2018–2019, 2020–2022	**Hilfswissenschaftler** TU Kaiserslautern Fachbereich Elektrotechnik und Informationstechnik, Lehrstuhl Nachrichtentechnik, Lehrstuhl Digitale Signalverarbeitung
2018–2020	**Studentische Hilfskraft** **Praktikum im Team Medienregulierung** Medienanstalt Rheinland-Pfalz

Akademische Ausbildung

2020–2022	**Master of Science: Medien-und Kommunikationstechnik** **Ø1,6** TU Kaiserslautern Masterarbeit: *Akustische Bildgebung mit einem irregulären Mikrophonarray*
2015–2021	**Bachelor of Science: Medien-und Kommunikationstechnik** **Ø2,2** TU Kaiserslautern Bachelorarbeit: *Entwicklung und Implementierung eines Nachhalleffekts als VST-Plugin*

Schulausbildung

2006-2015	**Abitur Ø2,4** Hohenstaufen-Gymasium, Kaiserslautern Leistungskurse: Physik, Mathematik, Englisch.