Robust and fast schemes in broadband active noise and vibration control

.

Robust and fast schemes in broadband active noise and vibration control

Rufus Fraanje

Ph.D. Thesis • University of Twente • The Netherlands

Thesis committee:

Prof. dr. ir. A. Bliek (chairman)	University of Twente
Dr. ir. N.J. Doelman (assistent promotor)	TNO Institute of Applied Physics
Prof. dr. S.J. Elliott	University of Southampton
Prof. dr. ir. P. Sas	Katholieke Universiteit Leuven
Prof. dr. ir. A.J. van der Schaft	University of Twente
Prof. dr. ir. C.H. Slump	University of Twente
Prof. dr. ir. M.H.G. Verhaegen (promotor)	University of Twente

Cover: Interfering waves.

ROBUST AND FAST SCHEMES IN
BROADBAND ACTIVE NOISE AND
VIBRATION CONTROL

PROEFSCHRIFT

ter verkrijging van

de graad van doctor aan de Universiteit Twente,

op gezag van de rector magnificus,

prof. dr. F.A. van Vught,

volgens besluit van het College voor Promoties

in het openbaar te verdedigen

op donderdag 6 mei 2004 om 13.15 uur

door

Petrus Rufus Fraanje

geboren op 18 juli 1976

te Ten Boer

Dit proefschrift is goedgekeurd door de promotor prof. dr. ir. M.H.G. Verhaegen en de assistent promotor dr. ir. N.J. Doelman.

(...) when I said,
'This far you may come, but no farther,
and here your proud waves must stop!'

Job 38 verse 11

.

ACKNOWLEDGMENTS

Though just one name is printed on the cover, this thesis could have been written only with the help of many others. In the first place, I want to thank my advisor prof. Michel Verhaegen for his many ideas, stimulating advise and the trust in me during the entire project. Your constructive comments helped me a lot to get the results more precise and present them in a more logical order. In the second place, I want to thank my second advisor dr. Niek Doelman for the many interesting discussions on adaptive filter algorithms, for keeping me focused on practice but who was also convinced of the necessity of a solid theoretical basis.

Furthermore, I want to thank prof. Ali Sayed for showing hospitality during my stay at his Adaptive Systems Lab at UCLA, for the many instructive discussions on robust filtering and fast array algorithms, and very carefully commenting my work. Also many thanks to Ananth Subramanian for his warm welcome, for having lunch together and not at least his continuing patience in explaining many difficult concepts in robust estimation. Thanks to the students of the Graduate Christian Fellowship for making my stay at UCLA complete.

Thanks to dr. Arthur Berkhoff for enabling me to do experiments on the smart panel system of TNO Institute of Applied Physics, and for his stimulating enthusiasm. Thanks to dr. Tom Basten for enabling me to use the Hybrid Isolation Construction (HIC) demonstrator and to get it up and running.

Thanks to prof. Rik Pintelon for the discussions and comments on frequency domain state-space model identification of causal/anti-causal systems. Also thanks to dr. Vincent Verdult for discussions and comments on this work as well as on many other things ranging from identification to education. Thanks to dr. Xavier Bombois for his illuminating comments on control-relevant identification and controller validation.

It was a pleasure to me, to work together with the former MSc. students, Ruben Haverkort, Martijn Krutzen, Karel Hinnen, Sara van der Hoeven, Martijn de Boer and Bart Loffeld. You helped me to prevent shutting up in my own work, and your questions and comments kept me sharp.

Thanks to Gerard Nijsse for explaining many concepts in adaptive filtering and active control at the start of the project, which saved me a lot of time. Now, I miss the creative discussions we had at the white-board at daytime or even at night.

Thanks to my (former) roommates, Gerard Nijsse, dr. Stoyan Kanev, dr. Andreas Hegyi, Karel Hinnen and Sippe Douma, for a good working climate, sharing LaTeX and MATLAB tricks and the nice discussions from Lagrange multipliers till good breathing techniques. Thanks to dr. Ichiro Jikuja and to dr. Hiroshi and Hiroko Oku for sharing their knowledge on difficult control problems and recipes from the Japanese kitchen. Thanks to Andreas and Suk-Han for dining together and correcting my English and even my Dutch. I guess you both are quite *integrated*, the question is whether you should be proud on this.

GLOSSARY

List of symbols

$\otimes$	Kronecker matrix product
$(.)^\dagger$	pseudo-inverse operator
$(.)^{-1}$	inverse operator
$(.)^T$	transpose operator
$(.)^*$	complex conjugate transpose operator
$[.]_+$	causality operator
$[.]_-$	anti-causality operator
$\|.\|_2$	H_2 norm operator
$\|.\|_\infty$	H_∞ norm operator
$\widehat{(.)}$	estimate of (.)
$\arg(.)$	argument of (.)
$^{10}\log(.)$	logarithm operator with base 10
$\min(.)$	minimizing operator
$\operatorname{tr}(.)$	trace operator
$\operatorname{var}(.)$	variance operator
$\mathbf{E}(.)$	stochastic expectation operator with respect to stochastic signals
$\overline{\mathbf{E}}(.)$	stochastic expectation operator with respect to uncertain systems with random parameters
$VAF(y_m, y_s)$	Variance Accounted For between the signals y_m and y_s
$\mathcal{A}/\mathcal{B}$	the set of elements contained in $\mathcal{A}$ which are not contained in $\mathcal{B}$
$\mathbb{R}^{m_x}$	set of all m_x dimensional vectors with real coefficients
$\mathbb{R}^{m_x \times m_y}$	set of all $m_x \times m_y$ dimensional matrices with real coefficients
$\mathcal{RH}_p^{m_x \times m_y}$	set of all proper rational $m_x \times m_y$ transferfunction matrices with real coefficients
$\mathcal{RH}_\infty^{m_x \times m_y}$	set of all proper and stable rational $m_x \times m_y$ transferfunction matrices with real coefficients
0_{m_x}	zero column vector of size m_x
$0_{m_x \times m_y}$	zero matrix of size $m_x \times m_y$
δ_{kl}	Kronecker delta function
γ	stepsize variable (time-varying)
ϵ	innovation signal or prediction error
θ	parameter vector, in Chapter 7 θ refers to the secondary path state
λ	exponential forgetting factor
μ	stepsize variable (constant)
π	the circle constant $\pi = 3.1415\cdots$
ρ	positive real-valued scalar tuning parameter
ϕ	date-vector used in regression vector of adaptive algorithms
ω	angular frequency (in rad/s)
d	disturbance signal to be suppressed, also called primary disturbance signal
e	error signal, also called residual signal (also refers to the constant

	$e = 2.7182\cdots$)
f	frequency (in Hz)
f_s	sampling rate (in Hz)
i	counter variable
j	complex-unit ($j^2 = -1$) or counter variable
k	discrete-time index
l	counter variable
m_x	number of elements of signal x, $x(k) \in \mathbb{R}^{m_x}$
n_x	order of system X, but note in some cases n_w refers to the number of taps in the FIR filter $W(q^{-1})$, which will be indicated
p	counter variable
q, q^{-1}	unit shift forward, shift back operator respectively.
r	reference signal
r_{fb}	reference signal distorted by feedback of r_u
r_u	feedback signal distorting reference signal
s	disturbance signal generated by the disturbance source
u	control signal
v	measurement noise signal
w	arbitrary noise signal, defined in context
x	state
y	counter acting disturbance signal, also called secondary disturbance signal
Δa_i	uncertain real-valued scalar a_i
$\Delta A(q^{-1})$	uncertain polynomial with uncertain coefficients a_i
ΔB	uncertain input matrix of state-space model
ΔC	uncertain output matrix of state-space model
ΔD	uncertain direct feedthrough matrix of state-space model
$\Delta G(q^{-1})$	uncertain transfer function $G(q^{-1})$
$\widetilde{\Delta G}$	(co-)spectral factor of $\Phi_{\Delta G}$
$\widehat{\widetilde{\Delta G}}$	estimate of $\widetilde{\Delta G}$
$\Phi_{\Delta G}$	power spectrum of ΔG
Ψ	pseudo shift matrix (Chapter 7)
Π_0	initial state-error covariance matrix
$A(q^{-1})$	nummerator polynomial
$B(q^{-1})$	denominator polynomial
A	state transition matrix of state-space model
B_u	input-matrix for input u of state-space model
B_s	input-matrix for input s of state-space model
C_e	output-matrix for output e of state-space model
C_r	output-matrix for output r of state-space model
D_{es}	direct-feedthrough from s to e of state-space model
D_{eu}	direct-feedthrough from u to e of state-space model
D_{rs}	direct-feedthrough from s to r of state-space model
D_{rs}	direct-feedthrough from u to r of state-space model
C	controller
G	arbitrary plant or system which meaning is clear from the context
G^o	nominal or optimal G
G_i	inner factor of G, $G_i^* G_i = I$
G_o	outer factor of G, $G_o^* G_o = G^* G$
G_{ci}	co-inner factor of G, $G_{ci} G_{ci}^* = I$
G_{co}	co-outer factor of G, $G_{co} G_{co}^* = G G^*$

$G_i^\perp$	inner transfer function perpendicular to G_i such that $G_i^{\perp*}G_i^\perp = I$ and $G_i^{\perp*}G_i = 0$
$G_{ci}^\perp$	co-inner transfer function perpendicular to G_{ci} such that $G_{ci}^\perp G_{ci}^{\perp*} = I$ and $G_{ci}G_{ci}^{\perp*} = 0$
G^{aug}	plant or system G which inputs and/or outputs are augmented
G_{es}	primary path, transferfunction between s and e
G_{eu}	secondary path, transferfunction between u and e
G_{rs}	detector path, transferfunction between s and r
G_{ru}	intrinsic feedback, also called acoustical feedback, transferfunction between u and r
H	measurement noise shaping
H_{rv}	shaping of measurement noise v on reference signal r
H_{ev}	shaping of measurement noise v on error signal e
I_m	identity matrix of size $m \times m$, sometimes m is dropped
J	cost function to be minimized (in Chapter 7 J also referers to a signature matrix)
J_{rob}	robust cost function to be minized for robust control
K	Kalman gain
N	number of samples
P	state-error covariance matrix (in Chapter 7 P refers to *computed* state-error covariance)
Q	process noise covariance matrix
R	output noise covariance matrix
S	cross-covariance between process and output noise
W	feedforward controller
Z	shift matrix

List of abbreviations

dB	decibel
rad	radians
sec	seconds
AC	Active Control
AD	Analogue to Digital (converter)
ANC	Active Noise Control
ANVC	Active Noise and Vibration Control
ASAC	Active Structural Acoustic Control
AVC	Active Vibration Control
DA	Digital to Analogue (converter)
DFT	Discrete Fourier Transform
DSP	Digital Signal Processor
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FRF	Frequency Response Function
FTF	Fast Transversal Filter
FuLMS	Filtered-U LMS
FxLMS	Filtered-X LMS
Hz	Hertz
IIR	Infinite Impulse Response
IMC	Internal Model Control
IN	Input-Normal (parameterization)
LMS	Least Means Squares
LQG	Linear Quadratic Gaussian
MIMO	Multiple Inputs Multiple Outputs
MOESP	Multivariable Output Error State sPace
MSE	Mean Squared Error
ODE	Ordinary Differential Equation
OE	Output Error
ON	Output-Normal (parameterization)
PDE	Partial Differential Equation
PID	Proportional Derivative Integral
PEM	Prediction Error Model
PFxLMS	Preconditioned FxLMS
PLR	Pseudo Linear Regression
PO-MOESP	Past Outputs MOESP
RFxLMS	Robust FxLMS
RLS	Recursice Least Squares
RPFxLMS	Robust PFxLMS
RSIANC	Robust SIANC
SER	Signal to Error Ratio (in dB)
SIANC	Subspace Identification for ANC
SISO	Single Input Single Output
SLICOT	Subroutine Library in Systems and Control Theory
SMI	Subspace Model Identification
SNR	Signal to Noise Ratio (in dB)
SPR	Strictly Positive Real
SVD	Singular Value Decomposition
VAF	Variance Accounted For

CONTENTS

CHAPTER 1

INTRODUCTION

This introductory chapter explains the basic principles of active noise control (ANC) and active vibration control (AVC). From the elementary outline, the key elements of the research are motivated, the control problem considered in this thesis and basic assumptions are described and an overview of the literature is given. At the end of the chapter, a summary of the personal contributions in this thesis and an outline of the remaining chapters is sketched.

1.1 Principle of active noise and vibration control

1.1.1 The need for noise and vibration reduction

Noise disturbance is a major threat to (mental) health as reported by the World Health Organization in [7,192]. Some important fields where the reduction of noise disturbance is important are:

- Take-off and landing of airplanes (e.g., [133]);

- Highways and railways close to urban and natural environments;

- Engine and airflow noise in interior of airplanes, cars, trains, etc.;

- Engine noise in crew/passenger cabins on ships;

- Industrial machines (e.g. metal/wood cutting, drilling);

- Military aircraft, vehicles, launching of projectiles;

- Air conditioning systems;

- Electronic transformers.

Often, noise disturbances are generated by vibrations in mechanical structures. In principle they are just disturbing waves propagating through different media. In our terminology we use *noise disturbance* for undesired sound, i.e. a disturbing wave propagating through air, and *vibration disturbance* for the undesired vibration of mechanical structures, i.e. a disturbing wave propagating through materials.

1

Vibration disturbance is not only a source for noise disturbances, but also complicates accurate positioning or damaging of materials. Some important sources for vibration disturbances are:

- External vibrations distorting positioning of (optical) precision instruments and lightweight manipulators;

- Vibrations in mechanical structures resulting in damage and fatigue, such as in (railway) bridges;

- Earthquake generated vibrations acting on civil structures (e.g. buildings and bridges).

The methods for attenuating noise and/or vibration disturbances may be classified into *passive* and *active* approaches. The principles of both approaches are sketched in the following subsections.

1.1.2   Passive noise and vibration control

The classical solution to reduce noise disturbances is for example by adding isolation or damping material, cf. Figure 1.1(a). Often, vibration disturbances are reduced by using damped springs and (extremely) heavy and stiff mechanical design. These approaches are called *passive* since no additional energy is injected, and the suppression is by means of reflection or by increasing the dissipation of the energy contained in the disturbing waves. The advantage of these passive approaches is that high frequency disturbances (above $\approx$1kHz) can be effectively suppressed. However, to isolate low frequency disturbances noise (below $\approx$1kHz) the thickness of the isolation material should be large, since the thickness of the isolation material is related to the wave length of the disturbance wave(s). Similarly, to isolate low frequency disturbing vibrations heavy and stiff isolation materials are needed. In many situations, adding big and heavy isolation material is undesirable, because of increased fuel consumption, reduced available space, reduced mobility (due to increase of weight), etc.

1.1.3   Active noise and vibration control

A fundamentally different approach is to reduce the disturbing noise or vibration by interference with (approximately) the same wave up to 180° phase shift (i.e. its negative), cf. Figure 1.1(b). An active noise or vibration suppression system consists of basically three components, cf. Figure 1.2. 1) Sensors which measure the disturbance or an *upstream* reference signal measured near the disturbance source. 2) Actuators which inject an anti-disturbance, or also called *secondary* disturbance, in the system, such that interference of the disturbance and the secondary disturbance yield a lower disturbance level at the so-called 'region of silence'. 3) An electronic device, the controller, to determine the control signal for the actuators producing the secondary disturbance given the measurements from the sensors, measuring the *upstream* reference and the residual disturbance at the region of silence. This approach is called *active* since, contrary to the passive approaches, energy is injected in the system.
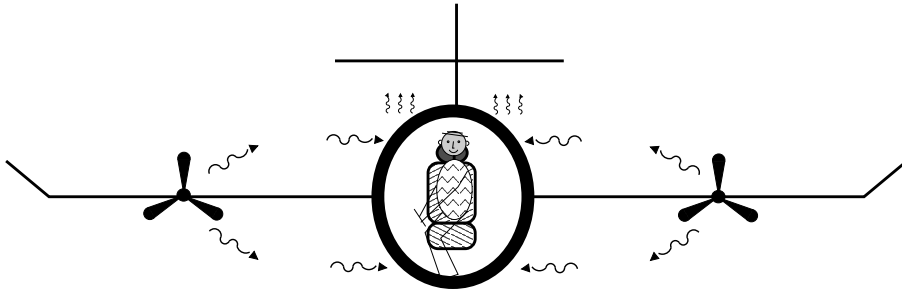
(a) Passive approach by isolation in an airplane.
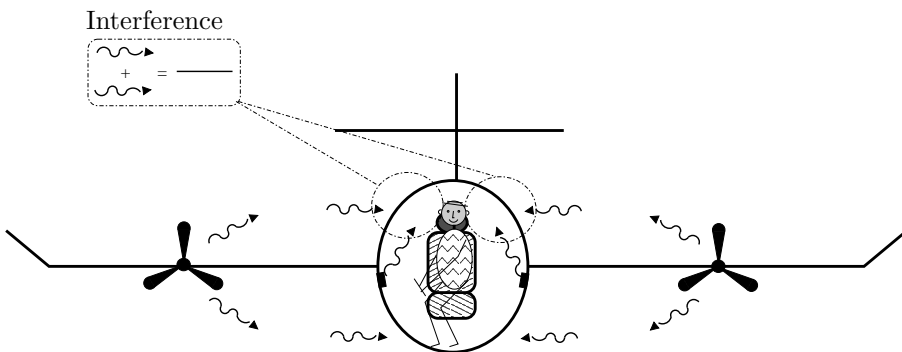
(b) Active approach by interference.

Figure 1.1: Passive and active suppression of engine noise in an airplane.

The active approach originates from the patents of Coanda in 1930 [28], Paul Lueg in 1936 [108] and Olson in 1953 [131]. Though the principle was correct, the practical implementation was naive and did not yield satisfying results [95]. In the sixties and seventies of the previous century, the digital signal processor (DSP) was developed, which yielded much more freedom for the control algorithm. Application of the DSP leaded to a breakthrough in active noise and vibration control (ANVC). Many researchers have been active in this field, and overviews of the results are provided in a number of textbooks, e.g., [46, 64, 90, 126, 140, 173].

The advantage of the active approach is that especially low frequency disturbances (below ≈1kHz) can be effectively suppressed without adding isolation material. In this sense, the active approach is an addition for passive approaches to improve disturbance suppression in the low frequency range. To actively suppress also high frequencies, not only the sampling rate needs to be increased to satisfy the Nyquist sampling criterion, see, e.g., [93], but also the number of resonance modes in the system will increase significantly. For example, in 3D acoustic enclosures the number of modes increases cubically in frequency, see [126]. This

Region of silence

PSfrag replacements

Tacho meter

DSP
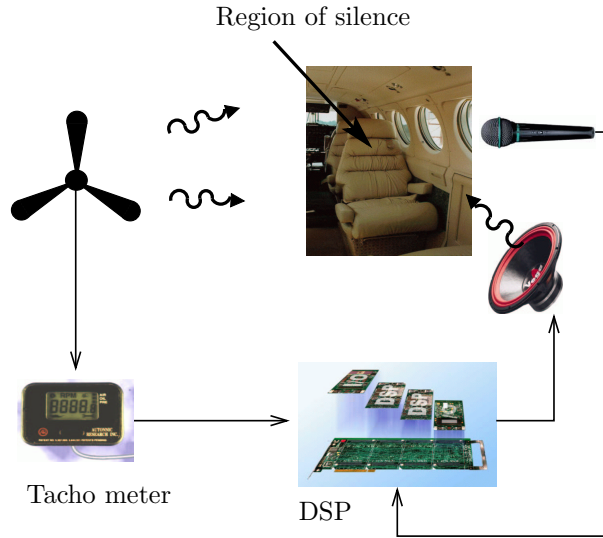
Figure 1.2: Basic setup of an active noise/vibration control system illustrated on an airplane, the tacho meter measures the rotation velocity of the engine, the DSP controls the loudspeaker to generate the anti-noise, the microphone measures the residual noise. Usually arrays of multiple loudspeakers and microphones are used in airplane applications.

increase of the number of modes will complicate the controller design. In general, to suppress higher frequencies also more sensors and actuators are needed, since the wavelength of the disturbance is decreased. With the current status of digital computation power, sampling rates between 1kHz and 10kHz are feasible for most active noise and vibration control (ANVC) problems.

Disturbing noise is often generated by radiation from vibrating mechanical structures. By reducing the vibration in the structure by means of active control, also the radiated noise will be suppressed. However, particular vibration modes in the mechanical structure are radiating the noise more efficiently than others, such that more weight has to be given to suppress these efficiently radiating modes. The active control of the vibration with the objective to reduce the power of the radiated noise, is known as Active Structural Acoustic Control (ASAC). The ASAC problem is currently extensively studied in the literature, see [8, 35, 66, 121, 161].

Usually, active control is applied in addition to passive control of noise and/or vibration. The integral design of joint active and passive methods is called *hybrid* controller design. The main question in the research on *hybrid* control design is to effectively combine active and passive methods, and has also been studied within the scope of the Knowledge Center "Sound and Vibration" of TNO Institute of Applied Physics and the University of Twente, see, e.g., [6].

Recently, active control methods are also applied in adaptive optics to counteract disturbances in the wavefront of light beams in astronomical observations. The actuator is usually a deformable mirror. For examples, we refer to the MSc.

theses [32, 88] which were also part of the subprogram Robust Active CONtrol (RACON) within the scope of the Knowledge Center "Sound and Vibration".

1.2 Motivation

Though ANVC is a successful method to suppress low frequency disturbances, it has (still) some drawbacks:

- Necessity of equipment: sensors, actuators, a DSP, power-supplies and cabling.

- Lack of robustness: most ANVC algorithms are using a model of the (acoustical or mechanical) plant. These models are always contained with model errors, e.g. caused by variations in the dynamics of the plant, and may yield degraded performance or even instability.

- Slow convergence: in most ANVC systems adaptive algorithms are used which are slowly converging in time, especially for broadband stochastic disturbances in multiple channel systems.

- Computational complexity: acoustical or mechanical systems are infinite dimensional systems and often relatively weakly damped, hence modeling these systems with finite dimensional models yields models with high orders (order $\approx 20 - 100$) and long impulse responses ($\approx 100 - 1000$ taps). Since most control algorithms are using these complex models the computational complexity is high.

- Tracking performance: most active control algorithms yield poor performance for non-stationary disturbances.

The first drawback is inherent to the active approach: instead of adding isolation or damping material, adding electronic equipment with sensors, actuators, a DSP, power supplies and cabling is necessary. One objective in the design of active control systems can be to reduce the amount of this equipment by optimizing sensor and actuator positions (see, e.g., [34]) and/or using instead of DSP's simple electronic devices (e.g., just single channel analogue integrators and amplifiers as in [66]). For example in applications to actively reduce noise disturbance in airplanes (see, e.g., [81]) and cars (see, e.g., [146]) reduction of the amount of equipment is important to reduce the financial costs and the complexity of the control problem. In this thesis, we will assume that the sensor and actuator positions are already optimized and no special constraints on the controller implementation are imposed (apart from that it should be possible to implement the control algorithm on nowadays DSP's).

Robustness should be taken into account in every active controller design. In normal operation the ANVC system should operate continuously within a period of some hours (a working day) till some days or even longer. In this period the dynamics of the acoustical or mechanical plant may change significantly, due to, e.g., moving of objects which changes the reflection in acoustical systems, mass

load variations which change vibration dynamics, temperature and humidity variations. These variations may yield undesired large performance reduction or even instability. Therefore the control algorithm should possess performance robustness and stability robustness, which means that the controller does not yield a significant loss of performance or yield instability respectively, in case of changes in the plant dynamics. However, increasing the performance and/or stability robustness is at the expense of lower performance for the situation the plant equals the model on which the controller is based, i.e. the nominal performance is reduced. Hence, there is a trade-off between performance and stability robustness on one side and nominal performance on the other side. One main objective of this thesis is to improve this trade-off.

The convergence time of LMS-based adaptive algorithms (see the Filtered-X LMS and Filtered-U LMS algorithms below that are widely used in ANVC) is very long for broadband disturbances, especially when the controller has multiple channels. To be useful in practice, it is desired that this convergence time is in the range of a few minutes rather than several hours (e.g. to obtain fast adaptation to variations in the system dynamics, and to ease the tuning of the controller parameters). On the other hand, most algorithms which have a short convergence time are computationally rather complex. Therefore, the second objective of this thesis is to improve the trade-off between the convergence time and the computational complexity.

Non-stationary disturbances are the result of e.g. variations in the rotation velocity of engines, changes in the velocity of airplanes, cars, trains, which change the airflow and/or the road-tire or road-wheel interaction. Also changes in the dynamics of the plant may yield non-stationary disturbances, e.g. opening/closing a door of a room changes the dynamics of the room and thus also the noise disturbance inside the room. Though, this thesis will not discuss the tracking of non-stationary disturbances in much detail, the algorithms proposed in Chapter 6 and Chapter 7 are also promising for non stationary applications.

1.3   Problem description and basic assumptions

This section briefly introduces the active control problem considered in this thesis from a practical point of view and outlines the basic assumptions. The problem will be described in more mathematical detail in Chapter 2. The end of this section sketches the strategy of the research of this thesis.

1.3.1   Control system configurations

In ANVC three main control system configurations can be distinguished:

- **Feedforward control problem** (depicted in Figure 1.3);

- **General feedforward/feedback problem** (depicted in Figure 1.4);

- **Feedback control problem** (depicted in Figure 1.5).

PSfrag replacements

PSfrag replacements

PSfrag replacements
s
d
r
u
y
PSfrag replacements
e
G_{es}
G_{rs}
G_{eu}
C
u
y
e
G_{es}
G_{rs}
G_{eu}
C
s
d
r
u
y
e
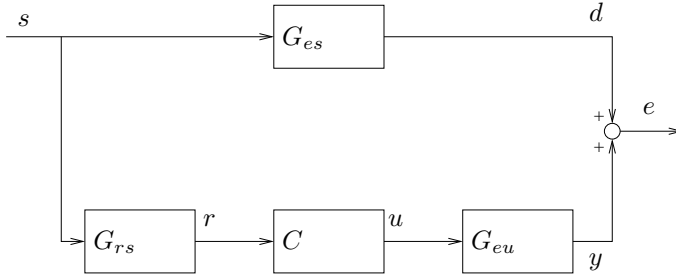
Figure 1.3: Block scheme of the feedforward active control problem, with s the signal from the disturbance source, d the disturbance to be cancelled, r the reference signal. The controller C computes the control signal u, which causes the secondary disturbance y and yields the residual disturbance e.
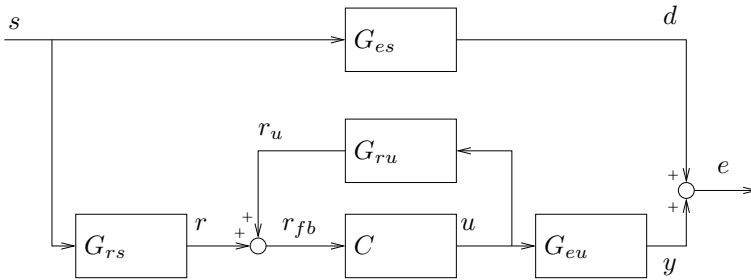
Figure 1.4: Block scheme of the general feedforward/feedback active control problem with the reference signal r is distorted by feedback of the control signal u.
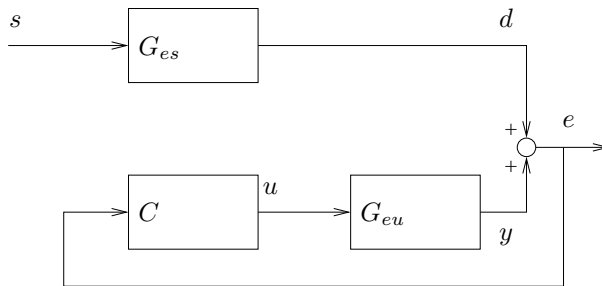
Figure 1.5: Block scheme of the feedback active control problem.

Feedforward control. First, consider the feedforward control problem as depicted in Figure 1.3. The disturbance signal generated by the disturbance source is denoted by s. For example, this signal represents the noise generated by vans in airconditiong systems, the noise or vibrations generated by engines in cars or airplanes or the vibrations induced by road-tire interaction. This signal will be a vector signal, when there are multiple disturbance sources. In general we are not able to effectively suppress this disturbance signal s directly at the source. But often, it is only needed to suppress the disturbance caused by s at some specific places, such as at the drivers, c.q. pilots, and passenger places in cars and airplanes or only at the end of exhaust pipes in cars or airconditioning systems. The disturbance caused by s, at the positions where disturbance suppression is obtained, is denoted by d. The path between s and d is called the *primary path* and denoted by G_{es}.

In feedforward control systems a reference signal r is available which is the input to the controller C. In applications with rotating engines r is usually chosen to be the rotation velocity measured by a tacho meter, c.f. the setup in Figure 1.2. In applications where s is a stochastic broadband signal, such as in case of vibrations generated by road-tire interaction, r is measured by a sensor positioned close to the disturbance source. For example, r is measured by accellerometer sensors mounted on the car chassis. The path between s and r is called the *detector path* and denoted by G_{rs}. The idea in feedforward control is that the reference signal r is strongly correlated with the disturbance signal d at future time instances, such that d can be predicted from r. Therefore r should be measured closer in distance to s than d, and r is referred to as the *upstream* reference signal and d the *downstream* disturbance signal.

The control signal u, computed by the controller, actuates loudspeakers or shakers to excite a secondary disturbance noise or vibration respectively. This secondary disturbance, denoted by y, interferes with the disturbance d and yields a residual disturbance e. The path between u and y is called the *secondary path* and denoted by G_{eu}.

The control signal should be determined such that e is 'as small as possible' (according to some criterium, see Chapter 2). This is ideally obtained when the series connection of G_{rs}, C and G_{eu} is exactly the oposite of G_{es}, which yields perfect cancellation. In case of stochastic disturbances, a necessary condition to obtain this perfect cancellation, is that the amount of delay in G_{rs} and G_{eu} is smaller than the amount of delay in G_{es}, which is due to the fact that C should be causal. From this observation, we conclude that it is desired that the reference sensors are positioned as close as possible to the disturbance sources, that the actuators are positioned as close as possible to the disturbances to be cancelled, and the distance between the disturbance sources and the disturbances to be cancelled should be as large as possible. Refer to Section 2.2 on page 31 for a more mathematical discussion of these system characteristics determing the performance.

General feedforward/feedback control. Contrary to applications with harmonic disturbances, in applications with broadband disturbances the reference signal is usually distorted by feedback of the control signal u, as is illustrated by Figure 1.4. Here, the feedback path is denoted by G_{ru} and is usually called the

acoustical feedback in acoustical applications. In general, we will refer to this feedback as the *intrinsic* feedback.

This feedback term does not (at least not *directly*) determine the optimal performance that can be obtained by a particular feedforward control system, as will be seen in Section 2.4 on page 45. But it determines the closed loop behavior together with the controller C, which will be *unstable* when C is not designed properly. This is an important aspect in controller design for systems with feedback and will also be addressed in this thesis, e.g. Section 4.5 on page 90.

Feedback control. In many applications with broadband disturbances, it is difficult or even impossible to measure an *upstream* reference signal which is strongly correlated with the disturbance d. For example, this is due to the fact that there are many disturbance sources at different locations, the disturbance d to be suppressed is already very close to the source (in this case $G_{rs} \approx G_{es}$ and $G_{ru} \approx G_{eu}$), or locating sensors close the disturbance source is difficult due to extreme (e.g., temperature) conditions. In these applications, the control signal u needs to be determined by feedback of the measurements of the residual disturbance e, as is illustrated by Figure 1.5.

The performance that can be obtained by a particular feedback problem, is (basically) determined by two aspects: 1) the delay in the secondary path G_{eu} and 2) the correlation in the disturbance signal d. When the delay is large, then the disturbance d needs to be predicted further in time to determine the current value of the control signal u. In addition, when the correlation between subsequent values of the disturbance signal d is very weak, it is not possible to accurately predict future values of the disturbance signal, which will degrade the performance.

Note, that the feedback control problem is a special case of the general feedforward/feedback control problem, where $G_{rs} = G_{es}$ and $G_{ru} = G_{eu}$.

Further note, that again a closed loop is involved which is determined by the controller C and the secondary path G_{eu}, such that stability of the closed loop needs to be taken into account when designing C.

In case of feedback control, where the objective is to minimize e in Figure 1.5, it is desired that the delay in G_{eu} is as small as possible, as mentioned before. Ideally this delay is zero, which corresponds physically with a *collocation* of actuators and sensors, i.e. each sensor is mounted rigidly to one actuator. In practice perfect collocation is not possible, though the system can be considered to be collocated up to a certain frequency. This means, that for higher frequencies the delay between the actuator and the sensor results in a phase shift which cannot be compensated by any causal filter, and thus cannot be neglected anymore. The design and control of these collocated systems is discussed in [140] for application in vibration control, also refer e.g. to the series of papers [11, 66, 67] for application in ASAC. For (approximately) collocated systems good performance and stability robustness can be obtained by relatively simple lead-, lag- and proportional-integral-derivative (PID) controllers, e.g. [140]. In this thesis, we will not focus on systems which can be considered as collocated, and thus do not put constraints on the positioning of sensors and actuators. But note, that actuator and sensor positioning is indeed an important issue in the design of the system. Further note, that when using discrete-time implementations of PID controllers (as is very common because of

the flexibility of DSP's) low-pass filters are needed to reconstruct a continuous-time control signal and to prevent aliasing effects. These low-pass filters have to be designed with care, since phase-lag is introduced, such that the bandwidth of the controller need to be reduced to preserve stability.

1.3.2 Basic assumptions

In this thesis, the following underlying assumptions will be made on the *physical* systems and signals:

1. **Stochastic broadband disturbances,** which are stationary or slowly non-stationary, in the sense they can be assumed stationary for the time necessary to compute the controller.

2. **Linear dynamic systems,** which are time-invariant or its time-variation small such that it can be considered as uncertainty.

Stochastic broadband disturbances.   The disturbance to be suppressed is considered to be a random noise process, which power spectrum is 'broad', i.e. the power of the disturbance signal is distributed over a particular *bandwidth* (e.g. typically $\approx 50 - 1kHz$) rather than located at single frequencies. Further, we will assume that the disturbance is stationary or semi-stationary, in the sense that over a period of a few seconds till a few minutes it can be considered to be stationary. So, here we will exclude non-stationary disturbances, such as speech signals or fast harmonic sweeps (e.g., due to fast acceleration of engines). As already said, though this thesis will not discuss the tracking of non-stationary disturbances some results are also promising for non-stationary applications, but needs further investigation.

So, the focus here is on stationary stochastic broadband disturbances. One can think of noise generated by jet engines [170], or road-tire (cars) [35] or rail-wheel (trains) [172] interaction or airflow around structures [169], just to name a few. Since the disturbances are stochastic the disturbance in the future is not (precisely) known, contrary to deterministic disturbances (single harmonic or broadband). In general there is correlation between the past and the future signal values of the disturbance, such that future signal values can be *predicted* given past signal values. The possibility to predict the future disturbance is crucial in active control to determine the control signal such that the disturbance will be counteracted.

Linear dynamic systems.   The propagation of acoustical waves in enclosures, such as an airplane cabin or an air conditioning duct, can be described by *linear* partial differential equations (PDE's), see e.g., [126]. Also, the mechanical vibrations in structures can often be described by *linear* PDE's, see e.g. [64]. Therefore, the first basic assumption we will make in this thesis is that the system can be considered as linear. Since, actuators and sensors are supposed to be part of the system, they are also assumed to be linear. Usually this assumption can be satisfied in practice very well. But care needs to be taken, especially in selecting the actuators to prevent non-linear hysteresis and/or saturation behavior, which may arise when using large voltages across piezoceramic actuators. We will also assume that the system is time invariant, or slowly varying around a particular average.

In this thesis we will consider possible time variations in the system as uncertainty for which the controller should be sufficiently robust.

1.3.3 Strategy of the research

Model-based controller design. A general approach to design controllers for feedforward or feedback control configurations is by minimizing a certain cost function, i.e. a performance measure, which is determined by an (accurate) model of the system. Since the acoustical or mechanical systems to be controlled are infinite dimensional systems, in fact also infinite dimensional models need to be determined. This can be done by determining the coefficients of the PDE's describing the system and the boundary conditions, which are determined by material and physical constants. However, in practice it is very difficult to determine the coefficients accurately, whereas the solutions of the PDE's are quite sensitive to the coefficients. In addition, the controller design is a difficult task since infinite dimensional systems are involved, c.f. e.g. [30]. But usually within a limited frequency band as considered in this thesis, these infinite dimensional systems can be approximated accurately by finite dimensional systems, though these finite dimensional systems may be of large order (orders ranging from 60-80 are not exceptional). Finite dimensional models can be estimated directly from input/output data by black-box model identification methods, such as Prediction Error Model (PEM) identification [104] or Subspace Model Identification (SMI) [177, 181], c.f. the discussion in Section 3.3 on page 53, which shows a preference for SMI methods. The obtained models are discrete-time models, which describe the dynamics between the sampled input and the sampled output signals obtained by digital to analogue (AD) and analogue to digital (DA) converters respectively. Because controllers can be easily implemented in DSP's, which are discrete-time digital signal processors, we will only consider discrete-time controllers in this thesis.

Model uncertainty. Models are always contained with model errors, e.g., due to measurement noise, limited number of samples, mismatch in the model structure or time variations in the system. Designing a controller by minimizing the cost function based on an imprecise model, may yield a controller which may provide significantly lower performance on the real system in comparison with the performance on the model, or the controller may even lead to instability in systems with a closed-loop. Hence, the model uncertainty needs to be taken into account in designing the controller. There are basically three approaches to account for the model uncertainty in designing the controller:

- **Adaptive control:** periodically, based on new measurements from the system, adjust the controller coefficients in the direction the cost function decreases;

- **Control-relevant identification:** design the controller by minimizing the cost function using the model as well as measured data from the system;

- **Robust control:** optimize an alternative (robust) or constrained cost function which is less sensitive to model errors.

These approaches are not fully independent, for example, there exists robust adaptive control algorithms and robust control-relevant identification methods.

The adaptive control algorithms are usually feedforward controller algorithms. These algorithms can be applied for applications with feedback using the so-called Internal Model Control approach as discussed in the next section and in more detail in Section 2.3. The adaptive algorithm that is mainly used for ANVC applications is the so-called Filtered-X Least Mean Squares (FxLMS) algorithm. This algorithm, which will be discussed in Section 1.4.1, has some nice properties such as easy implementation and its (convergence) behavior is well known in literature. However, for broadband disturbances the convergence of this algorithm may be very slow, especially for multichannel applications. Furthermore, the algorithm needs a model of the secondary path G_{eu}, but model errors may lead to instability of the update algorithm. Also the computational complexity of the algorithm may be quite severe since the controller consists of a tapped delay line which may be very long.

Robust control algorithms, such as the H_∞ optimal control [199], are not yet widely used in ANVC applications. The H_∞ control method using user defined weighting functions for the model uncertainty, has been used in [86] to control noise in a double wall panel. But here, it was concluded that the H_∞ design method introduces too much conservatism, such that no good performance could be obtained.

Research question. The question which we want to answer in this thesis, is whether new results in system, control and identification theory can solves these problems and yield robust control algorithms, which are quickly providing good, c.q. almost optimal performance, and are computationally feasible to be implemented on nowadays DSP's. In the first part of the thesis, we focus on accurate model identification, control-relevant identification, quantification of the model uncertainty and robust controller design. In the quantification of the model uncertainty and the robust controller design, we will follow and further develop a probabilistic robust control approach. Then, in the second part of the thesis we focus on adaptive control algorithms and analyze convergence properties and propose robust and fast converging adaptive algorithms motivated by (non-adaptive) concepts from the first part of the thesis. For a detailed overview of the thesis we refer to Section 1.6.

Though the analysis and development of feedforward and feedback control algorithms in this thesis, is primarily dedicated to disturbance suppression in active control applications, the results may also be usefull for other control applications, like the control of deformable mirrors in adaptive optics [174] or gas and liquid flow control [65].

1.4   Literature overview

1.4.1   Adaptive control

Filtered-X LMS and Filtered-U LMS.   Most practical active control systems are based on the Filtered-X LMS (FxLMS) adaptive algorithm —proposed

independently by Burgess [21] and Widrow et al. [191])— or the Filtered-U LMS (FuLMS) adaptive algorithm —proposed by Eriksson et al. [49]). As an illustration for the application to suppress interior aircraft noise we refer to [12, 17, 40]. Both algorithms are feedforward adaptive algorithms which need an 'upstream' reference signal (cf. the tacho meter signal in Figure 1.2), correlated with the disturbance to be canceled. This reference signal is the input to the filter which calculates the control signal. In the FxLMS algorithm this filter is a tapped-delay filter, which has a *finite* impulse response (FIR), and in the FuLMS algorithm a transfer function matrix, which has an *infinite* impulse response (IIR). The filter coefficients are adapted every sampling instant by a least means squares (LMS) algorithm using the measured residual disturbance [189]. The regressor is constructed by filtering the reference signal (and also the control signal for FuLMS) by a model of the secondary path, i.e. the system between the control actuators and the sensors which measure the residual disturbance (hence the names *"Filtered-X"* and *"Filtered-U"* LMS). This is because, the presence of the secondary path changes the surface of the cost function, and thus the gradient estimate needs to be adjusted accordingly. This is also one of the differences between adaptive (recursive) *identification* and adaptive *active control*.

The FuLMS algorithm can be seen as the IIR filter extension of the FxLMS algorithm, using Feintuch's LMS algorithm [50] to update the IIR filter coefficients. The FuLMS algorithm was proposed in [49] for applications where there is feedback from the control signal to the reference signal (in acoustical applications called *acoustical feedback*), but no convergence analysis had been given. The algorithms would also be useful in applications where the optimal filter contains weakly damped modes such that the choice of a very long FIR filter for FxLMS can be avoided.

The main advantage of the FxLMS and the FuLMS is their computational simplicity. The computational complexity of FxLMS for a single channel system is $O(4n_w + 2n_s)$ floating point operations (flops) per sampling instant, with n_w the number of taps in the FIR filter and n_s the number of taps in the FIR model of the secondary path. The computational complexity of FuLMS is $O(4n_a + 4n_b + 4n_s)$ flops per sampling instant, with n_a and n_b the number of coefficients in the controller numerator and denominator polynomials respectively.

Convergence and robustness of FxLMS and FuLMS. The algorithms are well studied in literature, especially the FxLMS algorithm, and some nice properties and variants are revealed. To start with, Hassibi et al. [76] showed that the LMS algorithm is H_∞ optimal. More precisely, they showed that, under some bound on the step size, the worst case energy of the error, obtained by LMS, is bounded by the energy of the measurement noise (no matter its source) and the squared norm of the initial estimate of the filter coefficients (weighted by the step size). Hence a performance robustness bound has been obtained for the LMS algorithm. Though the consequences of this analysis has not yet been fully analyzed for the FxLMS and FuLMS algorithms, it seems reasonable to assume that similar performance robustness bounds exists for these LMS-based algorithms. Inspired by the analysis in [76], Sayyar-Rodsari et al. [154] proposed a mathematical sound alternative for the FxLMS algorithm (and in [153] also for the FuLMS algorithm) for which a

similar performance bound as in [76] was derived. The algorithm shows improved convergence and robustness compared with the FxLMS algorithm, but its computational complexity is significantly larger; at least $O(n_w^2 + n_w n_s + n_s^2)^1$, which is not feasible in many high dimensional applications. Inspired by the systematic synthesis of [153] and results in fast-array algorithms [147, 151], this thesis proposes a computationally much more efficient algorithm in Chapter 7 (with complexity linear in n_w and n_s).

Conditions for the stability of the FxLMS update algorithm have been derived by Ren and Kumar [142] for the single channel case and Wang and Ren [188] for the multiple channel case. The analysis is based on the ordinary differential equation (ODE) approach to analyze the convergence of pseudo linear regressions (PLR's) developed by Ljung [101, 102]. Their result is that, under the condition that the secondary-path model is 'close' to the secondary path system in the sense that a particular *strictly positive real* (SPR) condition is satisfied, the step size vanishes and the input is persistently exciting, the filter coefficients asymptotically converge to their (unique) optimal value. Hence, the SPR condition is a bound on the *stability* robustness of the FxLMS update algorithm w.r.t. model uncertainty in the secondary-path model.

The analysis of the FuLMS algorithm is more difficult since an *IIR* filter is adapted, such that the control signal is *not* depending linearly on the filter coefficients. Hence, the cost function obtained to be minimized is not a convex function of the filter coefficients and local minima may exist, as was pointed out in the comments [84, 190] on Feintuch's paper [50]. Again using the ODE approach of [101, 102] (see also the textbook [106]) an SPR condition could be derived which guarantees that the cost function is *convex* in the IIR filter coefficients and the coefficients will asymptotically converge to their (unique) optimal value (also see the overview paper [157]). Based on this result Wang and Ren [187] also derived an SPR condition for global convergence of the FuLMS algorithm for the single channel case and Mosquera et al. [125] for the multiple channel case. Note, that the number of coefficients in the numerator and denominator of the optimal IIR filter need to be known and again the step size should vanish. However, these references, [125, 187], made a rather *strict assumption* that the system is such that all correlation between the reference signal and the residual noise could be removed, which is called perfect cancellation. In practical applications this is often not the case due to non-minimum phase behavior in the system. The assumption was made to directly adopt the results of [101, 102] on adaptive *identification*. This thesis reveals in Chapter 5 that this assumption is not necessary.

Drawbacks of FxLMS and FuLMS and alternative adaptive algorithms. Beside these nice results of computational efficiency, performance robustness w.r.t. measurement disturbances, stability robustness of the update rule w.r.t. secondary path model errors, FxLMS and FuLMS suffer to some important drawbacks:

---

[1]In [153] the secondary-path is modeled by a state-space model. Since the FIR model-structure is contained in the state-space structure, the algorithm can also be used for secondary-path models with FIR structure. The complexity of $O(n_w^2 + n_w n_s + n_s^2)$ is based on the FIR structure. In case a full state-space model has been used the complexity is at least $O(n_w^2 + n_w n_s + n_s^3)$, with n_s the dimension of the secondary-path state.

- Choice of the number of filter coefficients (for FuLMS the order of the numerator and denominator polynomials);

- Choice of step size;

- Divergence in case the SPR conditions are not satisfied;

- Slow convergence in case of broadband disturbances especially in multi-channel applications.

The choice of the number of filter coefficients in FxLMS and FuLMS is usually based on trial and error in combination with physical insight.

The choice of the step size is simplified by using the *normalized* LMS (NLMS) algorithm (see e.g., the textbooks on adaptive filtering [80, Sec. 9.11, p. 432-437] or [149, Sec. 5.6, p. 225-233]), in FxLMS and FuLMS, which improves the convergence rate too.

Divergence of FxLMS and FuLMS can be prevented by increasing the robustness of the adaptive algorithm by introducing a *leakage* term (see, e.g., [80, p. 441-442, 746-747], [149, e.g., p. 325] and for active control application [197] and [46, Sec. 3.4.7, p. 144-149]). Adding leakage to the adaptive algorithm can also be interpreted as adding a control-effort weighting to the cost function minimized by the adaptive algorithm. Increasing the leakage (control-effort weighting) will increase the stability robustness of the update algorithm, but also will *reduce* the performance. Therefore, in case the SPR condition is not satisfied in a particular frequency band, the leakage, or equivalently the control-effort weighting, can also be chosen to be *frequency dependent*, to increase robustness only in that specific frequency band, see, e.g., [48, 59].

The convergence rate is determined by the ratio between the largest and smallest eigenvalue of the auto-covariance matrix of the regression vector. This ratio is determined by the auto-spectrum of the reference signal and the dynamics of the secondary path model and can be very large for broadband disturbances. The convergence rate can be made independent of the auto-spectrum of the reference signal and the dynamics of the secondary path model by using the *Newton* LMS algorithm (see, e.g., [149, Sec. 4.5, p. 191-193]). In Newton LMS the update of the filter-coefficients is based on estimates of the auto-correlation of the regressor and the cross-correlation between the regressor and disturbance. The Newton LMS algorithm contains only a *single* convergence mode and (thus) yields a ratio between the largest and smallest eigenvalue of the regressor auto-covariance matrix which is equal to one. In this way the convergence rate is only depending on the step size, which can be chosen such that fast convergence is obtained. However, in general the Newton LMS algorithm requires the inversion of an $n_w \times n_w$ dimensional matrix, with n_w the number of coefficients of the adaptive filter, every sampling instant (or every blockperiod, when using a blockwise algorithm). The inversion takes at least $O(n_w^3)$ flops and thus the computational complexity of Newton LMS is too large for most practical active control applications, even if the update is not calculated every sampling instant. The computational complexity of the Newton LMS algorithm is reduced when using an *a priori* determined inverse of the regressor auto-covariance matrix, as, e.g., considered in [46].

Another approach, to obtain a ratio between the eigenvalues equal to one, is to prefilter the reference signal with a whitening filter, which is such that also all channels of the reference signal are independent (see, e.g., [149, Sec. 10.1, p.573-584]). In case there is also a secondary path, an inner-outer factorization of the secondary path model is used with the inner-factor, an all-pass filter (isometry), containing the non-minimum phase transmission zeros of the secondary path and the outer-factor minimum phase (thus stably invertible). The control signal —the output of the adaptive filter— is filtered with the inverse outer-factor, such that the transfer between the adaptive filter output and the error sensors is (approximately) equal to the inner-factor. For applying the Filtered-X LMS algorithm, the (pre-whitened) reference signal is filtered with the inner-factor of the secondary-path model. This algorithm is proposed in [45, 47] and also yields a ratio between the eigenvalues equal to one (see [160, 195] for similar preconditioning algorithms, and [58] and Chapter 5 for preconditioning of FuLMS). It appears that the robustness of FxLMS may deteriorate due to preconditioning, therefore Chapter 6 proposes a robustification of this preconditioned FxLMS algorithm.

Another alternative for the FxLMS algorithm to increase the convergence rate, is given by the affine projection algorithm (APA), see [41], see also [149, Sec. 5.8, p. 238-245], fast implementations of APA for active noise control are proposed in [18, 19]. The basic idea of the APA algorithm is to approximate the Newton LMS update direction using estimates of the auto-correlation of the regression and the cross-correlation between the regression and the disturbance by using a limited number of averages, say P. Using the matrix inversion lemma (MIL) the inversion of the auto-correlation matrix can be evaluated with complexity $O(P^3)$, which is much more efficient than $O(n_w^3)$ if $P \ll n_w$. For $P = 1$ the normalized FxLMS algorithm is obtained and for $P \to \infty$ the Newton LMS algorithm.

Finally, the recursive least squares (RLS) algorithm is known to provide the optimal (minimum-variance) estimate of the filter-coefficients every sampling instant, see, e.g., [85, Sec. 2.6, p. 55-58], [149, Sec. 5.9, p. 245-248]. However, the computational complexity of RLS is $O(n_w^2)$, which is too large for most practical applications. Fast-array and fast transversal filters (FTF) implementations of the RLS algorithm exists with much lower computational complexity ($O(16n_w)$ for the numerically better conditioned fast-array algorithms, and $O(8n_w)$ for the FTF algorithm). These fast algorithms are derived by exploiting shift structure which comes into the problem thanks to the tapped-delay structure of the filter, see, e.g., [149, Ch. 14, p. 816-873] for a good overview. The application of the fast implementations of the RLS algorithm to ANVC is discussed in, e.g., [4, 51, 156]. But for application in ANVC, a modification was necessary to compensate for delay introduced by filtering the reference signal by the secondary path model [51]. This modification was proposed earlier by Bjarnason in [15] for the FxLMS algorithm. Because using the RLS algorithm the filter-coefficients are varying faster, the effect of the delay is more dominant and the modification is more important. The derivation of the modified RLS algorithm for active control purposes, which we shall refer to as the modified Filtered-RLS algorithm, was quite ad hoc. In Chapter 7 a fast-array Kalman filter solution to the ANVC problem is proposed and it is shown that the modified Filtered-RLS algorithm is a special case of this Kalman filter solution.

Some interesting adaptive algorithms for (multiple) harmonic disturbances are given by Meurers and Veres [117,118] who propose an adaptive frequency selective feedback controller and by Bodson et al. [16] who propose an adaptive solution without assuming the disturbance frequencies to be known. In this thesis we do not go into further detail on harmonic disturbances, but will focus on stochastic broadband disturbances.

Internal Model Control (IMC) for applications with feedback. The adaptive algorithms just discussed in this subsection are all for feedforward active noise or vibration control problems where a reference signal is available which is not distorted by feedback from the control signal. In applications with harmonic disturbances such a reference signal is usually available, e.g., a tachometer signal from rotating engines which provide the frequencies (the base frequency and if necessary higher order harmonics) of the harmonic disturbances. However, in applications with broadband disturbances, such a reference signal is (usually) *not* available. For example in the widely used acoustical duct laboratory problem depicted in Figure 1.6 the reference microphone, which measures the broadband sound close to the disturbance source, also measures the sound generated by the secondary loudspeaker which is actuated by the controller. Hence a closed loop has been obtained, see the feedback via G_{ru} in the block scheme[2] of Figure 1.7.

To be able to continue using the adaptive algorithms discussed above, the feedback control problem can be reformulated into a feedforward control problem by means of the Youla parameterization [193, 194]. The Youla parameterization parameterizes all stabilizing controllers, by means of a *stable* filter W. Realizing W as an FIR filter, as is done in, e.g., the FxLMS algorithm, it is guaranteed that W is stable, and thus the controller is internally stabilizing. For stable[3] systems G_{ru} the Youla parameterization is illustrated by Figure 1.8. From this figure, we observe, that the influence of the feedback from the control signal on the reference signal is subtracted which yield a feedforward control problem

In practice G_{ru} is not known exactly and need to be replaced by a model $\widehat{G}_{ru}$, which is yields the well known Internal Model Control (IMC) approach, see [120]. Note, that because of model errors in $\widehat{G}_{ru}$ the stable filter W does not parameterize all stabilizing controllers anymore, which need to be taken into account in practice.

1.4.2 H_2 optimal and robust control

The research to algorithms for ANVC problems has been dominated by the signal processing community with *adaptive feedforward control* based algorithms. More recently the control systems community also addresses the broadband ANVC problem, but as an optimal or robust *output feedback control* problem, see, e.g., [22, 38, 39, 86, 87, 97, 121, 122, 135, 139, 141]. In this section, we discuss the use of H_2 optimal and robust control methods for the broadband ANVC problem. Since, most of these methods are model-based offline controller design methods, the next

---

[2]The notation and control problem will be discussed in more detail in Chapter 2.
[3]We only consider stable systems, which is not a limiting assumption, since most systems in active noise and vibration control problems are guaranteed to be stable.
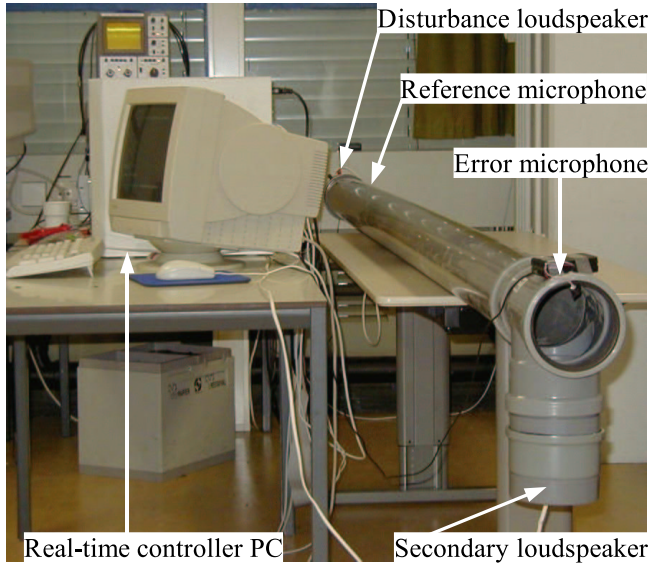
Figure 1.6: Experimental setup of the acoustical duct.

g replacements
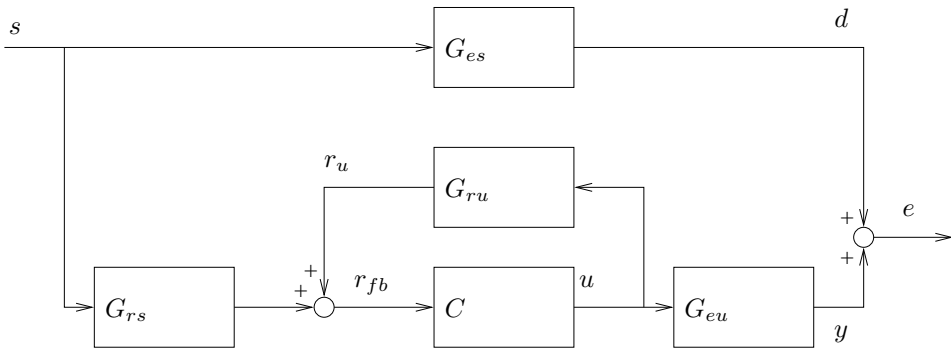
Figure 1.7: Block scheme of acoustical duct system[2].
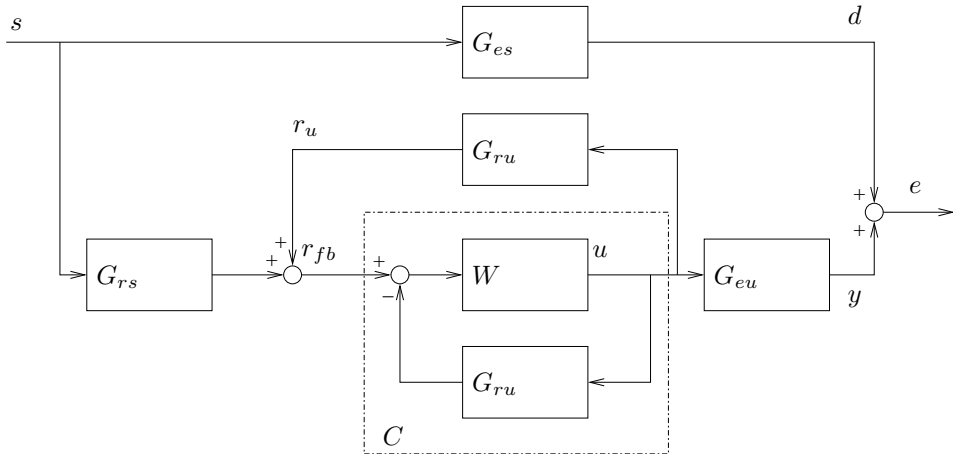
Figure 1.8: Block scheme of the controller C determined by the Youla-parameterization for stable G_{ru}, which parameterizes all stabilizing controllers by the *stable* filter W.

paragraph compares this approach with the adaptive approach of the previous section. Then, in the following paragraphs, various offline control design methods are discussed: the nominal LQG/LTR, the Worst-case robust design and the Cautious Wiener robust design methods.

Model-based offline controller design. The H_2 optimal and robust control methods proposed by researchers of the control systems community, are usually non-adaptive algorithms where the controller is designed using a model of the *complete* ANVC system, cf. the block scheme in Figure 1.7. The drawback of this approach is that the controller is not able to continuously adapt to disturbance variations. Still, the controller can be updated or redesigned after variations are detected and new models are updated or identified. Confer the recent contributions to iterative identification and controller design, see, e.g., [33, 69, 71, 96, 171, 175], though further research is necessary to develop reliable fully automatic procedures for high order systems without intervention of the user. The advantage of 'offline' designed controllers is that the number of real-time computations can be reduced considerably since no update of the controller coefficients is calculated every sampling instant. Slow convergence is prevented, and usually the time to compute the 'offline' controller is much smaller than the convergence time of FxLMS or FuLMS. But the controller should be designed to be robust for model errors in order to prevent the closed-loop to be de-stabilized. This is a necessary requirement in practice, since model errors cannot be avoided. The design of robust controllers which still provide good performance is the key issue in the 'offline' controller design approach.

In the following paragraphs, the H_2 optimal LQG controller and various robust design methods are discussed, which yield different trade-offs between performance and robustness.

LQG and LTR. The classical output feedback control solution to disturbance suppression is provided by the LQG/H_2 controller design method, see, e.g., [3, 85, 186]. The LQG controller minimizes the variance of the residual disturbance summed over all channels (i.e. the trace of the covariance matrix of the residual disturbance). However, in LQG/H_2 optimal design the model of the ANVC system, cf. Figure 1.7, is assumed to be perfect, which may lead to very poor stability robustness, see, e.g., Doyle's classical papers [42, 44]. At the expense of performance, the robustness can be improved by increasing the control-effort weighting and/or the variance of the measurement noise, cf. the loop transfer recovery (LTR) methods of [44, 83, 109]. For application to ANVC, see [13, 38, 39, 121, 122].

Worst-case robust design. Contrary to LQG/H_2, the H_∞ controller design methods explicitly take into account the presence of model uncertainty by (weighted) H_∞-norm bounds on the model error, see, e.g., [43, 166, 199]. For application to ANVC, see [22, 86, 87, 139]. However, in H_∞ design, the worst case model error condition is optimized, which may lead to too conservative controllers.

To enable a better trade-off between (robust) H_2 performance and H_∞ stability robustness, mixed H_2/H_∞ design methods are proposed, see, e.g., [10, 155]. For application to ANVC, see [97, 141].

More recently, the *minimax* LQG method was proposed, see [136], which minimizes the MSE for the worst case model error contained in a stochastic model uncertainty description. For application to ANVC see [134, 135]. For more on robust H_2 control, we refer to [132] and the references therein.

Cautious Wiener robust design. However, in all these robust design methods, the *likelihood* of the model errors is not taken into account. Such a design philosophy may be useful in critical applications where stability and a certain minimal (often low) level of performance should be guaranteed under all, including extremely rare, circumstances, e.g. in flight-by-wire control in aircrafts or biomedical control applications. Most active control problems are not that critical to pay a significant price on performance, and optimal performance on the *average* of all kind of model errors is desired. An additional reason is that most model identification methods give estimates of the likelihood of model errors, rather than hard bounds (see the discussion in [72] and [104]). In this line Sternad and Ahlén proposed a probabilistic robust filtering/feedforward control method, which minimizes the MSE averaged over the (estimated) *stochastic distribution of the model errors* [164]. The resulting robust filter is called a *Cautious* Wiener (CW) filter [129, 164]. This design approach can be interpreted as a frequency dependent control effort weighting, where the weighting is determined by the distribution of the model errors. This means, that the energy of the control signal is reduced in the frequency bands where the variation in the model uncertainty is large. In Chapter 4 of this thesis, the CW design philosophy is applied to feedforward and feedback ANVC. However, here a state-space approach has been taken, which is numerically better conditioned than the polynomial approach taken in [129, 164]. As an illustration of the cross-fertilization between control and signal processing, Chapter 6 of this thesis applies the CW design philosophy also to increase the robustness of FxLMS and Preconditioned FxLMS.

1.4.3 State-space based model identification and control

State-space models. Contrary to tapped delay (FIR) and transfer function models, state-space models are not common in ANVC applications. This is because the FxLMS (FuLMS) algorithm is based on an adaptive tapped delay (transfer function) filter. Since, the control signal obtained by a state-space filter is, contrary to a tapped delay filter, not depending linearly on the filter coefficients, a state-space filter is not very suitable to be used as adaptive filter. However, state-space filters can be used very well in

- Optimal / robust offline control: the controller can be just a single state-space filter;

- FxLMS (FuLMS): filtering of the reference signal (and control signal) with a state-space model of the secondary path;

- Preconditioned FxLMS (FuLMS): Pre-whitening of the reference signal and filtering the adaptive filter output with the inverse outer factor of the secondary path model;

- Reformulating adaptive filter problems as state-estimation problems: considering the adaptive filter problem as a state-estimation problem allows a unified analysis of numerous adaptive (RLS) filter algorithms as is demonstrated in [151], see also [154] for application to ANVC.

Hence, there might be some potential for state-space filters in ANVC. But, what are the advantages of a state-space filter over the tapped delay and transfer function filters? Advantages of the state-space structure is that it allows a compact and general model description, and it allows fast and numerically reliable filtering:

- **Compact model description:** If a model contains all observable and controllable modes of the system, increasing the number of inputs and outputs does not increase the model order; stated otherwise, the dynamics which is visible at multiple outputs is modeled just once. For FIR models or transfer-function matrices, the relation between the inputs and each output is modeled separately.

- **General model description:** The state-space model description is general, in the sense that it contains the tapped delay and transfer function model descriptions as special cases. Since, the state-space model description is unique up to a *similarity transformation* (i.e. a linear transformation of the state), there are (infinite) alternative choices for the state-space matrices (A, B, C, D), given by $(TAT^{-1}, TB, CT^{-1}, D)$ with T any non-singular matrix. For example one can choose (A, B, C, D) such that

$$AA^T + BB^T = I \qquad \text{(input-normal form)}$$

or

$$A^T A + C^T C = I \qquad \text{(output-normal form)}$$

with I the identity matrix. Other important alternative realizations are the balanced realization, the controller and observer canonical forms and the modal form [68].

- **Fast filtering:** There is a fundamental trade off between fast implementations and numerical reliability, see [68]. For example, the modal form and the observer and controller canonical forms enable very fast filtering especially for multiple channel filters, but are often sensitive to round-off errors due to finite word length arithmetic. The input- and output-normal form also allow fast filtering using unitary rotations (see Appendix C), but are less sensitive to round-off errors.

- **Numerical reliability:** The freedom in the state-space coefficients thanks to the similarity transformation, can be exploited to reduce the effect of round-off errors due to finite word length arithmetic. In [68] several methods are proposed to obtain the similarity transformation T which minimizes the sensitivity to these round-off errors. It is also shown that in general state-space filters in input- and output-normal form have low sensitivity to round-off errors.

  Reduced sensitivity to numerical round-off errors may allow to use reduced word length arithmetic, e.g. using single precision instead of double precision arithmetic, which can reduce the computation time and memory storage considerably.

State-space based model identification. State-space models can be estimated by black-box subspace model identification (SMI) methods, see, e.g., [177, 181, 185]. The subspace identification algorithms of the MOESP family [180, 183, 184] are all based on the QR decomposition and the SVD, which are numerical reliable algorithms. Though, the SMI methods do not explicitly solve an optimization problem as is obtained in prediction error methods (PEM) [104], it is observed in many practical applications (e.g. [1, 20, 114]) that accurate models are estimated, even in case the system has an Output Error (OE) structure for which PEM methods may converge to local optima. An additional nice property of SMI methods is that information of the model order is given by the rank of a particular matrix (i.e. the estimated extended observability matrix). Since in SMI algorithms no time consuming iterative search algorithm is necessary, these methods are well suited for the identification of multiple-input multiple-output (MIMO) systems with high order, such as most mechanical and acoustical systems have. Fast implementations of SMI algorithms [158] are provided in the SLICOT software library [159], see also [179].

The extension to the identification of mixed causal- anticausal systems is proposed in [182]. Subspace algorithms based on frequency domain data are proposed in [112, 113], also see [137]. The identification using frequency domain data can be preferred, since the number of data samples often can be reduced compared with using time-domain data. This is especially true for systems with (many) widely separated resonance modes (e.g., stiff systems and systems with a high number of resonances). The extension of discrete time-domain methods to continuous time-domain method has been given by [78]. More recently, also subspace algorithms for the identification of a class of *non-linear* systems are proposed, see, e.g., [178].

One drawback of SMI methods in comparison with PEM is, that no (stochastic) bound on the uncertainty of the parameter vector is provided. PEM methods

provide an estimate of the covariance matrix of the error in the parameter-vector [104], which can be used e.g., in robust controller design. In [14] a bootstrap method has been proposed to estimate the uncertainty for state-space models obtained by SMI. However, this approach is computationally complex and thus not very suitable for high order acoustical and mechanical systems as will be considered in this thesis. An alternative approach, proposed by Ljung in [103], to quantify the model uncertainty is to identify a model of the model error. The advantage of this approach is that it is independent of the identification algorithm or model structure, and will also be used in this thesis.

State-space based controller design. Most (offline) optimal and robust controller design methods are model based design methods, which are based on the *state-space* model description, see, e.g., [85, 199]. The use of state-space models often enables to solve the controller design problem by solving (a set of) Lyapunov equations, algebraic Riccati equations (ARE's) and/or linear matrix inequalities (LMI's), which can (usually) be done by numerically reliable algorithms. Often, control design methods based on polynomial model descriptions (see, e.g, [2, 91, 92, 163]) suffer to ill conditioned problems (polynomial spectral factorizations and Diophantine equations) especially for MIMO high order systems (cf. e.g. [196] on the condition of polynomials).

1.5 Contributions

The research strategy, which resulted in this thesis, was to obtain a cross-fertilization between the results from the adaptive/recursive adaptive control and identification together with results from the offline optimal and robust control design and identification. This resulted in various improvements of offline and adaptive controller design, which are proposed in order to meet the objectives given in Section 1.2.

By looking forward, the contributions of this thesis are:

1. A method to design a nominal broadband feedforward controller by solving a control-relevant identification problem using subspace model identification. This controller (partly) compensates for model errors in the stochastic disturbance model (Chapter 3).

2. Extension of this nominal design method to a probabilistic robust design method which is robust for model errors in the secondary path. The design method takes the likelihood of the model errors into account by means of a stochastic model error model (Chapter 4).

3. The nominal and robust methods can be used in applications with feedback by using the IMC approach (Chapter 3 and 4). For the robust design method, it is shown that stability robustness of the closed loop is increased by using a small gain theorem (Chapter 4).

4. By analysis of the optimal controller and the corresponding optimal residual signal (Chapter 2 and 3) it is shown that, under a technical assumption,

the optimal residual signal is uncorrelated with the regression vector of the Filtered-U LMS algorithm. As a consequence, for global convergence of the FuLMS algorithm it is not necessary to assume that perfect cancellation is achievable as is assumed in [187] (Chapter 5).

5. A method to increase the robustness of the (preconditioned) FxLMS algorithm by using a (stochastic) model error model of the model uncertainty in the secondary path (Chapter 6). The robustification approach is based on the probabilistic robust design method of Chapter 4.

6. A fast-array Kalman filter solution of the ANVC problem (Chapter 7):

   - By reformulating the feedforward ANVC problem to a state-estimation problem, the optimal minimum-variance estimate of the filter coefficients has been obtained by using a Kalman filter. The effect of unknown secondary-path initial state and measurement and process noise on the secondary-path state has been explicitly taken into account.

   - A fast-array implementation of the Kalman filter solution has been derived, which reduces computational complexity from $O((n_w + n_s)^2)$ to $O(23n_w + 16n_s)$ in the SISO case.

   - It is shown that the Kalman filter solution is a generalization of the modified Filtered-RLS algorithm. In this way, conditions for the optimality of the modified Filtered-RLS algorithm are derived. Additionally it has been shown that in case exponential forgetting is used in the modified Filtered-RLS algorithm, this exponential forgetting has to be taken into account in the generation of the Filtered-reference signal.

7. Application of the input- and output-normal parameterizations to obtain fast state-space filtering. The complexity of one state-space iteration is $O(2n(n + l) + 2(n + l)m)$, with l, m, n the number of outputs, the number of inputs and state-dimension respectively, when using the state-space matrices (A, B, C, D) and matrix-vector multiplications. Using the input- and output-normal parameterization the complexity of one iteration is reduced to $O(6nm + 2(n + m)l)$ and $O(6nl + 2(n + l)m)$ respectively, which enables real-time implementation of high order state-space controllers.
   In addition, using an input-normal parameterization of the secondary-path model a fast implementation of the generation of the regression vector in the multi-channel FxLMS algorithm has been obtained (Appendix C).

1.6 Outline of the thesis

The research strategy to obtain a cross-fertilization between offline (or blockwise) controller design and model identification on one hand and online adaptive filter design on the other hand, will also be visible in the structure of the thesis. The thesis is organized in two parts: Part I on Offline/blockwise algorithms and Part II on Online/sample-by-sample algorithms. Results on optimality and robustness obtained in Part I are used in Part II to analyze the convergence of the FuLMS algorithm and to propose a robust (preconditioned) FxLMS algorithm. Figure 1.9
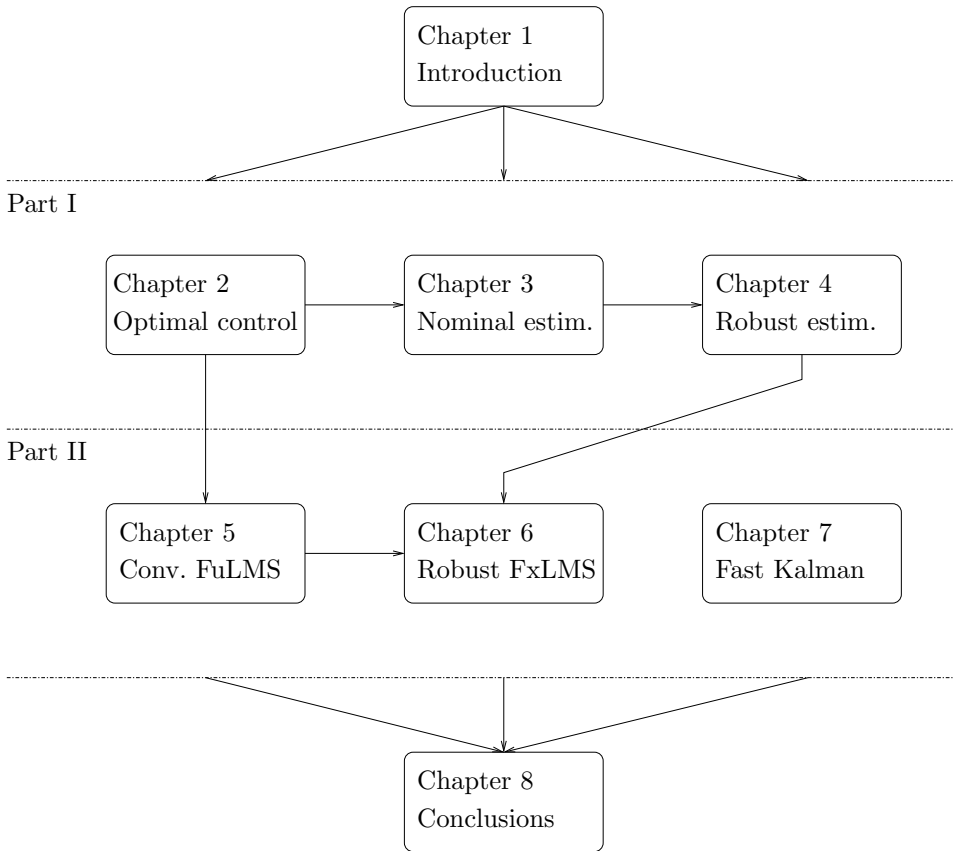
Figure 1.9: Dependencies among the chapters.

shows the dependencies among the chapters.

The outline of the thesis is given by:

Part I: Offline/blockwise algorithms consists of Chapter 2–4. **Chapter 2 Optimal feedforward and feedback control** introduces the notation, formulates the feedforward and feedback broadband active control problem and provides their H_2 optimal solutions. This chapter is based on the Technical Report [55].

Chapter 3 Nominal controller estimation builds further by developing a controller design procedure in which the H_2 optimal controller is estimated by solving a control relevant identification problem. The procedure is first derived based on Prediction Error identification with Output Error structure, which provides better theoretical insight. Second, the procedure is derived based on Subspace Model Identification, which provides a procedure which is better for practical use (especially for MIMO high-order systems). The

method is illustrated by experiments with an acoustical duct and a vibrating plate system. Parts of this chapter were previously published in the conference papers [54, 57, 61], and in Control Engineering Practice [62].

Chapter 4 Robust controller estimation extents the nominal controller estimation method to a robust controller estimation method, based on the Cautious Wiener design philosophy. The chapter starts with discussing the model uncertainty description. Then the robust feedforward controller, also called the Cautious Wiener filter, is derived. By using IMC the method can be applied for feedback systems too. The increased stability robustness is discussed using a small gain theorem. The method is illustrated by experiments with a vibrating plate system. Parts of this chapter were previously published in the conference paper [60] and Control Engineering Practice [62].

Part II: Online/sample-by-sample algorithms consists of Chapter 5–7. The results of Part I on offline/blockwise algorithms, provide also better understanding of online/sample-by-sample algorithms. This is illustrated by the analysis of Chapter 2, which is the basis of the new convergence result of FuLMS in Chapter 5 and by the robust design method of Chapter 4 which is the basis of the robust (preconditioned) FxLMS algorithm in Chapter 6.

The main result of **Chapter 5 Convergence analysis of Filtered-U LMS** is a relaxation of the conditions for global convergence of the FuLMS algorithm. Based on the analysis in Chapter 2, it is shown that the optimal residual signal is stochastically independent of the regression signal (under some technical conditions) and thus the ODE analysis can be applied even in case perfect cancellation is not achievable due to non-minimum phase zeros. The analysis of Chapter 2 also suggests a preconditioning of the FuLMS algorithm which, in general, improves the convergence rate considerably. This preconditioning was earlier proposed by Elliott and Cook in [47] for the FxLMS algorithm. The FuLMS convergence analysis and the preconditioned FuLMS algorithm are illustrated by simulation on an academic example and an acoustical duct system. Chapter 5 was previously published in Signal Processing [58] in a slightly different presentation.

Chapter 6 Robust preconditioned Filtered-X LMS proposes a robustitifaction of the ordinary FxLMS and the preconditioned FxLMS algorithm proposed by Elliott and Cook. The robustification is based on the robust design approach of Chapter 4 and mainly motivated by lack of robustness of the preconditioned FxLMS algorithm due to the inversion of a minimumphase spectral factor and outer factor. The robust preconditioned FxLMS algorithm is demonstrated on a acoustical duct system with delay variation in the secondary path. Chapter 6 was previously published in IEEE Signal Processing Letters [59].

In **Chapter 7 A Fast-array Kalman filter solution** another approach to adaptive filtering has been taken in which the adaptive ANVC problem is reformulated in a state-space estimation problem, which is solved by Kalman filtering. The main result of this chapter is the derivation of a fast-array implementation of this Kalman based algorithm. Also a comparison has

been made with the modified Filtered-RLS algorithm and it is shown that the Kalman based algorithm is a generalization of the modified Filtered-RLS algorithm. The performance of the fast-array Kalman filter method and the modified Filtered-RLS algorithm are compared by simulation on an acoustical duct system. Chapter 7 has been submitted for publication in the International Journal of Adaptive Control and Signal Processing [53].

Chapter 8 Conclusions concludes the thesis by evaluation of the contributions of the thesis and recommendations for further research.

The outline of the appendices is given by:

Appendix A Proof of the Causal Wiener Theorem: This appendix contains the proof of the Causal Wiener Theorem given in Chapter 2.

Appendix B The Causal Wiener state-space filter and relation with LQG: This appendix contains the derivation of a state-space expression of the Causal Wiener filter and relates the state-space solution to the LQG solution. Also state-space expressions of the inner-outer and outer-inner factorizations are given.

Appendix C Efficient state-space filtering using input- and output normal forms and application to FxLMS: Fast implementations of state-space filters in input- and output-normal form has been derived. It is shown how the filtering of the reference signal in the multiple-channel FxLMS algorithm can be evaluated efficiently using the input-normal form.

Appendix D A frequency domain subspace algorithm for mixed causal, anti-causal LTI systems: This appendix contains a paper on the identification of mixed causal, anti-causal systems using frequency domain data. Using the result of this paper, the algorithms of Chapter 3 and 4 can be performed by using frequency domain data as well. The paper was previously published in the Proc. of SYSID 2003, [63].

Appendix E: Robust Decision Feedback Equalizer design via the solution of a regularized least squares problem: This appendix contains a paper in which the design approach of Chapter 4 has been applied to the direct identification of a robust decision feedback equalizer using a probabilistic uncertainty model as in the Cautious Wiener approach. This chapter illustrates the spin off of the research to telecommunication applications. The chapter was previously published in the Proc. of the European Control Conference 2001, [56].

Part I

Offline/blockwise algorithms

CHAPTER 2

OPTIMAL FEEDFORWARD AND FEEDBACK CONTROL

2.1 Introduction

This introductory chapter presents the basic control configurations considered in ANVC and their H_2 optimal solutions. The results of this chapter can be found in different textbooks on control (see, e.g., [3, 186, 198]), or can be easily derived, and will be used in the following chapters.

Three different building blocks in the design and analysis of ANVC are reviewed. First, the pure feedforward control problem is considered, which is solved by the Causal Wiener filter. We also analyze the effect of noise on the reference signal and its dual case of control-effort weighting. The results obtained will play an important role in the derivation of robust algorithms in the following chapters.

Second, the pure feedback control problem is considered. Using the Youla parameterization of all stabilizing controllers this problem is reformulated into the pure feedforward control problem.

Third, we consider the feedforward control problem where the reference signal is distorted by the feedback control signal. This is the so called general feedforward/feedback control problem. The H_2 optimal solution also follows from the Youla-parameterization together with the Causal Wiener filter. We end the chapter with comments on robust performance and robust stability.

This chapter is primarily based on the Technical Report [55].

2.2 Feedforward problem

2.2.1 Without measurement noise

Consider Figure 2.1 which illustrates the feedforward active control problem, without measurement noise. Because we will focus on stationary broadband disturbances, we assume that the discrete-time signal $s(k) \in \mathbb{R}^{m_s}$, generated by the disturbance source, is a white and stationary random process with zero-mean and
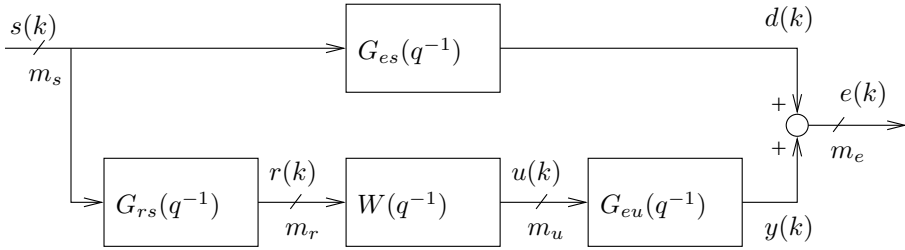
Figure 2.1: Block scheme of the multichannel feedforward active control problem, with m_s noise sources, m_r reference, m_u control and m_e error signals and no measurement noise.

unit covariance

$$\mathbf{E}\left(\left[\begin{array}{c} s(k) \\ 1 \end{array}\right] s^T(l)\right) = \left[\begin{array}{c} I_{m_s}\delta_{kl} \\ 0_{1\times m_s} \end{array}\right], \qquad (2.1)$$

with δ_{kl} the Kronecker delta function defined as

$$\delta_{kl} = \left\{ \begin{array}{ll} 1, & \forall k = l \\ 0, & \text{otherwise.} \end{array}\right.$$

The signal $s(k)$ propagates through the (discrete-time) primary path $G_{es}(q^{-1}) \in \mathcal{RH}_\infty^{m_e \times m_s}$, with $\mathcal{RH}_\infty^{m_e \times m_s}$ the set of all asymptotically stable rational $m_e \times m_s$ transfer function matrices with real coefficients[1] and q^{-1} the unit shift back operator:

$$q^{-1}s(k) = s(k-1).$$

The anti-causal, unit-shift forward operator is denoted by q:

$$qs(k) = s(k+1).$$

In the following, we will often skip the argument q^{-1} and will just write G_{es} when referring to $G_{es}(q^{-1})$. The output of the primary path is the disturbance signal $d(k) \in \mathbb{R}^{m_e}$, which needs to be canceled. This is done by the secondary disturbance signal $y(k) \in \mathbb{R}^{m_e}$, which is the output of the secondary path $G_{eu}(q^{-1}) \in \mathcal{RH}_\infty^{m_e \times m_u}$. The residual disturbance signal, also called error signal, is denoted by $e(k) \in \mathbb{R}^{m_e}$ and given by

$$e(k) = d(k) + y(k). \qquad (2.2)$$

where

$$d(k) = G_{es}s(k), \qquad (2.3)$$

$$y(k) = G_{eu}u(k). \qquad (2.4)$$

---

[1]In active noise and vibration control application the systems to be controlled are asymptotically stable, therefore we restrict to asymptotically stable systems. Note, that most results also generalize for unstable systems, though the generalization is not always straightforward.

The control signal $u(k) \in \mathbb{R}^{m_u}$ actuates G_{eu} and is generated by the feedforward controller $W(q^{-1}) \in \mathcal{RH}_\infty^{m_u \times m_r}$. The input to the controller is the reference signal $r(k) \in \mathbb{R}^{m_r}$, which is the output of the detector path $G_{rs}(q^{-1}) \in \mathcal{RH}_\infty^{m_r \times m_s}$

$$u(k) = Wr(k), \tag{2.5}$$
$$r(k) = G_{rs}s(k). \tag{2.6}$$

The idea of this feedforward control scheme, is that $r(k)$ is an *upstream* signal and $d(k)$ a *downstream* signal (the 'delay' in G_{rs} is smaller than the 'delay' in G_{es}). Ideally, the feedforward controller W should be such that the transfer function between $s(k)$ and $d(k)$ equals the negative of the transfer function between $s(k)$ and $y(k)$

$$G_{es} = -G_{eu}WG_{rs}. \tag{2.7}$$

If G_{eu} and G_{rs} can be inverted, solving W from (2.7) yields

$$W = -G_{eu}^{-1}G_{es}G_{rs}^{-1}. \tag{2.8}$$

However, W is constrained to be *causal* and *stable*, or else there will be no physical implementation of W be possible, which may prevent (2.7) to be satisfied. Furthermore, in the multiple channel case when e.g. $m_e > m_u$ or $m_s > m_r$ it is possible that (2.7) cannot be obtained due to under actuation or because the reference channels do not contain all information of the disturbance channels. Let us illustrate these causes, which prevent satisfying (2.7) for any causal W, by the following two examples respectively.

Example 2.1 *Consider the problem with*

$$G_{es} = q^{-1}\frac{1 - 1.1q^{-1}}{1 - 0.9q^{-1}}, \quad G_{eu} = q^{-1}, \quad G_{rs} = q^{-1}.$$

The only solution of W satisfying equation (2.7) *is given by*

$$W = -q\frac{1 - 1.1q^{-1}}{1 - 0.9q^{-1}} = q - 0.2 - 0.18q^{-1} - 0.162q^{-2} - \cdots$$

which is not causal.

Example 2.2 *Consider the problem with*

$$G_{es} = \begin{bmatrix} q^{-1} \\ q^{-2} \end{bmatrix}, \quad G_{eu} = \begin{bmatrix} q^{-1} \\ q^{-1} \end{bmatrix}, \quad G_{rs} = 1$$

Then, to satisfy the first row of the constraint (2.7) *we should choose $W = -1$, but to satisfy the second row of this constraint we should choose $W = -q^{-1}$. Thus it is clear that there does not exist a W (no matter causal or not) such that* (2.7) *holds.*

Hence, in some problems there does not exist a causal W, such that (2.7) holds and thus perfect cancellation of $d(k)$ such that $e(k) = 0$ can not be obtained. The

examples are mathematical examples, but illustrate system characteristics which
do occur in practical active control problems.

Therefore, we will not search for causal W satisfying (2.7), but search for the
best causal W which minimizes the error signal $e(k)$ according to a specific criterion.
The mean squared error (MSE) of the residual disturbance $e(k)$ given by

$$\text{tr}\mathbf{E}\left(e(k)e^T(k)\right) \tag{2.9}$$

is a very suitable criterion, because it is a measure of the power of the remaining
disturbance. Therefore, it is widely used in ANVC applications.

Because the signal $s(k)$ from the disturbance source is a white noise process
with unit covariance, the MSE value of $e(k)$ can also be written in terms of the H_2
norm of $G_{es} - G_{eu}WG_{rs}$ (using Parseval's Theorem), i.e.,

$$\text{tr}\mathbf{E}\left(e(k)e^T(k)\right) = ||G_{es} + G_{eu}WG_{rs}||_2^2$$

where the H_2-norm is defined in the next definition.

Definition 2.1 (H_2-norm) *The H_2-norm of the transfer function matrix $G(q^{-1}) \in \mathcal{RH}_\infty^{m_x \times m_y}$ is defined as*

$$||G(q^{-1})||_2 \triangleq \sqrt{\frac{1}{2\pi}tr\int\limits_{-\pi}^{\pi} G(e^{j\omega})G(e^{j\omega})^* d\omega} \tag{2.10}$$

where $(.)^*$ denotes the complex conjugate transpose.

Using the H_2-norm criterion the feedforward controller design problem is formu-
lated as follows.

Problem 2.1 (H_2 feedforward controller design problem) *Given $G_{es} \in \mathcal{RH}_\infty^{m_e \times m_s}$, $G_{eu} \in \mathcal{RH}_\infty^{m_e \times m_u}$, $G_{rs} \in \mathcal{RH}_\infty^{m_r \times m_s}$ determine $W \in \mathcal{RH}_\infty^{m_u \times m_r}$ such
that*

$$W = \arg\min_{\mathcal{W}\in\mathcal{RH}_\infty^{m_u \times m_r}} J(\mathcal{W}) \tag{2.11}$$

with the cost-function $J(\mathcal{W})$ defined by

$$J(\mathcal{W}) = ||G_{es} + G_{eu}\mathcal{W}G_{rs}||_2. \tag{2.12}$$

The following subsection will give the solution to this problem.

2.2.2   Causal Wiener filter

To solve Problem 2.1 we will use a factorization approach (see, e.g., [186]), which
provides some nice interpretations. Before stating the theorem which provides the
solution, we need to introduce the causality and anti-causality operators and the
inner-outer and outer-inner (or co-inner-outer) factorizations. We have already
introduced the set $\mathcal{RH}_\infty^{m_x \times m_y}$ of asymptotically stable transfer function matrices.
Now we will introduce the more general set $\mathcal{RH}^{m_x \times m_y} \supset \mathcal{RH}_\infty^{m_x \times m_y}$ of all $m_x \times m_y$

transfer function matrices with real coefficients and excluding singularities on the unit-circle. A singularity on the unit-circle is a systems with poles on the unit-circle in the complex-plane. Let us directly introduce another subset of $\mathcal{RH}^{m_x \times m_y}$, which is denoted as $\mathcal{RH}_p^{m_x \times m_y}$ and contains all *proper* transfer function matrices in $\mathcal{RH}^{m_x \times m_y}$. A transfer function is said to be *proper*, if the order of the denominator polynomial is greater or equal to the order of the numerator polynomial.

It can be proven that for any $G(q^{-1}) \in \mathcal{RH}^{m_x \times m_y}$, there exists a (convergent) Laurent series given by (see, e.g., [143])

$$G(q^{-1}) = \sum_{i=-\infty}^{\infty} G_i q^{-i}, \qquad G_i \in \mathbb{R}^{m_x \times m_y}. \tag{2.13}$$

Definition 2.2 (Causality and anti-causality operator) *Let the Laurent series of $G(q^{-1}) \in \mathcal{RH}^{m_x \times m_y}$ be given by (2.13), then the causal part and the anti-causal part of G are defined by*

$$[G]_+ \triangleq \sum_{i=0}^{\infty} G_i q^{-i}, \qquad \text{causal part,} \tag{2.14}$$

$$[G]_- \triangleq \sum_{i=-\infty}^{-1} G_i q^{-i}, \qquad \text{anti-causal part} \tag{2.15}$$

respectively, with $G_i \in \mathbb{R}^{m_x \times m_y}$.

Note, that our definition of anti-causality is strict in the sense that the term G_0 is excluded. It is straightforward to verify that

$$G = [G]_+ + [G]_-.$$

Furthermore, it can be shown that $[G]_+ \in \mathcal{RH}_\infty^{m_x \times m_y}$ and $[G]_- \in \mathcal{RH}^{m_x \times m_y}/\mathcal{RH}_\infty^{m_x \times m_y}$.

Lemma 2.1 (Inner-outer factorization [186]) *Let $G \in \mathcal{RH}_\infty^{m_x \times m_y}$. Then G has an inner-outer factorization*

$$G = G_i G_o, \tag{2.16}$$

with $G_i \in \mathcal{RH}_\infty^{m_x \times m}$ is an isometry ($G_i^* G_i = I_m$) and the outer factor $G_o \in \mathcal{RH}_\infty^{m \times m_y}$ has a stable right inverse, with $m \leq \min(m_x, m_y)$. If $G(q^{-1})$ does not loose rank $\forall |q| = 1$ (G has no zeros on the unit-circle), than G_o has an asymptotically *stable right inverse. Furthermore, there exists a $G_i^\perp \in \mathcal{RH}_\infty^{m_x \times m_x - m}$ such that $[G_i \; G_i^\perp]$ is unitary ($[G_i \; G_i^\perp]^*[G_i \; G_i^\perp] = [G_i \; G_i^\perp][G_i \; G_i^\perp]^* = I_{m_x}$).*

Lemma 2.2 (Outer-inner factorization [186]) *Let $G \in \mathcal{RH}_\infty^{m_x \times m_y}$. Then G has an outer-inner factorization*

$$G = G_{co} G_{ci}, \tag{2.17}$$

with $G_{ci} \in \mathcal{RH}_\infty^{m \times m_y}$ is a co-isometry ($G_{ci} G_{ci}^* = I_m$) and the co-outer factor $G_{co} \in \mathcal{RH}_\infty^{m_x \times m}$ has a stable left inverse, with $m \leq \min(m_x, m_y)$. If $G(q^{-1})$

does not loose rank $\forall |q| = 1$ (G has no zeros on the unit-circle), than G_{co} has an asymptotically stable left inverse. Furthermore, there exists a $G_{ci}^{\perp} \in \mathcal{RH}_{\infty}^{m_y - m \times m_y}$ such that $[G_{ci}^* \ G_{ci}^{\perp *}]^*$ is unitary ($[G_{ci}^* \ G_{ci}^{\perp *}][G_{ci}^* \ G_{ci}^{\perp *}]^* = [G_{ci}^* \ G_{ci}^{\perp *}]^* [G_{ci}^* \ G_{ci}^{\perp *}] = I_{m_y}$).

Note, that due to the isometry and co-isometry properties of G_i and G_{ci} respectively, we have

$$G^* G = G_o^* G_o,$$
$$GG^* = G_{co} G_{co}^*.$$

Appendix B shows how the inner-outer and the outer-inner factorization of G can be calculated given a *state-space* realization of G.

Because the right and left outer factors have a *stable* right and left inverse respectively, they are said to be *minimum phase*. Hence, G_{co} is a minimum phase spectral factor of GG^*.

Because $G_i^* G_i = I_m$, G_i^* is a *left-inverse* of G_i. On the other side, because $G_{ci} G_{ci}^* = I_m$, G_{ci}^* is a *right-inverse* of G_{ci}.

The basic idea behind the inner-outer and outer-inner factorization is to factorize dynamic systems in a part which is stably (causally) invertible (the (co-)outer factor) and a remaining part (the (co-)inner factor) which only yields a phase shift (e.g. due to delays) and thus does not affect the energy of signals.

Example 2.3 *Let the system G be given by*

$$G(q^{-1}) \ = \ \frac{q^{-1}(1 - 1.1q^{-1})}{1 - 0.9q^{-1}}$$

then the inner and outer factors are given by

$$G_i(q^{-1}) \ = \ q^{-1} \frac{1}{1.1} \frac{1 - 1.1q^{-1}}{1 - \frac{1}{1.1}q^{-1}}, \quad G_o(q^{-1}) \ = \ 1.1 \frac{1 - \frac{1}{1.1}q^{-1}}{1 - 0.9q^{-1}}$$

respectively.

Note, that for SISO systems the inner and outer factors are equal to the co-inner and the co-outer factors respectively.

The following theorem gives the solution to Problem 2.1.

Theorem 2.1 (Causal Wiener filter) *Given $G_{es}(q^{-1}) \in \mathcal{RH}_{\infty}^{m_e \times m_s}$, $G_{eu}(q^{-1}) \in \mathcal{RH}_{\infty}^{m_e \times m_u}$, $G_{rs}(q^{-1}) \in \mathcal{RH}_{\infty}^{m_r \times m_s}$. Assume that $G_{eu}(q^{-1})$ and $G_{rs}(q^{-1})$ do not loose rank $\forall |q| = 1$. Then, let*

$$G_{eu} = G_{eu,i} G_{eu,o} \tag{2.18}$$
$$G_{rs} = G_{rs,co} G_{rs,ci} \tag{2.19}$$

be the inner-outer and outer-inner factorization of G_{eu} and G_{rs} respectively and $G_{eu,o}^{\dagger}$ a right-inverse of $G_{eu,o}$ and $G_{rs,co}^{\dagger}$ a left-inverse of $G_{rs,co}$. Let $G_{eu,i}^{\perp}$ and $G_{rs,ci}^{\perp}$ be such that $[G_{eu,i} \ G_{eu,i}^{\perp}]$ and $[G_{rs,ci}^* \ G_{rs,ci}^{\perp *}]^*$ are unitary. Then

$$W \ = \ -G_{eu,o}^{\dagger} \left[G_{eu,i}^* G_{es} G_{rs,ci}^* \right]_+ G_{rs,co}^{\dagger} \tag{2.20}$$

minimizes

$$||G_{es} + G_{eu}\mathcal{W}G_{rs}||_2, \qquad subject\ to\ \mathcal{W} \in \mathcal{RH}_\infty^{m_u \times m_r} \qquad (2.21)$$

and its minimum value is given by

$$||G_{es} + G_{eu}WG_{rs}||_2 = $$
$$\sqrt{||G_{es}G_{rs,ci}^{\perp*}||_2^2 + ||G_{eu,i}^{\perp*}G_{es}G_{rs,ci}^*||_2^2 + || \left[G_{eu,i}^*G_{es}G_{rs,ci}^*\right]_-||_2^2} \qquad (2.22)$$

Proof: The proof is given by [186, Section 6.2]. We included the proof in a slightly different presentation in Appendix A. □

Appendix B shows how the Causal Wiener filter (2.20) can be calculated given state-space realizations of G_{es}, G_{eu} and G_{rs}.

We observe, that in case $G_{eu,i} = I_{m_u}$ and $G_{rs,ci} = I_{m_r}$ —which means that G_{eu} and G_{rs} do *not* have non-minimum phase zeros— the Causal Wiener filter (2.20) is given by

$$W = -G_{eu}^\dagger G_{es} G_{rs}^\dagger \qquad (2.23)$$

The expression is called the Wiener filter (without the preposition *Causal*) and equals (2.8) in case G_{eu} and G_{rs} are square.

Even in case $G_{eu,i} \neq I_{m_u}$ or $G_{rs,ci} \neq I_{m_r}$ *but*

$$G_{eu,i}^* G_{es} G_{rs,ci}^* = [G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ \qquad (2.24)$$

the Causal Wiener filter equation (2.20) simplifies to the Wiener equation (2.23). When equation (2.24) holds, non-minimum phase zeros in G_{eu} and G_{rs} cancel against non-minimum phase zeros in G_{es}.

Example 2.4 *For example, let*

$$G_{es} = q^{-p}, \quad G_{eu} = q^{-2}, \quad G_{rs} = q^{-1}$$

and thus, the inverses of G_{eu} and G_{rs} are not causal. Now, we have

$$G_{eu,o} = 1, \quad G_{eu,i} = q^{-2}, \quad G_{rs,co} = 1, \quad G_{rs,ci} = q^{-1}$$

and thus

$$G_{eu,i}^* G_{es} G_{rs,ci}^* = q^2 q^{-p} q^1 = q^{3-p}.$$

So for $p \geq 3$ we have

$$G_{eu,i}^* G_{es} G_{rs,ci}^* = [G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ = q^{3-p}$$

and the Causal Wiener (2.20) and the Wiener filter (2.23) are equal and given by

$$W = -q^{3-p}.$$

However, for $p < 3$ we have

$$[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ = 0$$

Hence, the Causal Wiener filter (2.20) is given by

$$W = 0$$

and is different from the Wiener filter $-q^{3-p}$, which is anti-causal for $p < 3$.

Further, we note that in case $m_r > m_s$ the left-inverse of $G_{rs,co} \in \mathcal{RH}_\infty^{m_r \times m_{rs}}$, with $m_{rs} \leq \min(m_r, m_s)$, is not unique. Analogue, in case that $m_u > m_e$ the right-inverse of $G_{eu,o} \in \mathcal{RH}_\infty^{m_{eu} \times m_u}$, with $m_{eu} \leq \min(m_e, m_u)$, is not unique. This non-uniqueness can be exploited, e.g., to minimize the control-effort.

In the case that $m_r < m_s$ it may happen —though not necessarily— that $G_{rs,ci}^\perp$ is non-zero, and thus the first term under the square-root in (2.22) contribute to the minimum value of the cost-function if also $G_{es}G_{rs,ci}^{\perp*}$ is non-zero. The interpretation is, that there is noise which contributes to the disturbance $d(k)$, but are not measured, or observed, at the reference sensors and thus not contained in $r(k)$. This is related to the concept of the concept of *unobservable* modes of the system, as is known in standard control literature [199].

Analogue, in the case that $m_u < m_e$, it may happen —though not necessarily— that $G_{eu,i}^\perp$ is non-zero, and thus the second term under the square-root in (2.22) contributes to the minimum value of the cost-function if also $G_{eu,i}^{\perp*}G_{es}G_{rs,ci}^*$ is non-zero. The interpretation is, that the actuator configuration (the number of actuators and their positioning) is too limited to counteract all disturbance channels in $d(k)$. This is related to the concept of *uncontrollable* modes of the system in standard control literature [199].

The first and second term under the square-root in (2.22) are related to the *geometry* of the active control system, the third term is related to the restriction that $W(q^{-1})$ should be causal. The third term is determined by delays and non-minimum phase zeros in G_{rs} and G_{eu}, which contribute to the anti-causal terms in $G_{rs,ci}^*$ and $G_{eu,i}^*$.

The error signal $e(k)$ obtained by using the Causal Wiener filter W given by (2.20) is given by

$$
\begin{aligned}
e(k) &= (G_{es} + G_{eu}WG_{rs})s(k), \\
&= \big(G_{eu,i}[G_{eu,i}^*G_{es}G_{rs,ci}^*]_- G_{rs,ci} + \\
&\quad + G_{eu,i}G_{eu,i}^*G_{es}G_{rs,ci}^{\perp*}G_{rs,ci}^\perp + G_{eu,i}G_{eu,i}^{\perp*}G_{es}\big)s(k)
\end{aligned}
\tag{2.25}
$$

where we have used

$$
G_{es} = \underbrace{\begin{bmatrix} G_{eu,i} & G_{eu,i}^\perp \end{bmatrix} \begin{bmatrix} G_{eu,i}^* \\ G_{eu,i}^{\perp*} \end{bmatrix}}_{=I_{m_e}} G_{es} \underbrace{\begin{bmatrix} G_{rs,ci}^* & G_{rs,ci}^{\perp*} \end{bmatrix} \begin{bmatrix} G_{rs,ci} \\ G_{rs,ci}^\perp \end{bmatrix}}_{=I_{m_s}},
$$

and

$$
G_{eu,i}^*G_{es}G_{rs,ci}^* - [G_{eu,i}^*G_{es}G_{rs,ci}^*]_+ = [G_{eu,i}^*G_{es}G_{rs,ci}^*]_-.
$$

Let us consider the examples of the previous section again.

Example 2.5 *Consider the system of Example 2.1, then the optimal causal filter W according to Theorem 2.1 is given by*

$$
W = \frac{0.2}{1 - 0.9q^{-1}}.
$$

Furthermore, the contributions to the cost-function are

$$
\|G_{es}G_{rs,ci}^{\perp*}\|_2^2 = 0, \quad \|G_{eu,i}^{\perp*}G_{es}G_{rs,ei}^*\|_2^2 = 0, \quad \|[G_{eu,i}^*G_{es}G_{rs,ci}^*]_-\|_2^2 = 1,
$$
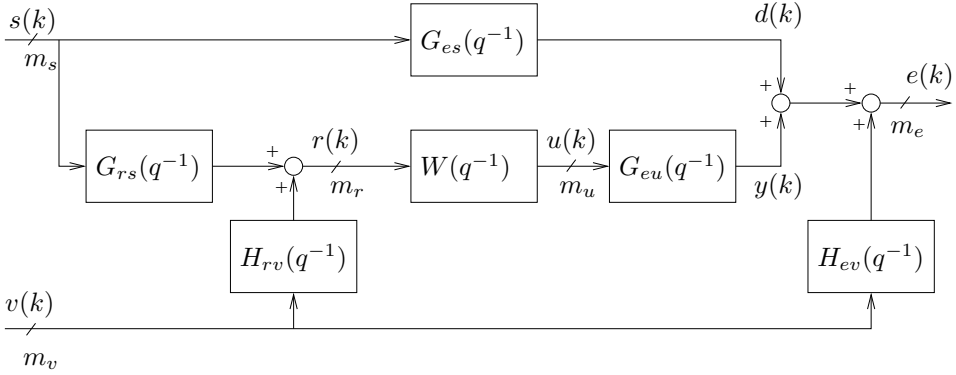
Figure 2.2: Block scheme of the multichannel feedforward active control problem, with m_s noise sources, m_r reference, m_u control and m_e error signals with measurement noise.

such that

$$||G_{es} + G_{eu}WG_{rs}||_2 \;=\; 1.$$

Note, that $||G_{es}||_2 \approx 1.1$ and because $||G_{es} + G_{eu}WG_{rs}||_2 < ||G_{es}||_2$ part of the disturbance is suppressed, but perfect cancellation is not achievable since $||G_{es} + G_{eu}WG_{rs}||_2 > 0$.

Example 2.6 *Consider the system of Example 2.2, then the optimal causal filter W according to Theorem 2.1 is given by*

$$W \;=\; -0.5 - 0.5q^{-1}.$$

Furthermore, the contributions to the cost-function are

$$||G_{es}G_{rs,ci}^{\perp *}||_2^2 \;=\; 0, \quad ||G_{eu,i}^{\perp *}G_{es}G_{rs,ei}^*||_2^2 \;=\; 1, \quad ||[G_{eu,i}^* G_{es}G_{rs,ci}^*]_-||_2^2 \;=\; 0,$$

such that

$$||G_{es} + G_{eu}WG_{rs}||_2 \;=\; 1.$$

Note, that $||G_{es}||_2 = \sqrt{2}$ and because $||G_{es} + G_{eu}WG_{rs}||_2 < ||G_{es}||_2$ part of the disturbance is suppressed, but perfect cancellation is not achievable since $||G_{es} + G_{eu}WG_{rs}||_2 > 0$.

Theorem 2.1 solves the feedforward optimal control problem without measurement noise of Problem 2.1, but we will use this theorem also in the following subsections to solve other control problems.

2.2.3 With measurement noise

Consider Figure 2.2, which illustrates the feedforward control problem with measurement noise $v(k) \in \mathbb{R}^{m_v}$. We assume, that $v(k)$, like $s(k)$, is a stationary

zero-mean white-noise signal which is independent of $s(k)$, and such that

$$
\mathbf{E}\left(\begin{bmatrix} s(k) \\ v(k) \\ 1 \end{bmatrix} \begin{bmatrix} s(l) \\ v(l) \end{bmatrix}^T\right) = \begin{bmatrix} I_{m_s}\delta_{kl} & 0 \\ 0 & I_{m_v}\delta_{kl} \\ 0 & 0 \end{bmatrix}. \tag{2.26}
$$

The transfer-function matrices $H_{rv} \in \mathcal{RH}_\infty^{m_r \times m_v}$ and $H_{ev} \in \mathcal{RH}_\infty^{m_e \times m_v}$ are noise shaping transfer-functions, which allow to consider various choices of measurement noise. For example, in case there is no noise on the reference signal $r(k)$, set $H_{rv} = 0$; for the case the noise on $r(k)$ and $e(k)$ is independent, set

$$
H_{rv} = \begin{bmatrix} H_{rv}^1 & 0_{m_r \times m_{v2}} \end{bmatrix}, \quad H_{ev} = \begin{bmatrix} 0_{m_e \times m_{v1}} & H_{ev}^2 \end{bmatrix}, \tag{2.27}
$$

with $H_{rv}^1 \in \mathcal{RH}_\infty^{m_r \times m_{v1}}$ and $H_{ev}^2 \in \mathcal{RH}_\infty^{m_e \times m_{v2}}$ with $m_{v1} + m_{v2} = m_v$.

From Figure 2.2, it is clear that the measurement noise $v(k)$ enters the system in exactly the same way as the disturbance source signal $s(k)$. Therefore, by making the following substitutions

$$
s(k) \leftarrow \begin{bmatrix} s(k) \\ v(k) \end{bmatrix}, \tag{2.28}
$$

$$
G_{rs} \leftarrow \begin{bmatrix} G_{rs} & H_{rv} \end{bmatrix}, \tag{2.29}
$$

$$
G_{es} \leftarrow \begin{bmatrix} G_{es} & H_{ev} \end{bmatrix}, \tag{2.30}
$$

the optimal feedforward control problem *without measurement noise*, Problem 2.1, equals the optimal feedforward control problem *with measurement noise*.

Using Theorem 2.1 together with the substitutions (2.28)-(2.30) the optimal feedforward controller for the case with measurement noise, is given by

$$
W = -G_{eu,o}^\dagger \left[G_{eu,i}^* G_{es}^{aug} G_{rs,ci}^{aug*} \right]_+ G_{rs,co}^{aug\dagger} \tag{2.31}
$$

with

$$
G_{es}^{aug} = \begin{bmatrix} G_{es} & H_{ev} \end{bmatrix}, \quad G_{rs}^{aug} = \begin{bmatrix} G_{rs} & H_{rv} \end{bmatrix}. \tag{2.32}
$$

and $G_{rs,co}^{aug} G_{rs,ci}^{aug}$ the outer-inner factorization of G_{rs}^{aug}.

In the case of *independent* measurement noise on $r(k)$ and $e(k)$, the expression (2.31) can be simplified in a way that the effect of the noise becomes more visible (note, that in the case of *dependent* measurement noise, the measurement noise can be considered as part of the disturbance). Let H_{rv} and H_{ev} be as in (2.27) and G_{rs}^{aug} and G_{es}^{aug} as in (2.32). Then, the outer-inner factorization of G_{rs}^{aug} is given by

$$
\begin{bmatrix} G_{rs} & H_{rv}^1 & 0_{m_r \times m_{v2}} \end{bmatrix} = G_{rs,co}^{aug} \begin{bmatrix} G_{rs,ci}^{1aug} & G_{rs,ci}^{2aug} & 0_{m_r \times m_{v2}} \end{bmatrix}.
$$

Using (2.31), yields

$$
W = -G_{eu,o}^\dagger \left[G_{eu,i}^* G_{es} G_{rs,ci}^{1aug*} \right]_+ G_{rs,co}^{aug\dagger} \tag{2.33}
$$

Note, that the solution is independent of H_{ev}^2, which means that measurement noise on $e(k)$ which is not correlated with the reference signal $r(k)$ cannot be suppressed,

as is intuitive. Furthermore, the structure of equation (2.33) is similar to the Causal Wiener filter (2.20) for the *no measurement* noise case, except for $G_{rs,ci}$ and $G_{rs,co}$ which are replaced by $G_{rs,ci}^{1aug}$ and $G_{rs,co}^{aug}$ respectively. We still have

$$G_{rs} = G_{rs,co}^{aug} G_{rs,ci}^{1aug}$$

but now, this factorization is not an outer-inner factorization of G_{rs} but is determined by the outer-inner factorization of G_{rs}^{aug}. Using the definition of G_{rs}^{aug} in (2.32), and the outer-inner factorization lemma, Lemma 2.2, we have

$$G_{rs,co}^{aug} G_{rs,co}^{aug*} = G_{rs}G_{rs}^* + H_{rv}^1 H_{rv}^{1*}.$$

Hence, the 'magnitude' of $G_{rs,co}^{aug}(e^{-j\omega})$ is increased, especially at the frequencies ω where the magnitude of $H_{rv}^1(e^{-j\omega})$ is 'large'. Since, the left-*inverse* of $G_{rs,co}^{aug}$ is a factor of the Causal Wiener filter (2.33), the control-action is reduced at those frequencies ω where $H_{rv}^1(e^{-j\omega})$ cannot be neglected. This is also intuitive, reducing the magnitude of W will reduce the contribution of the noise on $r(k)$ to the residual signal $e(k)$ and thus also to the cost-function (2.12).

2.2.4 With control effort weighting

In many practical problems, it is useful to include a control effort weighting in the cost-function, to increase robustness and/or prevent actuator saturation. Let us again, consider the feedforward control problem without measurement noise, illustrated by Figure 2.1. The cost-function we want to minimize now is given by

$$J(\mathcal{W}) = \text{tr}\mathbf{E}\left(e(k)e^T(k)\right) + \text{tr}\mathbf{E}\left(\tilde{u}(k)\tilde{u}^T(k)\right) \tag{2.34}$$

where the filtered control signal $\tilde{u}(k)$ is given by

$$\tilde{u}(k) = \widetilde{G}_{eu}u(k). \tag{2.35}$$

and $\widetilde{G}_{eu} \in \mathcal{RH}_\infty^{m_u \times m_u}$ a user chosen filter which gives the freedom to restrict the control signal more at specific frequencies. Note, that trade-off between minimizing the first and the second term in (2.34) is determined by the magnitude/gain of $\widetilde{G}_{eu}$.

Using Parseval's theorem, we can write (2.34) as

$$J(\mathcal{W}) = ||G_{es} + G_{eu}\mathcal{W}G_{rs}||_2^2 + ||\widetilde{G}_{eu}\mathcal{W}G_{rs}||_2^2, \tag{2.36}$$

$$= ||\begin{bmatrix} G_{es} \\ 0_{m_u \times m_u} \end{bmatrix} + \begin{bmatrix} G_{eu} \\ \widetilde{G}_{eu} \end{bmatrix}\mathcal{W}G_{rs}||_2^2, \tag{2.37}$$

$$= ||G_{es}^{aug} + G_{eu}^{aug}\mathcal{W}G_{rs}||_2^2. \tag{2.38}$$

with

$$G_{es}^{aug} = \begin{bmatrix} G_{es} \\ 0_{m_u \times m_u} \end{bmatrix}, \quad G_{eu}^{aug} = \begin{bmatrix} G_{eu} \\ \widetilde{G}_{eu} \end{bmatrix}.$$

The step from (2.36) to (2.37) can be verified using Definition 2.1 which defines the H_2-norm. Equation (2.37) is in the form of the cost-function (2.21) minimized

in the Causal Wiener filter theorem. Therefore, we directly can write the solution minimizing the cost-function (2.37) as:

$$W = -G_{eu,o}^{aug\dagger} \left[G_{eu,i}^{aug*} G_{es}^{aug} G_{rs,ci}^* \right]_+ G_{rs,co}^\dagger, \qquad (2.39)$$

where

$$G_{eu}^{aug} = G_{eu,i}^{aug} G_{eu,o}^{aug}$$

the inner-outer factorization of G_{eu}^{aug}.

Using the definition of G_{eu}^{aug}, we can write

$$\left[\begin{array}{c} G_{eu} \\ \widetilde{G}_{eu} \end{array} \right] = \left[\begin{array}{c} G_{eu,i}^{1aug} \\ G_{eu,i}^{2aug} \end{array} \right] G_{eu,o}^{aug}.$$

Together with the definition of G_{es}^{aug}, the expression of the Causal Wiener filter (2.39) can be simplified to

$$W = -G_{eu,o}^{aug\dagger} \left[G_{eu,i}^{1aug} G_{es} G_{rs,ci}^* \right]_+ G_{rs,co}^\dagger, \qquad (2.40)$$

which structure is similar to the Causal Wiener filter (2.20) for the case of *no control effort weighting*, except for $G_{eu,i}$ and $G_{eu,o}$ which are replaced by $G_{eu,i}^{1aug}$ and $G_{eu,o}^{aug}$ respectively. We still have

$$G_{eu} = G_{eu,i}^{1aug} G_{eu,o}^{aug}$$

but now, this factorization is not an inner-outer factorization of G_{eu} but determined by the inner-outer factorization of G_{eu}^{aug}. Using the definition of G_{eu}^{aug}, and the inner-outer factorization lemma, Lemma 2.1, we have

$$G_{eu,o}^{aug*} G_{eu,o}^{aug} = G_{eu}^* G_{eu} + \widetilde{G}_{eu}^* \widetilde{G}_{eu}.$$

This adjustment of the Causal Wiener filter in Theorem 2.1 is the dual form of the adjustment of the Causal Wiener in case of measurement noise on the reference signal in the previous subsection, Section 2.2.3.

Analogue to the discussion at the end of Subsection 2.2.3 we can conclude that the magnitude of the Causal-Wiener filter is reduced especially at the frequencies ω where the magnitude of $\widetilde{G}_{eu}(e^{-j\omega})$ is 'large'. This is of course also what is obtained by control-effort weighting.

These adjustments will also play an important role in the following chapters, in the development of robust control algorithms.

2.3   Feedback problem

Consider Figure 2.3 which illustrates the active control problem for feedback systems. Note, that the measurement noise $v(k)$ enters the system similar as $s(k)$, and thus $v(k)$ can be included in $s(k)$, as will be done below by augmenting G_{es} with H_{ev}.

In feedback systems, no reference signal $r(k)$ is available, and the control action $u(k)$ should be determined fully from the residual signal $e(k)$. The controller is
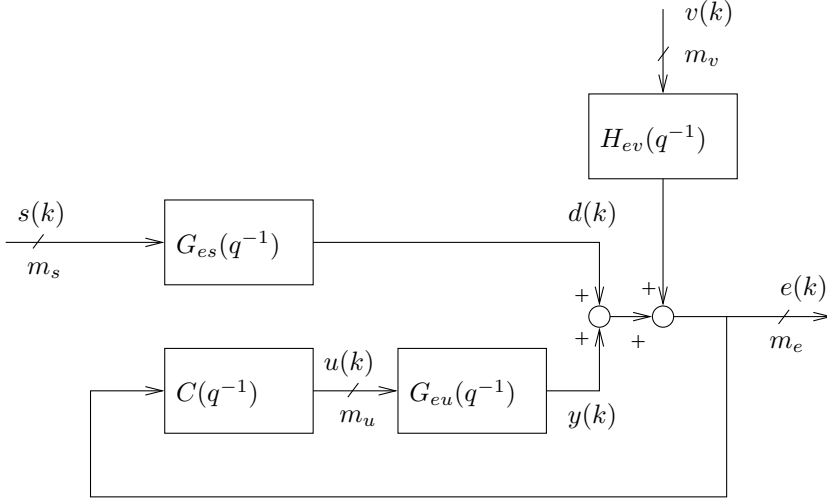
Figure 2.3: Block scheme of the multichannel feedback active control problem, with m_s noise sources, m_u control and m_e error signals with measurement noise.
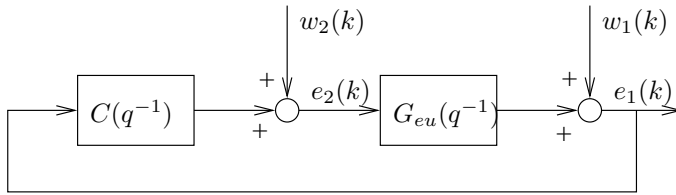
Figure 2.4: Block scheme of the closed-loop system for the definition of internal stability.

now indicated by $C(q^{-1}) \in \mathcal{RH}_p^{m_u \times m_e}$ and is not restricted to be asymptotically stable ($\mathcal{RH}_p^{m_x \times m_y}$ is the set of all proper rational transfer-functions matrices of size $m_x \times m_y$ with real coefficients). However, the closed-loop should be internally stable, which means that for bounded $s(k)$ and $v(k)$, all signals in Figure 2.3 should remain bounded. The formal definition of internal stability is given by the following definition using the closed-loop of Figure 2.4.

Definition 2.3 *Consider the feedback control loop of Figure 2.4 with controller $C \in \mathcal{RH}_p^{m_u \times m_e}$ applied to a plant $G_{eu} \in \mathcal{RH}_p^{m_e \times m_u}$, then the feedback control loop is internally stable, if the system*

$$\left[\begin{array}{c} e_1(k) \\ e_2(k) \end{array} \right] = \left[\begin{array}{cc} I_{m_e} & -G_{eu} \\ -C & I_{m_u} \end{array} \right] \left[\begin{array}{c} w_1(k) \\ w_2(k) \end{array} \right]$$

is bounded-input bounded-output (BIBO) stable in the l_2 sense.

Comparing Figure 2.3 and 2.4, we infer that the closed-loop in Figure 2.3 is a special
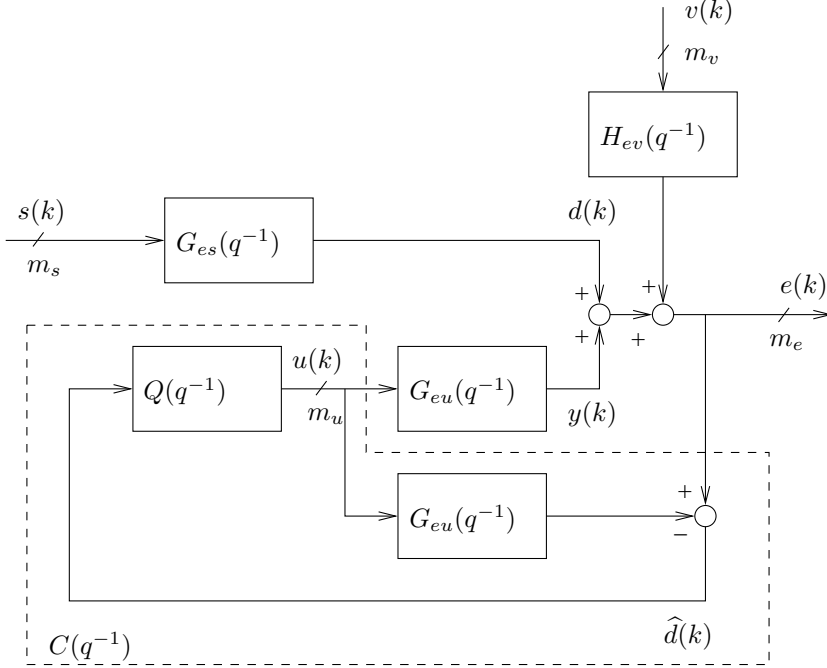
Figure 2.5: Block scheme of the Youla parameterization for stable $G_{eu}(q^{-1})$ with $Q(q^{-1}) \in \mathcal{RH}_\infty^{m_u \times m_e}$ the Youla parameter, also denoted by $W(q^{-1})$.

case (i.e. without $w_2(k)$) of the closed-loop in Figure 2.4. Thus, if C and G_{eu} are such that the closed-loop is of Figure 2.4 is internally stable, the closed-loop of Figure 2.3 is also internally stable.

The Youla parameterization [193, 194] provides a parameterization of all stabilizing controllers. For *stable* G_{eu}, the Youla parameterization of all stabilizing controllers is given by

$$ C = (I_{m_u} + QG_{eu})^{-1}Q, \qquad \forall Q \in \mathcal{RH}_\infty^{m_u \times m_e} \tag{2.41} $$

with $W \in \mathcal{RH}_\infty^{m_u \times m_e}$ the so-called Youla parameter. Figure 2.5 illustrates the structure of the controller (2.41). From this figure, we directly infer that the influence of $u(k)$ on $e(k)$ is subtracted, resulting in $\widehat{d}(k)$ which is independent of $u(k)$:

$$ \widehat{d}(k) \;=\; G_{es}s(k) + H_{ev}v(k). $$

Hence, the feedback problem is reformulated into a feedforward control problem and therefore we will design $Q(q^{-1})$ by solving a feedforward control problem. Since $Q(q^{-1})$ represents a feedforward controller, we will write $W(q^{-1})$ instead of $Q(q^{-1})$ in the sequal (so $Q(q^{-1}) = W(q^{-1})$). This strategy, to reformulate the feedback problem in a feedforward problem, is well known as the so-called Internal Model Control approach, see, e.g., [120]. Now, the residual signal $e(k)$ can be rewritten

in terms of W (which appears linearly!) as follows

$$e(k) = \left(\begin{bmatrix} G_{es} & H_{ev} \end{bmatrix} + G_{eu}W \begin{bmatrix} G_{es} & H_{ev} \end{bmatrix}\right) \begin{bmatrix} s(k) \\ v(k) \end{bmatrix} \qquad (2.42)$$

The resulting optimization problem in W is to minimize

$$\mathrm{tr}\mathbf{E}\left(e(k)e^T(k)\right) = \left|\left| \begin{bmatrix} G_{es} & H_{ev} \end{bmatrix} + G_{eu}W \begin{bmatrix} G_{es} & H_{ev} \end{bmatrix}\right|\right|_2^2 \qquad (2.43)$$

$$= ||G_{es}^{aug} + G_{eu}WG_{es}^{aug}||_2^2 \qquad (2.44)$$

with

$$G_{es}^{aug} = \begin{bmatrix} G_{es} & H_{ev} \end{bmatrix}.$$

subject to $W \in \mathcal{RH}_\infty^{m_u \times m_e}$. Again Theorem 2.1 provides the solution to this problem, which is given by

$$W = -G_{eu,o}^\dagger \left[G_{eu,i}^* G_{es}^{aug} G_{es,ci}^{aug*}\right]_+ G_{es,co}^{aug\dagger},$$

$$= -G_{eu,o}^\dagger \left[G_{eu,i}^* G_{es,co}^{aug}\right]_+ G_{es,co}^{aug\dagger}. \qquad (2.45)$$

with

$$G_{es}^{aug} = G_{es,co}^{aug} G_{es,ci}^{aug}$$

the outer-inner factorization of G_{es}^{aug}. Hence, the optimal internally stabilizing feedback controller is given by (2.41) with W given by (2.45). Note, that W only depends on the co-outer factor (i.e. the minimum-phase spectral factor) of G_{es}^{aug}, and thus delays in G_{es}^{aug} will not affect W.

Furthermore, note that in case $G_{es,co}^{aug} = I_{m_e}$ and G_{eu} contains one sample pure delay (e.g. due to discretization, which is always present in practical active control systems), we have $G_{eu,i}^*$ is *strictly* anti-causal and $[G_{eu,i}^* G_{es,co}^{aug}]_+ = [G_{eu,i}^*]_+ = 0$. Hence, in this case, we have $W = 0_{m_u \times m_e}$ and thus $C = 0_{m_u \times m_e}$, which means that the residual signal $e(k)$ cannot be reduced (but only increased) by any control action. This is also intuitive, since $G_{es}^{aug} = I_{m_e}$ means that the disturbance is white, and no information from the disturbance at k can be used to counteract the disturbance at $k+1$.

2.4 General feedforward/feedback problem

We are now ready to solve the general feedforward/feedback problem illustrated by Figure 2.6. The plant to be controlled is described by

$$e(k) = G_{es}s(k) + G_{eu}u(k) + H_{ev}v(k), \qquad (2.46)$$

$$r(k) = G_{rs}s(k) + G_{ru}u(k) + H_{rv}v(k), \qquad (2.47)$$

and the control-law by

$$u(k) = Cr(k). \qquad (2.48)$$
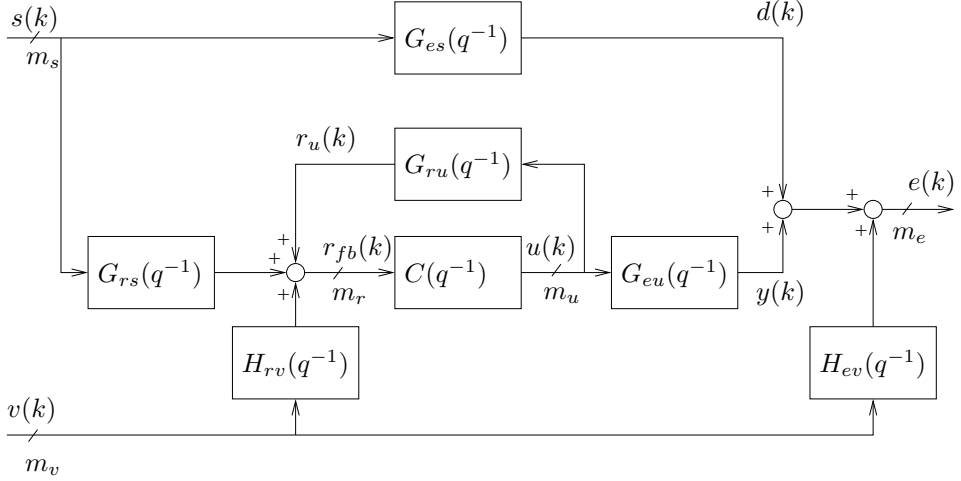
Figure 2.6: Block scheme of the general multichannel feedforward/feedback active control problem, with m_s noise sources, m_r reference, m_u control and m_e error signals with measurement noise.

As in Section 2.3, we will use the Youla parameterization to parameterize all internally stabilizing controllers by

$$C = (I_{m_u} + WG_{ru})^{-1}W, \qquad \forall W \in \mathcal{RH}_\infty^{m_u \times m_r}. \tag{2.49}$$

With this controller structure, the control problem is reformulated in a feedforward control problem which is solved by Theorem 2.1. To see this, note that the feedback of G_{ru} is canceled by (2.49), which yields

$$e(k) \;=\; \left(\begin{bmatrix} G_{es} & H_{ev} \end{bmatrix} + G_{eu}W \begin{bmatrix} G_{rs} & H_{rv} \end{bmatrix}\right) \begin{bmatrix} s(k) \\ v(k) \end{bmatrix}.$$

Hence, the trace of the covariance of $e(k)$ is given by

$$\text{tr}\mathbf{E}\left(e(k)e^T(k)\right) \;=\; ||G_{es}^{aug} + G_{eu}WG_{rs}^{aug}||_2^2, \tag{2.50}$$

where

$$G_{es}^{aug} \;=\; \begin{bmatrix} G_{es} & H_{ev} \end{bmatrix}, \quad G_{rs}^{aug} \;=\; \begin{bmatrix} G_{rs} & H_{rv} \end{bmatrix}.$$

Hence, minimizing (2.50) subject to $W \in \mathcal{RH}_\infty^{m_u \times m_e}$ is solved by Theorem 2.1 and its solution is given by

$$W = -G_{eu,o}^\dagger [G_{eu,i}^* G_{es}^{aug} G_{rs,ci}^{aug*}]_+ G_{rs,co}^{aug\dagger} \tag{2.51}$$

with

$$G_{rs}^{aug} \;=\; G_{rs,co}^{aug} G_{rs,ci}^{aug},$$

the outer-inner factorization of G_{rs}^{aug}.

Using the expression for the minimum value of the cost-function (2.22) in Theorem 2.1, we have for W given by (2.51)

$$\|G_{es}^{aug} + G_{eu} W G_{rs}^{aug}\|_2^2 =$$
$$\|G_{es}^{aug} G_{rs,ci}^{aug\perp*}\|_2^2 + \|G_{eu,i}^{\perp*} G_{es}^{aug} G_{rs,ci}^{aug*}\|_2^2 + \|[G_{eu,i}^{aug*} G_{es}^{aug} G_{rs,ci}^{aug*}]_-\|_2^2.$$

Note, that this value does not depend on G_{ru}.

2.5 State-space LQG solution

In the previous sections, we have solved various feedforward and feedback control problems using the Causal Wiener filter Theorem 2.1 and the Internal Model Control approach. The obtained expressions for the optimal controllers can be evaluated using models given in, e.g., transfer function or state-space notation. See Appendix B, which derives state-space expressions for the controllers.

However, to derive a state-space expression for the controllers, another approach is often followed in literature, see, e.g., [3, 24, 85], via *Kalman filtering* and *Linear Quadratic (LQ) control*. Let us consider, the general feedforward/feedback problem of Section 2.4, with no measurement noise, $v(k) = 0$ (or, the measurement noise $v(k)$ is contained in $s(k)$). Furthermore, let us assume a state-space realization of the system $\begin{bmatrix} G_{es} & G_{eu} \\ G_{rs} & G_{ru} \end{bmatrix}$ is given by

$$\left[\begin{array}{c} x(k+1) \\ \hline e(k) \\ r(k) \end{array} \right] = \left[\begin{array}{c|cc} A & B_s & B_u \\ \hline C_e & D_{es} & D_{eu} \\ C_r & D_{rs} & 0_{m_r \times m_u} \end{array} \right] \left[\begin{array}{c} x(k) \\ \hline s(k) \\ u(k) \end{array} \right] \tag{2.52}$$

with $x(k) \in \mathbb{R}^n$ the state, n the system order and A, B_s, B_u, C_e, C_r, D_{es}, D_{eu} and D_{rs} the state-space matrices, which are real-valued and of appropriate dimensions.

Then, the state-space controller minimizing the cost-function (2.12) is given by linear quadratic Gaussian (LQG) controller given by the following theorem.

Theorem 2.2 (LQG control [25]) *Given the state-space description* (2.52). *Further let*

- *the pair (C_r, A) be detectable and the pair (A, B_u) be stabilizable;*

- $R_{rs} = D_{rs} D_{rs}^T > 0$, $R_{eu} = D_{eu}^T D_{eu} > 0$;

- *the matrices*

$$\begin{bmatrix} A - \lambda I & B_u \\ C_e & D_{eu} \end{bmatrix}, \quad \begin{bmatrix} A - \lambda I & B_s \\ C_r & D_{rs} \end{bmatrix}$$

have full rank for all complex λ such that $|\lambda| = 1^2$.

---

[2] This condition is equivalent to the condition in Theorem 2.1 that $G_{eu}(q^{-1})$ and $G_{rs}(q^{-1})$ do not loose rank for all $|q| = 1$

For ease of notation, we also define

$$Q_s = B_s B_s^T, \quad Q_e = C_e^T C_e, \quad S_{rs} = B_s D_{rs}^T, \quad and \quad S_{eu} = C_e^T D_{eu}.$$

Then, there exist unique $P_{rs} = P_{rs}^T > 0$ and $P_{eu} = P_{eu}^T > 0$ which are stabilizing solutions to the following Riccati equations

$$P_{rs} = AP_{rs}A^T - (AP_{rs}C_r^T + S_{rs})(C_r P_{rs} C_r^T + R_{rs})^{-1}(AP_{rs}C_r^T + S_{rs})^T + Q_s,$$
$$P_{eu} = A^T P_{eu} A - (A^T P_{eu} B_u + S_{eu})(B_u^T P_{eu} B_u + R_{eu})^{-1}(A^T P_{eu} B_u + S_{eu})^T + Q_e.$$

Let

$$F_{eu} = (B_u^T P_{eu} B_u + R_{eu})^{-1}(B_u^T P_{eu} A + S_{eu}^T), \qquad (2.53)$$
$$F_{eu}^o = (B_u^T P_{eu} B_u + R_{eu})^{-1}(B_u^T P_{eu} B_s + D_{eu}^T D_{es}), \qquad (2.54)$$

and

$$K_{rs} = (AP_{rs}C_r^T + S_{rs})(C_r P_{rs} C_r^T + R_{rs})^{-1}, \qquad (2.55)$$
$$K_{rs}^o = (F_{eu} P_{rs} C_r^T + F_{eu}^o D_{rs}^T)(C_r P_{rs} C_r^T + R_{rs})^{-1}, \qquad (2.56)$$

Then, the optimal state-space controller, which minimizes (2.12) *is given by*

$$\left[\begin{array}{c} \widehat{x}(k+1|k) \\ \hline u(k) \end{array} \right] = \left[\begin{array}{c|c} A + B_u K_{rs}^o C_r - B_u F_{eu} - K_{rs} C_r & B_u K_{rs}^o - K_{rs} \\ \hline F_{eu} - K_{rs}^o C_r & -K_{rs}^o \end{array} \right]$$
$$\cdot \left[\begin{array}{c} \widehat{x}(k|k-1) \\ \hline y(k) \end{array} \right] \qquad (2.57)$$

with $\widehat{x}(k|k-1)$ the estimate of $x(k)$ given the measurements $y(i)$, $i \le k - 1$.

Proof: For the proof see [25] and the technical report [26]. $\square$

In Appendix B we proof that this controller is equivalent to the controller obtained in Section 2.4 by the Causal Wiener filter and IMC.

2.6   Robustness problem

The controllers of the previous sections are designed based on the underlying assumption that the model equals the system. This assumption is known as the *certainty equivalence principle*. In practice, this assumption is never satisfied, because the model is always an approximation of the plant. Therefore, the controller which is designed to be optimal for the model, will (usually) not be optimal for the plant. In case of feedback systems, the closed-loop may even be unstable.

Basically, there are two approaches, which may also be combined, to deal with this problem:

- Control-relevant identification: the *control* cost-function is minimized during model identification;

- Robust controller design based on model uncertainty description.

Control-relevant identification reduces the effect of model errors, since the model necessary for controller design is accurately estimated only at the frequency bands where the controller should be active. This approach has been taken in Chapter 3.

The other approach is to design the controller such that it is robust (in the sense of performance and stability robustness) for model uncertainty. This approach has been taken in Chapter 4.

Chapter 3

Nominal controller estimation

3.1 Introduction

In the previous chapter, the design equations for the H_2 optimal feedforward and feedback controllers have been given. In this chapter, we consider the question to estimate these optimal controllers given measured data, which is the nominal controller estimation problem.

The standard approach is to estimate a full model of the system, i.e. estimate $G_{es}, G_{eu}, G_{rs}, G_{ru}$ and H_{ev}, H_{rv}, and use these models in the optimal controller expressions. The idea is that, a 'good' approximation of the models also yields a 'good' approximation of the controller. However, this is not necessarily the case, since the optimal controller expression may be very sensitive to (particular) model errors.

Another approach to calculate the controller, is by solving a control-relevant identification problem using the measured-data. A control-relevant identification problem is an identification problem where the cost function to be minimized for estimating the parameters, equals the cost function to be minimized by the *controller*. Hence, solving a control-relevant identification problem to determine the controller or a factor of the controller, *explicitly* minimizes the control cost function and thus usually gives better performance than the standard model based approach.

Control-relevant identification problems often arise in closed-loop identification problems, for example in the famous windsurfer approach proposed in [96] (also see [33]). Here, we will propose an *open-loop* (feedforward) control-relevant identification method to estimate the Causal Wiener filter.

The chapter is organized as follows. Section 3.2 discusses the standard model based controller design procedure. Section 3.3 briefly discusses the choice of prediction error or subspace model identification methods. Section 3.4 discusses two modeling approaches, a one step and a two step approach. In the two step approach two sets of measurements are collected, one with the disturbance signal turned off and the control signal turned on, and one the other way around. Then, the main

contribution of the chapter is given in Section 3.5 which proposes a control-relevant identification approach to estimate the H_2 optimal controller. The method has been developed for PEM as well as SMI, which differs slightly as will be explained in the sequel. Section 3.6 illustrates the estimation of the controller by experiments with an 1-dimensional acoustical duct and a 4×4 vibrating plate system.

The content of this chapter was published before at the ISMA conference 2002 in Leuven, [61], and was also part of the publication in Control Engineering Practice [62].

3.2   Standard model based controller design

Consider the basic feedforward control problem of Section 2.2.1 to minimize the cost-function

$$J(W) = \mathrm{tr}\mathbf{E}\left(e(k)e^T(k)\right) = ||G_{es} + G_{eu}WG_{rs}||_2^2. \tag{3.1}$$

This problem was solved by the Causal Wiener filter

$$W = -G_{eu,o}^{\dagger}[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ G_{rs,co}^{\dagger} \tag{3.2}$$

in Theorem 2.1.

Using identification methods implemented in e.g. the MATLAB Identification Toolbox [105], the Subspace Model Identification toolbox [77] or the SLICOT software package [159], models of G_{es}, G_{rs} and G_{eu} can be identified, which are denoted by $\widehat{G}_{es}$, $\widehat{G}_{rs}$ and $\widehat{G}_{eu}$ respectively. Then the Causal Wiener filter can be estimated by replacing the systems by their estimates, which yields

$$\widehat{W} = -\widehat{G}_{eu,o}^{\dagger}[\widehat{G}_{eu,i}^* \widehat{G}_{es} \widehat{G}_{rs,ci}^*]_+ \widehat{G}_{rs,co}^{\dagger}. \tag{3.3}$$

In this step, it is assumed that the models are accurate and provide an exact description of the true systems, which is called the principle of *certainty equivalence*. However, this assumption is easily violated, due to

- Measurement noise;

- The use of a finite number of samples (such that measurement noise does not vanish by averaging);

- Unmodeled dynamics (the true system is not contained in the model set);

- Nonlinear behavior of the system (e.g. saturation);

- Variations in the system (e.g. due to temperature variations);

- Sub-optimality of the identification method (e.g. the identification of an Output Error model is a non-linear optimization problem and can only be approximated by iterative nonlinear optimization methods which may converge to local optima).

Hence, the estimated controller (3.3) will not minimize the cost-function (3.1), and the question arises whether we can determine a controller using the measured data which better approaches the minimum of (3.1). This problem will be considered in the following section.

The extension to the feedback systems of Section 2.3 and 2.4 is very similar. For the internal models $\widehat{G}_{eu}$ and $\widehat{G}_{ru}$ are used in for the feedback system of Section 2.3 and 2.4 respectively. Note, that the estimated internal models which are corrupted by uncertainty are not parameterizing all stabilizing controllers anymore (cf. (2.41) and (2.49)), and thus the closed loop may be *unstable*. We will refer to this problem in Section 4.5 on page 90 where we discuss the stability robustness of the closed loop when using the robust Cautious Wiener filter.

3.3 Choice of the model identification method

There are basically two main stream approaches to parametric model identification using input/output data: prediction error modeling (PEM) and subspace model identification (SMI). PEM methods yield a transfer function model, whereas SMI methods yield a state-space model of the system. In the following, we briefly discuss some differences between PEM and SMI and motivate our preference for SMI to identify models for ANVC. We only discuss the identification in open-loop configuration, where the control input is not determined by feedback from the measured output. For closed-loop system identification, we refer to, e.g., [175].

3.3.1 Prediction error model identification

In PEM identification, the transfer-function from the inputs to each output is identified separately. Hence, common dynamics which determine multiple outputs will be modeled multiple times. In SMI there is no distinction between the SISO and the MIMO case. This is a main difference between PEM identification and SMI. Therefore, let us consider the SISO case when focusing on PEM. In PEM, the system to be identified has the following structure (see e.g. [104])

$$y(k) = G(q^{-1}, \theta)u(k) + H(q^{-1}, \theta)v(k) \tag{3.4}$$

with $u(k)$ the deterministic input and $v(k)$ the unmeasured zero-mean white-noise input and

$$
\begin{array}{rcccl}
G(q^{-1}, \theta) & = & \frac{A^g(q^{-1}, \theta)}{B^g(q^{-1}, \theta)} & = & \frac{a_0^g + a_1^g q^{-1} + \cdots + a_{n_g}^g q^{-n_g}}{1 + b_1^g q^{-1} + \cdots + b_{n_g}^g q^{-n_g}}, \\
H(q^{-1}, \theta) & = & \frac{A^h(q^{-1}, \theta)}{B^h(q^{-1}, \theta)} & = & \frac{a_0^h + a_1^h q^{-1} + \cdots + a_{n_h}^h q^{-n_h}}{1 + b_1^h q^{-1} + \cdots + b_{n_h}^h q^{-n_h}}
\end{array}
$$

and the parameter-vector

$$\theta = \begin{bmatrix} a_0^g & \cdots & a_{n_g}^g & b_1^g & \cdots & b_{n_g}^g & a_0^h & \cdots & a_{n_h}^h & b_1^h & \cdots & b_{n_h}^h \end{bmatrix}^T.$$

which needs to be estimated.

The one-step ahead predictor is given by

$$\widehat{y}(k, \theta) = (1 - H^{-1}(q^{-1}, \theta))y(k) + H^{-1}(q^{-1}, \theta)G(q^{-1}, \theta)u(k) \tag{3.5}$$

and the (one-step ahead) prediction error by

$$\epsilon(k,\theta) = y(k) - \widehat{y}(k,\theta). \tag{3.6}$$

It is constrained that $H(q^{-1},\theta)$ is *minimum-phase* which means that $H^{-1}(q^{-1},\theta)$ is stable. Then, the objective is to estimate θ by minimizing

$$J(\theta) = \frac{1}{N} \sum_{k=1}^{N} \epsilon^2(k,\theta) \tag{3.7}$$

where N is the number of samples which is available for identification. However, in general, $\widehat{y}(k,\theta)$, and thus $\epsilon(k,\theta)$, is a non-linear function of θ, and the cost-function (3.7) will be non-convex in θ which highly complicates solving the minimization problem. In some special cases the structure of G and H is such that $\widehat{y}(k,\theta)$ is linear in θ such that the minimization problem is quadratic in θ and can be solved by linear least squares. This happens e.g. in case of the ARX (auto-regressive exogenous input) model structure in which

$$G(q^{-1},\theta) \;=\; \frac{A^g(q^{-1},\theta)}{B^g(q^{-1},\theta)}, \qquad H(q^{-1},\theta) \;=\; \frac{1}{B^g(q^{-1},\theta)}.$$

In all other cases where $\widehat{y}(k,\theta)$ is nonlinear in θ, the minimization will be done by using iterative nonlinear optimization algorithms, such as Gauss-Newton or Levenberg-Marquardt methods, which may converge to *local minima*. In the implementation provided by the MATLAB Identification toolbox [105] it is also possible to include the constraint that G should be stable, which is implemented by means of a projection on the set of stable systems in every iteration step. This will be useful in Subsection 3.5.1 where a control-relevant identification method is proposed for using PEM.

3.3.2 Subspace Model Identification

In SMI the objective is to estimate the state-space matrices (A, B, C, D, K) and the covariance R of the following state-space system, given in innovation form:

$$x(k+1) = Ax(k) + Bu(k) + Kv(k) \tag{3.8}$$
$$y(k) = Cx(k) + Du(k) + v(k) \tag{3.9}$$

with $\mathrm{E}[v(k)v^T(l)] = R\delta_{kl}$. To guarantee minimum-phase relation between $v(k)$ and $y(k)$, we constrain that the eigenvalues of $A - KC$ are inside the unit-circle in the complex-plane. The state-space system (3.8),(3.9) can be written in the form of (3.4) by setting

$$G(q^{-1},\theta) \;=\; C(zI-A)^{-1}B + D,$$
$$H(q^{-1},\theta) \;=\; C(zI-A)^{-1}K + I.$$

The state-space description is unique up to a similarity transformation, which means that for nonsingular T, the same system is described by $(TAT^{-1}, TB, CT^{-1}, D, TK)$. Hence, the problem is to find (A, B, C, D, K, R) up to a similarity transformation.

Note, that $y(k)$ is (again) nonlinear in the coefficients of A, B, C, D and K. But if the state $x(k)$ would be known we obtain linear relations in (A, B, C, D), and thus (A, B, C, D) can be solved by linear least squares. Then, to estimate K and R the following signals are calculated using $u(k), x(k), y(k)$ and the estimated (A, B, C, D):

$$
\begin{array}{rcll}
w(k) & = & x(k+1) - Ax(k) - Bu(k), & k = 1, \cdots N \\
v(k) & = & y(k) - Cx(k) - Du(k), & k = 1, \cdots N.
\end{array}
$$

Using these sequences, estimates of

$$
\begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} = \mathbf{E} \left(\begin{bmatrix} w(k) \\ v(k) \end{bmatrix} \begin{bmatrix} w(k) \\ v(k) \end{bmatrix}^T \right)
$$

can be made. Then K, such that the eigenvalues of $A - KC$ are within the unit circle, is calculated by solving the stabilizing solution $P = P^T > 0$ to the discrete Riccati equation

$$
P = APA^T - (APC^T + S)(CPC^T + R)^{-1}(APC^T + S)^T + Q,
$$

and

$$
K = (APC^T + S)(CPC^T + R)^{-1}.
$$

There are several subspace identification algorithms proposed in the literature, see e.g. [177, 181] for further details. A common property of these algorithms is, that they are based on numerically reliable algorithms such as the QR decomposition and the SVD. In this thesis we will use the PO-MOESP[1] algorithm [181].

3.3.3 A simulation example of identifying a vibrating plate model

To illustrate the identification of systems for ANVC using the PEM and SMI methods, we performed a series of simulation experiments using a 20th order 4×4 vibrating plate model. This model has once been identified using measured data from a real vibrating plate system. We have chosen to perform simulation experiments, such that a series of system identification tests can be performed under exactly the same conditions with only other realizations of the measurement noise. The measurement noise is chosen to be a zero-mean white-noise process, which variance is such that the signal-to-noise ratio (SNR) is 20dB. In total 100 identification experiments are performed. In each experiment the excitation signal is chosen to be zero-mean white noise process of $N = 4000$ samples. The output is calculated using the 20th order 4×4 vibrating plate model, and is distorted by the measurement noise.

Using PEM a transfer-function model with Output-Error structure has been identified. To save computation time, the computation of the covariance matrix of

---

[1]MOESP stands for Multivariable Output Error State sPace, the preposition PO stands for Past Outputs and refers to the construction of the Instrumental Variable to cancel state- and measurement-noise by projection.

Table 3.1: Comparison of PEM-OE (MATLAB Identification toolbox implementation) and PO-MOESP (SLICOT Optimized Fortran implementation) for identification of a 4×4 vibrating plate of order $n = 20$ and signal-to-noise ratio $SNR = 20$dB. Results are averaged over 100 experiments with different noise realizations.

Property	PEM-OE	PO-MOESP
Number of parameters:	656	576
Calculation time (s):	319	53
Signal to simulation error ratio (dB):	15.8	20.1
VAF channel 1 (%):	94.3	99.0
VAF channel 2 (%):	96.9	99.0
VAF channel 3 (%):	97.4	99.0
VAF channel 4 (%):	97.6	99.0
% experiments signal to error ratio >15dB:	83	100

the parameter vector has been turned off. The number of parameters of the 4×4 transfer-function matrix, with each transfer-function of order 20, is 656. The subspace identification algorithm, more precisely the PO-MOESP algorithm, see [181], estimates a 4×4 state-space model of order 20, which has 576 parameters (counting all elements of the state-space matrices A, B, C, D). The computation time and the model accuracy have been averaged over the 100 experiments and given in Table 3.1. The model accuracy has been expressed in the signal to simulation error ratio (in dB) and the Variance Accounted For (VAF, in %). The signal to simulation error (SER) is defined as

$$SER(y_{measured}, y_{simulated}) \triangleq 10^{10} \log \left(\frac{\sum\limits_{i=1}^{m_y} \text{var}(y_{measured,i})}{\sum\limits_{i=1}^{m_y} \text{var}(y_{measured,i} - y_{simulated,i})} \right) \text{(dB)}$$

with $y_{measured,i}$ and $y_{simulated,i}$ the i-th channel of the measured output and the simulated output, obtained by using the model, respectively. The VAF is defined for each channel $i = 1, \cdots, m_y$ separately:

$$VAF(y_{measured,i}, y_{simulated,i}) \triangleq \left(1 - \frac{\text{var}(y_{measured,i} - y_{simulated,i})}{\text{var}(y_{measured,i})} \right) \times 100\%$$

We observe, that the model obtained by the PO-MOESP algorithm is accurately identifying the system (simulation error is at the level of the measurement noise), whereas the PEM-OE yields a less accurate model and its computation takes much more time (about 5 minutes). We have to note, that for the PO-MOESP algorithm a fast implementation in optimized Fortran code has been used, which is contained in the SLICOT package [159]. It may be clear that we prefer to use this fast and

accurate implementation of the PO-MOESP algorithm also in the sequel of this thesis.

3.4 One step and two step modeling approach

Let us now have a closer look at the identification of the models. We will distinguish two approaches: (1) a one step approach where the full model of the system which contains all transfer-functions (explicitly in transfer-function description such as the left-coprime factorization or the common denominator form or in a state-space description) which are necessary to design the controller, and (2) a two step approach where in each step a part of the system is estimated: one step uses measured data, if possible, under no disturbance condition $s(k) = 0$, and the other step used measured data under no control condition $u(k) = 0$. The two approaches are explained in more detail in the sequel of this subsection. We also explain that the second approach usually yields a more accurate model description of the system than the first approach, and thus the controller calculated using this model usually gives better performance.

3.4.1 Full model identification using one experiment

Let us consider the general feedforward/feedback configuration of Section 2.4 on page 45. Other, more special, cases can be considered simply by setting specific transfer-functions to zero. The system to be identified is given by

$$
\left[\begin{array}{c} e(k) \\ r(k) \end{array} \right] = \left[\begin{array}{c} G_{eu} \\ G_{ru} \end{array} \right] u(k) + \left[\begin{array}{cc} G_{es} & H_{ev} \\ G_{rs} & H_{rv} \end{array} \right] \left[\begin{array}{c} s(k) \\ v(k) \end{array} \right].
\tag{3.10}
$$

Usually $s(k)$ cannot be measured and therefore can be considered together with the measurement noise $v(k)$ as part of the unmeasured noise. Using SMI or PEM methods the transfer-functions can be identified in state-space or rational polynomial description as described in the previous paragraph using the input/output data

$$
\{u(k), \left[\begin{array}{c} e(k) \\ r(k) \end{array} \right] \}_{k=1}^N
$$

with $u(k)$ persistently exciting. The condition that $u(k)$ should be persistently exciting means that, loosely speaking, all dynamics which needs to be modeled should also be excited by $u(k)$. Usually $u(k)$ is chosen to be low-pass filtered zero-mean Gaussian white-noise, with the cut-off frequency chosen such that the Nyquist criterion is satisfied. When using identification methods based on frequency domain data, it is advised to use pseudo-random multi-sines such that leakage can be prevented [138].

When using a subspace identification method, the main advantage of this full model identification approach, is that *one compact state-space* model of the system is identified. This state-space model can be directly used to design an H_2 optimal controller, e.g. according to the design equations of Section 2.5.

However, an important disadvantage of this approach is that the model accuracy can be poor. This is illustrated in the simulation example of Section 3.4.3 on page 59 and the experiments on the acoustical duct system in Section 3.6.1 on page 64. In these examples, it is observed that the model accuracy of G_{eu} and G_{ru} is significantly less accurate then in the two step approach discussed in Section 3.4.2 below. This two step method requires that the disturbance source can be turned off, such that $s(k) = 0$.

3.4.2   Separate model identification using two experiments

In many practical situations the disturbance source can be turned off, $s(k) = 0$, during start-up time. For example in an air-conditioning system, the van can be turned off for some time, such that no disturbing sound is generated. Also for applications in e.g. airplanes, experiments can be performed when the airplane is on ground with the engines turned off, such that no disturbing sound is generated. In practice, note that the system conditions will in general be different in flight condition, but is usually a less important factor than the disturbance which distorts the modelling of G_{eu} and G_{ru} in flight condition.

The property that the disturbance can be turned off, enables to prevent the problem of interference of $s(k)$ and $u(k)$ in the identification of $\begin{bmatrix} G_{eu} \\ G_{ru} \end{bmatrix}$ and $\begin{bmatrix} G_{es} \\ G_{rs} \end{bmatrix}$ respectively. The following identification steps are proposed:

1. Identify $\begin{bmatrix} G_{eu} \\ G_{rs} \end{bmatrix}$ and, if necessary, the noise models $\begin{bmatrix} H_{ev} \\ H_{rv} \end{bmatrix}$, using the following input/output data under the condition that $s(k) = 0$:

$$\{u(k), \begin{bmatrix} e(k) \\ r(k) \end{bmatrix}\}_{k=1}^N, \qquad s(k) = 0, \ k = 1, \cdots, N$$

    and $u(k)$ should be persistently exciting.

2. Identify the minimum-phase spectral factor of $\begin{bmatrix} G_{es} \\ G_{rs} \end{bmatrix}$ using the following output data under the condition that $u(k) = 0$:

$$\{\begin{bmatrix} e(k) \\ r(k) \end{bmatrix}\}_{k=1}^N, \qquad u(k) = 0, \ k = 1, \cdots, N.$$

Note, that in the second step the *minimum-phase spectral factors* of G_{es} and G_{rs} are identified because usually $s(k)$ cannot be measured and only the output data is available. Also note that the effect of the measurement noise $v(k)$ is present in the measurements of $e(k)$ and $r(k)$, such that the noise-models H_{ev} and H_{rv} are (implicitly) modeled together with the spectral factors of G_{es} and G_{rs}.

The identification of the minimum-phase spectral factor, using output-only data, can also be done using SMI methods, see e.g. [111].

3.4.3   A simulation example

Consider the following state-space realization of the general ANVC system (3.10) on page 57:

$$x(k+1) = Ax(k) + B_u u(k) + B_s s(k)$$

$$\begin{bmatrix} e(k) \\ r(k) \end{bmatrix} = \begin{bmatrix} C_e \\ C_r \end{bmatrix} x(k) + \begin{bmatrix} D_{eu} & D_{es} \\ D_{ru} & D_{rs} \end{bmatrix} \begin{bmatrix} u(k) \\ s(k) \end{bmatrix} + \begin{bmatrix} 0.1 v_e(k) \\ 0.1 v_r(k) \end{bmatrix}$$

where

$$A = \begin{bmatrix} 0.9 & 0.1 \\ -0.1 & 0.9 \end{bmatrix}, \quad B_u = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad B_s = \begin{bmatrix} 0.2 \\ 0.2 \end{bmatrix}$$

$$C_e = \begin{bmatrix} 1 & 1 \end{bmatrix}, \quad C_r = \begin{bmatrix} -0.5 & 1.5 \end{bmatrix}$$

$$D_{eu} = D_{es} = D_{ru} = D_{rs} = 0,$$

and $s(k)$, $v_e(k)$ and $v_r(k)$ all independent zero-mean white noise processes with unit variance.

The one step approach of Section 3.4.1 and the two step approach of Section 3.4.2 have been used to estimate the system. In the one step approach the PO-MOESP algorithm has been used together with the data batch

$$\{u(k), \begin{bmatrix} e(k) \\ r(k) \end{bmatrix} \}_{k=1}^{N}$$

which estimates the A, B_u, C_e, C_r, D_{eu} and D_{ru} state-space matrices as well as the covariance matrices Q, R and S of the noise on the state and the output due to $s(k)$ (and $v(k)$). Using these estimates of Q, R and S together with A, C_e and C_r the Kalman gain K can be computed to obtain a full model description in innovation form, cf. Section 3.3.2 on page 54.

In the two step approach $s(k)$ has been set to zero: $s(k) = 0, \; k = 1, \cdots, N$ for the identification of G_{eu} and G_{ru}. Again using the PO-MOESP algorithm, A, B_u, C_e, C_r, D_{eu} and D_{ru} have been estimated, but the Q, R and S matrices do not need to be estimated since $s(k) = 0$.

The identification experiments are repeated 100 times with in each experiment different realizations of $s(k)$ (only in the one step approach), $v(k)$ and the zero-mean white noise excitation signal $u(k)$, which has also unit variance.

Table 3.2 shows the average H_2 norm of the model errors of the models $\widehat{G}_{eu}$ and $\widehat{G}_{ru}$ obtained by the two methods for different numbers of samples. We clearly observe, that the identification of G_{eu} and G_{ru} is much more accurate for the two step approach where G_{eu} and G_{ru} are identified using measured data which is not distorted by $s(k)$.

Table 3.2: H_2 norm of the model errors in $\widehat{G}_{eu}$ and $\widehat{G}_{ru}$ obtained by the one step approach where $s(k) \neq 0$ and the two step approach where $s(k) = 0$, for different numbers of samples. The H_2 norms are averaged over 100 simulation experiments.

Number of	$\|G_{eu} - \widehat{G}_{eu}\|_2$		$\|G_{ru} - \widehat{G}_{ru}\|_2$	
samples	One step	Two step	One step	Two step
$N = 100$	0.367	0.0253	0.219	0.0250
$N = 500$	0.145	0.0090	0.092	0.0100
$N = 1000$	0.105	0.0065	0.066	0.0070

3.5   Control-relevant identification

3.5.1   Prediction error modeling

In control-relevant identification, the key idea is to identify the model by minimizing the control cost-function. Here, we discuss the control-relevant identification using PEM, and more specifically, with Output-Error (OE) model structure, abbreviated as PEM-OE. In the PEM-OE tool as implemented in the MATLAB Identification toolbox the constraint that the model should be stable can be included. As will be explained in the sequel, this is a nice feature which yields that the explicit identification of the spectral factor of G_{rs} (and G_{es}) is not necessary. In SMI this stability constraint cannot be taken into account, and an estimate of the spectral factor of G_{rs} is necessary in the control-relevant identification method using SMI algorithms, see Section 3.5.2. However, this disadvantage of SMI can be neglected compared with the advantage of SMI over PEM as illustrated by Table 3.1 on page 56. Therefore, in this subsection we derive the control-relevant identification method using PEM to explain the concept. In the next subsection we show how the control-relevant identification method can be adjusted for SMI methods.

In both approaches (PEM and SMI), we assume that G_{eu} and G_{ru} are perfectly modeled (i.e. no model errors), which means that the control-relevant identification cannot compensate for model errors in G_{eu} and G_{ru}, but only in G_{rs} and G_{es}. In the following chapter on robustness, Chapter 4, we will deal with uncertainty in G_{eu} and G_{ru}. In fact, we will only adjust the second step of the "Separate model identification using two experiments" in the previous section.

The key observation is that the factor $[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ G_{rs,co}^\dagger$ of the Causal Wiener filter (3.2), can be solved from

$$[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ G_{rs,co}^\dagger = \arg \min_{X \in \mathcal{RH}_\infty^{m_{eu} \times m_r}} J_{id}(X), \qquad (3.11)$$

with

$$J_{id}(X) = \mathrm{tr}\mathbf{E}\left((G_{eu,i}^* d(k) - Xr(k))(G_{eu,i}^* d(k) - Xr(k))^T \right) \qquad (3.12)$$

which will be proven by Lemma 3.1 below.

Since, the constrained optimization problem to solve $X = [G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ G_{rs,co}^\dagger$ is exactly the optimization problem considered in

PEM-OE identification with the constraint that the model should be stable, the factor $[G^*_{eu,i}G_{es}G^*_{rs,ci}]_+G^\dagger_{rs,co}$ can be estimated using PEM-OE with the following input/output data

$$\{r(k), G^*_{eu,i}d(k)\}^N_{k=1} \tag{3.13}$$

with $d(k) = e(k)$ measured under the condition that $u(k) = 0$. Note, that the sequence $\{G^*_{eu,i}d(k)\}^N_{k=1}$ is generated by filtering *backward* in time from $k = N$ to $k = 1$ because $G^*_{eu,i}$ is unstable.

The following lemma shows that the identification of the factor $X = [G^*_{eu,i}G_{es}G^*_{rs,ci}]_+G^\dagger_{rs,co}$ is indeed minimizing J_{id} subject to the stability constraint and that the identification is control-relevant. By control-relevant we mean, that in the limit $N \to \infty$ (under reasonable conditions, see [104]) the cost-function to be minimized in the identification (3.12) equals the cost-function to be minimized during control (3.1).

Lemma 3.1 *Let $G_{eu} \in \mathcal{RH}^{m_e \times m_u}_\infty$ be known, let*

$$W(X) \;=\; -G^\dagger_{eu,o}X, \qquad with \; X \in \mathcal{RH}^{m_{eu} \times m_r}_\infty$$

then minimizing (3.1) over all $X \in \mathcal{RH}^{m_{eu} \times m_r}_\infty$ is equivalent to minimizing (3.12) over all $X \in \mathcal{RH}^{m_{eu} \times m_r}_\infty$.

Proof: Using the definition of the H_2-norm (Definition 2.1 on page 34) and $W = -G^\dagger_{eu,o}X$ the expression for the control cost-function (3.1) is written as

$$J = \frac{1}{2\pi}\mathrm{tr}\int\limits_{-\pi}^{\pi} \big(G_{es} - G_{eu,i}XG_{rs}\big)\big(G_{es} - G_{eu,i}XG_{rs}\big)^* d\omega$$

Because $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ and $\begin{bmatrix} G_{eu,i} & G^\perp_{eu,i} \end{bmatrix}$ is unitary, this can be rewritten as

$$J =$$

$$= \frac{1}{2\pi}\mathrm{tr}\int\limits_{-\pi}^{\pi} (G_{es} - G_{eu,i}XG_{rs})^* \begin{bmatrix} G_{eu,i} & G^\perp_{eu,i} \end{bmatrix} \begin{bmatrix} G^*_{eu,i} \\ G^{\perp *}_{eu,i} \end{bmatrix} (G_{es} - G_{eu,i}XG_{rs}))d\omega$$

$$= \frac{1}{2\pi}\mathrm{tr}\int\limits_{-\pi}^{\pi} (G^*_{eu,i}G_{es} - XG_{rs})^*(G^*_{eu,i}G_{es} - XG_{rs})d\omega + \frac{1}{2\pi}\mathrm{tr}\int\limits_{-\pi}^{\pi} G^*_{es}G^\perp_{eu,i}G^{\perp *}_{eu,i}G_{es}d\omega$$

$$= J_{id}(X) + \frac{1}{2\pi}\mathrm{tr}\int\limits_{-\pi}^{\pi} G^*_{es}G^\perp_{eu,i}G^{\perp *}_{eu,i}G_{es}d\omega$$

where we used Parseval's equality to write (3.12) in the frequency domain. Hence, minimizing the identification cost-function (3.12) is equivalent to minimizing the control cost-function (3.1) (subject to $X \in \mathcal{RH}^{m_{eu} \times m_u}_\infty$) which had to be proven.

$\square$

Let us discuss the advantage of estimating the factor $[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ G_{rs,co}^\dagger$ by PEM-OE over explicitly calculating $[G_{eu,i}^* \widehat{G}_{es} \widehat{G}_{rs,ci}^*]_+ \widehat{G}_{rs,co}^\dagger$. Note, that $\widehat{G}_{es}$ and $\widehat{G}_{rs}$ are minimum-phase models which are estimated using output-only data, thus $\widehat{G}_{rs,co} = \widehat{G}_{rs}$ and $\widehat{G}_{rs,ci} = I_{m_r}$. Often, it is difficult to estimate $\widehat{G}_{es}$ and $\widehat{G}_{rs}$ accurately, especially their zeros. The model-errors will propagate in $[G_{eu,i}^* \widehat{G}_{es} \widehat{G}_{rs,ci}^*]_+ \widehat{G}_{rs,co}^\dagger$ and lead to degradation of the performance. Also the estimate of $[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ G_{rs,co}^\dagger$ obtained by the identification approach of this section is contaminated with model errors and thus yields suboptimal performance. However, here $[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ G_{rs,co}^\dagger$ is estimated *directly* by minimizing the control cost-function, and will usually lead to better performance compared with the *indirectly*, model-based approach.

3.5.2   Subspace model identification

An attractive alternative for multi variable system identification problems is to apply the subspace identification methods, as e.g. implemented in the SLICOT library [159]. These methods can be used only for causal/stable systems, but with the method of [182] the tools can be adjusted easily to identify mixed causal/anti-causal systems.

As noted above, in the SMI methods no stability constraint on the model can be taken into account. Further research should give insight, whether the stability constraint can be incorporated by means of an iterative search algorithm with projection on the stable set of system (e.g., by means of the methods proposed in [27,111]). Since the stability constrained is not incorporated in the SMI methods, when using SMI with input/output data (3.13) the anti-causal system $G_{eu,i}^*$ would be modeled instead of $[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ G_{rs,co}^\dagger$. Suppose $G_{rs,co}$ is known, then a state-space realization of the factor $[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+$ can be estimated by the mixed causal/anti-causal SMI method of [182] using input/output data

$$\{G_{rs,co}^\dagger r(k), G_{eu,i}^* d(k)\}_{k=1}^N. \tag{3.14}$$

In case only a model $\widehat{G}_{rs,co}$ of $G_{rs,co}$ is known $[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+$ is estimated using input/output data

$$\{\widehat{G}_{rs,co}^\dagger r(k), G_{eu,i}^* d(k)\}_{k=1}^N.$$

and the estimate of the Causal Wiener filter is given by

$$W = -G_{eu,o}^\dagger [G_{eu,i}^* \widehat{G_{es} G_{rs,ci}^*}]_+ \widehat{G}_{rs,co}^\dagger \tag{3.15}$$

with $[G_{eu,i}^* \widehat{G_{es} G_{rs,ci}^*}]_+$ denotes the estimate of $[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+$. Note, that even in case of model errors in $\widehat{G}_{rs,co}$, which result in non-white input sequence $\{\widehat{G}_{rs,co}^\dagger r(k)\}_{k=1}^N$, the estimate $[G_{eu,i}^* \widehat{G_{es} G_{rs,ci}^*}]_+$ is determined such that

$$[G_{eu,i}^* \widehat{G_{es} G_{rs,ci}^*}]_+ \left(\widehat{G}_{rs,co}^\dagger r(k)\right) \approx G_{eu,i}^* d(k), \qquad \text{for } k = 1, \cdots, N.$$

Again, it can be shown that the identification problem is control-relevant. The proof is similar to the proof of Lemma 3.1 and therefore not included here. This advantage over the *indirectly* model-based approach of Section 3.2 will be illustrated in the experiments of Section 3.6.2 on page 68.

Table 3.3: The SIANC algorithm

1. Estimate the secondary path and the acoustical feedback models $\widehat{G}_{eu}$ and $\widehat{G}_{ru}$ and calculate the inner-outer factorization $\widehat{G}_{eu} = \widehat{G}_{eu,i}\widehat{G}_{eu,o}$ and $\widehat{G}^{\dagger}_{eu,o}$ the right inverse of $\widehat{G}_{eu,o}$

2. Measure data $\{r(k), d(k)\}_{k=1}^{N}$

3. Estimate $\widehat{G}_{rs,co}$ (minimum-phase spectral factor) and calculate its left inverse $\widehat{G}^{\dagger}_{rs,co}$

4. Estimate the causal part of $[G^{*}_{eu,i}G_{es}G^{*}_{rs,ci}]$ denoted as $[G^{*}_{eu,i}\widehat{G_{es}G^{*}_{rs,ci}}]_{+}$ by mixed causal/anti-causal subspace identification using input/output data:
   $\{\widehat{G}^{\dagger}_{rs,co}r(k), \widehat{G}^{*}_{eu,i}d(k)\}_{k=1}^{N}$

5. Using a realization of an (arbitrary) white-noise process $\zeta(k)$ construct the sequences
   $$\eta(k) = \widehat{G}_{rs,co}\zeta(k)$$
   $$\xi(k) = [G^{*}_{eu,i}\widehat{G_{es}G^{*}_{rs,ci}}]_{+}\zeta(k)$$

6. Estimate the causal H_2 optimal controller W by using input/output data:
   $\{\eta(k), -\widehat{G}^{\dagger}_{eu,o}\xi(k)\}_{k=1}^{M}$

7. Implement the controller $\widehat{C}$ by including intrinsic feedback compensation:
   $$\widehat{C} = (I_{m_u} + \widehat{W}\widehat{G}_{ru})^{-1}\widehat{W}$$
   as demonstrated in Figure 2.5

3.5.3 Order reduction

By explicitly evaluating the multiplications of the factors $G^{\dagger}_{eu,o}$, $[G^{*}_{eu,i}\widehat{G_{es}G^{*}_{rs,ci}}]_{+}$ and $\widehat{G}^{\dagger}_{rs,co}$ in (3.15) (and analogue in (3.2)) and in case of feedback also the calculations for the IMC controller (2.49), yield a very high order controller. An alternative approach is to generate the (optimal) control signal $u_{opt}(k)$ using the individual factors and the reference signal $r(k)$. Then, the multiplications to connect the individual factors can be evaluated implicitly by identifying the optimal controller using input/output data

$$\{r(k), u_{opt}(k)\}_{k=1}^{N}.$$

The procedure thus obtained to estimate the Causal Wiener filter is indicated by the acronym SIANC (Subspace Identification for Active Noise Control) and summarized in Table 3.3. The underlying assumption in the derivation of the SIANC algorithm in Table 3.3 is that G_{eu} and G_{ru} can be perfectly estimated, i.e. $\widehat{G}_{eu} = G_{eu}$ and $\widehat{G}_{ru} = G_{ru}$.

3.6   Practical demonstration on an acoustical duct and a vibrating plate

3.6.1   Acoustical duct

In the experiment on the acoustical duct we compare the SIANC algorithm, with LQG based on a full model of the duct identified in one step (cf. Section 3.4.1) and the Filtered-U LMS algorithm. We refer to Section 5.3 on page 106 for the derivation of Filtered-U LMS algorithm.

Description of the setup.   The experimental setup is illustrated by Figure 3.1 and 3.2. The acoustical duct, of 3m in length and 12cm diameter, is controlled at a sampling rate of f_s =1kHz using a Pentium 467MHz Celeron PC running Real-Time Linux version 1.2. Real-time control with a fixed sampling rate can be guaranteed while other applications, like monitoring with MATLAB can be performed at a lower priority[2]. The disturbance signal $s(k)$ is a white noise disturbance generated by the computer and filtered by an anti-aliasing first order low-pass filter with cutoff frequency at 500Hz. The control signal $u(k)$ actuates the loudspeaker via a DA-converter and an anti-aliasing low-pass filter again of first order and cutoff frequency at 500Hz.

SIANC.   The SIANC algorithm summarized in Table 3.3 has been used to estimate the Causal Wiener filter by a 48th order model $\widehat{W}$. Because of using the two-step approach in which the input/output data (after removing 2 samples delay in G_{eu} and 9 in G_{ru})

$$\{u(k), \begin{bmatrix} y(k) \\ r_u(k) \end{bmatrix} \}_{k=1}^N$$

is measured under the condition that $s(k) = 0$, $k = 1, \cdots, N$ (hence, $e(k) = y(k)$ and $r(k) = r_u(k)$) a very accurate 28th order state-space model of $\begin{bmatrix} G_{eu} \\ G_{ru} \end{bmatrix}$ could be identified using $N = 4000$ samples. The model accuracy of the estimated state-space model $(\widehat{A}, \widehat{B}_u, \begin{bmatrix} \widehat{C}_e \\ \widehat{C}_r \end{bmatrix}, \begin{bmatrix} \widehat{D}_{eu} \\ \widehat{D}_{er} \end{bmatrix})$ is evaluated using the variance accounted for (VAF), which is defined as

$$VAF(y_{measured}, y_{simulated}) \triangleq \left(1 - \frac{\text{var}(y_{measured} - y_{simulated})}{\text{var}(y_{measured})} \right) \times 100\% \quad (3.16)$$

where the scalar signals $y_{measured}$ and $y_{simulated}$ are the measured and the simulated output respectively. For the identified model of $\begin{bmatrix} G_{eu} \\ G_{ru} \end{bmatrix}$ the simulated

---

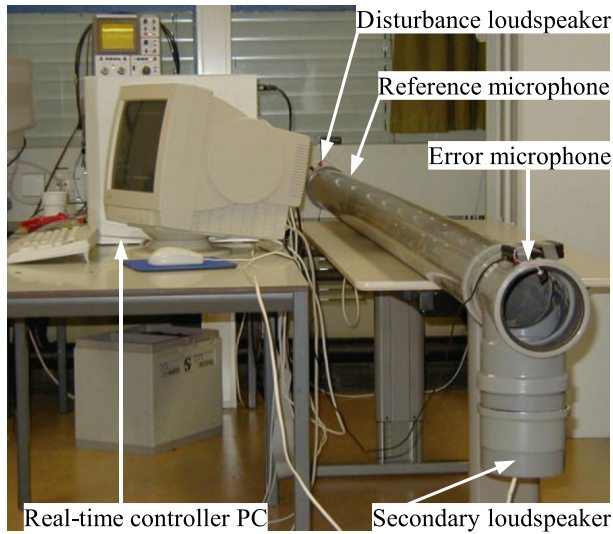[2]The software was provided by TNO Institute of Applied Physics, Delft, The Netherlands.

Figure 3.1: Experimental setup of the acoustical duct which is controlled by a PC running Real-Time Linux.
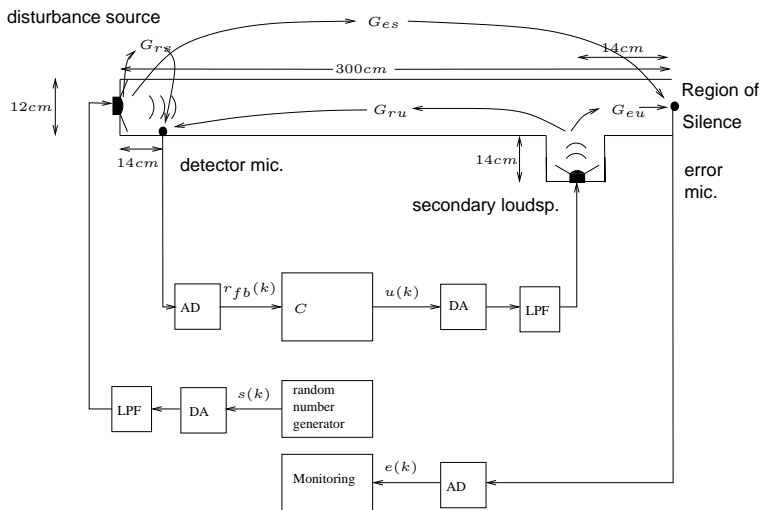
Figure 3.2: Block scheme of the acoustical duct experiment.

outputs are determined by

$$\xi(k+1) = \widehat{A}\xi(k) + \widehat{B}_u u(k), \qquad \xi(0) = 0$$

$$\left[\begin{array}{c} y_{simulated}(k) \\ r_{u,simulated}(k) \end{array}\right] = \left[\begin{array}{c} \widehat{C}_e \\ \widehat{C}_r \end{array}\right] \xi(k) + \left[\begin{array}{c} \widehat{D}_{eu} \\ \widehat{D}_{ru} \end{array}\right] u(k)$$

The VAF values obtained with the identified model is 99.7% and 99.9% for the model of G_{eu} and G_{ru} respectively using the same data as used for the identification. After collecting new data, the VAF values are 99.3% and 99.7% for the model of G_{eu} and G_{ru} respectively, which indicates that G_{eu} and G_{ru} are accurately modeled. We refer to [179] for a more detailed description of the use of subspace identification algorithms to identify these acoustical models.

Using a stochastic subspace identification algorithm an accurate 30th order model of the spectral factor $G_{rs,co}$ is estimated using the batch $\{r(k)\}_{k=1}^{N}$ of samples of the reference signal collected under the no-control condition, i.e. $u(k) = 0$ for $k = 1, \cdots, N$. Since, the disturbance signal $s(k)$ is generated by the computer, the complete detector path G_{rs} can be identified too, which shows that the detector path has non-minimum phase zeros at 347Hz, 407Hz and 468Hz, which will reduce the performance at these frequencies. These non-minimum phase zeros are due to aliasing since the anti-aliasing filter is just a simple first-order low-pass filter with relatively high cutoff at 500Hz. Choosing higher order aliasing filters and/or with lower cutoff frequency will reduce the effect of aliasing but also add more delay, which also reduces performance. By trial and error one can find the best choice of aliasing filters, but here we keep using the first order filters to illustrate the effect of aliasing at the frequencies.

Then, the factor $[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+$ of the Causal Wiener filter is identified using causal-/anti-causal subspace identification resulting in a 38th order model. The anti-causal modes are at 347Hz, 408Hz and 484Hz which are approximately the frequencies where G_{rs} has non-minimum phase zeros. Finally a 48th order model of W has been estimated. The obtained performance is illustrated by the power spectrum of the residual signal, given by the solid curve in Figure 3.3 on page 68.

LQG. The LQG controller has been designed using a full state-space model of the system with $s(k)$ considered as an unknown white noise stochastic process as is assumed in the SIANC approach. Using the PO-MOESP subspace identification algorithm the deterministic/stochastic model which consists of the state-space model $(A, B_u, \left[\begin{array}{c} C_e \\ C_r \end{array}\right], \left[\begin{array}{cc} D_{eu}, & D_{ru} \end{array}\right])$ as well as the covariance matrices

$$\left[\begin{array}{cc} Q_{rs} & S_{rs} \\ S_{rs}^T & R_{rs} \end{array}\right] = \left[\begin{array}{c} B_s \\ D_{rs} \end{array}\right] \left[\begin{array}{cc} B_s^T & D_{rs}^T \end{array}\right] \qquad \text{and} \qquad \left[\begin{array}{cc} Q_e & S_{eu} \\ S_{eu}^T & R_{eu} \end{array}\right] = \left[\begin{array}{c} C_e^T \\ D_{eu}^T \end{array}\right] \left[\begin{array}{cc} C_e & D_{eu} \end{array}\right]$$

are estimated, cf. the state-space model (2.52), where $D_{ru} = 0_{m_r \times m_u}$. Then, using the result of Theorem 2.2 the H_2 optimal controller of order 50 can be calculated. To determine the accuracy of the deterministic/stochastic model, we will evaluate the VAF based on the measured output and the one-step ahead prediction obtained

using the Kalman state-estimator, which is given by

$$\xi(k+1) = \widehat{A}\xi(k) + \widehat{B}_u u(k) + \widehat{K}_{rs}\epsilon(k)$$

$$\begin{bmatrix} \widehat{e}(k|k-1) \\ \widehat{r}(k|k-1) \end{bmatrix} = \begin{bmatrix} \widehat{C}_e \\ \widehat{C}_r \end{bmatrix} \xi(k) + \begin{bmatrix} \widehat{D}_{eu} \\ \widehat{D}_{ru} \end{bmatrix} u(k)$$

$$\epsilon(k) = \begin{bmatrix} e(k) \\ r(k) \end{bmatrix} - \begin{bmatrix} \widehat{e}(k|k-1) \\ \widehat{r}(k|k-1) \end{bmatrix}$$

with $\widehat{K}_{rs}$ the estimate of the Kalman gain given by Theorem 2.2 and $\epsilon(k)$ is called the innovation. Then, the VAF values between the measured and the predicted outputs are 98.1% and 97.4% for the output $e(k)$ and $r(k)$ respectively based on measured data used for identification. Using new measured data, the VAF values are 98.0% and 97.2% which are approximately the same and indicates that no modeling of e.g. noise has been occurred. However, these values are significantly lower than obtained by the models of G_{eu} and G_{ru} estimated with $s(k) = 0$ in the two-step approach as considered in the SIANC approach. This is explained by the presence of the unmeasured disturbance $s(k)$ which acts as a disturbance on the measurements.

The lower accuracy of the model is more clear when using a logarithmic measure between the simulation or prediction error and the measured output. Therefore let us express the ratio between the measured output variance and the error variance in deci-Bell's (dB) according to

$$10^{10}\log\left(\frac{\text{var}(y_{measured})}{\text{var}(y_{measured} - y_{simulated/predicted})}\right), \qquad (dB)$$

Then, a VAF of 99.5% yields a ratio of 23dB (the level of the model accuracy obtained by identifying $\begin{bmatrix} G_{eu} \\ G_{ru} \end{bmatrix}$ alone, with $s(k) = 0$), whereas the ratio of the prediction of $e(k)$ and $r(k)$ using the Kalman filter are 17dB and 15.5dB, which are significantly lower! The obtained performance is illustrated by the power spectrum of the residual signal, given by the dashed curve in Figure 3.3.

FuLMS. For the FuLMS algorithm we used the same model of G_{eu} and G_{ru} as determined in the SIANC approach, for generating the regression vector and compensating the acoustical feedback respectively. The step-size was chosen to be 0.1 and the regression vector was normalized with its power, estimated using an exponential forgetting algorithm. The order of the numerator and the denominator polynomials was chosen to be 50, which has been chosen on the basis of the controllers obtained with the SIANC and LQG methods. The obtained performance is illustrated by the power spectrum of the residual signal, given by the dash-dotted curve in Figure 3.3.

Results. Figure 3.3 shows the power spectra of the measured disturbance (no control) and the residuals obtained with control based on the SIANC, the LQG and the FuLMS algorithms. The power spectrum of the residual signal obtained
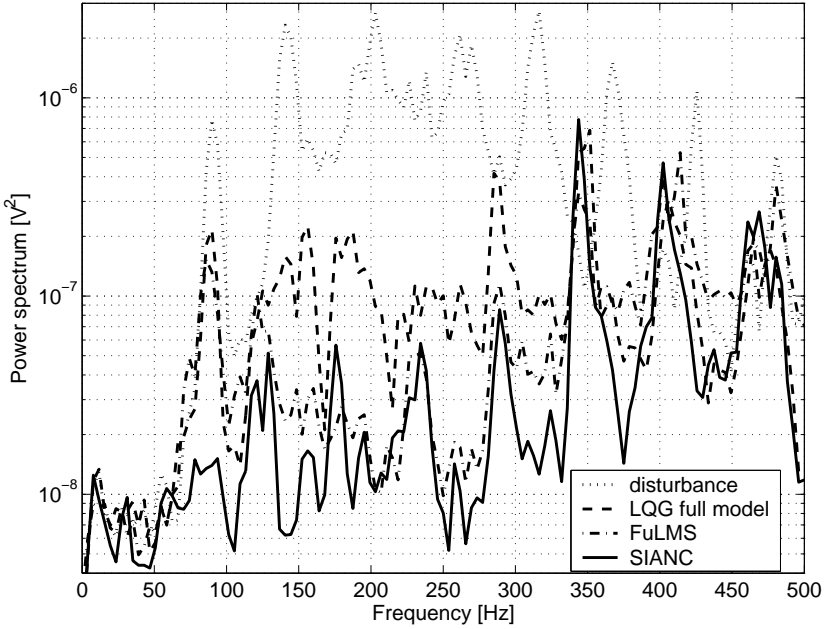
Figure 3.3: Spectra of the disturbance (dotted) and the residual signals obtained by the controller estimated by the LQG approach (dashed), the FuLMS algorithm after 8 minutes (dash-dotted) and the SIANC approach (solid).

by FuLMS is obtained after 8 minutes, after which the algorithm converged. We observe that all methods yields resonances at about 347Hz, 407Hz and 468Hz which are explained by aliasing. Furthermore, we observe that the LQG controller calculated using the full deterministic/stochastic model does not yield optimal performance. This can be explained by the fact that the model error is relatively large.

The FuLMS algorithm yields better performance, but after convergence of 8 minutes the performance is still not the same as obtained with the SIANC approach in which the controller is calculated within 45 seconds. Furthermore, it is possible that the FuLMS algorithm has been converged to a local minimum, which may explain that after ≈ 8 minutes the algorithm does not converge anymore. In Chapter 5 we will analyze the convergence of the FuLMS algorithm in more detail, and will also indicate how to improve the convergence rate of FuLMS.

3.6.2 Practical demonstration on a vibrating plate

In the experiment on the vibrating plate we illustrate the control of a pure feedback system, cf. the block scheme of Figure 2.3, using the SIANC and the H_2 optimal approach by calculating the Causal Wiener filter (2.20) explicitly. In Chapter 4 we will use the same system to demonstrate the robust version of the SIANC algorithm.
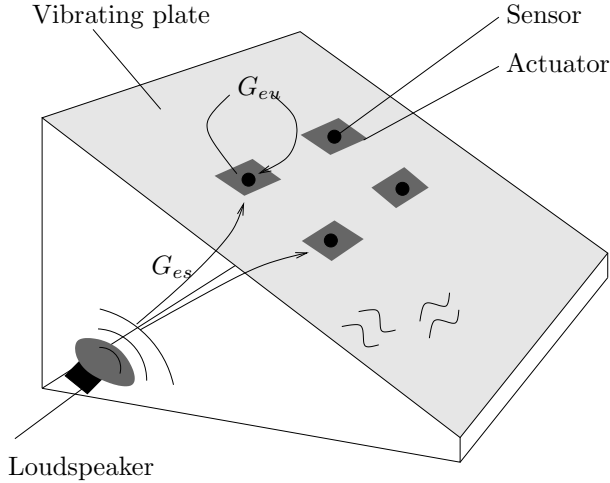
Figure 3.4: Schematic picture of the vibrating plate setup.

Description of the experimental setup. Consider Figure 3.4, which shows a schematic picture of the vibrating plate setup. The plate is an aluminium sandwich plate of 6mm thickness, 60cm width and 75cm height. The plate density is 870kg/m^3 and the Young's modulus $3.6 \cdot 10^{10}$Pa. Under the vibrating plate a loudspeaker generates a broadband disturbance sound, which propagates through the vibrating plate due to vibration of the plate. The resonance of the plate should be counteracted by piezoelectric actuators mounted on the lower side of the plate. The residual vibration of the plate is measured by piezoelectric accellerometers mounted at the upper side of the plate, collocated with the piezoelectric actuators. In the experiments 4 sensors and 4 actuators are used as indicated in Figure 3.4. The sampling frequency was f_s =2000Hz, and 4th order low-pass filters with cut-off frequency at 600Hz are used to prevent aliasing. Though the actuators and sensors are collocated, there is significant delay between the discrete actuator and the discrete sensors signals available at the DSP (i.e. PC) of two collocated actuator/sensor pairs, c.f. Figure 3.7 in the next paragraph. This delay is mainly due to the aliasing filters, and prevents using simple lead-, lag- or PID controllers to obtain good disturbance rejection.

The block scheme of the system and the IMC based controller is depicted in Figure 3.5 and equals the pure feedback configuration as discussed in Section 2.3. Using the IMC approach and assuming that G_{eu} is perfectly modeled, $\widehat{G}_{eu} = G_{eu}$, the feedback control problem reformulates into a feedforward control problem, as depicted in Figure 3.6, which was explained before in Section 2.3. Note, that in the following chapter, Chapter 4, we will account for the model errors in G_{eu}. Hence, the SIANC algorithm summarized in Table 3.3 on page 63 can be used again to estimate W, by making the following identifications

$$G_{rs} \rightarrow G_{es} \quad G_{ru} \rightarrow G_{eu} \quad r(k) \rightarrow d(k).$$

PSfrag replacements
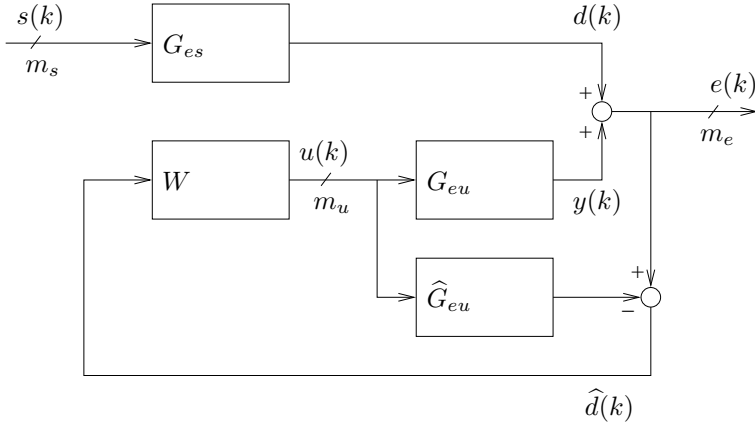
Figure 3.5: Block scheme of the vibrating plate feedback problem
with IMC controller.
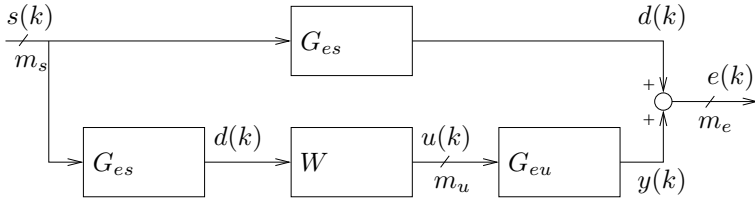
Figure 3.6: Feedforward equivalent block scheme of vibrating plate
feedback problem with perfect IMC controller.

Identification of G_{eu} and $G_{es,co}$. As in the case of the acoustical duct system, the 4×4 system G_{eu} is accurately estimated using SMI and 14000 samples (VAF values of about 99.7% for all channels) by a state-space model of order 80. Figure 3.7 shows the magnitude and the phase of the identified secondary path model $\widehat{G}_{eu}$ for input and output channels 1 and 2. We clearly see that the phase between actuator-sensor pair 11 and 22 drops significantly, which corresponds with a delay of about 2 samples, which cannot be effectively compensated by simple lead filters over a sufficiently large bandwidth.

Also the 4×4 spectral factor $G_{es,co}$ has been identified, resulting in an 40th order model. Though this model has found to be the best that could be obtained after trying different model orders based on a batch of 14000 samples of the measured disturbance $d(k)$, the inverse spectral factor $G_{es,co}^{-1}$ does not *perfectly* whiten $d(k)$. This is illustrated by Figure 3.8 which shows the spectrum of the measured disturbance (dashed) and the whitened disturbance obtained by the filtering $\widehat{G}_{es,co}^{-1}d(k)$ (solid) for output channel 1. We will see that the model errors in $\widehat{G}_{es,co}$ will be (partly) compensated using the control-relevant SIANC method.

(a) Input 1 to output 1.

(b) Input 2 to output 1.

(c) Input 1 to output 2.

(d) Input 2 to output 2.
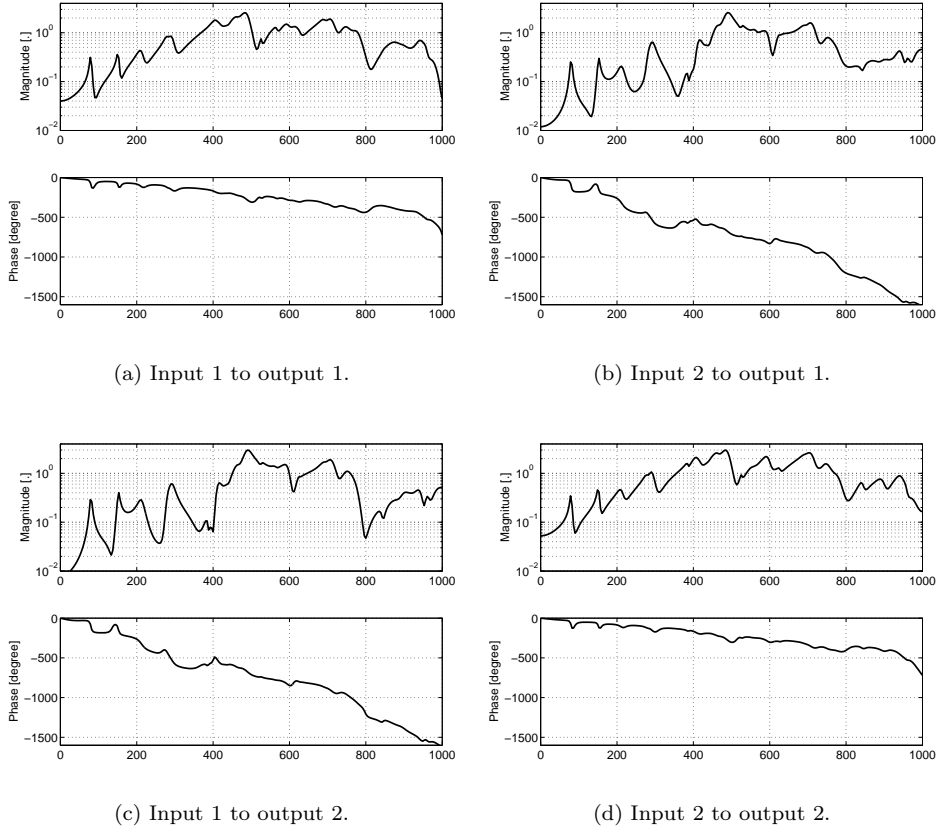
Figure 3.7: Maginute and phase of the identified secondary path model $\widehat{G}_{eu}$ for input and output channels 1 and 2.

Controller estimation and results. Using $\widehat{G}_{es,co}$ and $\widehat{G}_{eu}$, the causal Wiener filter (2.20) can be calculated explicitly. After model reduction, a 120th order feed forward controller has been obtained. The average performance over the whole frequency band based on simulation (hence there is no measurement noise and the secondary path is perfectly modeled) is given in Table 3.4.

However, the H_2 optimal controller can also be estimated using the identification approach described in Section 3.5.2. The factor $[G^*_{eu,i}G_{es}G^*_{es,ci}]_+ = [G^*_{eu,i}G_{es,co}]_+$ was estimated by causal/anti-causal subspace identification, using 14000 samples and 100 block rows, the estimate of the order is 48. Because 10 anti-causal poles of $G^*_{eu,i}G_{es,co}$ are rejected, the order of $[\widehat{G^*_{eu,i}G_{es,co}}]_+$ is 38. The controller obtained by explicitly evaluating the multiplications in (2.20) is of order 158. The time to calculate this controller was about 140sec.'s. The performance of this controller is given in the second row in Table 3.4. A reduced order feed forward controller of order 80 could be estimated by formulating the controller order reduction problem as an identification problem (see Section 3.5.3). The simulated
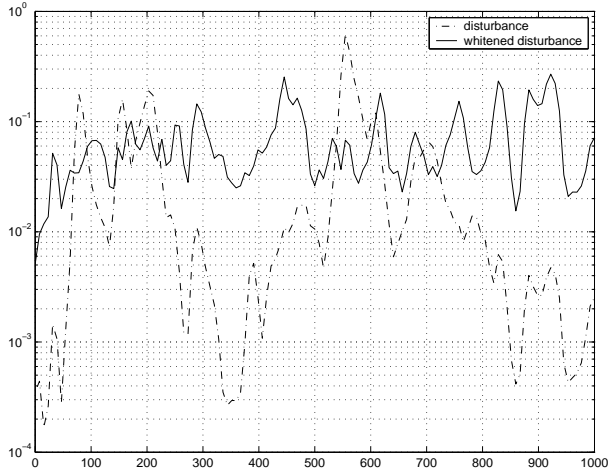
Figure 3.8: Spectrum of measured disturbance (dash-dotted) and the whitened disturbance (solid) obtained using the 4×4 spectral factor $\widehat{G}_{es,co}$ for output 1 only.

performance of this controller is given by the third row in Table 3.4. From Table 3.4 we infer that the performance of the identified Causal Wiener filter (based on simulation) is slightly better than the performance of the explicitly calculated Causal Wiener filter using (2.20). This can be explained by the fact that in the identification approach, a control-relevant identification problem is solved which compensates for model errors in $\widehat{G}_{es,co}$. Furthermore, the order of the identified controller could be reduced by about a factor 2 (from 158 to 80), without (significant) performance loss.

The IMC controller, consisting of the 80th order identified Causal Wiener filter and the 80th order secondary path model $\widehat{G}_{eu}$ (so the controller is of order 160!), has been plugged into the real-time setup to control the vibrating plate. Actually, the state-space controller was transformed into *output-normal* form to obtain a very efficient implementation of the controller iterations, see Appendix C. Figure 3.9(a)-(d) show the measured disturbance and the measured residual signals obtained by using this controller for output 1-4 respectively. We conclude that good disturbance suppression has been obtained. Only at $\approx$ 110Hz performance reduction cannot be obtained, which is due to a deep anti-resonance in the secondary path at this frequency, c.f. Figure 3.7. The average performance of this controller is given in the last row of Table 3.4.

3.7   Conclusions

In this chapter the Causal Wiener filter has been estimated using an identification approach based on prediction error modeling (PEM) or subspace model identifica-

Table 3.4: Average reduction in dB's of obtained by simulation of the Causal Wiener based on (2.20), the estimated Causal Wiener using the SIANC method without and with order reduction. The last row contains the measured reduction obtained by this reduced order controller determined by SIANC method.

	output 1	output 2	output 3	output 4
Causal Wiener based on (2.20) (order=120)	8.3	9.3	6.6	9.7
Identified Causal Wiener: (order=158)	8.7	9.6	7.0	10.0
Order reduced: (order=80)	8.5	9.5	6.9	9.8
Measured performance: (order reduced)	7.0	8.8	6.1	9.3

tion (SMI). The advantage of SMI over PEM has been illustrated by the identification of a 4×4 vibrating plate system of order 20, which is accurately modeled by SMI (VAF of 99%, signal to error ratio is 20dB) within less than 1 minutes computation time, whereas PEM modeling yields less accurate models (VAF of 97%, signal to error ratio is 15dB) in more than 5 minutes.

Using SMI a two stage approach has been proposed, in which first the secondary path G_{eu} and (acoustical) feedback G_{ru} has been accurately identified (preferred with no disturbance $s(k) = 0$) and second the causal wiener filter is estimated using the measured reference and disturbance signals. The approach is called SIANC (subspace identification for active noise control) and demonstrated on an acoustical duct. The method outperforms the LQG controller based on the full deterministic/stochastic model identified in one experiment. The SIANC approach also resulted in better performance than the FuLMS algorithm after 8 minutes converging time.

The second contribution of the chapter is, that it has been shown that the cost-function minimized for identification is the same as which has to be minimized during control. Hence, the identification is said to be control-relevant. In this way, model errors on the spectral factor of the reference signal (which is the disturbance signal in pure feedback applications) are partly compensated, hence better performance can be obtained as was demonstrated in a vibrating plate experiment.
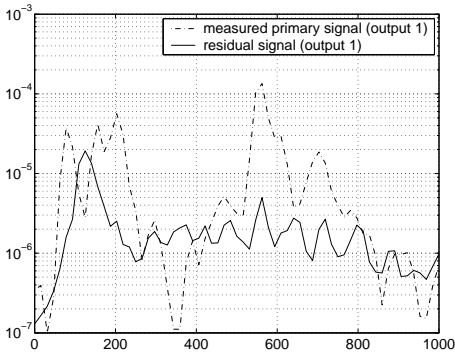
(a) Channel 1.
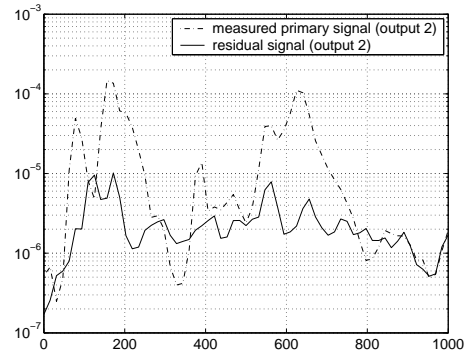
(b) Channel 2.
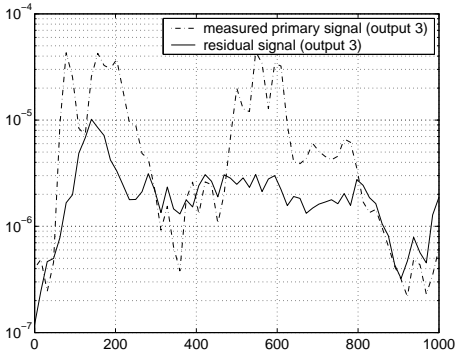
(c) Channel 3.
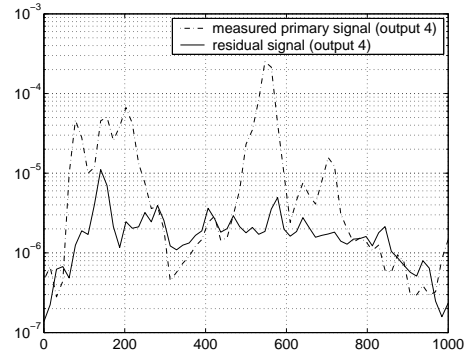
(d) Channel 4.

Figure 3.9: Spectrum of measured disturbance (dash-dotted) and the measured residual signal (solid) for output channel 1–4.

Chapter 4

Robust controller estimation

4.1 Introduction

In the previous chapter, a nominal controller has been estimated by solving a control-relevant identification problem. It has been shown, that model errors in the detector-path G_{rs} and the primary path G_{es} (of G_{es} no model is even necessary) can be partly compensated for, since the control cost function is explicitly minimized during the design. However, model uncertainty due to system variations and model uncertainty in the secondary path G_{eu} and the feedback G_{ru} have not been accounted for.

In this chapter a probabilistic robust feedforward design approach will be used, which explicitly accounts for uncertainty in G_{es}, G_{rs}, G_{eu}. We will also show, that robust stability is obtained for feedback systems when using the robust feedforward controller with the Internal Model Control approach.

In robust control literature, much attention has been paid to minimizing the H_∞ norm, in which model errors are taken into account explicitly (see e.g. [199]). Though, stability robustness can be increased significantly, too often the performance is poor when optimizing for the worst case condition. Better performance is obtained by mixed H_2/H_∞ control design, where the H_2 performance measure is optimized subject to H_∞ constraints to guarantee user determined stability/performance robustness margins (see e.g. [10]). More recently, a *minimax* LQG method was proposed in [136], which minimizes the MSE for the worst case model error (contained in a stochastic model uncertainty description). For more on robust H_2 control, see, e.g., [132] and the references therein.

However, in all these robust design methods, the *likelihood* of the model errors is not taken into account. Such a design philosophy may be useful in critical applications where stability and a certain minimal (often low) level of performance should be guaranteed under all, including extremely rare, circumstances, e.g. in flight-by-wire control in aircrafts or biomedical control applications. Most active control problems are not that critical to pay a significant price on performance, and optimal performance on the *average* of all kind of model errors is desired. An

additional reason is that most model identification methods give estimates of the likelihood of model errors, rather than hard bounds [104]. In [72] Goodwin et al. write:

> "A hard-bound noise model is a very coarse (worst case) model for phys-
> ical reality since every value within a compact domain is considered as
> likely as any other. A distribution, with noncompact support, is a model
> of reality in which the noise values are assumed to be on average cen-
> tered around some mean value without precluding the possibility of the
> occasional outlier."

These notions where the motivation of the stochastic embedding approach to quan-tify model uncertainty. In the same line Sternad et al. [130, 164] proposed a prob-abilistic robust filtering/feedforward control design method, which minimizes the MSE averaged over the (estimated) stochastic distribution of the model errors. The resulting robust filter is called a *Cautious* Wiener (CW) filter.

The approach of [164] has been taken up in [9], which nicely relates the proba-bilistic robust solution to standard LQG solutions. However, a particular open-loop uncertainty description has been adopted which prevents to consider independent uncertainty on the models of the primary path G_{es}, secondary path G_{eu} and de-tector path G_{rs}. As a result, the Cautious Wiener Theorem 4.3 on page 84 given below, will be a more general result than Theorem 1 and Theorem 2 in [9].

The feedback control design problem based on stochastic model errors has been considered before by Goodwin et al. in [73]. Due to the model uncertainty, the feedback cannot be perfectly compensated by IMC, which yields that the sensitiv-ity function is non-linear in the controller. In [73] a linear approximation of the sensitivity function has been made, which is similar to the linearization approach for feedback systems taken in Section 4.5 of this chapter. However, in [73] it is a priory assumed that a stabilizing controller is available which yields a desired sen-sitivity in closed-loop with the nominal model. Then, their objective is to design a new controller by minimizing the H_2 norm of the difference between the desired sensitivity function and the actual sensitivity function averaged over the stochastic model uncertainty.

In this chapter, we extend the first proposed polynomial based probabilistic robust feedforward of [130, 164] in a state-space framework. Also, the control-relevant identification approaches to estimate the nominal controller as presented in Chapter 3 are extended to estimate the robust controller. The control-relevant estimation of the feedback control design approach of Goodwin et al. [73] is left for further research.

The chapter is organized as follows. Section 4.2 presents the probabilistic model uncertainty description and ways to estimate the uncertainty. Section 4.3 derives the robust feedforward controller, which is called the Cautious Wiener filter, and is designed by minimizing the mean squared error averaged over the distribution of the model uncertainty. Section 4.4 extends the nominal controller estimation approach taken in Chapter 3 to the estimation of the robust controller. Section 4.5 discusses the feedback configuration using IMC compensation (Section 2.3) together with the Cautious Wiener filter. It is shown that using the Cautious Wiener filter the stability robustness is improved compared with using the Causal Wiener filter

of equation (2.20) on page 36. Finally, Section 4.6 demonstrates the robust design method on the vibrating plate setup which has been considered in Section 3.6.2 too. In the experiment the uncertainty is caused by variation in an additional mass mounted on the plate.

4.2   Model uncertainty description

4.2.1   Polynomial uncertainty description

The robust filtering and feedforward control, proposed in [164], is based on a *probabilistic* description of the model error. Here we briefly discuss the SISO case, for the MIMO case, see [129, 130]. The nominal model $G^o(q^{-1})$ is considered to be in transfer-function form

$$G^o(q^{-1}) \;=\; \frac{A^o(q^{-1})}{B^o(q^{-1})}$$

with $A^o(q^{-1}) \;=\; a_0^o + a_1^o q^{-1} + \cdots + a_{n_a^o}^o q^{-n_a^o}$ and $B^o(q^{-1}) \;=\; 1 + b_1^o q^{-1} + \cdots + b_{n_b^o}^o q^{-n_b^o}$. Then, the structure of the uncertainty model is given by

$$G(q^{-1}) \;=\; \frac{A^o(q^{-1})}{B^o(q^{-1})} + \frac{A^1(q^{-1})\Delta A(q^{-1})}{B^1(q^{-1})} \tag{4.1}$$

with $A^1(q^{-1})$ and $B^1(q^{-1})$ are polynomials in q^{-1} and determine the weighting of the uncertainty. The polynomial $\Delta A(q^{-1})$ is an 'uncertain' polynomial which coefficients are stochastic variables:

$$\Delta A(q^{-1}) \;=\; \Delta a_0 + \Delta a_1 q^{-1} + \cdots + \Delta a_{\delta a} q^{-\delta a} \tag{4.2}$$

such that

$$\overline{\mathbf{E}}\left(\Delta a_i\right) = 0, \qquad i = 0, \cdots, \delta a \tag{4.3}$$

$$\overline{\mathbf{E}}\left(\Delta a_i \Delta a_j^*\right) = r_{ij}, \qquad i,j = 0, \cdots, \delta a \tag{4.4}$$

with $r_{ij} \in \mathbb{R}$ given and $\overline{\mathbf{E}}\left(.\right)$ expectation over the model uncertainty.

The uncertainty model (4.1) can describe parametric (structured) as well as non-parametric (unstructured) uncertainty (see [129, 164] for illustrations), and is related to the stochastic embedding approach of Goodwin et al., that also accounts for unmodeled dynamics, see, e.g., [72]. Note, that no assumptions are made on the particular *distribution* of the coefficients Δa_i in (4.2).

Furthermore, note, that it is reasonable to set $\overline{\mathbf{E}}\left(\Delta a_i\right) = 0$ in probabilistic model uncertainty description, since in case a nonzero mean value $\overline{\mathbf{E}}\left(\Delta a_i\right)$ would be known it can be contained in the nominal model $G^o(q^{-1})$.

Hence, the uncertainty model is determined by the weighting polynomials $A^1(q^{-1}), B^1(q^{-1})$ and the correlations coefficients r_{ij}. Further note, that only uncertainty in the zeros is considered. This is to guarantee stability of the model for all possible model errors and to prevent non-linear relations in the uncertainty parameters which cause severe difficulties in the controller design [164]. Uncertainty in the poles has to be reformulated in terms of uncertainty in nummerator

coefficients (and thus zeros) by means of Taylor expansions which will determine the weightings $A^1(q^{-1})$ and $B^1(q^{-1})$, c.f. [129, Sec. 3.7].

It is clarifying to consider the uncertain transfer-function ΔG

$$\Delta G(q^{-1}) \;=\; \frac{A^1(q^{-1})\Delta A(q^{-1})}{B^1(q^{-1})}$$

in the frequency domain. Therefore, let us define

$$\phi_{\delta a}(q^{-1}) = \begin{bmatrix} 1 & q^{-1} & \cdots & q^{-\delta a} \end{bmatrix}$$

$$R_{\delta a} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1\delta a} \\ r_{21} & r_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ r_{\delta a 1} & \cdots & \cdots & r_{\delta a \delta a} \end{bmatrix}$$

Then, we have

$$\overline{\mathbf{E}}\left(\Delta G(e^{-j\omega})\right) = 0, \qquad -\pi \le \omega < \pi \tag{4.5}$$

$$\overline{\mathbf{E}}\left(\Delta G(e^{-j\omega})\Delta G(e^{-j\omega})^*\right) = \Phi_{\Delta G}(e^{-j\omega}), \qquad -\pi \le \omega < \pi \tag{4.6}$$

with

$$\Phi_{\Delta G}(e^{-j\omega}) = \frac{A^1(e^{-j\omega})\phi_{\delta a}(e^{-j\omega})R_{\delta a}\phi_{\delta a}^T(e^{j\omega})A^1(e^{j\omega})}{B^1(e^{-j\omega})B^1(e^{j\omega})}. \tag{4.7}$$

From these expressions, we observe that ΔG can be considered as a random variable in the frequency domain, with zero mean and variance $\Phi_{\Delta G}(e^{-j\omega})$. This interpretation will be useful in the sequel.

4.2.2   State-space uncertainty description

The probabilistic model uncertainty description is not limited to polynomial model descriptions, but can also be used e.g. for uncertain state-space model descriptions. This can be seen directly from the fact that the polynomial model description is a special case of a state-space model description. Stated otherwise, the polynomial description is contained in the state-space description. In the sequel, the MIMO case will be considered directly.

Consider the state-space description $(A^1, \Delta B, C^1, \Delta D)$ of $\Delta G(q^{-1})$, denoted as

$$\Delta G(q^{-1}) \sim \left[\begin{array}{c|c} A^1 & \Delta B \\ \hline C^1 & \Delta D \end{array} \right] \tag{4.8}$$

with $A^1 \in \mathbb{R}^{n_\delta \times n_\delta}$ and $C^1 \in \mathbb{R}^{l \times n_\delta}$ known matrices and $\Delta B \in \mathbb{R}^{n_\delta \times m}$ and $\Delta D^{l \times m}$ real-valued stochastic variables, such that

$$\overline{\mathbf{E}}\left(\begin{bmatrix} \Delta B \\ \Delta D \end{bmatrix}\right) = 0_{n_\delta + l \times m} \tag{4.9}$$

$$\overline{\mathbf{E}}\left(\begin{bmatrix} \Delta B \\ \Delta D \end{bmatrix}\begin{bmatrix} \Delta B \\ \Delta D \end{bmatrix}^T\right) = \begin{bmatrix} \Phi_{BB} & \Phi_{BD} \\ \Phi_{DB} & \Phi_{DD} \end{bmatrix} \tag{4.10}$$

with $\Phi_{BB} \in \mathbb{R}^{n_\delta \times n_\delta}$, $\Phi_{BD} = \Phi_{DB}^T \in \mathbb{R}^{n_\delta \times l}$ and $\Phi_{DD} \in \mathbb{R}^{l \times l}$. Then, in the frequency domain ΔG is characterized by

$$\overline{\mathbf{E}}\left(\Delta G(e^{-j\omega})\right) = 0, \qquad -\pi \leq \omega < \pi \qquad (4.11)$$

$$\overline{\mathbf{E}}\left(\Delta G(e^{-j\omega})\Delta G(e^{-j\omega})^*\right) = \begin{bmatrix} C^1(e^{j\omega}I - A^1)^{-1} & I_l \end{bmatrix} \begin{bmatrix} \Phi_{BB} & \Phi_{BD} \\ \Phi_{DB} & \Phi_{DD} \end{bmatrix} \cdot$$
$$\begin{bmatrix} (e^{-j\omega}I - A^{1T})^{-1}C^{1T} \\ I_l \end{bmatrix}, \quad -\pi \leq \omega < \pi \quad (4.12)$$

Using the canonical spectral factorization theorem (see, e.g., [85, Theorem 8.3.2, page 277]) the spectrum at the right-hand side of (4.12) can be written in a form which will be useful for the robust controller design.

Theorem 4.1 (Canonical spectral factorization) *Let A^o be stable, $(A^o - \Phi_{BD}\Phi_{DD}^{-1}C^o, \Phi_{BB} - \Phi_{BD}\Phi_{DD}^{-1}\Phi_{DB})$ be controllable on the unit-circle and*

$$\begin{bmatrix} \Phi_{BB} & \Phi_{BD} \\ \Phi_{DB} & \Phi_{DD} \end{bmatrix} \geq 0, \qquad \Phi_{DD} > 0$$

Then, the discrete-time algebraic Riccati equation (DARE)

$$P = A^1 P A^{1T} + \Phi_{BB} - K_p R_e K_p^T \qquad (4.13)$$

with

$$R_e = \Phi_{DD} + C^1 P C^{1T}, \qquad K_p = (A^1 P C^{1T} + \Phi_{BD})R_e^{-1}$$

has a unique solution P such that $A^1 - K_p C^1$ is stable, $P \geq 0$ and $R_e > 0$.
  Moreover, the right-hand side of (4.12) *can be factorized as*

$$\begin{bmatrix} C^1(e^{j\omega}I - A^1)^{-1} & I_l \end{bmatrix} \begin{bmatrix} \Phi_{BB} & \Phi_{BD} \\ \Phi_{DB} & \Phi_{DD} \end{bmatrix} \begin{bmatrix} (e^{-j\omega}I - A^{1T})^{-1}C^{1T} \\ I_l \end{bmatrix} =$$

$$\widetilde{\Delta G}(e^{-j\omega})\widetilde{\Delta G}(e^{-j\omega})^*$$

with $\widetilde{\Delta G}(q^{-1}) \in \mathcal{RH}_\infty^{l \times l}$ given by

$$\widetilde{\Delta G}(q^{-1}) = C^1(q^{-1}I - A^1)^{-1}K_p R_e^{1/2} + R_e^{1/2}, \qquad R_e^{1/2}R_e^{T/2} = R_e$$

and since, $A^o - K_p C^1$ is stable, $\widetilde{\Delta G}(q^{-1})$ is minimum-phase.

Proof: For the proof we refer to [85, Theorem 8.3.2, page 277]. $\qquad \square$

Note, that under the conditions given in Theorem 4.1, the number of inputs of the spectral factor $\widetilde{\Delta G}$ is equal to the number of outputs l, no matter the number of inputs, m, of ΔG.

In some situations, another uncertainty description, which is the dual of (4.8), will be more useful. Consider, the dual uncertain state-space description $(A^1, B^1, \Delta C, \Delta D)$ of ΔG, denoted as

$$\Delta G(q^{-1}) \sim \left[\begin{array}{c|c} A^1 & B^1 \\ \hline \Delta C & \Delta D \end{array} \right] \tag{4.14}$$

with $A^1 \in \mathbb{R}^{n_\delta \times n_\delta}$, $B^1 \in \mathbb{R}^{n_\delta \times m}$ known matrices and $\Delta C \in \mathbb{R}^{l \times n_\delta}$ and $\Delta D \in \mathbb{R}^{l \times m}$ real-valued stochastic variables, such that

$$\overline{\mathbf{E}} \left(\left[\begin{array}{cc} \Delta C & \Delta D \end{array} \right] \right) = 0_{l \times n_\delta + m}, \qquad -\pi \leq \omega < \pi \tag{4.15}$$

$$\overline{\mathbf{E}} \left(\left[\begin{array}{cc} \Delta C & \Delta D \end{array} \right]^T \left[\begin{array}{cc} \Delta C & \Delta D \end{array} \right] \right) = \left[\begin{array}{cc} \Phi_{CC} & \Phi_{CD} \\ \Phi_{DC} & \Phi_{DD} \end{array} \right], -\pi \leq \omega < \pi \tag{4.16}$$

with $\Phi_{CC} \in \mathbb{R}^{n_\delta \times n_\delta}$, $\Phi_{CD} = \Phi_{DC} \in \mathbb{R}^{n_\delta \times m}$ and $\Phi_{DD} \in \mathbb{R}^{m \times m}$. Then, in the frequency domain ΔG is characterized by

$$\overline{\mathbf{E}} \left(\Delta G(e^{-j\omega}) \right) = 0_{l \times m}, \qquad -\pi \leq \omega < \pi \tag{4.17}$$

$$\overline{\mathbf{E}} \left(\Delta G(e^{-j\omega})^* \Delta G(e^{-j\omega}) \right) = \left[\begin{array}{cc} B^{1T}(e^{-j\omega}I - A^{1T})^{-1} & I_m \end{array} \right] \left[\begin{array}{cc} \Phi_{CC} & \Phi_{CD} \\ \Phi_{DC} & \Phi_{DD} \end{array} \right] \cdot$$
$$\left[\begin{array}{c} (e^{j\omega}I - A^1)^{-1}B^1 \\ I_m \end{array} \right] \tag{4.18}$$

Again, a useful factorization of the spectrum given by (4.18) exists, which is dual to the factorization of Theorem 4.1.

Theorem 4.2 (Canonical *co*-spectral factorization) *Let A^o be stable, $(A^o - B^o\Phi_{DD}^{-1}\Phi_{DC}, \Phi_{CC} - \Phi_{CD}\Phi_{DD}^{-1}\Phi_{DC})$ be observable on the unit-circle and*

$$\left[\begin{array}{cc} \Phi_{CC} & \Phi_{CD} \\ \Phi_{DC} & \Phi_{DD} \end{array} \right] \geq 0, \qquad \Phi_{DD} > 0$$

Then, the discrete-time algebraic Riccati equation (DARE)

$$P = A^{1T}PA^o + \Phi_{CC} - K_p^T R_e K_p \tag{4.19}$$

with

$$R_e = \Phi_{DD} + B^{1T}PB^o, \qquad K_p = R_e^{-1}(B^{1T}PA^o + \Phi_{DC})$$

has a unique solution P such that $A^1 - B^1 K_p$ is stable, $P \geq 0$ and $R_e > 0$. Moreover, the right-hand side of (4.18) can be factorized as

$$\left[\begin{array}{cc} B^{1T}(e^{-j\omega}I - A^{1T})^{-1} & I_m \end{array} \right] \left[\begin{array}{cc} \Phi_{CC} & \Phi_{CD} \\ \Phi_{DC} & \Phi_{DD} \end{array} \right] \left[\begin{array}{c} (e^{j\omega}I - A^1)^{-1}B^1 \\ I_m \end{array} \right] =$$

$$\widetilde{\Delta G}(e^{-j\omega})^* \widetilde{\Delta G}(e^{-j\omega})$$

with $\widetilde{\Delta G}(q^{-1}) \in \mathcal{RH}_\infty^{m \times m}$ given by

$$\widetilde{\Delta G}(q^{-1}) = R_e^{1/2} K_p (q^{-1} I - A^1)^{-1} B^1 + R_e^{1/2}, \qquad R_e^{T/2} R_e^{1/2} = R_e$$

and since, $A^o - B^1 K_p$ is stable, $\widetilde{\Delta G}(q^{-1})$ is minimum-phase.

Proof: The proof is very similar to the proof of Theorem 4.1, but when making use of the following substitutions

$$A^1 \to A^{1T}, \quad C^1 \to B^{1T},$$
$$\Phi_{BB} \to \Phi_{CC}, \quad \Phi_{BD} \to \Phi_{CD}, \quad \Phi_{DB} \to \Phi_{DC}.$$

Cf. [85, Theorem.8.3.2, page 277]. □

Now, note, that under the conditions given in Theorem 4.2, the number of outputs of the spectral factor $\widetilde{\Delta G}$ is equal to the number of inputs m, no matter the number of outputs, l, of ΔG.

4.2.3 Obtaining the uncertainty model

The stochastic uncertainty models (4.8) and (4.14) are determined by a state-space realization of $\widetilde{\Delta G}$. However, the question is, how to obtain a state-space realization of $\widetilde{\Delta G}$ which captures the uncertainty. As in [129], we can distinguish several approaches to quantify the model errors:

- Pragmatic tuning: *in case it is known that in a particular frequency band the model uncertainty is large, $\widetilde{\Delta G}$ can be chosen to be a band-pass filter with high gain (determined by trial and error) in this frequency band.*

- First principles: *in case a white-box model is available, the uncertainty models can be derived by variation of physical parameters (e.g., using Taylor expansion).*

- System identification / Stochastic embedding: *using PEM and neglecting bias errors the covariance of the parameter error vector can be estimated too [104]. When the model to be identified is linear in the parameters (e.g. FIR, Laguerre or Kautz expansions), from this covariance matrix the uncertainty model with the structure (4.1) can be derived directly. When bias errors cannot be neglected, the stochastic embedding approach can be used in which the bias error is also considered as a random variable, see, e.g., [72].*

additionally we distinguish:

- Model error modeling: *in [103] it was proposed to estimate an error model to quantify the model error. By estimating a series of error models a sample estimate of the variance can be estimated, as will be discussed below in more detail. The advantage of this approach, is that it is independent of the identification method and thus subspace identification can be used as well.*

The pragmatic tuning approach is quite *ad hoc*, but may often be useful, e.g., to model uncertainty at low and high frequencies just outside or at the border of the frequency band of interest for control.

Derivation of the uncertainty model from first principles is often complicated, since it is difficult to obtain an accurate white-box model of the system to be controlled.

The identification of the uncertainty model using PEM system identification or stochastic embedding, is only restricted to model structures which are linear in the parameters, and thus cannot be used in combination with subspace identification methods. In [14] uncertainty of state-space models has been estimated using bootstrap-based methods. However, these approaches are computationally very complex, and thus not feasible for the high dimensional models considered in active noise and vibration control.

The approach, taken in this thesis, to estimate the model uncertainty is partly based on model error modeling. The model of the nominal system G^o and the uncertainty spectral factor $\widetilde{\Delta G}$ are estimated by means of a series of p experiments. In each experiment the (environmental) conditions are different resembling realistic situations in which the system can be (for example a vibrating plate with different mass loads). In each experiment G is modeled which yields a series of p models

$$\{\widehat{G}^i\}_{i=1}^p$$

In case the bias and variance errors in each model $\widehat{G}^i$ can be neglected compared with the differences $\widehat{G}^i - \widehat{G}^j$ $(i \neq j)$, the average $\widehat{G}^o$ which is the model of the *nominal* system G^o is determined by the average

$$\widehat{G}^o(e^{-j\omega}) = \frac{1}{p}\sum_{i=1}^p \widehat{G}^i(e^{-j\omega}), \qquad \pi \leq \omega < \pi$$

Furthermore, the spectral factor $\widehat{\Delta G}$ which models $\widetilde{\Delta G}$ is determined as a spectral factor of

$$\widehat{\Delta G}(e^{-j\omega})^*\widehat{\Delta G}(e^{-j\omega}) = \frac{1}{p}\sum_{i=1}^p (\widehat{G}^i(e^{-j\omega}) - \widehat{G}^o(e^{-j\omega}))^*(\widehat{G}^i(e^{-j\omega}) - \widehat{G}^o(e^{-j\omega}))$$
$$\pi \leq \omega < \pi$$

Since, $\widehat{G}^o$ is estimated by an average over p models which states may be in different subspaces, the order of $\widehat{G}^o$ will be very large (i.e. the sum of the orders of the p models) when the average is calculated explicitly. To circumvent a very high order model for $\widehat{G}^o$ and a high order spectral factorization problem for $\widehat{\Delta G}$, we approximate $\widehat{G}^o$ and $\widehat{\Delta G}$ by solving identification problems. The restricted order model $\widehat{G}^o$ can be estimated using input/output data $\{u(k), y(k)\}_{k=1}^N$ with $u(k) \in \mathbb{R}^m$ zero mean white noise and

$$y(k) = \frac{1}{p}\sum_{i=1}^p \widehat{G}^i u(k)$$
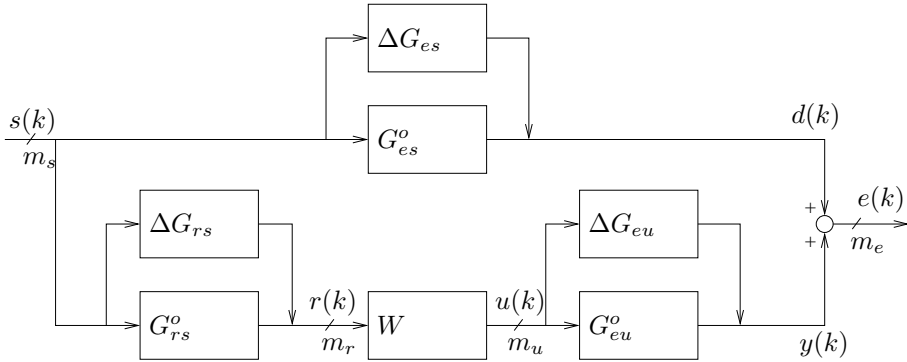
Figure 4.1: Block scheme of the multichannel feedforward active control problem with uncertainty, with m_s noise sources, m_r reference, m_u control and m_e error signals and no measurement noise.

The restricted order spectral factor $\widehat{\Delta G}$ can be determined by estimating the spectrum of

$$\xi(k) \;=\; \frac{1}{\sqrt{p}} \sum_{i=1}^{p} (\widehat{G}^i - \widehat{G}^o)\zeta^i(k)$$

with $\zeta^i(k),\; i = 1, ..., p$ such that $\mathrm{E}[\zeta^i(k)] = 0$ and $\mathrm{E}[\zeta^i(k)\zeta^i(l)] \;=\; I_m \delta(k - l)$ and ζ^i independent of ζ^j for $i \neq j$.

A state-space realization of $\{\widehat{G}^i\}_{i=1}^p$, $\widehat{G}^o$ as well as the spectral factor $\widehat{\Delta G}$ can be estimated using subspace identification methods. The spectral factor $\widehat{\Delta G}$ needs to be estimated by *stochastic* subspace identification methods, see, e.g., [111].

4.3 Derivation of the Cautious Wiener filter

Let us now return to the feedforward control problem considered in Section 2.2.1 on page 31. However, now the systems are perturbed with uncertainty as illustrated in Figure 4.1, cf. Figure 2.1. The uncertain systems ΔG_{es}, ΔG_{rs} and ΔG_{eu} are assumed to be stochastic variables as considered in Section 4.2, and independent of each other, independent of the nominal systems G_{es}^o, G_{rs}^o and G_{eu}^o and the stochastic process $s(k)$. This means that for u, x referring both to e or r and v, y to s or r, we also have

$$\overline{\mathbf{E}}\left(\Delta G_{uv}^* \Delta G_{xy}\right) = 0, \qquad \text{for } uv \neq xy \tag{4.20}$$

$$\overline{\mathbf{E}}\left(\Delta G_{uv} \Delta G_{xy}^*\right) = 0, \qquad \text{for } uv \neq xy. \tag{4.21}$$

These assumptions are usually not satisfied in practice, since the uncertainties in G_{es}, G_{rs} and G_{eu} are in general not independent. For example, temperature variations will affect all transfers. However, without assuming this simplifying assumption, the modelling and analysis of the uncertainty as well as the robust controller design will become very difficult, if even possible at all. In addition, in the

following we show that even though these assumptions (4.20),(4.21) are simplifying reality, controllers are obtained which are indeed *robust* and in Section 4.6 good robust performance is obtained in real-life experiments.

Furthermore, we assume that the uncertainties has zero mean and known co-variance, i.e. let

$$\overline{\mathbf{E}}\left(\Delta G_{es}\right) = 0_{m_e \times m_s}, \quad \overline{\mathbf{E}}\left(\Delta G_{es} \Delta G_{es}^*\right) = \widetilde{\Delta G}_{es} \widetilde{\Delta G}_{es}^* \tag{4.22}$$

$$\overline{\mathbf{E}}\left(\Delta G_{rs}\right) = 0_{m_r \times m_s}, \quad \overline{\mathbf{E}}\left(\Delta G_{rs} \Delta G_{rs}^*\right) = \widetilde{\Delta G}_{rs} \widetilde{\Delta G}_{rs}^* \tag{4.23}$$

$$\overline{\mathbf{E}}\left(\Delta G_{eu}\right) = 0_{m_e \times m_u}, \quad \overline{\mathbf{E}}\left(\Delta G_{eu}^* \Delta G_{eu}\right) = \widetilde{\Delta G}_{eu}^* \widetilde{\Delta G}_{eu} \tag{4.24}$$

with $\widetilde{\Delta G}_{rs}$ and $\widetilde{\Delta G}_{eu}$ given. In the following, we will see that the solution of the robust feedforward control problem is independent of $\widetilde{\Delta G}_{es}$, and thus it is not necessary to assume $\widetilde{\Delta G}_{es}$ to be known.

The objective in the Cautious Wiener design approach is to minimize the MSE averaged over the distribution of the uncertain systems [164], i.e. to minimize

$$J_{rob} = \text{tr}\overline{\mathbf{E}}\left(\mathbf{E}\left(e(k)e^T(k)\right)\right). \tag{4.25}$$

In this expression the residual disturbance $e(k)$ and the reference signal $r(k)$ are given by

$$e(k) = (G_{es}^o + \Delta G_{es})s(k) + (G_{eu}^o + \Delta G_{eu})u(k) \tag{4.26}$$

$$r(k) = (G_{rs}^o + \Delta G_{rs})s(k) \tag{4.27}$$

and the feedforward control-law by

$$u(k) = Wr(k). \tag{4.28}$$

The robust feedforward controller problem is summarized as follows.

Problem 4.1 (Robust feedforward controller design problem) *Given the nominal systems $G_{es}^o \mathcal{RH}_\infty^{m_e \times m_r}$, $G_{rs}^o \in \mathcal{RH}_\infty^{m_r \times m_s}$, $G_{eu}^o \in \mathcal{RH}_\infty^{m_e \times m_u}$ and the spectral-factors $\widetilde{\Delta G}_{rs} \in \mathcal{RH}_\infty^{m_r \times m_r}$, $\widetilde{\Delta G}_{eu} \in \mathcal{RH}_\infty^{m_u \times m_u}$ which determine the uncertainty ΔG_{rs} and ΔG_{eu} according to (4.23),(4.24). Then, solve $W \in \mathcal{RH}_\infty^{m_u \times m_r}$ such that the robust cost-function (4.25) is minimized, for the uncertain system given by (4.26),(4.27) and the control-law (4.28).*

The problem is solved by the following theorem.

Theorem 4.3 (Cautious Wiener filter) *Given the nominal systems $G_{es}^o \in \mathcal{RH}_\infty^{m_e \times m_r}$, $G_{rs}^o \in \mathcal{RH}_\infty^{m_r \times m_s}$, $G_{eu}^o \in \mathcal{RH}_\infty^{m_e \times m_u}$ and the spectral-factors $\widetilde{\Delta G}_{rs} \in \mathcal{RH}_\infty^{m_r \times m_r}$, $\widetilde{\Delta G}_{eu} \in \mathcal{RH}_\infty^{m_u \times m_u}$. Assume that $G_{rs}^{aug}(q^{-1}) = \left[\begin{array}{cc} G_{rs}^o(q^{-1}) & \widetilde{\Delta G}_{rs}(q^{-1}) \end{array}\right]$ and $G_{eu}^{aug}(q^{-1}) = \left[\begin{array}{c} G_{eu}^o(q^{-1}) \\ \widetilde{\Delta G}_{eu}(q^{-1}) \end{array}\right]$ do not loose rank $\forall |q| = 1$. Then, let*

$$\left[\begin{array}{cc} G_{rs}^o & \widetilde{\Delta G}_{rs} \end{array}\right] = G_{rs,co}^{aug}\left[\begin{array}{cc} G_{rs,ci1}^{aug} & G_{rs,ci2}^{aug} \end{array}\right] \tag{4.29}$$

$$\left[\begin{array}{c} G_{eu}^o \\ \widetilde{\Delta G}_{eu} \end{array}\right] = \left[\begin{array}{c} G_{eu,i1}^{aug} \\ G_{eu,i2}^{aug} \end{array}\right] G_{eu,o}^{aug} \tag{4.30}$$

be the outer-inner and inner-outer factorizations of the augmented systems G_{rs}^{aug} and G_{eu}^{aug} respectively and $G_{rs,co}^{aug\dagger}$ a left-inverse of $G_{rs,co}^{aug}$ and $G_{eu,o}^{aug\dagger}$ a right-inverse of $G_{eu,o}^{aug}$. Then the Cautious Wiener filter is given by

$$W \;=\; -(G_{eu,o}^{aug})^{\dagger}\left[G_{eu,i1}^{aug*}\,G_{es}^o\,G_{rs,ci1}^{aug*}\right]_{+}(G_{rs,co}^{aug})^{\dagger} \tag{4.31}$$

and minimizes (4.25), which is equivalent to

$$J_{rob} \;=\; \left\|\begin{bmatrix} G_{es}^o & 0_{m_e \times m_r} \\ 0_{m_u \times m_s} & 0_{m_u \times m_r} \end{bmatrix} + \begin{bmatrix} G_{eu}^o \\ \widetilde{\Delta G}_{eu} \end{bmatrix} W \begin{bmatrix} G_{rs}^o & \widetilde{\Delta G}_{rs} \end{bmatrix}\right\|_2^2 + \|\widetilde{\Delta G}_{es}\|_2^2 \tag{4.32}$$

Proof: The proof is obtained by rewriting (4.25) in the form that is minimized by the Causal Wiener filter Theorem 2.1. Using Parseval's rule, the stationarity of $s(k)$ and its independence with the uncertainties ΔG_{es}, ΔG_{rs} and ΔG_{eu}, (4.25) can be written in the frequency domain as follows

$$J_{rob} \;=\; \frac{1}{2\pi}\overline{\mathbf{E}}\left(\mathrm{tr}\int_{-\pi}^{\pi}\Big((G_{es}^o+\Delta G_{es})+(G_{eu}^o+\Delta G_{eu})W(G_{rs}^o+\Delta G_{rs})\Big)(.)^{*}d\omega\right)$$

where $(.)^*$ denotes the complex-conjugate of the preceding factor. Because of the zero-mean property of ΔG_{es}, ΔG_{rs}, ΔG_{eu} and the fact they are mutually independent and independent with G_{es}^o, G_{rs}^o and G_{eu}^o the cost-function J_{rob} can be further rewritten as

$$J_{rob}=\frac{1}{2\pi}\mathrm{tr}\int_{-\pi}^{\pi}\Big\{(G_{es}^o+G_{eu}^oWG_{rs}^o)(.)^{*}+\widetilde{\Delta G}_{es}\widetilde{\Delta G}_{es}^{*}+G_{eu}^oW\widetilde{\Delta G}_{rs}\widetilde{\Delta G}_{rs}^{*}W^{*}G_{eu}^{o*}+$$

$$+\,\widetilde{\Delta G}_{eu}WG_{rs}^oG_{rs}^{o*}W^{*}\widetilde{\Delta G}_{eu}^{*}+\widetilde{\Delta G}_{eu}W\widetilde{\Delta G}_{rs}\widetilde{\Delta G}_{rs}^{*}W^{*}\widetilde{\Delta G}_{eu}^{*}\Big\}d\omega \tag{4.33}$$

$$=\frac{1}{2\pi}\mathrm{tr}\int_{-\pi}^{\pi}\Big\{\Big(\underbrace{\begin{bmatrix} G_{es}^o & 0_{m_e \times m_r} \\ 0_{m_u \times m_s} & 0_{m_u \times m_r} \end{bmatrix}}_{=G_{es}^{aug}}+\underbrace{\begin{bmatrix} G_{eu}^o \\ \widetilde{\Delta G}_{eu} \end{bmatrix}}_{=G_{rs}^{aug}}W\underbrace{\begin{bmatrix} G_{rs}^o & \widetilde{\Delta G}_{rs} \end{bmatrix}}_{=G_{eu}^{aug}}\Big)(.)^{*}+$$

$$+\,\widetilde{\Delta G}_{es}\widetilde{\Delta G}_{es}^{*}\Big\}d\omega \tag{4.34}$$

which is equivalent to (4.32). Since the second term is independent of W, (4.32) is minimized by minimizing only its first term, which has the same structure as the cost-function (2.21) on page 37 minimized in Theorem 2.1. Since it is assumed that $G_{rs}^{aug}(q^{-1})$ and $G_{eu}^{aug}(q^{-1})$ do not loose rank $\forall |q|=1$, according to Theorem 2.1 the filter W which minimizes J_{rob} is given by

$$W \;=\; -G_{eu,o}^{aug\dagger}[G_{eu,i}^{aug*}G_{es}^{aug}G_{rs,ci}^{aug*}]_{+}G_{rs,co}^{aug\dagger}.$$

Because of the zero blocks in G_{es}^{aug}, this expression can be reduced to (4.31). $\square$

From this proof it is inferred, that the effect of the uncertainty ΔG_{rs} is the same

as adding measurement noise on the reference signal, with $H_{rv} = \widetilde{G}_{rs}$ and $H_{ev} = 0_{m_e \times m_v}$, as considered in Section 2.2.3 on page 39. Dual to this, the effect of the uncertainty ΔG_{eu} is the same as the control-effort weighting, with $\widetilde{G}_{eu} = \widetilde{\Delta G}_{eu}$, as considered in Section 2.2.4 on page 41.

Furthermore, (4.33) the uncertainty terms $\widetilde{\Delta G}_{rs}$ and $\widetilde{\Delta G}_{eu}$ push W to zero. This means that in the frequency bands where the uncertainty in G_{rs} and G_{eu} is large, the gain of W is reduced, compared with the Causal Wiener solution based on the nominal system. This is also intuitive, because by reducing the gain of W the effect of the uncertainty on $e(k)$ is also reduced. When choosing $W = 0_{m_u \times m_r}$ the uncertainties ΔG_{rs} and ΔG_{es} has no effect on $e(k)$, but on the other side no performance will be obtained. Choosing W to be the Causal Wiener filter optimal performance is obtained under the nominal condition, but there exists ΔG_{rs} and ΔG_{eu} which severely deteriorate the performance. Between these utmost conditions, choosing W to be the Cautious Wiener filter (4.31) yields the best performance on the average.

Note, that the solution (4.31) is independent of the uncertainty factor $\widetilde{\Delta G}_{es}$. This is also intuitive since W and ΔG_{es} are parallel to each other and ΔG_{es} has zero-mean. Hence, by changing the value of W the effect of the uncertainty ΔG_{es} on $e(k)$ cannot be varied.

The solution (4.31) is obtained by minimizing the cost function (4.34) which equals the H_2 norm of the system with augmented inputs and outputs. Confer this solution with Theorem 1 and Theorem 2 in [9], where only the inputs or outputs of the system are augmented.

The structure of the Cautious Wiener filter (4.31) is very similar to the structure of the Causal Wiener filter (2.20) on page 36, but now the factorizations

$$G_{rs}^o = G_{rs,co}^{aug} G_{rs,ci1}^{aug} \tag{4.35}$$

$$G_{eu}^o = G_{rs,i1}^{aug} G_{eu,o}^{aug} \tag{4.36}$$

are no more outer-inner and inner-outer factorizations anymore. From the outer-inner factorization of G_{rs}^{aug} and the inner-outer factorization of G_{eu}^{aug} in Theorem 4.3, it is observed that

$$G_{rs,co}^{aug} G_{rs,co}^{aug*} = G_{rs}^o G_{rs}^{o*} + \widetilde{\Delta G}_{rs} \widetilde{\Delta G}_{rs}^{*}$$

$$G_{eu,o}^{aug*} G_{eu,o}^{aug} = G_{eu}^{o*} G_{eu}^o + \widetilde{\Delta G}_{eu}^{*} \widetilde{\Delta G}_{eu}.$$

There are situations in which the uncertainty is such that

$$G_{rs}^o G_{rs}^{o*} + \widetilde{\Delta G}_{rs} \widetilde{\Delta G}_{rs}^{*} \approx G_{rs}^o G_{rs}^{o*} + \rho_{rs}^2 I \tag{4.37}$$

$$G_{eu}^{o*} G_{eu}^o + \widetilde{\Delta G}_{eu}^{*} \widetilde{\Delta G}_{eu} \approx G_{eu}^{o*} G_{eu}^o + \rho_{eu}^2 I \tag{4.38}$$

with $\rho_{rs}, \rho_{eu} \in \mathbb{R}$. In this case the uncertainty perturbs especially the zeros of the system and all at the same extent. The uncertainty is (approximately) frequency independent and determined by the scalar values ρ_{rs} and ρ_{eu} only. This case is equivalent to increasing the variance of measurement noise $v(k)$ on the reference signal to ρ_{rs}^2 and adding control-effort weighting to the cost-function with a weight of ρ_{eu}^2. Tuning the measurement noise and the control-effort weighting are well

known tools in literature to increase robustness (cf. the LTR approach, e.g., [42], also see [121, Section 8.5]), but here the tuning values ρ_{rs}^2 and ρ_{eu}^2 are chosen with explicit relation with the uncertainty. It may be clear that in case (4.37),(4.38) do not hold, but still values for ρ_{rs}^2 and ρ_{eu}^2 are used to increase robustness, the performance to robustness trade-off may be poor.

4.4 Estimation of the Cautious Wiener filter

4.4.1 Uncertainty in G_{eu}

In the case there is only uncertainty in the secondary path system G_{es} (and thus $G_{es} = G_{es}^o$ and $G_{rs} = G_{rs}^o$), the *control-relevant* estimation of the Cautious Wiener filter is very similar to the control-relevant approach described for the Causal Wiener filter in Section 3.5. The models of $G_{eu,i}$ and $G_{eu,o}$ need to be replaced by models of $G_{eu,i1}^{aug}$ and $G_{eu,o}^{aug}$. The models of $G_{eu,i1}^{aug}$ and $G_{eu,o}^{aug}$ can be derived (using inner-outer factorization, cf. Section 4.3) from the nominal model $\widehat{G}_{eu}^o$ and the uncertainty spectral factor $\widehat{\Delta G}_{eu}$, which are obtained as described in Section 4.2.3.

The control-relevant identification can be performed, because the factor $[G_{eu,i1}^{aug*}G_{es}G_{rs,ci}^*]_+G_{rs,co}^{\dagger}$ of the Cautious Wiener filter (4.31), can be solved from

$$[G_{eu,i1}^{aug*}G_{es}G_{rs,ci}^*]_+G_{rs,co}^{\dagger} = \arg \min_{X\in\mathcal{RH}_\infty^{m_{eu}\times m_r}} J_{id}(X), \qquad (4.39)$$

with

$$J_{id}(X) = \text{tr}\mathbf{E}\left((G_{eu,i1}^{aug*}d(k) - Xr(k))(G_{eu,i1}^{aug*}d(k) - Xr(k))^T\right) \qquad (4.40)$$

which is proven by the next lemma, Lemma 4.1.

Lemma 4.1 *Let* $G_{eu}^{aug} = \begin{bmatrix} G_{eu}^o \\ \widetilde{\Delta G}_{eu} \end{bmatrix} \in \mathcal{RH}_\infty^{m_e+m_u\times m_u}$ *be known, let*

$$W(X) = -G_{eu,o}^{aug\dagger}X, \qquad with\ X \in \mathcal{RH}_\infty^{m_{eu}\times m_r}$$

then minimizing (4.25) over all $X \in \mathcal{RH}_\infty^{m_{eu}\times m_r}$ *is equivalent to minimizing (4.40) (with* $\Delta G_{es} = 0_{m_e\times m_s}$ *and* $\Delta G_{rs} = 0_{m_r\times m_s}$*) over all* $X \in \mathcal{RH}_\infty^{m_{eu}\times m_r}$.

Proof: The proof is along the same lines as the proof of Lemma 3.1. Using the definition of the H_2-norm (Definition 2.1 on page 34) and $W = -G_{eu,o}^{aug\dagger}X$ the expression for the robust control cost-function (4.25), i.e. (4.34), is written as

$$J_{rob} = \frac{1}{2\pi}\text{tr}\int_{-\pi}^{\pi}\left(\begin{bmatrix} G_{es} \\ 0_{m_u\times m_s} \end{bmatrix} - \begin{bmatrix} G_{eu,i1}^{aug} \\ G_{eu,i2}^{aug} \end{bmatrix}XG_{rs}\right)(.)^*d\omega$$

Because $\text{tr}(AB) = \text{tr}(BA)$ and $\begin{bmatrix} G_{eu,i1}^{aug} & G_{eu,i1}^{aug\perp} \\ G_{eu,i2}^{aug} & G_{eu,i2}^{aug\perp} \end{bmatrix}$ is unitary, this can be rewritten

as

$$
\begin{aligned}
J_{rob} &= \frac{1}{2\pi}\mathrm{tr}\int_{-\pi}^{\pi}\left(\begin{bmatrix} G_{es} \\ 0_{m_u\times m_s} \end{bmatrix} - \begin{bmatrix} G_{eu,i1}^{aug} \\ G_{eu,i2}^{aug} \end{bmatrix}XG_{rs}\right)^{*}\begin{bmatrix} G_{eu,i1}^{aug} & G_{eu,i1}^{aug\perp} \\ G_{eu,i2}^{aug} & G_{eu,i2}^{aug\perp} \end{bmatrix}. \\
&\qquad \begin{bmatrix} G_{eu,i1}^{aug*} & G_{eu,i2}^{aug*} \\ G_{eu,i1}^{aug\perp*} & G_{eu,i2}^{aug\perp*} \end{bmatrix}\left(\begin{bmatrix} G_{es} \\ 0_{m_u\times m_s} \end{bmatrix} - \begin{bmatrix} G_{eu,i1}^{aug} \\ G_{eu,i2}^{aug} \end{bmatrix}XG_{rs}\right)d\omega \\
&= \frac{1}{2\pi}\mathrm{tr}\int_{-\pi}^{\pi}\left(G_{eu,i1}^{aug*}G_{es} - XG_{rs}\right)^{*}\left(G_{eu,i1}^{aug*}G_{es} - XG_{rs}\right)d\omega + \\
&\qquad\qquad\qquad\qquad\qquad +\frac{1}{2\pi}\mathrm{tr}\int_{-\pi}^{\pi}G_{es}^{*}G_{eu,i1}^{aug\perp}G_{eu,i1}^{aug\perp*}G_{es}\,d\omega \\
&= J_{id}(X) + \frac{1}{2\pi}\mathrm{tr}\int_{-\pi}^{\pi}G_{es}^{*}G_{eu,i1}^{aug\perp}G_{eu,i1}^{aug\perp*}G_{es}\,d\omega
\end{aligned}
$$

where we used Parseval's equality to write (4.40) in the frequency domain. Hence, minimizing the identification cost-function (4.40) is equivalent to minimizing the control cost-function (4.25) (subject to $X \in \mathcal{RH}_{\infty}^{m_{eu}\times m_u}$) which had to be proven.
□

Hence, using PEM-OE and the input/output data

$$\{r(k), G_{eu,i1}^{*}d(k)\}_{k=1}^{N}$$

the factor $[G_{eu,i1}^{aug*}G_{es}G_{rs,ci}^{*}]_{+}G_{rs,co}^{\dagger}$ of the Cautious Wiener filter can be determined, cf. Section 3.5.1. Similarly, it can be shown that using SMI (cf. Section 3.5.2), and the input/output data

$$\{G_{rs,co}^{\dagger}r(k), G_{eu,i1}^{aug*}d(k)\}_{k=1}^{N}$$

the factor $[G_{eu,i1}^{aug*}G_{es}G_{rs,ci}^{*}]_{+}$ can be determined.

4.4.2   Uncertainty on G_{es}, G_{rs} and G_{eu}

In case there is also uncertainty on G_{es} and G_{rs} the control-relevant identification of the factors $[G_{eu,i1}^{aug*}G_{es}^{o}G_{rs,ci1}^{aug*}]_{+}G_{rs,co}^{aug\dagger}$ (with PEM-OE) and $[G_{eu,i1}^{aug*}G_{es}^{o}G_{rs,ci1}^{aug*}]_{+}$ (with SMI) needs a bit more care. First of all, we assume that the measured data

$$\{r(k), d(k)\}_{k=1}^{N},$$

with $u(k) = 0$, $k = 1, \cdots N$, is measured under the *nominal* system condition, such that

$$
\begin{aligned}
r(k) &= G_{rs}^{o}s(k) \\
d(k) &= G_{es}^{o}s(k).
\end{aligned}
$$

This is a rather strong assumption, which cannot always be achieved in practice, since the nominal (i.e. average) condition cannot always be guaranteed. In case

$s(k)$ can be chosen freely, one can choose the *same realization* of $\{s(k)\}_{k=1}^N$ for all p conditions (cf. Section 4.2.3) and determine $r(k)$ and $d(k)$ by averaging:

$$r(k) \;=\; \frac{1}{p}\sum_{i=1}^p r^i(k), \qquad \text{and } d(k) \;=\; \frac{1}{p}\sum_{i=1}^p d^i(k), \qquad \text{for } k = 1, \cdots, N$$

and with $\{r^i(k)\}_{k=1}^N$ and $\{d^i(k)\}_{k=1}^N$ the reference signal and the disturbance measured under condition i.

Second, let us introduce

$$r_\delta(k) \;=\; r(k) + \nu_\delta(k) \tag{4.41}$$

with

$$\nu_\delta(k) \;=\; \widetilde{\Delta G}_{rs}(q^{-1})\nu(k) \tag{4.42}$$

and $\nu(k) \in \mathbb{R}^{m_u}$ is zero-mean white noise and uncorrelated with $s(k)$, such that

$$\mathbf{E}\left(\begin{bmatrix} \nu(k) \\ 1 \\ s(k) \end{bmatrix} \nu^T(l)\right) = \begin{bmatrix} I_{m_u}\delta_{kl} \\ 0_{1 \times m_u} \\ 0_{m_s \times m_u} \end{bmatrix}.$$

Then, using Parseval's equation, it follows that the power-spectrum of $r_\delta(k)$, denoted by $\Phi_{r_\delta}(e^{-j\omega})$, is given by

$$\begin{aligned}
\Phi_{r_\delta}(e^{-j\omega}) &= G_{rs}^o(e^{-j\omega})G_{rs}^o(e^{-j\omega})^* + \widetilde{\Delta G}(e^{-j\omega})\widetilde{\Delta G}_{rs}(e^{-j\omega})^* \\
&= G_{rs,co}^{aug}(e^{-j\omega})G_{rs,co}^{aug}(e^{-j\omega})^*,
\end{aligned}$$

and thus $G_{rs,co}^{aug\dagger}$ is a whitening filter for $r_\delta(k)$.

Third, let us define the cost-function

$$J_{id}(X) \;=\; \text{tr}\mathbf{E}\left((G_{eu,i1}^{aug*}d(k) - Xr_\delta(k))(G_{eu,i1}^{aug*}d(k) - Xr_\delta(k))^T\right). \tag{4.43}$$

Then, from the next lemma, Lemma 4.2 it follows that the factor $[G_{eu,i1}^{aug*}G_{es}^o G_{rs,ci1}^{aug*}]_+G_{rs,co}^{aug\dagger}$ can be solved from

$$[G_{eu,i1}^{aug*}G_{es}^o G_{rs,ci1}^{aug*}]_+G_{rs,co}^{aug\dagger} \;=\; \min_{X \in \mathcal{RH}_\infty^{m_{eu} \times m_r}} J_{id}(X)$$

This optimization problem can be solved using PEM-OE using the constraint that X should be stable and the input/output data

$$\{r_\delta(k), G_{eu,i1}^{aug*}d(k)\}_{k=1}^N.$$

Lemma 4.2 *Let* $G_{eu}^{aug} = \begin{bmatrix} G_{eu}^o \\ \widetilde{\Delta G}_{eu} \end{bmatrix} \in \mathcal{RH}_\infty^{m_e+m_u \times m_u}$ *be known, let*

$$W(X) \;=\; -G_{eu,o}^{aug\dagger}X, \qquad \text{with } X \in \mathcal{RH}_\infty^{m_{eu} \times m_r}$$

then minimizing (4.25) over all $X \in \mathcal{RH}_\infty^{m_{eu} \times m_r}$ *is equivalent to minimizing (4.43) over all* $X \in \mathcal{RH}_\infty^{m_{eu} \times m_r}$.

Proof: The proof is along the same lines as the proof of Lemma 3.1 and 4.1. Using the definition of the H_2-norm (Definition 2.1 on page 34) and $W = -G_{eu,o}^{aug\dagger} X$ the expression for the robust control cost-function (4.25), i.e. (4.34), is written as

$$J_{rob} = \frac{1}{2\pi}\text{tr}\int_{-\pi}^{\pi}\left(\begin{bmatrix} G_{es}^o & 0_{m_e\times m_r} \\ 0_{m_u\times m_s} & 0_{m_u\times m_r} \end{bmatrix} - \begin{bmatrix} G_{eu,i1}^{aug} \\ G_{eu,i2}^{aug} \end{bmatrix} X \begin{bmatrix} G_{rs}^o & \widetilde{\Delta G}_{rs} \end{bmatrix}\right)(\cdot)^* d\omega +$$

$$+ \widetilde{\Delta}_{es}$$

with

$$\widetilde{\Delta}_{es} = ||\widetilde{\Delta G}_{es}||_2^2.$$

Because $\text{tr}(AB) = \text{tr}(BA)$ and $\begin{bmatrix} G_{eu,i1}^{aug} & G_{eu,i1}^{aug\perp} \\ G_{eu,i2}^{aug} & G_{eu,i2}^{aug\perp} \end{bmatrix}$ is unitary, this can be rewritten as

$$J_{rob} =$$
$$= \frac{1}{2\pi}\text{tr}\int_{-\pi}^{\pi}\left(\begin{bmatrix} G_{es}^o & 0_{m_e\times m_r} \\ 0_{m_u\times m_s} & 0_{m_u\times m_r} \end{bmatrix} - \begin{bmatrix} G_{eu,i1}^{aug} \\ G_{eu,i2}^{aug} \end{bmatrix} X \begin{bmatrix} G_{rs}^o & \widetilde{G}_{rs} \end{bmatrix}\right)^* \begin{bmatrix} G_{eu,i1}^{aug} & G_{eu,i1}^{aug\perp} \\ G_{eu,i2}^{aug} & G_{eu,i2}^{aug\perp} \end{bmatrix}.$$
$$\begin{bmatrix} G_{eu,i1}^{aug*} & G_{eu,i2}^{aug*} \\ G_{eu,i1}^{aug\perp*} & G_{eu,i2}^{aug\perp*} \end{bmatrix}\left(\begin{bmatrix} G_{es} & 0_{m_e\times m_r} \\ 0_{m_u\times m_s} & 0_{m_u\times m_r} \end{bmatrix} - \begin{bmatrix} G_{eu,i1}^{aug} \\ G_{eu,i2}^{aug} \end{bmatrix} X \begin{bmatrix} G_{rs}^o & \widetilde{\Delta G}_{rs} \end{bmatrix}\right)d\omega + \widetilde{\Delta}_{es}$$
$$= \frac{1}{2\pi}\text{tr}\int_{-\pi}^{\pi}\left\{\left(G_{eu,i1}^{aug*}G_{es} - XG_{rs}\right)\left(G_{eu,i1}^{aug*}G_{es} - XG_{rs}\right)^* + X\widetilde{\Delta G}_{rs}\widetilde{\Delta G}_{rs}^* X^*\right\}d\omega +$$

$$+ \frac{1}{2\pi}\text{tr}\int_{-\pi}^{\pi} G_{es}^* G_{eu,i1}^{aug\perp}G_{eu,i1}^{aug\perp*}G_{es}d\omega + \widetilde{\Delta}_{es}$$

$$= J_{id}(X) + \frac{1}{2\pi}\text{tr}\int_{-\pi}^{\pi} G_{es}^* G_{eu,i1}^{aug\perp}G_{eu,i1}^{aug\perp*}G_{es}d\omega + \widetilde{\Delta}_{es}$$

where we used Parseval's equality to write (4.40) in the frequency domain and using the independence between $s(k)$ and $\nu(k)$. Hence, minimizing the identification cost-function (4.43) is equivalent to minimizing the control cost-function (4.25) (subject to $X \in \mathcal{RH}_\infty^{m_{eu}\times m_u}$) which had to be proven. □

Now, it follows that using SMI (cf. Section 3.5.2), and the input/output data

$$\{G_{rs,co}^{aug\dagger}r_\delta(k), G_{eu,i1}^{aug*}d(k)\}_{k=1}^N$$

the factor $[G_{eu,i1}^{aug*}G_{es}G_{rs,ci1}^{aug*}]_+$ can be determined.

4.5   Stability robustness in feedback applications

In the previous sections, Section 4.3 and 4.4, the feedforward controller configuration has been considered. Now, we will consider the feedback configuration
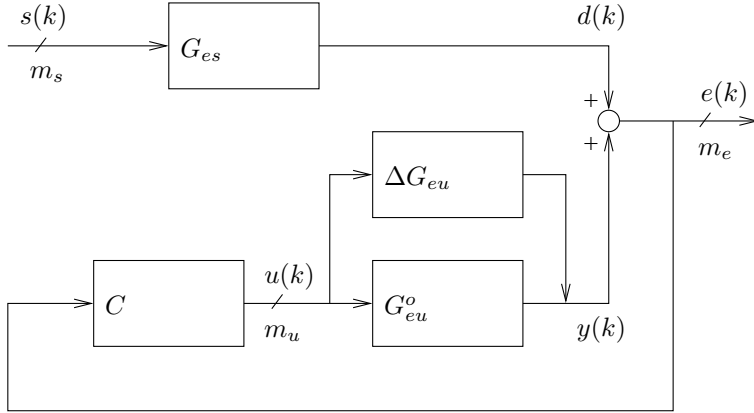
Figure 4.2: Block scheme of the multichannel feedback active control problem with uncertainty in $G_{eu} = G_{eu}^o + \Delta G_{eu}$.

discussed in Section 2.3, but with uncertainty in the secondary path G_{eu}, cf. Figure 4.2. As in Section 4.3 on page 83 we will assume that the model uncertainty $\Delta G_{eu}(e^{-j\omega})$ is a stochastic variable with

$$\overline{\mathbf{E}}\left(\Delta G_{eu}(e^{-j\omega})\right) = 0_{m_e \times m_u}, \qquad -\pi \leq \omega < \pi,$$
$$\overline{\mathbf{E}}\left(\Delta G_{eu}(e^{-j\omega})^* \Delta G_{eu}(e^{-j\omega})\right) = \widetilde{\Delta G}_{eu}(e^{-j\omega})^* \widetilde{\Delta G}_{eu}(e^{-j\omega}), \quad -\pi \leq \omega < \pi$$

and the nominal system G_{eu}^o and the uncertainty spectral factor $\widetilde{\Delta G}_{eu}$ are known.

However, contrary to the case without uncertainty the IMC controller (cf. (2.41) on page 44)

$$C = (I_{m_u} + W G_{eu}^o)^{-1} W \tag{4.44}$$

illustrated in Figure 4.3(a) does *not* internally stabilize the closed loop for all $W \in \mathcal{RH}_\infty^{m_u \times m_e}$ and $\Delta G_{eu} \neq 0_{m_e \times m_u}$! The presence of $\Delta G_{eu} \neq 0_{m_e \times m_u}$ highly complicates the controller design, because ΔG_{eu} yields a feedback, and thus $e(k)$ is not linear in W. In fact

$$e(k) = \left(G_{es} + (G_{eu}^o + \Delta G_{eu})(I_{m_u} - W\Delta G_{eu})^{-1} W G_{es}\right) s(k) \tag{4.45}$$

which is equivalent to $e(k)$ generated by the system depicted in Figure 4.3(b). From (4.45) it is derived, that the loop-gain of the closed-loop is given by

$$L(q^{-1}) = W(q^{-1})\Delta G_{eu}(q^{-1}) \tag{4.46}$$

cf. Figure 4.3(b). Then, according to the small gain theorem, see, e.g., [199], stability is guaranteed if $W \in \mathcal{RH}_\infty^{m_u \times m_e}$, $\Delta G_{eu} \in \mathcal{RH}_\infty^{m_e \times m_u}$ and the following inequality is satisfied

$$\|W(e^{-j\omega})\|_\infty < \frac{1}{\|\Delta G_{eu}(e^{-j\omega})\|_\infty}, \qquad -\pi \leq \omega < \pi. \tag{4.47}$$

PSfrag replacements
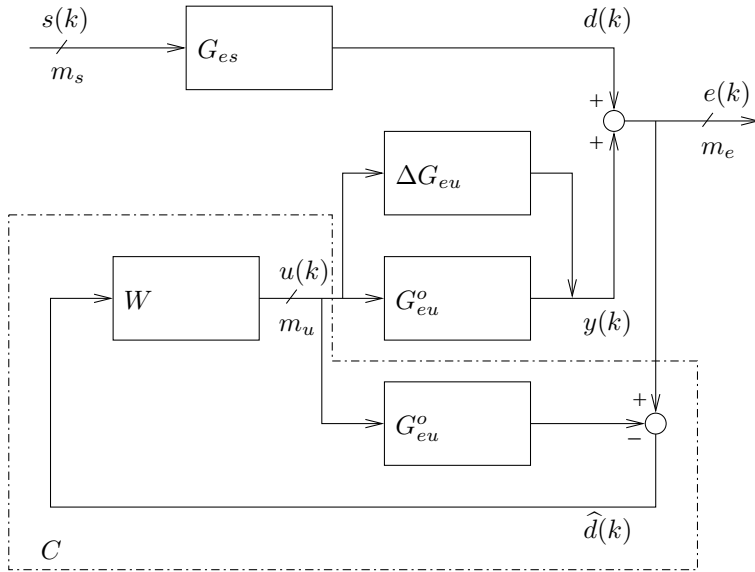
(a) IMC controller based on G_{eu}^o.

PSfrag replacements
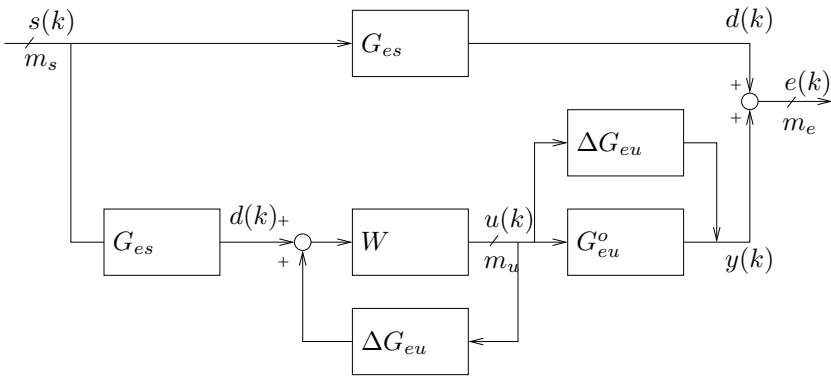
(b) Simplified block scheme obtained using the IMC controller based on G_{eu}.

Figure 4.3: Block schemes of the system with the IMC controller based on the nominal system G_{eu}^o and the simplified block scheme which illustrates the feedback due to ΔG_{eu}.

with $||A||_\infty$ is the maximum singular value of the complex-valued matrix $A \in \mathbb{C}^{m_x \times m_y}$. From (4.47) we infer that —loosely speaking— the gain of W has to be decreased at those frequencies where ΔG_{eu} is 'large' on the average. This is exactly what is achieved by the Cautious Wiener filter, cf. Section 4.3, where the uncertainty factor $\widetilde{\Delta G_{eu}}$ pushes W to zero at the frequencies ω where $\widetilde{\Delta G}_{eu}(e^{-j\omega})$ is significant. However, when using the Cautious Wiener filter (4.31) (with $G_{rs} = G_{es}$ for the pure-feedback case, cf. Section 2.3) the condition (4.47) is still not *guaranteed* to be satisfied. Therefore, we will introduce a nonnegative tuning parameter $\rho \in \mathbb{R}$, which allows to push W stronger to zero, by replacing

$$\widetilde{\Delta G_{eu}} \to \sqrt{\rho}\widetilde{\Delta G_{eu}} \tag{4.48}$$

in Theorem 4.3. Choosing $\rho = 1$ is equivalent to the Cautious Wiener. By choosing $\rho > 1$ the gain of W will be more reduced such that (4.47) will be satisfied for some ρ.

Note, that here the frequency dependency of ΔG_{eu} is explicitly taken into account, whereas in H_∞-theory the frequency dependency of model errors is implicitly taken into account by means of choosing the proper weighting functions. Furthermore, the introduction of the tuning parameter ρ is similar to the introduction of a user chosen ∞-norm bound (often indicated by γ) on the uncertainty; both parameters can be increased to improve stability robustness.

In the design of W, based on the Cautious Wiener filter, it is implicitly assumed that there is no feedback, which is not a correct assumption as is clear from (4.45) and Figure 4.3(b) which shows indeed feedback from ΔG_{eu}. However, assuming that

$$||W\Delta G_{eu}||_\infty \ll 1$$

and thus (4.47) is certainly satisfied, we can write

$$(I_{m_u} - W\Delta G_{eu})^{-1} = I_{m_u} + W\Delta G_{eu} + (W\Delta G_{eu})^2 + (W\Delta G_{eu})^3 + \cdots \tag{4.49}$$
$$\approx I_{m_u}. \tag{4.50}$$

Hence (4.45) can be approximated by

$$e(k) \approx \big(G_{es} + (G_{eu}^o + \Delta G_{eu})WG_{es}\big)s(k) \tag{4.51}$$

which justifies neglecting the feedback term ΔG_{eu}. This approximation of $e(k)$ to be linear in W, is also obtained by taking the first and second terms of the Taylor series expansion of 4.45 with respect to W. The linear approximation is comparable with the linearization of the (closed-loop) sensitivity function in [73].

Table 4.1 summarizes the obtained robust algorithm based on subspace model identification, which we call the robust SIANC algorithm.

4.6 Practical demonstration on a vibrating plate

The vibrating plate considered in Section 3.6.2 on page 68 has been used again, now to validate the Cautious Wiener design method in combination with IMC to

Table 4.1: Robust SIANC algorithm to estimate the robust (feedback) controller

1. Estimate models $\{\widehat{G}_{eu}^k\}_{k=1}^p$ under p conditions;

2. Estimate average model $\widehat{G}_{eu} = \frac{1}{p}\sum_{k=1}^p \widehat{G}_{eu}^k$ and the spectral factor $\widehat{\Delta G}_{eu}$ such that $\widehat{\Delta G}_{eu}^* \widehat{\Delta G}_{eu} = \frac{1}{p}\sum_{k=1}^p (\widehat{G}_{eu}^k - \widehat{G}_{eu}^o)^* (\widehat{G}_{eu}^k - \widehat{G}_{eu}^o)$;

3. Estimate model $\widehat{G}_{es,co}^o$ (using measurements under p conditions);

4. Choose ρ and calculate inner-outer factorization of $\widehat{G}_{eu}^{aug} = \begin{bmatrix} \widehat{G}_{eu}^o \\ \sqrt{\rho}\widehat{\Delta G}_{eu} \end{bmatrix}$;

5. Estimate model $[\widehat{G_{eu,i1}^{aug*}G_{es,co}^o}]_+$;

6. Estimate (reduced order) filter $\widehat{G}_{eu,o}^{aug-1}[\widehat{G_{eu,i1}^{aug*}G_{es,co}^o}]_+ \widehat{G}_{es,co}^{-1}$;

7. Apply controller which consists of Wiener filter $\widehat{G}_{eu,o}^{aug-1}[\widehat{G_{eu,i1}^{aug*}G_{es,co}}]_+ \widehat{G}_{es,co}^{-1}$ in closed loop with internal model $\widehat{G}_{eu}^o$ (in output-normal form).

account for the feedback. The vibrating plate is controlled under two different operating conditions: without and with additional mass, of $\approx 6\%$ of the plate mass, mounted on the plate at the place between the 4 sensors, see Figure 4.4. The mass variation was such that using the nominal design of Chapter 3 did not stabilize the closed loop.

Identification of the model G_{eu}^o and the uncertainty $\widetilde{\Delta G}_{eu}$. Under both conditions a state-space model $\widehat{G}_{eu}^i$, $i = 1, 2$ of the secondary path was estimated using the PO-MOESP subspace model identification method [181] using band limited white noise as the excitation signal. Here, a short description of the identification procedure is given, for a more detailed description of the procedure to identify models for active control applications see [179]. Two input/output data sequences each of 14000 samples (i.e. a measurement of 7 sec.) were recorded, one for identification and the other for validation of the model. Both 4-inputs 4-outputs models $\widehat{G}_{eu}^1$ and $\widehat{G}_{eu}^2$ are of order 80 and accurately model the dynamics of the secondary path without and with additional mass mounted on the plate respectively. A measure of the accuracy of the model is the Variance Accounted For (VAF), which was defined by (3.16) on page 64. Table 4.2 gives the VAF values obtained by $\widehat{G}_{eu}^i$, $i = 1, 2$ and $\widehat{G}_{eu}^o \approx (\widehat{G}_{eu}^1 + \widehat{G}_{eu}^2)/2$ on validation data measured under both conditions with and without additional mass. It can be concluded, that the model obtained under *no additional mass* condition is not accurate anymore under the *additional mass* condition and visa versa. The average model $\widehat{G}_{eu}^o$ models both conditions much better, but still has significant model errors for which the controller should be robust.
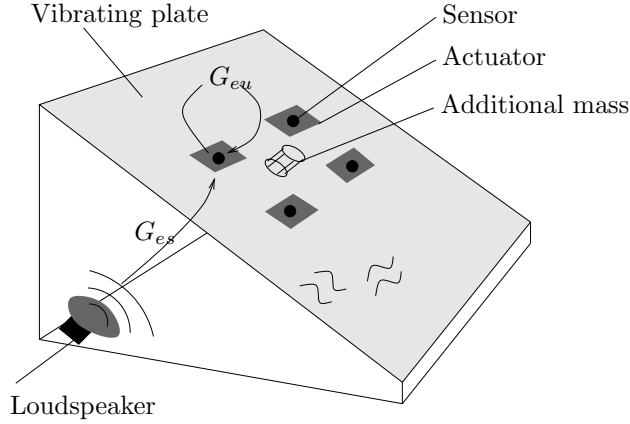
Figure 4.4: Schematic picture of the vibrating plate setup with mass variation.

Table 4.2: VAF values (%) obtained by $\widehat{G}_{eu}^{i}$, $i = 1, 2$ and $\widehat{G}_{eu}^{o}$ on validation data for the conditions with and without additional mass.

	output 1	output 2	output 3	output 4
$VAF(y_{no\ mass}, \widehat{G}_{eu}^{1}u)$:	99.76	99.79	99.73	99.77
$VAF(y_{with\ mass}, \widehat{G}_{eu}^{1}u)$:	81.77	87.58	82.88	88.28
$VAF(y_{no\ mass}, \widehat{G}_{eu}^{2}u)$:	83.10	87.68	84.74	88.09
$VAF(y_{with\ mass}, \widehat{G}_{eu}^{2}u)$:	99.72	99.78	99.73	99.77
$VAF(y_{no\ mass}, \widehat{G}_{eu}^{o}u)$:	95.39	96.66	95.76	96.71
$VAF(y_{with\ mass}, \widehat{G}_{eu}^{o}u)$:	95.03	96.63	95.24	96.76

Figure 4.5(a) shows the largest singular value of $\widehat{G}_{eu}^{i}(e^{-j2\pi f/f_s})$, $i = 1, 2$, which shows that the first and the fourth resonance frequencies significantly change in case an additional mass is mounted. The average secondary path model $\widehat{G}_{eu}^{o}$ and the model error model $\widehat{\Delta G}_{eu}$ are estimated as described in Section 4.2.3. The average model $\widehat{G}_{eu}^{o}$ has order 80, and $\widehat{\Delta G}_{eu}$ has order 50. Figure 4.5(b) shows the largest singular value $\|\widehat{\Delta G}_{eu}(e^{-j2\pi f/f_s})\|_{\infty}$. It is clearly seen, that the model error is depending on the frequency.

Estimation of the average spectral factor of the disturbance. The spectral factor $G_{es,co}$ of the disturbance signal changes when an additional mass is mounted on the vibrating plate. Therefore, like in modeling the secondary path, an average model of the spectral factor $G_{es,co}^{o}$ has been estimated by averaging the measured disturbances under both conditions. The estimated spectral factor model $\widehat{G}_{es,co}^{o}$ has order 40. The filter $\widehat{G}_{es,co}^{o-1}$ is an approximation of the whitening filter for the disturbance signal. Figure 4.6 shows the spectra of the disturbance signal of output 1 (dashed) and the spectrum obtained by using the whitening filter $\widehat{G}_{es,co}^{o-1}$ (solid)
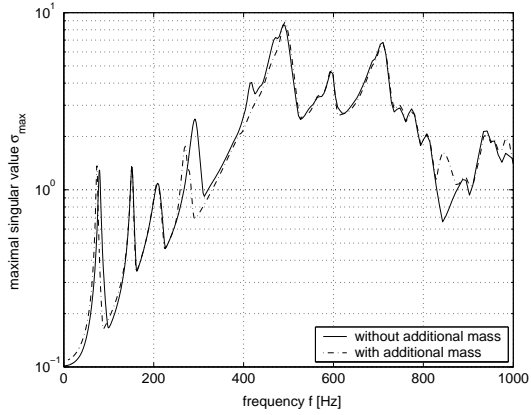
(a) $\|\widehat{G}^1_{eu}(e^{-j2\pi f/f_s})\|_\infty$ (without mass added, solid) and $\|\widehat{G}^2_{eu}(e^{-j2\pi f/f_s})\|_\infty$ (mass added, dashed)
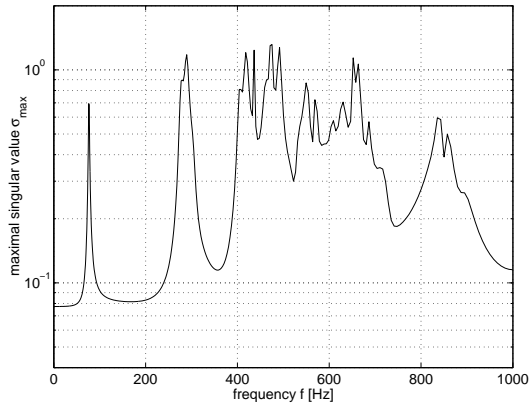
(b) $\|\widehat{\Delta G}_{eu}(e^{-j2\pi f/f_s})\|_\infty$

Figure 4.5: Maximal singular value of estimated model with and without additional mass (a) and their difference (b).

for the average disturbance, the measured disturbance without additional mass and the measured disturbance with additional mass respectively.

From these figures it can be concluded that $\widehat{G}_{es,co}^{o-1}$ indeed approximately whitens the disturbance signal under both conditions. However at 300Hz and 600Hz there are still resonances in the whitened disturbance signal, which are due to the fact that the spectrum of the disturbance signal differs under both conditions. At high frequencies, above ≈ 700Hz, the spectrum of the disturbance signal falls down.

Robust controller estimation and validation.   The robust controller was determined in two different ways. One way is to minimize the nominal cost-function with additional input weighting, hence without additionally filtering the input. This is equivalent to minimizing the robust cost-function (4.33) on page 85 and setting $\widetilde{\Delta G}_{es} = \widetilde{\Delta G}_{rs} = 0$ and replacing $\widetilde{\Delta G}_{eu}$ by just $\sqrt{\rho}I_{m_u}$. The other way is the method is based on the Cautious Wiener approach for feedback systems as outlined in Section 4.5. In this approach, the control signal is filtered with $\sqrt{\rho}\widehat{\Delta G}_{eu}$ to reduce the control effort in the frequency regions where the uncertainty is large. In both methods the value of $\rho > 0$ was varied to increase (and decrease) robustness.

Like in the design of the nominal controller, the order of the robust controllers was reduced to 80 by means of solving an identification problem. Figure 4.7(a) shows the measured average reduction in dB's over all 4 outputs versus ρ obtained by the controller determined without frequency dependent regularization for the case without (o's) and with (+'s) additional mass. For $\rho \geq 0.13$ it is observed that the closed loop was stabilized by the controller. Figure 4.7(b) shows the measured average reduction in dB's over all 4 outputs versus ρ obtained by the controller determined with frequency dependent regularization also for the case without (o's) and with (+'s) additional mass. In this figure, ρ is scaled with a factor 33 for visualization reasons. The maximum performance is obtained for $33\rho = 5, ..., 10$ for the cases with and without additional mass. Note, that for maximum performance, the weighting $33\rho > 1$ to increase stability robustness.

Comparing both figures, it is concluded that using the frequency dependent regularization as in the Cautious Wiener design method may yield better performance. This is also what is expected, because the reduction of the control effort is emphasized only at the critical frequencies (i.e. where the uncertainty is large).

4.7   Conclusions

In this chapter, the model uncertainty description has been considered as a stochastic variable, with zero mean and known, frequency dependent, covariance. The covariance is modeled by means of a (co-)spectral factor, given in state-space description. Based on this model uncertainty description the robust feedforward controller —Cautious Wiener filter— is design by minimizing the mean-squared error averaged over the distribution of the model uncertainty. In this way, the average performance is optimized.

It has been shown that the magnitude of the controller is reduced at those frequencies where the uncertainty in the detector path G_{rs} and secondary path G_{eu} is large, such that the contribution of the uncertainty to the residual disturbance
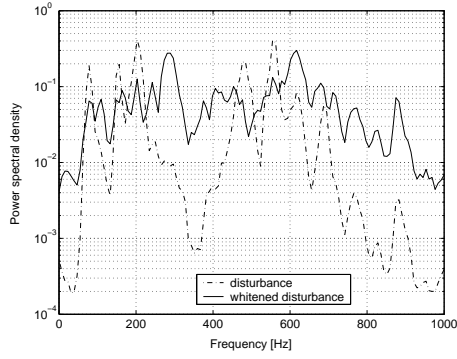
(a) Average disturbance and whitened.
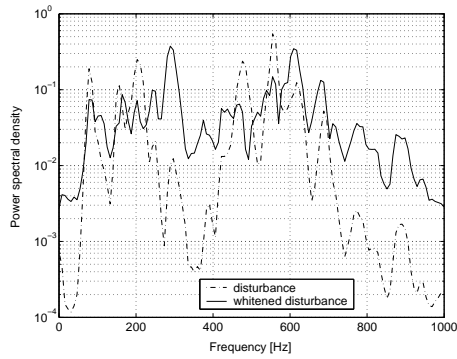
(b) Disturbance without additional mass and whitened.
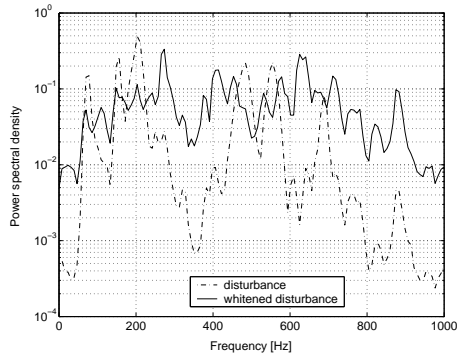
(c) Disturbance with additional mass and whitened.

Figure 4.6: Spectrum of disturbance at output 1 under various conditions (dashed) and pre-whitened (solid).
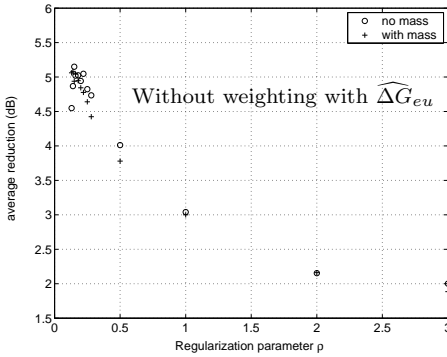
(a) Performance versus ρ *without* weighting.
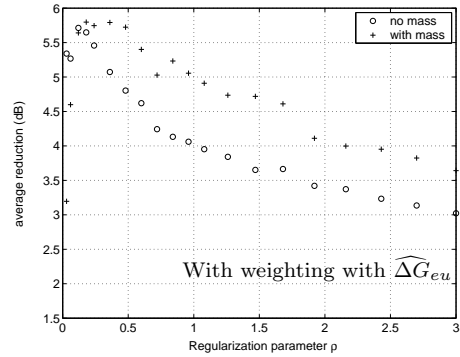
(b) Performance versus ρ *with* weighting.

Figure 4.7: Average performance over all outputs (dB) versus regularization parameter ρ *without* (a) and *with* (b) weighting with $\widehat{\Delta G}_{eu}$ for the case without additional mass (o's) and with additional mass (+'s).

is reduced. Furthermore, the uncertainty on G_{rs} acts in the same way as measurement noise on the reference signal which power spectrum is equal to the frequency dependent covariance of ΔG_{rs}. Dually, the uncertainty in G_{eu} acts in the same way as adding control effort weighting (prefiltered with the co-spectral factor of the frequency dependent covariance of ΔG_{eu}) to the nominal cost-function.

The control-relevant identification methods of Chapter 3 have been extended to estimate the robust feedforward controller. Furthermore, for the feedback application using the IMC compensation, it has been shown, using a small gain argument, that the stability robustness of the closed loop has been improved.

Finally, the robust controller design method was illustrated on a MIMO vibrating plate experimental setup without and with additional mass. The nominal controller of Chapter 3 did not stabilize the closed loop. By introducing sufficient control weighting the closed-loop was stabilized. It has also been shown that weighting the control signal with the probabilistic model error model could lead to a less conservative controller and hence better performance could be obtained.

In Chapter 6 the robust controller design approach taken in this chapter, has been used to improve the robustness of the Filtered-X LMS and the Preconditioned Filtered-X LMS algorithms.

Part II

Online/sample-by-sample algorithms

.

CHAPTER 5

CONVERGENCE ANALYSIS OF FILTERED-U LMS

5.1 Introduction

The Filtered-U LMS (FuLMS) algorithm, proposed by Eriksson et al. [49] for active noise control applications with acoustical feedback, adapts the coefficients of an infinite impulse response (IIR) controller. The motivation of Eriksson et al. was using an adaptive transferfunction model, the autoregressive behavior of the denominator would compensate for the acoustical feedback. The algorithms can be applied for feedforward applications as well. For example, an IIR controller is preferred over an FIR controller in case the optimal controller has one or more poles close to the unit circle, which yields a very long impulse response of the optimal controller. The FuLMS algorithm is based on the LMS algorithm for adaptive IIR filtering proposed by Feintuch [50]. However, due to the non-quadratic nature of the cost function to be minimized in case of updating an IIR filter, convergence to a local minimum might occur, as illustrated in [84] (see also [157]).

The convergence of the algorithm proposed by Feintuch, which is a pseudo linear regression (PLR), is analyzed by various approaches, see [106] for stochastic approaches and [144] for a deterministic approach using the small-gain theorem. A well known approach to derive conditions for global convergence of PLR's was proposed by Ljung in [101, 102] and is referred to as the ordinary differential equation (ODE) method. Using this method, the main condition for global convergence of PLR's is that a certain transfer function should be strictly positive real (SPR). In [187] the ODE method has been applied to derive conditions for global convergence of the FuLMS algorithm (see also [123] and for the multiple-channel case see [125]) for feedforward control applications[1]. To use the ODE method it was assumed that there exists a controller which yields perfect cancellation. This assumption is equivalent to the assumption that the system is in the model set, which is assumed in the analysis of recursive identification methods using the ODE

---

[1]Note, that the derivation of conditions for global convergence of the FuLMS algorithm for *feedback* applications is still an open question, as pointed out in [168], c.f. the discussion in Section 5.3 below.

method [106]. However, in most AC applications perfect cancellation is not achievable, due to pure delays and non-minimum phase zeros in the system, which are often present in real situations.

The main contribution of this chapter is to show, that the assumption requiring perfect cancellation is not necessary to analyze the convergence of the FuLMS algorithm via the ODE method. The relaxation of the assumptions, made in the convergence analysis in [123,125,187], supports the practical evidence known by the practical engineer, that the FuLMS algorithm can even converge when canceling the primary noise is not perfectly possible. This new global convergence result is based on the analysis of the Causal Wiener filter, which yields optimal performance in the mean square error sense (in case of no intrinsic feedback). The main step in the derivation is to show that the residual disturbance obtained by the Causal Wiener filter is stochastically independent of the regression vector, which is necessary to apply the ODE method.

Though the FuLMS algorithm converges to the global minimum under certain conditions presented in this chapter, its convergence rate can be very slow. The structure of the Causal Wiener filter suggests methods of preconditioning to increase the convergence speed of the FuLMS algorithm. For FxLMS it can be shown that preconditioning increases the convergence rate by prewhitening the reference signal and exploiting the inner-outer factorization of the secondary path transfer function [45, 47, 127] (see also [160, 195] for comparable algorithms to improve the convergence rate of FxLMS). Similar preconditioning can be used to increase the convergence rate of the FuLMS algorithm. Therefore, the second contribution of this chapter is the presentation of a preconditioned FuLMS algorithm, which can increase the convergence rate significantly as demonstrated by simulation examples.

The chapter is organized as follows. Section 5.2 recapitulates the feedforward active control problem, and discusses the cases of perfect cancellation by the Wiener filter and optimal cancellation by the Causal Wiener filter. These results give insight in the conditions for which perfect cancellation is not achievable, and on the characteristics of the remaining residual signal. Section 5.3 recapitulates briefly the existing conditions for global convergence of the FuLMS algorithm for feedforward control [187]. Section 5.4 shows that the non-trivial assumption of perfect cancellation is not necessary and presents the new theorem for global convergence of the FuLMS algorithm. Section 5.5 presents the preconditioned FuLMS algorithm which can increase the convergence rate significantly as is demonstrated in Section 5.6 by two simulation examples. Finally, Section 5.7 concludes the chapter.

The content of this chapter was published before in the publication in Signal Processing [58].

5.2 Problem formulation

Feedforward control problem. Recall the feedforward control problem of Section 2.2 on page 31, with measurement noise on the residual signal only, which is again depicted in Figure 5.1. For reasons of simplicity, in this chapter we consider only the single channel case, as considered in the convergence analysis in [187].
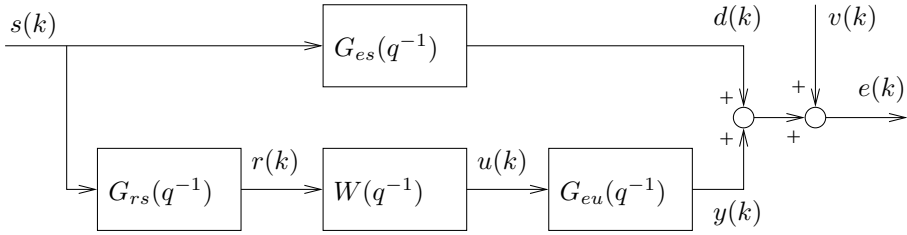
Figure 5.1: Block scheme of the feedforward active control problem with measurement noise.

Thus, the residual signal is given by

$$e(k) \;=\; G_{es}s(k) + G_{eu}u(k) + v(k) \tag{5.1}$$

with the primary path $G_{es}(q^{-1}) \in \mathcal{RH}_\infty^{1\times1}$ (in the following, we will denote $\mathcal{RH}_\infty^{1\times1}$ just by $\mathcal{RH}_\infty$) and $G_{eu}(q^{-1}) \in \mathcal{RH}_\infty$, $s(k)$ the zero-mean white noise disturbance with unit variance, $u(k)$ the control signal and $v(k)$ zero-mean (possibly colored) measurement noise uncorrelated with $s(l)$ (for all l) and with variance σ_v^2.

The reference signal is given by

$$r(k) \;=\; G_{rs}s(k) \tag{5.2}$$

with the detector path $G_{rs} \in \mathcal{RH}_\infty$. Note that it is assumed that $r(k)$ is not distorted by feedback of $u(k)$ via G_{ru}. This means that the transfer function G_{ru} is not present in the system to be controlled, or it is compensated.

The control-law is given by

$$u(k) \;=\; Wr(k), \tag{5.3}$$

with $W \in \mathcal{RH}_\infty$. Then, the problem is to minimize the following cost function

$$J \;:=\; \mathbf{E}\left(e^2(k)\right), \tag{5.4}$$

Perfect and optimal cancellation.   If $G_{eu}^{-1}G_{es}G_{rs}^{-1}$ is stable, then choosing W to be

$$W \;=\; -G_{eu}^{-1}G_{es}G_{rs}^{-1}$$

(i.e. the Wiener filter) would yield perfect cancellation, i.e.,

$$e(k) \;=\; v(k). \tag{5.5}$$

However, often $G_{eu}^{-1}G_{es}G_{rs}^{-1}$ is *not* stable, as illustrated in Section 2.2.1 on page 31, due to non-minimum phase zeros in G_{eu} and G_{rs} which are not canceled by non-minimum phase zeros in G_{es}. In this case, the constraint that W should be stable has to be explicitly taken into account in the optimization problem.

In Section 2.2.1 on page 31 the solution of this feedforward control problem has been given by Theorem 2.1 on page 36 which provides the *Causal* Wiener filter.

Here, we briefly recall the results for the single channel case. Let $G_{es}, G_{rs}, G_{eu} \in \mathcal{RH}_\infty$ and both $G_{rs}(q^{-1}) \neq 0$ and $G_{eu}(q^{-1}) \neq 0$ for any $|q| = 1$ (i.e. G_{rs} and G_{eu} do not have zeros on the unit-circle in the complex plane). Then, the outer-inner factorization of G_{rs} and the inner-outer factorization of G_{eu} exist and are given by

$$G_{rs} = G_{rs,co}G_{rs,ci}$$
$$G_{eu} = G_{eu,i}G_{eu,o}$$

respectively. Because G_{rs} and G_{eu} do not have zeros on the unit circle we have that $G_{rs,co}$ and $G_{eu,o}$ are minimum phase and thus $G_{rs,co}^{-1}, G_{eu,o}^{-1} \in \mathcal{RH}_\infty$. Then the Causal Wiener filter which minimizes (5.4) is given by

$$W \;=\; -G_{eu,o}^{-1}[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ G_{rs,co}^{-1} \tag{5.6}$$

(see Theorem 2.1).

By filling in the Causal Wiener filter in (5.3), and using (5.1),(5.2), we obtain the expression for the *optimal* residual signal

$$e(k) \;=\; G_{eu,i}[G_{eu,i}^* G_{es} G_{rs,ci}^*]_- G_{rs,ci}s(k) + v(k) \tag{5.7}$$

Hence, in case perfect cancellation is *not* achievable by a causal controller, all the contribution of $s(k-i)$, $-\infty < i < \infty$ to the value of J can *not* be removed.

The existing convergence result of the FuLMS algorithm in [187] (and [125] for the multiple channel case), that will be discussed in the next section, is under the condition of perfect cancellation, such that there exists a controller W such that (5.5) will be obtained. In Section 5.4 we will show that the assumption of perfect cancellation can be dropped and the convergence result can be extended to the case perfect cancellation is not achievable.

5.3   Derivation Filtered-U LMS and existing convergence result

Derivation Filtered-U LMS.   The FuLMS algorithm has been proposed by Eriksson et al. [49] to minimize (5.4) by an *adaptive* IIR controller. However, no proof of convergence to the optimal controller has been given. In [187] conditions were derived for the FuLMS algorithm to converge to the optimal controller in case perfect cancellation is achievable. Also the case with acoustical feedback G_{ru} has been considered in [187]. However in the analysis in [187] for the case with acoustical feedback, a wrong *wrong* regression vector has been used, which consists of the reference signal *not* distorted by feedback of the control signal. This *un*distorted reference signal is not available (except for the case the feedback is perfectly canceled), as was pointed out by Sun and Meng in [168]. The wrong convergence result for systems with acoustical feedback was cited in [46, 125] as well as the convergence analysis presented in [58]. Though, the result in [58] has been formulated for the general case *with* acoustic feedback G_{ru}, and thus is not correctly, the result still holds for the case *without* acoustic feedback, $G_{ru} = 0$, which will be discussed in Section 5.4. The problem of deriving sufficient conditions

for global convergence of the FuLMS algorithm for the case $G_{ru} \neq 0$ is still an open and left for further research.

Therefore, in this section, we only recall this convergence result for the FuLMS algorithm in [187] for the case *without* acoustic feedback.

Let the controller W in (5.3) be replaced by the time-varying controller

$$W(q^{-1}, k) \ := \ \frac{A(q^{-1}, k)}{B(q^{-1}, k)} \ = \ \frac{a_0(k) + a_1(k)q^{-1} + \cdots + a_N(k)q^{-N}}{1 + b_1(k)q^{-1} + \cdots + b_M(k)q^{-M}}. \quad (5.8)$$

Let us define the controller parameter vector $\theta(k)$ and the vector $\phi(k)$ by

$$\theta(k) = \begin{bmatrix} a_0(k) & a_1(k) & \cdots & a_N(k) & b_1(k) & \cdots & b_M(k) \end{bmatrix}^T \quad (5.9)$$

$$\phi(k) = \begin{bmatrix} r(k) & r(k-1) & \cdots & r(k-N) & u(k-1) & \cdots & u(k-M) \end{bmatrix}^T \quad (5.10)$$

respectively. In the following, we will assume that N and M are, such that the optimal controller (5.6) is contained in the structure given by (5.8). This means, that the optimal controller can be written as

$$W(q^{-1}) = \frac{A^o(q^{-1})}{B^o(q^{-1})} \ = \ \frac{a_0^o + a_1^o q^{-1} + \cdots + a_N^o q^{-N}}{1 + b_1^o q^{-1} + \cdots + b_M^o q^{-M}}. \quad (5.11)$$

Then, the optimal parameter vector is defined as

$$\theta^o \ = \ \begin{bmatrix} a_0^o & a_1^o & \cdots & a_N^o & b_1^o & \cdots b_M^o \end{bmatrix}^T.$$

Now, the control signal $u(k)$ can be written as

$$u(k) \ = \ \phi^T(k)\theta(k).$$

Further, it can be shown, see e.g. [187], that the gradient (5.4) with respect to the controller parameters $\theta(k)$ is (approximately) given by

$$\nabla_{\theta(k)}J \ = \ 2\mathbf{E}\left(\left\{ G_{eu}(q^{-1})\frac{1}{B(q^{-1}k)}\phi(k) \right\} e(k) \right).$$

Note, that this gradient is clearly a non-linear function of the controller coefficients $\theta(k)$, hence local minima might exist. In the FuLMS algorithm the gradient is approximated by neglecting the expectation operator $\mathbf{E}(.)$ which results in an LMS-type algorithm and neglecting the filtering with $1/\big(B(q^{-1}, k)\big)$ which results in a PLR-type algorithm. Hence, the gradient is approximated as

$$\widehat{\nabla}_{\theta(k)}J \ = \ 2\big(\widehat{G}_{eu}\phi(k)\big)e(k), \quad (5.12)$$

where $\widehat{G}_{eu}$ models the secondary path G_{eu}. Now, the FuLMS update-equation for the controller-coefficients is given by

$$\theta(k+1) \ = \ \theta(k) - \gamma(k)\big(\widehat{G}_{eu}\phi(k)\big)e(k), \quad (5.13)$$

with $\gamma(k)$ a positive adaptation gain.

The approximation of the gradient (5.12) significantly reduces the complexity of the update-equation. However, this approximation is valid only if the negative gradient is pointing downward on the average, such that on the average the cost function will be minimized and convergence will be obtained. To satisfy this, sufficient conditions are given in the following (Theorem 5.1 and 5.2).

Existing convergence result. Ljung's ODE method to prove convergence of PLR's to their global optimum was used in [187] to analyze the convergence of the FuLMS algorithm. In order to guarantee that the residual signal $e(k)$ is uncorrelated with the regression vector $\widehat{G}_{eu}\phi(k)$, which is a necessary assumption to apply Ljung's ODE method, it was assumed that perfect cancellation is achievable. The FuLMS convergence theorem proven in [187] is summarized in the following theorem.

Theorem 5.1 (Convergence of FuLMS [187]) *Let, N and M be such that the optimal controller (5.11) is contained in the adaptive filter structure (5.8). Then, under the following conditions:*

1. *perfect cancellation is achievable, i.e. $G_{eu}^{-1}G_{es}G_{rs}^{-1} \in \mathcal{RH}_{\infty}$;*

2. $\dfrac{G_{eu}}{B^0\widehat{G}_{eu}}$ *is SPR*

3. *$\gamma(k) \to 0$ for $n \to \infty$;*

4. *persistence of excitation in the reference signal $r(k)$ specified in [101];*

5. *regularity and boundedness conditions specified in [101];*

by Ljung's ODE Theorem [101], the FuLMS algorithm (5.13) converges with probability one to the unique equilibrium point θ^o.

In [187] assumption *(1)* was formulated slightly differently. Namely, it was stated that the primary path G_{es} needs to be given by $G_{eu}(-W)G_{rs}$, with $W \in \mathcal{RH}_{\infty}$. Then [187] shows that the adaptive controller will converge to the optimal controller A^0/B^0. This controller yields perfect cancellation, because A^0 and B^0 are such that $G_{eu}(-W)G_{rs} = G_{es}$ and $W \in RH_{\infty}$, and thus $G_{eu}^{-1}G_{es}G_{rs}^{-1} \in \mathcal{RH}_{\infty}$.

Assumption *(4)* states that $\gamma(k)$ vanishes. This assumption is common in asymptotic convergence analysis, see also [101].

In the next section we will show, that the first assumption in Theorem 5.1 can be omitted.

5.4 Convergence of FuLMS when perfect cancellation is not achievable

In Section 5.2 we inferred that the optimal performance was completely determined by the Causal Wiener filter (the intrinsic feedback was compensated for by IMC). By exploiting the structure of the Causal Wiener filter, global convergence of the FuLMS algorithm can be proven also for the case perfect cancellation is not achievable, as shown in the proof of the next theorem.

Theorem 5.2 (General convergence of FuLMS) *Let, N and M be such that the optimal controller (5.11) is contained in the adaptive filter structure (5.8). Then, under the following conditions:*

1. *$\widehat{G}_{eu} = G_{eu,i}\widehat{G}_{eu,o}$ with $\widehat{G}_{eu,o} \in \mathcal{RH}_{\infty}$;*
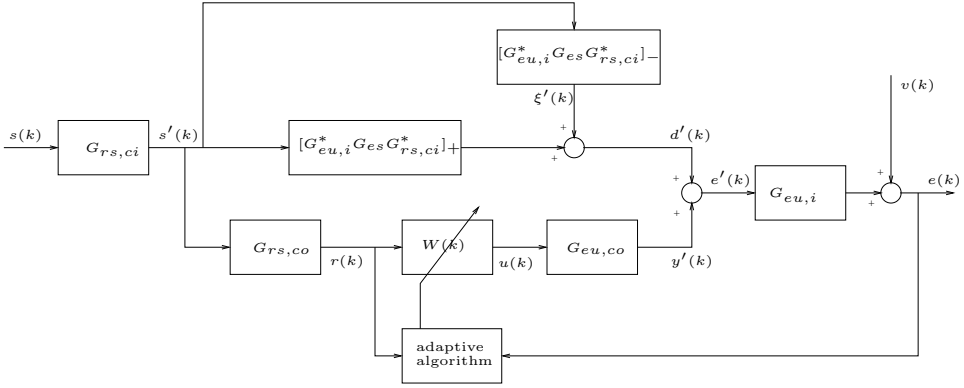
Figure 5.2: Block scheme of the adaptive Active Control problem with decomposed plant.

2. $\dfrac{G_{eu}}{B^0 \widehat{G}_{eu}}$ is SPR;

3. $\gamma(k) \to 0$ for $n \to \infty$;

4. persistence of excitation in the reference signal $r(k)$ as specified in [101];

5. regularity and boundedness conditions specified in [101];

by Ljung's ODE Theorem [101], the FuLMS algorithm (5.13) converges with probability one to the unique equilibrium point θ^o.

Proof: We show that the ODE method can be used to prove global convergence of $W(q^{-1}, k)$. Crucially, to apply the ODE method is that the *optimal residual signal* $e(k)$ given by (5.7) is *uncorrelated* with the regression vector $\widehat{G}_{eu}\phi(k)$. To satisfy this constraint in the recursive system identification configuration, it is usually assumed that the system is in the model set and there is no correlation between the measurement noise and the input of the system. That the optimal residual signal is uncorrelated to the regression vector, also for the case perfect cancellation is not achievable, is shown in the following. The proof can be completed similarly to the proof given in [187].

First, note that because $G_{rs,ci}$ and $G_{eu,i}$ are unitary, we can decompose G_{es} in the following way

$$G_{es} = G_{eu,i}[G^*_{ei,i}G_{es}G^*_{rs,ci}]_{+}G_{rs,ci} + G_{eu,i}[G^*_{eu,i}G_{es}G^*_{rs,ci}]_{-}G_{rs,ci},$$

and thus the residual signal $e(k)$ can be written as (cf. Figure 5.2)

$$e(k) = G_{eu,i}[G^*_{eu,i}G_{es}G^*_{rs,ci}]_{+}G_{rs,ci}s(k) + G_{eu,i}[G^*_{eu,i}G_{es}G^*_{rs,ci}]_{-}G_{rs,ci}s(k) + \\ + G_{eu,i}G_{eu,o}WG_{rs,co}G_{rs,ci}s(k) + v(k). \tag{5.14}$$

By filling in the Causal Wiener filter (5.6) in (5.14), the first and the third terms in (5.14) cancel each other, which yields the optimal residual error (cf. Eq. (5.7))

$$e(k) = \xi(k) + v(k), \tag{5.15}$$

with $\xi(k) = G_{eu,i}[G_{eu,i}^* G_{es} G_{rs,ci}^*]_- G_{rs,ci} s(k)$.

It is obvious that $v(k)$ and $\widehat{G}_{eu}\phi(k)$ are uncorrelated. To show that $\xi(k)$ and $\widehat{G}_{eu}\phi(k)$ are independent, i.e.

$$\mathbf{E}\left((\widehat{G}_{eu}\phi(k))\xi(k)\right) = 0,$$

we have to show that $\mathbf{E}\left(\{\widehat{G}_{eu}r(k-i)\}\xi(k)\right) = 0, \forall i \geq 0$.

Because $\widehat{G}_{eu} = G_{eu,i}\widehat{G}_{eu,o}$ and $\mathbf{E}\left(s^2(k)\right) = 1$, we have

$$
\begin{aligned}
\mathbf{E}\left(\{\widehat{G}_{eu}r(k-i)\}\xi(k)\right) &= \\
&= \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} e^{-j\omega i} G_{eu,i}\widehat{G}_{eu,o}G_{rs,co}G_{rs,ci}G_{rs,ci}^*[G_{eu,i}^* G_{es} G_{rs,ci}^*]_-^* G_{eu,i}^* d\omega, \\
&= \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} e^{-j\omega i} \widehat{G}_{eu,o}G_{rs,co}[G_{eu,i}^* G_{es} G_{rs,ci}^*]_-^* d\omega.
\end{aligned}
$$

Because $[G_{eu,i}^* G_{es} G_{rs,ci}^*]_-$ is anti-causal $[G_{eu,i}^* G_{es} G_{rs,ci}^*]_-^*$ is causal, but has no direct feed through. Thus there exists a causal filter $F(q^{-1})$ such that

$$[G_{eu,i}^* G_{es} G_{rs,ci}^*]_-^* = q^{-1}F.$$

Hence

$$\mathbf{E}\left(\{\widehat{G}_{eu}r(k-i)\}\xi(k)\right) = \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} e^{-j\omega(i+1)}\widehat{G}_{eu,o}G_{rs,co}F d\omega.$$

Since $G_{es}, G_{rs}, G_{eu}, \widehat{G}_{eu} \in \mathcal{RH}_\infty$ and thus have no poles on the unit-circle, $\widehat{G}_{eu,o}G_{rs,o}F$ has no poles on the unit-circle. Hence, $\widehat{G}_{eu,o}G_{rs,o}F$ can be written as a Laurent series (see, e.g., [143])

$$\widehat{G}_{eu,o}G_{rs,o}G = \sum_{i=-\infty}^{\infty} \gamma_i q^i.$$

Because $\widehat{G}_{eu,o}G_{rs,o}F$ is causal, $\gamma_i = 0, \forall i > 0$. Furthermore, we have

$$\frac{1}{2\pi} \int\limits_{-\pi}^{\pi} e^{-j\omega(i+1)}\widehat{G}_{eu,o}G_{rs,co}F d\omega = \gamma_{i+1}.$$

Combining these results, we get

$$\mathbf{E}\left(\{\widehat{G}_{eu}r(k-i)\}\xi(k)\right) = \gamma_{i+1} = 0, \quad \forall i \geq 0.$$

Thus, the optimal residual signal (5.15) is uncorrelated with the regression vector $\widehat{G}_{eu}\phi(k)$. Therefore, the ODE method can be used as in [187] by considering $\xi(k)$ just as additional measurement noise. By following the same reasoning as in [187] a differential equation can be derived which describes the evolution of the adaptive controller coefficients $\theta(k)$. Under conditions *(2)–(4)* in the theorem, especially the

SPR condition, it is shown that the differential equation has a *unique* stationary point in θ^o which is also stable. Hence, $\theta(k)$ converges to θ^o. $\qquad\square$

Modification of $\widehat{G}_{eu,o}$ can be used as an instrument to attempt to get the SPR condition in assumption *(2)* satisfied. This was also stated in [187]. An alternative is to filter the residual signal (see, e.g., [157]), because then the secondary path is also altered. However, in this case the variance of the filtered residual signal is minimized. This is not equivalent to the minimization of the variance of the actual (unfiltered) residual signal, and thus may yield a biased solution.

5.5 The preconditioned FuLMS algorithm

By Theorem 5.2 we know under which conditions the FuLMS algorithm converges to the Causal Wiener filter. However, the rate of convergence can be extremely slow, which is illustrated by the simulation examples in Section 5.6. For the Filtered-X LMS (FxLMS) algorithm, in [45, 47] a prewhitening and the use of the inner-outer factorization of the secondary path transfer function were proposed in a FIR/frequency domain context to increase the convergence rate. Similar preconditioning can be applied to increase the convergence rate of the FuLMS algorithm. In the following, we will assume there is no intrinsic feedback or it has been compensated.

Because it is assumed that $G_{rs}(q^{-1})$ does not have any zeros on the unit circle, we have that the power-spectrum of $r(k)$ is given by

$$\Phi_r(e^{-j\omega}) \;=\; G_{rs,co}(e^{-j\omega})G_{rs,co}(e^{-j\omega})^* > 0, \qquad -\pi \le \omega \le \pi,$$

and $G_{rs,co}$ minimum phase. Hence, $G_{rs,co}^{-1} \in \mathcal{RH}_\infty$ is a whitening filter for $r(k)$. Let

$$s'(k) \;=\; G_{rs,co}^{-1}r(k) \;=\; G_{rs,ci}s(k),$$

then $s'(k)$ is a white noise signal. By using the prewhitening filter $G_{rs,co}^{-1}$ the convergence rate has become independent of the detector path. To make the convergence rate also independent of the secondary path, the following adaptive controller structure is proposed

$$W(q^{-1}, k) \;=\; G_{eu,o}^{-1}(q^{-1})W'(q^{-1}, k)G_{rs,co}^{-1}(q^{-1}), \qquad (5.16)$$

with

$$W'(q^{-1}, k) \;=\; \frac{a'_0(k) + a'_1(k)q^{-1} + \cdots a'_K(k)q^{-K}}{1 + b'_1(k)q^{-1} + \cdots b'_L(k)q^{-L}}.$$

K and L are the orders of the numerator and the denominator polynomials of $[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+$ respectively. Note, that $-[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+$ is the optimal value of $W'(q^{-1}, k)$, cf. (5.6). Furthermore, note that the order of the numerator and denominator polynomials of $W'(q^{-1}, k)$ and the Causal Wiener filter (5.6) may

PSfrag replacements

$r(k)$
$u(k)$ $s(k)$

$v'(k)$
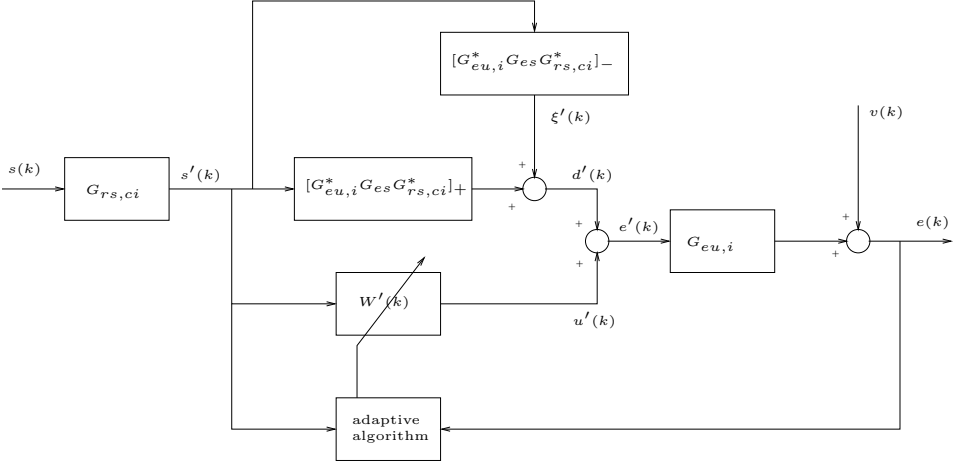
$G^*_{eu,i}$

Figure 5.3: Block scheme of the preconditioned adaptive Active Control problem with decomposed plant.

differ (it can be shown, that in general the order of the Causal Wiener filter given by (5.6) is two times larger than the order of $[G^*_{eu,i}G_{es}G^*_{rs,ci}]_+$, cf. Appendix B). Hence the number of adaptive filter coefficients for the preconditioned FuLMS algorithm may be different for the FuLMS algorithm.

Applying the controller (5.16), the control problem is reduced to the control problem illustrated in Figure 5.3. Let

$$\theta'(k) \; := \; \begin{bmatrix} a'_0(k) & a'_1(k) & \cdots & a'_K(k) & b'_1(k) & \cdots & b'_L(k) \end{bmatrix}^T,$$
$$\phi'(k) \; := \; \begin{bmatrix} s'(k) & s'(k-1) & \cdots & s'(k-K) & -u'(k-1) & \cdots & -u'(k-L) \end{bmatrix}^T,$$

with

$$u'(k) = W'(k)s'(k)$$
$$u(k) = G^{-1}_{eu,o}u'(k).$$

Then the update equation of the preconditioned FuLMS algorithm is given by

$$\theta'(k+1) \; = \; \theta'(k) - \gamma(k)[G_{eu,i}\phi'(k)]e(k).$$

Hence, the vector $\phi'(k)$ has to be filtered only by the all-pass filter $G_{eu,i}$.

Due to the nonlinearity of the update law given by the FuLMS algorithm, it is very difficult to prove that the convergence rate of the preconditioned FuLMS algorithm has been increased. Therefore, the increase of the convergence rate for the preconditioned FuLMS algorithm has been demonstrated by simulation examples in the next section.

5.6 Simulations and experimental validation

5.6.1 An illustrative example

In the first example, the AC system is given by

$$G_{es}(q^{-1}) = \frac{1+0.9q^{-1}}{1+0.5q^{-1}}, \quad G_{rs}(q^{-1}) = \frac{1-1.1q^{-1}}{1+0.5q^{-1}}, \quad G_{eu}(q^{-1}) = \frac{1+1.1q^{-1}}{1+0.5q^{-1}},$$

and $G_{ru} = 0$. Note, that G_{rs} and G_{eu} have non-minimum phase zeros at 1.1 and -1.1 respectively and G_{es} has *no* non-minimum phase zeros at 1.1 and -1.1, hence perfect cancellation is not achievable. The Causal Wiener filter (5.6) to which the FuLMS should converge and the factor $[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+$ to which the preconditioned FuLMS should converge are given by

$$-G_{eu,o}^{-1}[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ G_{rs,co}^{-1} = \frac{0.7491 + 0.9892q^{-1} + 0.3074q^{-2}}{1 - 0.8264q^{-2}},$$

$$-[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ = \frac{0.9064 + 0.7438q^{-1}}{1 + 0.5q^{-1}}.$$

The variance of the measurement noise is zero, i.e. $\sigma_v^2 = 0$. Then the optimal value of the cost function (5.4) is $J \approx 0.2792$ and the variance of the disturbance is $\mathbf{E}\left(d^2(k)\right) \approx 1.2133$. With $s(k)$ a white noise process, using $\widehat{G}_{eu} = G_{eu}$, all conditions in Theorem 5.2 for global convergence are satisfied, for FuLMS as well as preconditioned FuLMS. For the SPR conditions this is illustrated by the Nyquist plots of $1/B^o$ for FuLMS and preconditioned FuLMS in Figure 5.4(a) and 5.4(b) respectively.

In the (preconditioned) FuLMS experiments the normalized (preconditioned) FuLMS algorithm has been used, i.e. $\gamma(k)$ in (5.13) is replaced by

$$\gamma(k) = \frac{\mu}{||\widehat{G}_{eu}\phi(k)||_2^2},$$

with the step size $\mu = 0.01$. Figure 5.5(a) and 5.5(b) show the convergence of the filter-coefficients over $1 \cdot 10^6$ and $1 \cdot 10^4$ samples for FuLMS and preconditioned FuLMS respectively.

From the figures, it can be inferred that the filter coefficients using the FuLMS and the preconditioned FuLMS indeed converge to the optimal values. This is in agreement with Theorem 5.2. Note, that Theorem 5.1 does not hold in this case, because perfect cancellation is not achievable. The value of the cost function obtained by (preconditioned) FuLMS is calculated using the last $1 \cdot 10^4$ samples of the residual signal. For the standard FuLMS algorithm we get

$$J_{\text{FuLMS}} = \mathbf{E}\left(e^2(k)\right) \approx \frac{1}{10^4} \sum_{n=10^6-10^4+1}^{10^6} e^2(k) = 0.3016,$$

and for the preconditioned FuLMS algorithm we get

$$J_{\text{precon. FuLMS}} = \mathbf{E}\left(e^2(k)\right) \approx \frac{1}{10^4} \sum_{n=10^6-10^4+1}^{10^6} e^2(k) = 0.2957.$$

(a) $1/B^o$ for FuLMS.

(b) $1/B^o$ for preconditioned FuLMS.
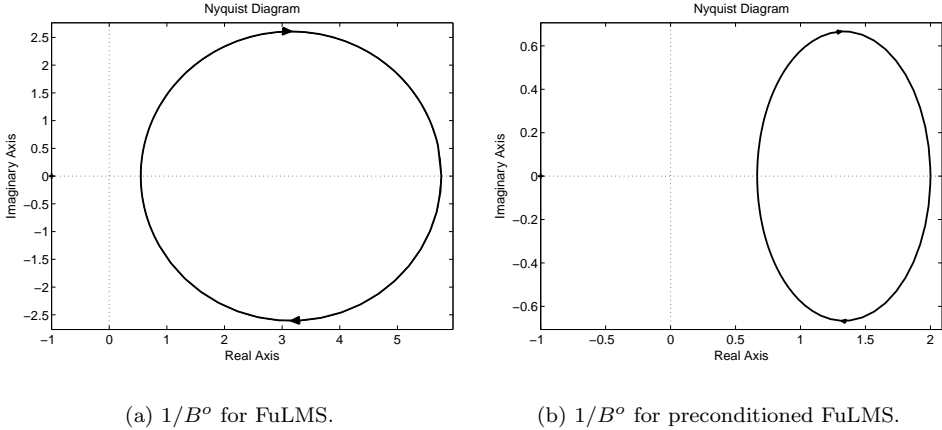
Figure 5.4: Nyquist plots of $1/B^o$ for FuLMS and preconditioned FuLMS illustrating the SPR condition.

These values are higher than the minimal value of the cost function, which can be explained by the fact that the step size μ does not converge to zero but is fixed, which results in misadjustment.

Furthermore, it can be concluded from Figure 5.5(a) and 5.5(b), that the convergence rate of the preconditioned FuLMS algorithm is much faster than the convergence rate of the standard FuLMS algorithm.

5.6.2 Acoustical duct system

In the second simulation experiment, the AC system is a 30th order model of the acoustical duct used in Section 3.6.1 on page 64, see Figure 3.1 on page 65. The model has been identified using the Subspace Model Identification method [181]. In the simulations the acoustical feedback has been compensated perfectly.

The Causal Wiener filter given by (5.6), for this system is a 60th order filter. The reduction in decibels (dB's) obtained by using this filter is 14.2dB. The optimal filter to which the preconditioned FuLMS should converge to, i.e. $-[G^*_{eu,i}G_{es}G^*_{rs,ci}]_+$, is a 30th order filter.

For the standard FuLMS and the preconditioned FuLMS algorithm, the SPR condition in Theorem 5.2 is not satisfied, hence both algorithms may not converge to the global minimum. In the simulations the filter order used in the standard and the preconditioned FuLMS algorithm are chosen to be the same, both numerator and denominator are of order 30. The cost function is estimated by the following recursive filter

$$\widehat{J}(k) = \alpha\widehat{J}(k-1) + (1-\alpha)e^2(k), \qquad \widehat{J}(0) = e^2(0),$$

with $\alpha = 0.999$. Figure 5.6(a) shows the learning curves $\widehat{J}(k)$ averaged over 100 experiments for the (normalized) FuLMS and preconditioned (normalized) FuLMS

(a) Filter coefficients of FuLMS for 10^6 samples.

(b) Filter coefficients of preconditioned FuLMS for 10^4 samples.
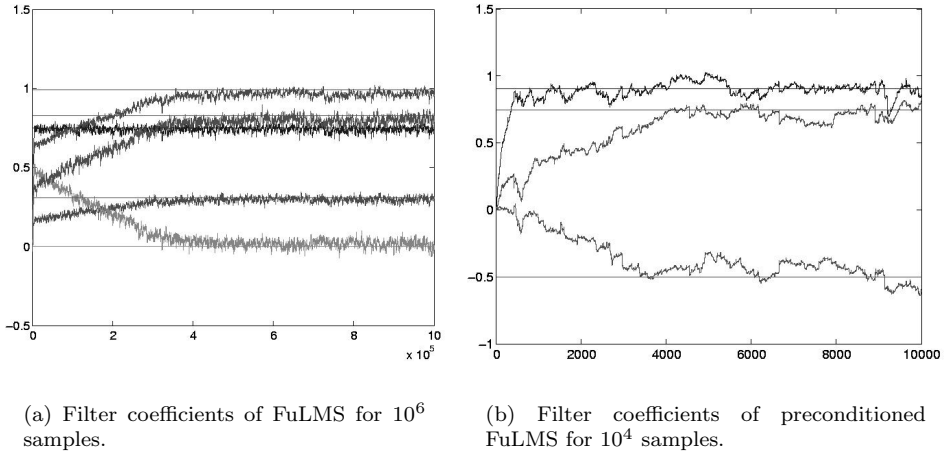
Figure 5.5: Convergence of the filter coefficients for FuLMS and preconditioned FuLMS with $\mu = 0.01$.

algorithm for $2 \cdot 10^5$ samples using a step size $\mu = 0.1$. Figure 5.6(b) zooms in on the first $1 \cdot 10^4$ samples. From these figures, it is also clear that the convergence rate of the preconditioned FuLMS algorithm is improved compared with the standard FuLMS algorithm.

An interesting fact is that after $1.4 \cdot 10^5$ samples the performance obtained by the FuLMS algorithm is better than the performance of the preconditioned FuLMS algorithm. Because the SPR conditions for global convergence of the FuLMS and the preconditioned FuLMS algorithm are not satisfied, both may converge to local minima. Hence, we may conclude that the preconditioned FuLMS algorithm has been converged to a local minimum, which results in less disturbance rejection than the (possibly local) minimum to which the FuLMS algorithm has converged.

Figure 5.6(a) and 5.6(b) also show the learning curves of the FxLMS algorithm and the preconditioned FxLMS algorithm proposed in [47]. The order of the FIR filters was chosen to be 60 such that the number of filter coefficients used in the (preconditioned) FxLMS algorithm is the same as in the (preconditioned) FuLMS algorithm. From these figures we may conclude, that the convergence rate of the preconditioned FxLMS algorithm is largest.

Also the reduction obtained by the preconditioned FxLMS algorithm (10.9dB after convergence) is better than the reduction of the other adaptive algorithms. Because the preconditioned FxLMS algorithm minimizes a quadratic criterion, it converges to the global minimum which is determined by the length of the FIR filter. However, because the Causal Wiener filter has an IIR structure, optimal performance (14.2dB reduction) could not be obtained by the preconditioned FxLMS algorithm as is also clear from Figure 5.6(a).

(a) Convergence curves for $2 \cdot 10^5$ samples.

(b) Convergence curves for the first 10^4 samples.
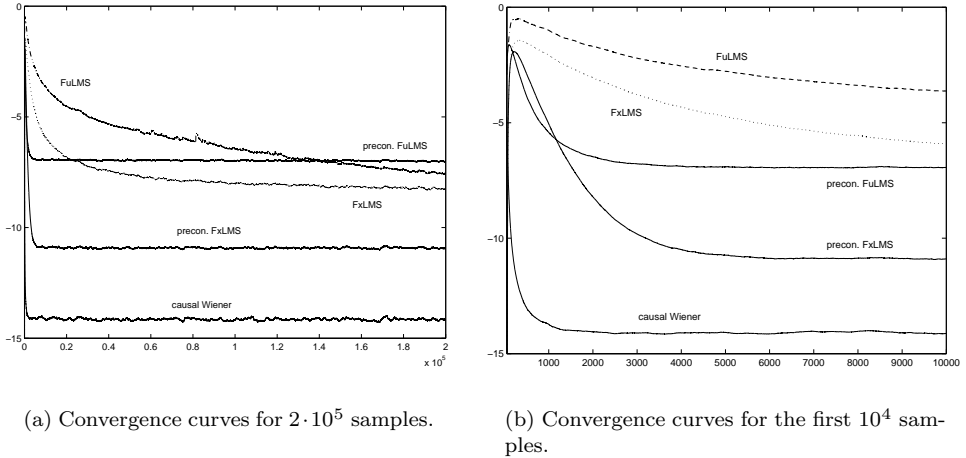
Figure 5.6: Learning curves (in dB) averaged over 100 experiments for the (normalized) FuLMS and preconditioned FuLMS algorithm using a IIR filter with numerator and denominator of order 30, and FxLMS and preconditioned FxLMS using a FIR filter of order 60, all with $\mu = 0.1$.

5.7 Conclusions

The structure of the optimal controller and the remaining residual noise have been analyzed for Active Control (AC) applications. Insights from this analysis have been used to generalize the convergence results of the Filtered-U LMS (FuLMS) algorithm which were derived in [187] for the case perfect cancellation is not achievable. It is shown that the regression vector used in the FuLMS algorithm is always uncorrelated with the residual signal obtained by using the optimal controller (in case the secondary path model contains the inner factor of the real secondary path). Hence, SPR conditions for global convergence derived in [187] using Ljung's ODE method in the case with and without intrinsic (or acoustical) feedback, also hold in the case perfect cancellation is not achievable.

We also proposed a preconditioned FuLMS algorithm which yields a considerable increase of the convergence speed in an academic and a realistic acoustical duct simulation experiment. The preconditioned FuLMS algorithm contains a prewhitening filter of the reference signal and applies the inner-outer factorization of the secondary path, as was proposed for the Filtered-X LMS algorithm in [47].

CHAPTER 6

ROBUST PRECONDITIONED FILTERED-X LMS

6.1 Introduction

The Filtered-X LMS (FxLMS) algorithm is a very popular algorithm for feedforward active noise and vibration control, because the implementation is simple and its recursions are well studied (see e.g. [29, 46, 90, 145, 188], just to name a few). In broadband applications the convergence rate of FxLMS may be poor due to correlation in the regression vector. To overcome this problem, [45, 47] proposes a preconditioning of the FxLMS (PFxLMS) algorithm, which removes all correlation in the regression vector. This can increase the convergence rate significantly as shown in [176] for a realistic active control problem.

However, in [45, 47, 176] it was also noted that regularization is necessary in case the system has zeros on and/or close to the unit circle to reliably calculate the prefilters and prevent large amplification of the preconditioning filters, which may yield oversteering of e.g. the DA converters. An even more important problem is, that undermodeling of and variations in the secondary path may yield instability of the filter update algorithm if a particular well known strictly positive real (SPR) condition is not satisfied [125, 188].

The main focus of this chapter is to adjust the PFxLMS algorithm, without allowing too much performance degradation, such that the stability of the filter update algorithm is less sensitive to errors in the secondary path model. Stated otherwise, our objective is to increase the stability robustness of the PFxLMS update algorithm w.r.t. model errors.

In literature, two approaches are proposed to improve the robustness of the update algorithm: 1. on-line secondary path modeling and 2. adjusting the adaptive algorithm to relax the SPR condition. Both approaches have their advantages and drawbacks. On-line secondary path modeling (see e.g. [90]) may keep track of variations in the secondary path and may thus yield optimal performance even if the secondary path varies. However, the computational complexity is increased and injection of an auxiliary dither signal is usually necessary with the consequence of reduced performance.
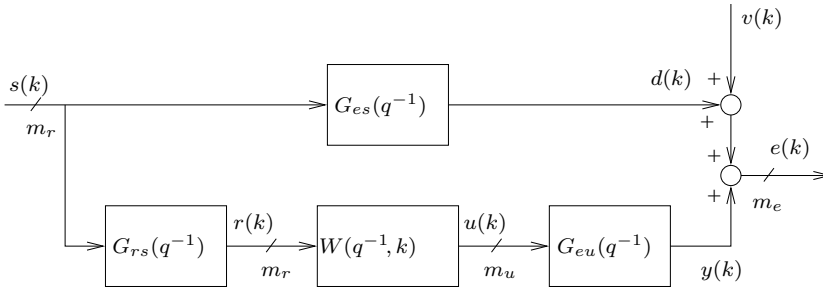
Figure 6.1: Block scheme of the general multichannel feedforward active control system, with m_r reference, m_u control and m_e residual signals.

An example of the second approach, is proposed in [124] where a model is derived which satisfies the SPR condition for *multiple* secondary plant systems by solving the so-called *robust SPR* problem. The method focuses on IIR filtering, but can also be applied to FIR filtering. However, the set of multiple secondary plant systems should satisfy a particular condition for solving the robust SPR problem [124]. Furthermore, for every secondary plant system a different precondition filter would be necessary.

An alternative method which relaxes the SPR condition is control effort weighting. In [45, 46] this was done by tuning a scalar parameter which weights the trace of the control effort covariance matrix, and results in Leakage FxLMS/PFxLMS. Besides the necessity of tuning a scalar regularization parameter, the method may be too conservative.

The contribution of this chapter, is the derivation of the robust versions of both FxLMS and PFxLMS in the framework of *probabilistic robust filtering* approach of Chapter 4. The robust method uses a model uncertainty model of the secondary path, which acts as a *frequency dependent* control effort weighting. As such the method results in a generalization of standard control effort weighting and hence a generalization of Leakage FxLMS/PFxLMS (cf. e.g. [46, 90]). It is shown that the SPR condition is relaxed in a well motivated manner, and hence the stability robustness of the update algorithm is increased. A simulation example shows that this method yields better performance than Leakage FxLMS/PFxLMS.

The chapter is organized as follows. Section 6.2 derives the Robust FxLMS (RFxLMS) algorithm and its new SPR condition for global convergence. Section 6.3 derives the Robust PFxLMS (RPFxLMS) algorithm, shows that large amplification of the precondition filter is prevented and derives the SPR condition for RPFxLMS. Section 6.4 illustrates the method by a simulation example.

This chapter has been published before in Signal Processing Letters, see [59].

6.2   Robust Filtered-X LMS

Consider Figure 6.1, which illustrates the feedforward active control problem (acoustical feedback is neglected or assumed to be perfectly compensated by Internal Model Control). Here, $s(k) \in \mathbb{R}^{m_r}$ represents the signal from the distur-

bance source and is assumed to be a zero mean white noise stochastic process
with $\mathrm{E}[s(k)s^T(l)] = I_{m_r}\delta(k-l) \ \forall k,l$, where $\delta(0) := 1$, $\delta(k) := 0$, $n \neq 0$. Let
$RH_\infty^{m_e \times m_u}$ the set of all stable proper rational $m_e \times m_u$ transfer functions matri-
ces in the unit delay operator q^{-1} with real coefficients. Then the primary path,
the detector path and the secondary path are denoted by $G_{es}(q^{-1}) \in RH_\infty^{m_e \times m_r}$,
$G_{rs}(q^{-1}) \in RH_\infty^{m_r \times m_r}$ and $G_{eu}(q^{-1}) \in RH_\infty^{m_e \times m_u}$ respectively. The adaptive feed-
forward controller is an $m_u \times m_r$ matrix with FIR filters of length n_w and its ℓ, m^{th}
element is given by

$$W_{\ell m}(q^{-1}, k) = \sum_{i=0}^{n_w - 1} w_{\ell m}^i(k) q^{-i}$$

with $w_{\ell m}^i(k) \in \mathbb{R}$. For ease of notation, we define $\mathbf{w}_{\ell m}(k) :=$
$[w_{\ell m}^0(k) \cdots w_{\ell m}^{n_w-1}(k)]^T \in \mathbb{R}^{n_w}$, $\mathbf{w}_\ell(k) := [\mathbf{w}_{\ell 1}^T(k) \cdots \mathbf{w}_{\ell m_r}^T(k)]^T \in \mathbb{R}^{n_w m_r}$ and
$\mathbf{W}(k) := [\mathbf{w}_1(k) \cdots \mathbf{w}_{m_u}(k)] \in \mathbb{R}^{n_w m_r \times m_u}$ and the vector stacking of all con-
troller coefficients $\theta(k) := \mathrm{vec}(\mathbf{W}(k)) \in \mathbb{R}^{m_u n_w m_r}$. The input to the adaptive
filter is the reference signal $r(k) \in \mathbb{R}^{m_r}$, let $\phi(k) := [x_1(k) \cdots x_1(k - n_w + 1) \cdots x_{m_r}(k) \cdots x_{m_r}(k - n_w + 1)]^T \in \mathbb{R}^{n_w m_r}$. Then the control signal $u(k) \in \mathbb{R}^{m_u}$
is given by

$$u(k) := W(q^{-1}, k) r(k) = \mathbf{W}^T(k) \phi(k)$$

The objective is to determine $u(k)$ such that $y(k) \in \mathbb{R}^{m_e}$, counteracts the distur-
bance signal $d(k) \in \mathbb{R}^{m_e}$. The measured residual signal is corrupted with a zero
mean stochastic noise process $v(k) \in \mathbb{R}^{m_e}$, with intensity $\sigma_v^2 := \mathrm{tr}\, \mathrm{E}[v(k)v^T(k)]$,
which is independent of $s(k)$, i.e. $\mathrm{E}[v(k)s^T(l)] = 0_{m_e \times m_r}$, $\forall k, l$. The measured
residual $e(k) \in \mathbb{R}^{m_e}$ is given by

$$e(k) = G_{es}(q^{-1}) s(k) + G_{eu}(q^{-1}) u(k) + v(k)$$

Then, the FxLMS algorithm, which objective is to minimize $\xi = \mathrm{tr}\, \mathrm{E}[e(k)e^T(k)]$
is given by

$$\theta(k+1) = \theta(k) - \gamma(k)[\widehat{G}_{eu}^T(q^{-1}) \otimes \phi(k)] e(k)$$

with $\otimes$ denotes the Kronecker matrix product, $\gamma(k) \geq 0$ the step size. Using
Ljung's [101] ordinary differential equation (ODE) method, [188] (see also [125])
shows that if $\gamma(k)$ suitably vanishes, $r(k)$ is persistently exciting and the following
SPR condition is satisfied

$$G_{eu}^T(e^{j\omega})\widehat{G}_{eu}(e^{-j\omega}) + \widehat{G}_{eu}^T(e^{j\omega})G_{eu}(e^{-j\omega}) > 0, \quad -\pi \leq \omega \leq \pi \qquad (6.1)$$

then the associated ODE, which describes the asymptotic behavior of $\theta(k)$, is
asymptotically stable. Hence, $\theta(k)$ converges, with probability one, to its unique
global optimum [188]

$$\theta_{opt} = -\mathbf{E}\left(\left(\widehat{G}_{eu}^T \otimes \phi(k) \right) \left(G_{eu}^T \otimes \phi(k) \right)^T \right)^{-1} \mathbf{E}\left(\left(\widehat{G}_{eu}^T \otimes \phi(k) \right) d(k) \right)$$

To increase the robustness of the FxLMS algorithm w.r.t. uncertainty in $\widehat{G}_{eu}$,
we may want to have a (probabilistic) model of the uncertainty. Here, we will
follow the idea of the probabilistic robust filtering approach proposed in [164]. We
assume that G_{eu} can be modeled as a stochastic variable, such that

$$G_{eu}(q^{-1}) = G_{eu}^o(q^{-1}) + \Delta G_{eu}(q^{-1})$$

with $\overline{\mathbf{E}}\left(\Delta G_{eu}(e^{-j\omega})\Delta G_{eu}^T(e^{j\omega})\right) = \Phi_{\Delta G_{eu}}(e^{-j\omega})$ and $\overline{\mathbf{E}}\left(G_{eu}(e^{-j\omega})\right) = G_{eu}^o(e^{-j\omega})$, for $-\pi \leq \omega \leq \pi$. $\overline{\mathbf{E}}(.)$ denotes expectation over stochastic systems. Further, let ΔG_{eu} be independent of G_{eu}^o, G_{es}, G_{rs}, $s(k)$ and $v(k)$ $\forall n$. The objective of the robust filtering approach is to minimize

$$\xi_{rob} = \text{tr } \overline{\mathbf{E}}\left(\mathbf{E}\left(e(k)e^T(k)\right)\right) \tag{6.2}$$

By Parseval's equality and the independence between ΔG_{eu} and the other factors,

$$\xi_{rob} = \frac{1}{2\pi}\text{tr}\int_{-\pi}^{\pi}(.)^*(G_{es} + G_{eu}^o W G_{rs}) + G_{rs}^* W^* \Phi_{\Delta G_{eu}} W G_{rs} d\omega + \sigma_v^2$$

$$= \frac{1}{2\pi}\text{tr}\int_{-\pi}^{\pi}(.)^*(G_{es}^{aug} + G_{eu}^{aug} W G_{rs})d\omega + \sigma_v^2$$

here $(.)^*$ indicates complex conjugate transpose of the following factor, $G_{es}^{aug} = [G_{es}^T \; 0_{m_u \times J}]^T$, $G_{eu}^{aug} = [G_{eu}^{oT} \; \widetilde{\Delta G}_{eu}^T]^T$ and $\widetilde{\Delta G}_{eu} \in RH_\infty^{m_e \times m_u}$ such that $\widetilde{\Delta G}_{eu}^* \widetilde{\Delta G}_{eu} = \Phi_{\Delta G_{eu}}$. Now, let $\widehat{G}_{eu} \in RH_\infty^{m_e \times m_u}$ and $\widehat{\Delta G}_{eu} \in RH_\infty^{m_e \times m_u}$ be models of G_{eu}^o and $\widetilde{\Delta G}_{eu}$ respectively and $\widehat{G}_{eu}^{aug} = [\widehat{G}_{eu}^T \; \widehat{\Delta G}_{eu}^T]^T$. Then the Robust FxLMS (RFxLMS) algorithm is obtained by updating the filter coefficients in the direction of (the LMS estimate of) the negative gradient direction of the *robust cost function* ξ_{rob}. This yields the following RFxLMS update rule

$$\theta(k+1) = \theta(k) - \gamma(k)[\widehat{G}_{eu}^{aug \; T}(q^{-1}) \otimes \phi(k)]\begin{bmatrix} e(k) \\ \widehat{\Delta G}_{eu}(q^{-1})u(k) \end{bmatrix}$$

We observe, that the RFxLMS algorithm is identical to the FxLMS algorithm with the secondary path model augmented by $\widehat{\Delta G}_{eu}$ and the performance channels $e(k)$ augmented by $\widehat{\Delta G}_{eu}u(k)$. This additional term reduces the energy of the control signal at frequencies where the uncertainty, i.e. $\Phi_{\Delta G_{eu}}(e^{-j\omega})$, is large. The uncertainty model $\widehat{\Delta G}_{eu}$ can be obtained from e.g. identification of $\widehat{G}_{eu}$, see also [103, 129, 164]. An other approach is by performing a series of identification experiments under different secondary path conditions which yields $\widehat{G}_{eu}$ as the average model and $\widehat{\Delta G}_{eu}$ as a stable spectral factor of the estimated covariance $\widehat{\Phi}_{\Delta G_{eu}}$. In case, $\widehat{\Delta G}_{eu} = \beta I_{m_u}$ with $\beta > 0$ a constant real scalar, the RFxLMS algorithm can be reduced to the Leakage FxLMS algorithm (cf. e.g. [46, 90]).

To derive the SPR condition for the RFxLMS algorithm, we have to rewrite the FxLMS SPR condition (6.1) for the augmented system, which yields

$$G_{eu}^T(e^{j\omega})\widehat{G}_{eu}(e^{-j\omega}) + \widehat{G}_{eu}^T(e^{j\omega})G_{eu}(e^{-j\omega}) + 2\widehat{\Delta G}_{eu}^T(e^{j\omega})\widehat{\Delta G}_{eu}(e^{-j\omega}) > 0,$$
$$-\pi \leq \omega \leq \pi \tag{6.3}$$

Because, $\widehat{\Delta G}_{eu}^T(e^{j\omega})\widehat{\Delta G}_{eu}(e^{-j\omega}) = \widehat{\Phi}_{\Delta G_{eu}}(e^{-j\omega}) \geq 0$ for $-\pi \leq \omega \leq \pi$, the SPR condition is relaxed, especially at frequencies where the magnitude of the uncertainty model is large.

6.3 Robust Preconditioned Filtered-X LMS

The robustness of the PFxLMS algorithm can be increased too by minimizing the *robust* cost function (6.2). Like the preconditioning filters for the FxLMS algorithm are factors of the Causal Wiener filter (see [45, 47]), the robust preconditioning filters are factors of the Robust Wiener filter —called the *Cautious* Wiener filter in [164]— which minimizes (6.2) and is given by

$$W_{rob,opt} = -(G_{eu,o}^{aug})^{-1}[G_{eu,i}^{aug*}G_{es}^{aug}G_{rs,ci}^*]_+G_{rs,co}^{-1}$$

with $[.]_+$ the causality operator, $G_{rs} = G_{rs,co}G_{rs,ci}$ is the outer-inner factorization of G_{rs} and $G_{eu}^{aug} = G_{eu,i}^{aug}G_{eu,o}^{aug}$ is the inner-outer factorization of G_{eu}^{aug}. Note, that $G_{rs,co}^{-1}$ is a whitening filter for the reference signal and $(G_{eu,o}^{aug})^{-1}$ inverts the minimum phase part of the *augmented* secondary path (if $G_{eu,o}^{aug}$ is non square $(G_{eu,o}^{aug})^{-1}$ denotes a right inverse).

Models of $G_{rs,co}^{-1}$ and $(G_{eu,o}^{aug})^{-1}$ can be used to precondition the RFxLMS problem by removing the correlation in the regression vector, which yields the RPFxLMS algorithm:

RPFxLMS algorithm: *The control law is given by*

$$u(k) = (\widehat{G}_{eu,o}^{aug}(q^{-1}))^{-1}\tilde{u}(k)$$
$$\tilde{u}(k) = W(q^{-1},k)\tilde{r}(k)$$
$$\tilde{r}(k) = (\widehat{G}_{rs,co}(q^{-1}))^{-1}r(k)$$

and the update algorithm by

$$\theta(k+1) = \theta(k) - \gamma(k)[(\widehat{G}_{eu,i}^{aug}(q^{-1}))^T \otimes \tilde{\phi}(k)]\begin{bmatrix} e(k) \\ \widehat{G}_{eu,i2}^{aug}(q^{-1})\tilde{u}(k) \end{bmatrix} \quad (6.4)$$

with $\tilde{\phi}(k)$ is defined similar to $\phi(k)$ but $r(k)$ is replaced by $\tilde{r}(k)$, and $\widehat{\Delta G}_{eu}(\widehat{G}_{eu,o}^{aug})^{-1} = \widehat{G}_{eu,i2}^{aug}$ equals the last m_e rows of $\widehat{G}_{eu,i}^{aug}$.

Note, that $\widehat{G}_{eu,o}^{aug*}\widehat{G}_{eu,o}^{aug} = \widehat{G}_{eu}^*\widehat{G}_{eu} + \widehat{\Delta G}_{eu}^*\widehat{\Delta G}_{eu}$ and thus the gain of $(\widehat{G}_{eu,o}^{aug})^{-1}$ will be reduced where $\widehat{\Delta G}_{eu}^*\widehat{\Delta G}_{eu} > 0$, which may prevent oversteering of e.g. the DA converters.

Assuming $\widehat{G}_{rs,co}^{-1} = G_{rs,co}^{-1}$, which is such that $\mathrm{E}[\tilde{r}(m)\tilde{r}^T(k)] = I_{m_r}\delta(m-k)$, it can be proven that the autocorrelation matrix of the regression vector,

$$R = \mathbf{E}\left(\left(\widehat{G}_{eu,i}^{aug\ T} \otimes \tilde{\phi}(k)\right)\left(\widehat{G}_{eu,i}^{aug\ T} \otimes \tilde{\phi}(k)\right)^T\right)$$

equals the identity matrix $I_{m_uIm_r}$. Therefore, under this condition all modes converge at the same rate, which is determined by the step size $\gamma(k)$.

Using the ordinary differential equation (ODE) method as in [188] the following theorem on the convergence of RPFxLMS is obtained.

Theorem 6.1 (Convergence RPFxLMS) *If $\gamma(k)$ suitably vanishes, $r(k)$ is persistently exciting, the regularity conditions of the ODE theorem [101] are satis-*

fied and the following SPR condition holds

$$\left(\widehat{G}_{eu,o}^{aug}(e^{j\omega})\right)^{-T}\left(G_{eu}^{T}(e^{j\omega})\widehat{G}_{eu}(e^{-j\omega})+\widehat{G}_{eu}^{T}(e^{j\omega})G_{eu}(e^{-j\omega})+\right.$$
$$\left.+2\widehat{\Delta G}_{eu}^{T}(e^{j\omega})\widehat{\Delta G}_{eu}(e^{-j\omega})\right)\left(\widehat{G}_{eu,o}^{aug}(e^{-j\omega})\right)^{-1}>0,$$
$$for\ -\pi\leq\omega\leq\pi\ (6.5)$$

then the associated ODE, which describes the asymptotic behavior of $\theta(k)$, is asymptotically stable. Furthermore, $\theta(k)$ converges, with probability one, to its unique global optimum

$$\theta_{opt}=$$
$$-\mathbf{E}\left(\left(\widehat{G}_{eu,i1}^{augT}\otimes\widetilde{\phi}(k)\right)\left((G_{eu}\widehat{G}_{eu,o}^{aug-1})^{T}\otimes\widetilde{\phi}(k)\right)^{T}+\left(\widehat{G}_{eu,i2}^{augT}\otimes\widetilde{\phi}(k)\right)\left(\widehat{G}_{eu,i2}^{augT}\otimes\widetilde{\phi}(k)\right)^{T}\right)^{-1}$$
$$\mathbf{E}\left(\left(\widehat{G}_{eu,i1}^{augT}\otimes\widetilde{\phi}(k)\right)d(k)\right)$$

Proof: The proof is along the same lines as in [188], but with augmented secondary path and precondition filters. □

The SPR condition (6.5) for RPFxLMS is a weighted version of the SPR condition (6.3) for RFxLMS, with weighting function $(\widehat{G}_{eu,o}^{aug})^{-1}$. If $(\widehat{G}_{eu,o}^{aug})^{-1}$ is square and full rank (which is usually the case), then the SPR condition for RPFxLMS (6.5) can be simplified further to the SPR condition of RFxLMS (6.3). In the case $(\widehat{G}_{eu,o}^{aug})^{-1}$ is tall (i.e. if $\widehat{G}_{eu}^{aug}$ has more columns than rows), (6.5) is less strict than (6.3). Hence, if RFxLMS converges then RPFxLMS converges, provided the step size is small enough.

Uncertainty in the detector path G_{rs} can be taken into account similarly. But, instead of augmenting the performance channels to deal with uncertainty in S, the *reference* channels have to be augmented with an additional noise signal uncorrelated with $s(k)$ and $v(k)$ (cf. [48]). Furthermore, the same robustification method can be used in the Adjoint FxLMS algorithms as in [45,47] and the (preconditioned) FuLMS algorithm considered in Chapter 5.

6.4   Simulation example

The RPFxLMS algorithm is tested on a 1-dimensional acoustical duct simulation model, discretized using a sampling rate of $f_s = 1\cdot 10^3$Hz. The magnitude of the secondary path G_{eu} is depicted in Figure 6.2.

By varying the delay between 0 and $2\cdot 10^{-3}$s uncertainty is introduced in the system such that the SPR condition does not hold anymore. Figure 6.3 shows the value of $G_{eu}(e^{j2\pi f/f_s})^*\widehat{G}_{eu}(e^{j2\pi f/f_s}) + \widehat{G}_{eu}(e^{j2\pi f/f_s})^*G_{eu}(e^{j2\pi f/f_s})$ where the difference in the delay between G_{eu} and $\widehat{G}_{eu}$ is $1\cdot 10^{-3}$s, from which it is clear that the SPR condition does not hold for all frequencies. The covariance $\Phi_{\Delta G_{eu}}$ of the uncertainty has been estimated by uniformly varying the delay between 0 and
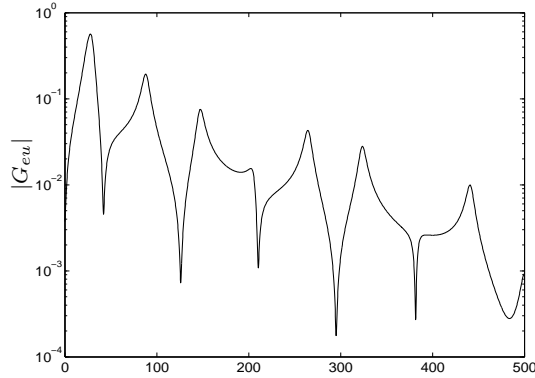
Figure 6.2: Magnitude of the secondary path system G_{eu}.
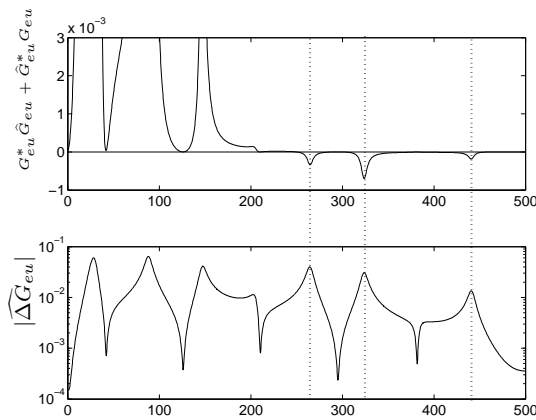
Figure 6.3: Upper: $G_{eu}(e^{j2\pi f/f_s})^* \widehat{G}_{eu}(e^{j2\pi f/f_s}) + \widehat{G}_{eu}(e^{j2\pi f/f_s})^* G_{eu}(e^{j2\pi f/f_s})$ for $f = 0, \cdots, 500$Hz. Lower: magnitude of the estimated uncertainty $\widehat{\Delta G}_{eu}$.

$2 \cdot 10^{-3}$s. The magnitude of the estimated spectral factor $\widehat{\Delta G}_{eu}$ is also depicted in Figure 6.3.

In the experiments the delay in the secondary path *model* has been varied, from 0 to $4 \cdot 10^{-3}$s additional delay. Depending on the amount of additional delay the SPR condition (6.1) does not hold anymore as has been illustrated in Figure 6.3. The RPFxLMS algorithm has been applied for various choices of $\widehat{\Delta G}_{eu}$: 1. $\widehat{\Delta G}_{eu} = 0$, i.e. the nominal case, 2. $\widehat{\Delta G}_{eu} = 1.87 \cdot 10^{-2}$, which is such that (6.5) just holds for a delay of $1 \cdot 10^{-3}$s, 3. $\widehat{\Delta G}_{eu} = 4.27 \cdot 10^{-2}$, which is such that (6.5) just holds for a delay of $2 \cdot 10^{-3}$s and finally 4. $\widehat{\Delta G}_{eu}$ estimated via the covariance of the model error $\Phi_{\Delta G_{eu}}$ due to a delay uniformly distributed from 0 to $2 \cdot 10^{-3}$s. In all experiments, the normalized step size is chosen to be 0.1, the number of filter coefficients $n_w = 400$ and the measurement noise is absent ($\sigma_v^2 = 0$).

Figure 6.4 shows the reduction after 160s (if the algorithm converges it is usually
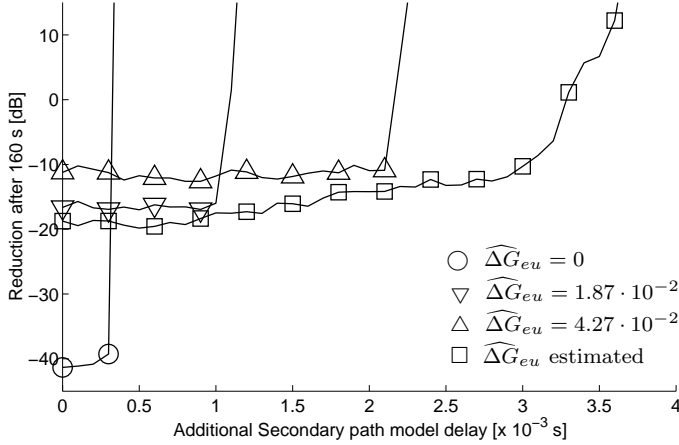
Figure 6.4: Reduction of RPFxLMS after 160s obtained for various choices of $\widehat{\Delta G}_{eu}$, versus additional delay in $\widehat{G}_{eu}$.

converged after ≈ 30s, but 160s has been chosen to fully guarantee the algorithm is converged). The nominal case (marked with $\bigcirc$) yields best performance between $0 - 0.3 \cdot 10^{-3}$s, however the adaptive algorithm diverges for larger delays. Using scalar regularization (marked with $\triangledown$ and $\triangle$), the robustness can be improved, but at the expense of significant performance. By estimating the uncertainty model $\widehat{\Delta G}_{eu}$ via the covariance $\Phi_{\Delta G_{eu}}$ with delay uniformly distributed between 0 and $2 \cdot 10^{-3}$s (marked with $\square$), the robustness of the update algorithm is increased significantly without paying too much performance.

6.5   Conclusions

The robustness of the preconditioned FxLMS algorithm, proposed by Elliott et al. is increased by following a probabilistic robust filtering method. The SPR condition is relaxed by taking the model uncertainty in the secondary path model explicitly into account. Furthermore, the gain of the precondition filter is reduced, which may prevent oversteering problems.

The approach to increase the robustness of the (preconditioned) FxLMS algorithm, can also be used to increase the robustness of the Adjoint FxLMS algorithms as considered in [45, 47] and the (preconditioned) FuLMS algorithm considered in Chapter 5.

Chapter 7

A Fast-array Kalman filter solution

7.1 Introduction

Active noise and vibration control (ANVC) systems usually deal with a large amount of dominant, relatively weakly damped, resonance modes (in the order of 10 to 50) and need controllers with a large impulse response to obtain good disturbance rejection (in the order of 100-1000 taps). Furthermore, sampling rates are often in the order of 1-10kHz to have sufficient control bandwidth. Besides, the controller should be able to adjust for variations in the system, like temperature variations. These constraints make ANVC a challenging control problem even in a time of fast increasing computer power.

Because of its computational efficiency and robustness properties, the Filtered-X LMS (FxLMS) algorithm is very popular in ANVC systems (see, e.g., [46]). However, in applications with broadband disturbances, and especially in the case of multiple channels, the convergence and tracking capacity of FxLMS is poor. This problem has encouraged many researchers to develop alternatives to the FxLMS algorithm, see, e.g., [18, 45, 51, 107, 118, 145, 154].

Almost all algorithms proposed for ANVC are of the so-called filtered-X or filtered-reference type. They rely on the assumption that the adaptive filter and the controlled system, the so-called secondary path, may be interchanged. Neglecting transients from initial states, this is true for systems that are constant in time. However, since the adaptive filter varies in time this is not true anymore and algorithms based on this assumption are inherently slowly converging. This observation has motivated the introduction of so-called *modified* filtered-X algorithms, in which the disturbance is estimated from the residual signal and an internal model of the secondary path, see [15, 51]. It is interesting to note that in [37] (especially equation (13)) for deterministic disturbances, the modified filtered-X algorithms, though not named this way, were already derived from a self tuning regulator point of view. Although this is still an approximation in nonstationary applications where the optimal controller is varying in time, the modified filtered-X algorithms yield better convergence than the filtered-X algorithms, at the expense

of increased computational load.

The problem of interchanging the adaptive filter and the secondary path was addressed in [154] by reformulating the ANVC control problem as a state estimation problem. The state to be estimated contains the unknown filter coefficients and the unknown state of the secondary path system. However, this approach required an estimate of the disturbance signal.

In this chapter, we reformulate the ANVC control problem also as a state-estimation problem, but without using an estimate of the disturbance. The state-estimation problem is solved by a Kalman filter, which has been chosen from an optimality point of view. Uncertainty in the secondary-path state, due to initial state uncertainty and/or noise, can be taken into account explicitly with improved convergence. We also show that in case there is no uncertainty in the secondary-path state, then the Kalman algorithm is equivalent to a modified filtered-RLS algorithm. The analysis of the equivalence of both algorithms also shows that the exponential forgetting needs to be applied to filtering the reference signal as well. Furthermore, a fast-array implementation of the Kalman filter algorithm is derived, which enables practical application of the algorithm. The derivation of this fast algorithm is based on the observation that although the underlying state-space model is *not* time-invariant, it is nevertheless a structured model in the sense defined in [150, 151].

The chapter is organized as follows. Section 7.2 formulates the estimation problem and presents the Kalman algorithm to solve this problem. Section 7.3 derives the new fast-array implementation of the Kalman algorithm. Section 7.4 compares the Kalman algorithm and a modified filtered-RLS algorithm and presents the conditions for equivalence of both algorithms. Section 7.5 illustrates the (fast-array) Kalman algorithm by simulation and Section 7.6 concludes the chapter.

7.2   The Kalman filter solution

7.2.1   The state estimation problem

Consider the active feedforward control problem illustrated in Figure 7.1. The objective is to counteract the disturbance signal $d(k)$ by a secondary signal $y(k)$, which yields the residual signal $e(k)$. It is assumed, that the primary path can be decomposed into a series connection of an optimal feedforward controller $W^o(q^{-1})$ and a secondary path $G_{eu}(q^{-1})$, which contains the dynamics between the actuators and the sensors. Usually this assumption is not satisfied, but the error due to imposing this assumption can often be neglected. Assuming that the noise signals $v_s(k)$ and $v_m(k)$ are uncorrelated with the disturbance reference signal $r(k)$, it can be verified easily that the residual signal is minimized if the feedforward controller $\widehat{W}_k(q^{-1})$ equals $W^o(q^{-1})$. In this way, the feedforward control problem is reformulated in an estimation context [154].

We assume that the unknown optimal controller is a FIR filter of length n_w

$$W^o(q^{-1}) = w_0 + w_1 q^{-1} + \cdots w_{n_w-1} q^{-n_w+1}. \tag{7.1}$$
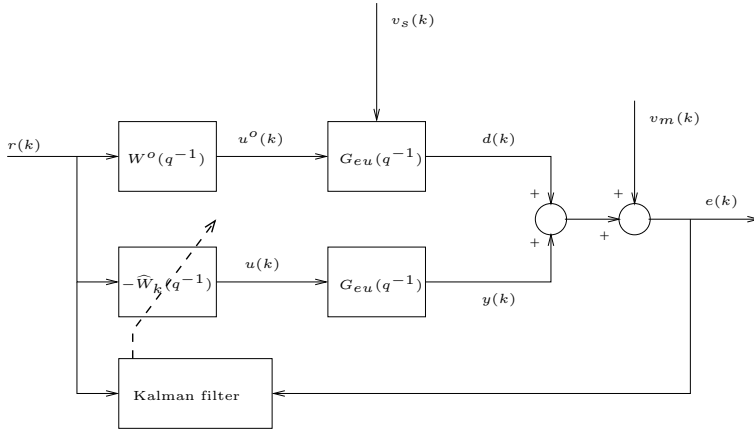
Figure 7.1: A block diagram representation of a feedforward active control configuration with a feedforward controller $\widehat{W}_k(q^{-1})$ estimated by a Kalman filter to minimize the residual disturbance $e(k)$. The primary path is assumed to consist of a series connection of the (unknown) optimal feedforward controller $W^o(q^{-1})$ and the secondary path $G_{eu}(q^{-1})$.

and that $\widehat{W}_k(q^{-1})$ has the same structure

$$\widehat{W}_k(q^{-1}) = \widehat{w}_0(k) + \widehat{w}_1(k)q^{-1} + \cdots \widehat{w}_{n_w-1}(k)q^{-n_w+1}. \tag{7.2}$$

Let

$$w^o = \begin{bmatrix} w_0 & w_1 & \cdots & w_{n_w-1} \end{bmatrix}^T, \tag{7.3}$$

$$\widehat{w}(k) = \begin{bmatrix} \widehat{w}_0(k) & \widehat{w}_1(k) & \cdots & \widehat{w}_{n_w-1}(k) \end{bmatrix}^T, \tag{7.4}$$

$$r_{n_w}(k) = \begin{bmatrix} r(k) & r(k-1) & \cdots & r(k-n_w+1) \end{bmatrix}^T \tag{7.5}$$

then optimal control signal is given by

$$u^o(k) = r_{n_w}^T(k)w^o, \tag{7.6}$$

and the actual control signal by

$$u(k) = -r_{n_w}^T(k)\widehat{w}(k). \tag{7.7}$$

The secondary path $G_{eu}(q^{-1})$ will be described in state-space form. Usually, the state dimension n_{eu} is high (in the range of 20 - 100) for acoustical or vibrational systems. For this reason, using a FIR model of sufficient length to model the dynamics of $G_{eu}(q^{-1})$ can help lower the computational complexity, especially for well damped systems. Because the FIR model structure is contained in the state-space model structure, i.e., a FIR model is a state-space model with particular structure, it is just a matter of straightforward computation to derive the expressions for secondary path models with FIR structure. Other (canonical) parameterizations contained in the state-space structure can be used as well.

The noise signal $v_s(k) \in \mathbb{R}^{n_v}$ distorts the secondary-path state and $v_m(k)$ distorts the measured output $e(k)$. We assume $v_s(k)$ and $v_m(k)$ are both stationary zero-mean white-noise signals that are independent of $r(k)$ and satisfy

$$\mathbf{E}\left(\begin{bmatrix} v_s(k) \\ v_m(k) \end{bmatrix}\begin{bmatrix} v_s(l) \\ v_m(l) \end{bmatrix}^T\right) = \begin{bmatrix} Q & 0_{n_v \times 1} \\ 0_{1 \times n_v} & R \end{bmatrix}\delta_{kl}, \quad Q \geq 0, \quad R > 0 \quad (7.8)$$

with δ_{kl} the Kronecker delta function (if $k = l$ then $\delta_{kl} = 1$, otherwise $\delta_{kl} = 0$). Note that the assumption that $v_s(k)$ is white is not restrictive, since this can always be assured by incorporating the noise-shaping filter into the secondary path system. We will not assume any conditions on the reference disturbance signal $r(k)$. For example, it may be white or colored noise, a sinusoid, stationary or nonstationary.

Let (A, B_u, C_e, D_{eu}) be the state-space matrices that model the secondary path $G_{eu}(q^{-1})$, then the disturbance $d(k)$ is written as

$$\theta^1(k+1) = A\theta^1(k) + B_u u^o(k) + G_s v_s(k), \quad \theta^1(0) = \theta_0^1, \quad (7.9)$$
$$d(k) = C_e\theta^1(k) + D_{eu}u^o(k), \quad (7.10)$$

and the secondary signal $y(k)$ as

$$\theta^2(k+1) = A\theta^2(k) + B_u u(k), \quad\quad\quad \theta^2(0) = \theta_0^2, \quad (7.11)$$
$$y(k) = C_e\theta^2(k) + D_{eu}u(k) \quad (7.12)$$

with $\theta^1(k), \theta^2(k) \in \mathbb{R}^{n_s}$. Since

$$e(k) = d(k) + y(k) + v_m(k) \quad (7.13)$$

and the state-space matrices in (7.9)-(7.10) and (7.11)-(7.12) are the same, we can write

$$\theta(k+1) = A\theta(k) + B_u(u^o(k) + u(k)) + G_s v_s(k), \quad\quad \theta(0) = \theta_0 \quad (7.14)$$
$$e(k) = C_e\theta(k) + D_{eu}(u^o(k) + u(k)) + v_m(k) \quad (7.15)$$

with $\theta_0 = \theta_0^1 + \theta_0^2$ and $\theta(k) = \theta^1(k) + \theta^2(k)$.

Substituting (7.6) and (7.7) into the state-space equations (7.14)-(7.15) gives the final state-space description of the active control system considered in this chapter

$$\begin{bmatrix} w(k+1) \\ \theta(k+1) \end{bmatrix} = \begin{bmatrix} \lambda^{-1/2}I_{n_w} & 0_{n_w \times n_s} \\ B_u r_{n_w}^T(k) & A \end{bmatrix}\begin{bmatrix} w(k) \\ \theta(k) \end{bmatrix} - \begin{bmatrix} 0_{n_w \times n_w} \\ B_u r_{n_w}^T \end{bmatrix}\widehat{w}(k)+$$

$$+ \begin{bmatrix} 0_{n_w \times n_v} \\ G_s \end{bmatrix}v_s(k), \quad\quad \begin{bmatrix} w(0) \\ \theta(0) \end{bmatrix} = \begin{bmatrix} w^o \\ \theta_0 \end{bmatrix} \quad (7.16)$$

$$e(k) = \begin{bmatrix} D_{eu}r_{n_w}^T(k) & C_e \end{bmatrix}\begin{bmatrix} w(k) \\ \theta(k) \end{bmatrix} - D_{eu}r_{n_w}^T(k)\widehat{w}(k) + v_m(k) \quad (7.17)$$

where λ bounded by

$$0 \ll \lambda \le 1$$

is an exponential forgetting factor; it is included to model variations in the optimal filter coefficients to obtain better tracking capacity (see, e.g., [151]). In the stationary case, where $w(k) = w^o$, for all k, we set $\lambda = 1$. We could have included (white-)noise into $w(k+1)$ with known covariance, which would often yield an improved tracking capacity.

For ease of notation, we define

$$A_k = \begin{bmatrix} \lambda^{-1/2} I_{n_w} & 0_{n_w \times n_s} \\ B_u r_{n_w}^T(k) & A \end{bmatrix} \tag{7.18}$$

$$B_k = \begin{bmatrix} 0_{n_w \times n_w} \\ -B_u r_{n_w}^T(k) \end{bmatrix} \tag{7.19}$$

$$G = \begin{bmatrix} 0_{n_w \times n_v} \\ G_s \end{bmatrix} \tag{7.20}$$

$$C_k = \begin{bmatrix} D_{eu} r_{n_w}^T(k) & C_e \end{bmatrix} \tag{7.21}$$

$$D_k = -D_{eu} r_{n_w}^T(k) \tag{7.22}$$

and the augmented state

$$x(k) = \begin{bmatrix} w(k) \\ \theta(k) \end{bmatrix} \tag{7.23}$$

With these definitions, (7.16) and (7.17) are rewritten more compactly as

$$x(k+1) = A_k x(k) + B_k \widehat{w}(k) + G v_s(k) \tag{7.24}$$

$$e(k) = C_k x(k) + D_k \widehat{w}(k) + v_m(k). \tag{7.25}$$

7.2.2 The Kalman filter

Since the Kalman filter provides a minimum variance estimate of the state at every sampling instant given the model of the system and the covariances of the white-noise signals $v_s(k)$ and $v_m(k)$, we will use this filter to estimate $w(k)$ as well as $\theta(k)$.

To apply the Kalman filter, we assume that the initial state is uncorrelated with $v_s(k)$ and $v_m(k)$, i.e.,

$$\mathrm{E}\left(\begin{bmatrix} x(0) \\ v_s(k) \\ v_m(k) \end{bmatrix} \begin{bmatrix} x(0) \\ v_s(l) \\ v_m(l) \end{bmatrix}^T \right) = \begin{bmatrix} \Pi_0 & 0_{n_w+n_s \times n_v} & 0_{n_w+n_s \times 1} \\ 0_{n_v \times n_w+n_s} & Q\delta_{kl} & 0_{n_v \times 1} \\ 0_{1 \times n_w+n_s} & 0_{1 \times n_v} & R\delta_{kl} \end{bmatrix},$$

where

$$\Pi_0 = \begin{bmatrix} \Pi_0^{ww} & \Pi_0^{w\theta} \\ \Pi_0^{\theta w} & \Pi_0^{\theta\theta} \end{bmatrix} > 0,$$

$$\mathrm{E}(w(0)w^T(0)) = \Pi_0^{ww}, \quad \mathrm{E}(w(0)\theta^T(0)) = \Pi_0^{w\theta} = \Pi_0^{\theta w T}, \quad \mathrm{E}(\theta(0)\theta^T(0)) = \Pi_0^{\theta\theta}.$$

The Kalman filter can be described in at least two forms: the time/measurement update form and the prediction form [85]. The output of the time/measurement update form is an estimate of $x(k)$ given the measurements $\{e(0), e(1), \cdots, e(k)\}$, denoted as $\widehat{x}(k|k)$, together with its error covariance matrix $P_{k|k}$. The output of the prediction form is an estimate of $x(k+1)$ given the same measurements $\{e(0), e(1), \cdots, e(k)\}$, denoted as $\widehat{x}(k+1|k)$ or just $\widehat{x}(k+1)$, together with its error covariance matrix $P_{k+1|k}$ or just P_{k+1}. Because, we need an estimate of $w(k+1)$ to calculate the control signal at the (next) iteration $k+1$, we will use the prediction form in the sequel.

The Kalman filter in prediction form is given by the following equations for $k \geq 0$:

$$\widehat{x}(0) = 0_{n_w + n_s \times 1}, \tag{7.26}$$

$$P_0 = \Pi_0, \tag{7.27}$$

$$\epsilon(k) = e(k) - C_k \widehat{x}(k) - D_k \widehat{w}(k), \tag{7.28}$$

$$R_{e,k} = R + C_k P_k C_k^T, \tag{7.29}$$

$$K_k = A_k P_k C_k^T, \tag{7.30}$$

$$\widehat{x}(k+1) = A_k \widehat{x}(k) + B_k \widehat{w}(k) + K_k R_{e,k}^{-1} \epsilon(k), \tag{7.31}$$

$$P_{k+1} = A_k P_k A_k^T - K_k R_{e,k}^{-1} K_k^T + GQG^T. \tag{7.32}$$

For further reference, we partition P_k similarly to Π_0 as

$$P_k = \begin{bmatrix} P_k^{ww} & P_k^{w\theta} \\ P_k^{\theta w} & P_k^{\theta\theta} \end{bmatrix}. \tag{7.33}$$

Note that using the definitions (7.18)-(7.23) and partitioning

$$K_k = \begin{bmatrix} K_k^w \\ K_k^\theta \end{bmatrix}, \quad K_k^w \in \mathbb{R}^{n_w \times 1}, \quad K_k^\theta \in \mathbb{R}^{n_s \times 1}, \tag{7.34}$$

then expression (7.28) for the innovation $\epsilon(k)$, and expression (7.31) for the state-estimate update equation $\widehat{x}(k+1)$ can be simplified to

$$\epsilon(k) = e(k) - C_e \widehat{\theta}(k), \tag{7.35}$$

$$\begin{bmatrix} \widehat{w}(k+1) \\ \widehat{\theta}(k+1) \end{bmatrix} = \begin{bmatrix} \lambda^{-1/2} \widehat{w}(k) \\ A \widehat{\theta}(k) \end{bmatrix} + \begin{bmatrix} K_k^w \\ K_k^\theta \end{bmatrix} R_{e,k}^{-1} \epsilon(k). \tag{7.36}$$

The resulting Kalman filter algorithm, to solve the active control problem, is listed in the first column of Table 7.1.

Now, the implementation of the Kalman filter using expressions (7.29), (7.30) and (7.32) is computationally complex for most practical applications. For instance, recursion (7.32) has at least $O((n_w + n_s)^2)$ complexity, assuming that A is just a shift matrix, which is the case when an FIR model is used for the secondary path (if A has no structure at all, the complexity will be at least $O(n_s^3 + (n_w + n_s)^2)$).

In a future Section 7.4 we will show that by setting

$$\Pi_0^{w\theta} = \Pi_0^{\theta w T} = 0_{n_w \times n_s}, \quad \Pi_0^{\theta\theta} = 0_{n_s \times n_s}, \quad \text{and} \quad Q = 0_{n_v \times n_v},$$

which assumes perfect knowledge of the secondary path initial state and $v_s(k) = 0$ for all $k \geq 0$, the Kalman algorithm simplifies to a modified Filtered-RLS algorithm. However, this assumption is rather strong, and may degrade the performance of the algorithm severely in case it is not satisfied, as will be illustrated by simulation in Section 7.5.

For now we proceed to derive a fast-array implementation of the Kalman filter algorithm by exploiting structure in the state-space matrices, thus reducing the computational complexity down to $O(n_w + n_s)$ per iteration.

7.3 The fast-array Kalman filter

7.3.1 The structure in the state-space model

We will base our derivation on [150], where a fast implementation of the Kalman filter for certain *time-varying* systems with structure was derived. Consider the definition of $r_{n_w}(k)$ in (7.5). It is clear that $r_{n_w}(k)$ and $r_{n_w}(k+1)$ have $n_w - 1$ entries in common but at shifted positions of each other. Let us define the shift-matrix $Z_{n_w} \in \mathbb{R}^{n_w \times n_w}$ as the matrix with ones on its first subdiagonal and zeros elsewhere. Then, we can write

$$r_{n_w}^T(k) \;=\; r_{n_w}^T(k+1)Z_{n_w} + \left[\begin{array}{cc} 0_{1 \times n_w - 1} & r(k - n_w + 1) \end{array}\right].$$

Using this result, we are able to relate A_k and A_{k+1} to each other as well as C_k and C_{k+1}. In our case, G does not depend on k, but should satisfy a particular condition given below. Though the state-space matrices B_k and D_k are also related to B_{k+1} and D_{k+1} we do not need this relation in the derivation of the fast-array Kalman filter, since they determine the deterministic part of the state update which does not influence the Kalman filter expressions (7.29)-(7.32). Note, that we already exploited the structure in B_k and D_k in the equations (7.35) and (7.36).

To relate A_k and A_{k+1}, and C_k and C_{k+1}, we will first introduce the augmented state-space system, which is equivalent to (7.24)-(7.25):

$$\tilde{x}(k+1) = \tilde{A}_k \tilde{x}(k) + \tilde{B}_k \widehat{w}(k) + \tilde{G} v_s(k) \tag{7.37}$$

$$e(k) = \tilde{C}_k \tilde{x}(k) + D_k \widehat{w}(k) + v_m(k) \tag{7.38}$$

where

$$\tilde{A}_k = \left[\begin{array}{cc} \lambda^{-1/2} I_{n_w+1} & 0_{n_w+1 \times n_s} \\ B_u r_{n_w+1}^T(k) & A \end{array}\right], \tag{7.39}$$

$$\tilde{B}_k = \left[\begin{array}{c} 0_{n_w+1 \times n_w} \\ -B_u r_{n_w}^T(k) \end{array}\right], \tag{7.40}$$

$$\tilde{G} = \left[\begin{array}{c} 0_{n_w+1 \times n_v} \\ G_s \end{array}\right], \tag{7.41}$$

$$\tilde{C}_k = \left[\begin{array}{cc} D_{eu} r_{n_w+1}^T(k) & C_e \end{array}\right], \tag{7.42}$$

and the augmented initial state is given by

$$\tilde{x}(0) = \left[\begin{array}{c} w^o \\ 0 \\ \theta(0) \end{array} \right] \in \mathbb{R}^{n_w+1+n_s}. \tag{7.43}$$

Because the $(n_w + 1)^{th}$ entry of $\tilde{x}(k)$ is uncontrollable from the deterministic input $\widehat{w}(k)$ as well as from the stochastic input $v_s(k)$, it will keep its initial zero value. It can be verified easily using (7.37) that $\tilde{x}(k)$ has a similar form, namely

$$\tilde{x}(k) = \left[\begin{array}{c} w(k) \\ 0 \\ \theta(k) \end{array} \right], \qquad \forall k \geq 0$$

and thus the output given by (7.38) is exactly the same as the output given by (7.25). Because there is no uncertainty in the $(n_w + 1)^{th}$ entry of $\tilde{x}(0)$ we define

$$\text{E}\left(\tilde{x}(0)\tilde{x}^T(0) \right) = \tilde{\Pi}_0 = \left[\begin{array}{ccc} \Pi_0^{ww} & 0_{n_w \times 1} & \Pi_0^{w\theta} \\ 0_{1 \times n_w} & 0 & 0_{1 \times n_s} \\ \Pi_0^{\theta w} & 0_{n_s \times 1} & \Pi_0^{\theta\theta} \end{array} \right],$$

and it is clear that $\text{E}(\tilde{x}(0)[v_m^T(k) \quad v_s(k)]) = 0_{n_w+1+n_s \times n_v+1}$.

Now, let us define

$$\Psi = \left[\begin{array}{cc} Z_{n_w+1} & 0_{n_w+1 \times n_s} \\ 0_{n_s \times n_w+1} & I_{n_s} \end{array} \right] \tag{7.44}$$

then it is straightforward to verify that

$$\tilde{A}_{k+1}\Psi + \Delta_k^a = \Psi\tilde{A}_k, \tag{7.45}$$

$$\tilde{G} = \Psi\tilde{G}, \tag{7.46}$$

$$\tilde{C}_k = \tilde{C}_{k+1}\Psi + \Delta_k^c, \tag{7.47}$$

$$\Delta_k^a = \left[\begin{array}{cc} 0_{n_w+1 \times n_w+1} & 0_{n_w+1 \times n_s} \\ \left[0_{n_s \times n_w} \quad B_u r(k-n_w) \right] & 0_{n_s \times n_s} \end{array} \right], \tag{7.48}$$

$$\Delta_k^c = \left[0_{1 \times n_w} \quad D_{eu} r(k-n_w) \quad 0_{1 \times n_s} \right], \tag{7.49}$$

where (7.45)-(7.47) are, up to the Δ-terms, equal to (a special case of) the relations in [150].

7.3.2   The fast-array iterations

The Kalman filter equations of the augmented system (7.37)-(7.38) are given by

$$\widehat{\tilde{x}}(0) = 0_{n_w+1+n_s \times 1}, \tag{7.50}$$

$$\tilde{P}_0 = \tilde{\Pi}_0, \tag{7.51}$$

$$\tilde{\epsilon}(k) = e(k) - \tilde{C}_k\widehat{\tilde{x}}(k) - D_k\widehat{w}(k), \tag{7.52}$$

$$\tilde{R}_{e,k} = R + \tilde{C}_k \tilde{P}_k \tilde{C}_k^T, \tag{7.53}$$

$$\tilde{K}_k = \tilde{A}_k \tilde{P}_k \tilde{C}_k^T, \tag{7.54}$$

$$\widehat{\tilde{x}}(k+1) = \tilde{A}_k \widehat{\tilde{x}}(k) + \tilde{B}_k \widehat{w}(k) + \tilde{K}_k \tilde{R}_{e,k}^{-1} \epsilon(k), \tag{7.55}$$

$$\tilde{P}_{k+1} = \tilde{A}_k \tilde{P}_k \tilde{A}_k^T - \tilde{K}_k \tilde{R}_{e,k}^{-1} \tilde{K}_k^T + \tilde{G} Q \tilde{G}^T. \tag{7.56}$$

Because the augmented system (7.37)-(7.38) is equivalent to the original system (7.24)-(7.25), the Kalman filter of the augmented system should provide the same state-estimate, state-error covariance, and innovations. By straightforward computation, it can be verified (e.g., by induction) that for all $k \geq 0$:

$$\tilde{P}_k = \begin{bmatrix} P_k^{ww} & 0_{n_w \times 1} & P_k^{w\theta} \\ 0_{1 \times n_w} & 0 & 0_{1 \times n_s} \\ P_k^{\theta w} & 0_{n_s \times 1} & P_k^{\theta\theta} \end{bmatrix},$$

$$\tilde{R}_{e,k} = R_{e,k},$$

$$\tilde{K}_k = \begin{bmatrix} K_k^w \\ 0 \\ K_k^\theta \end{bmatrix}.$$

Hence, we also have for all $k \geq 0$:

$$\widehat{\tilde{x}}(k) = \begin{bmatrix} \widehat{w}(k) \\ 0 \\ \widehat{\theta}(k) \end{bmatrix},$$

$$\tilde{\epsilon}(k) = \epsilon(k).$$

The idea behind fast-array algorithms is to update the difference

$$d\tilde{P}_k = \tilde{P}_k - \Psi \tilde{P}_{k-1} \Psi^T$$

rather than $\tilde{P}_k$ itself. In many cases, depending on the choice of Π_0, it can be shown that $d\tilde{P}_k$ has a low rank α with $\alpha \ll (n_w + 1 + n_s)$ (in the next subsection, we will exhibit a choice for Π_0 such that $\alpha = 2$). Hence, $d\tilde{P}_k$, or a factorization for it, can be updated very efficiently [150, 151].

Let us define also the difference quantities

$$d\tilde{R}_{e,k} = \tilde{R}_{e,k} - \tilde{R}_{e,k-1},$$

$$d\tilde{K}_k = \tilde{K}_k - \Psi \tilde{K}_{k-1}.$$

Then, using (7.53), (7.54) and (7.56) together with relations (7.45)-(7.47) we get

$$d\tilde{R}_{e,k} = \tilde{C}_k d\tilde{P}_k \tilde{C}_k^T, \tag{7.57}$$

$$d\tilde{K}_k = \tilde{A}_k d\tilde{P}_k \tilde{C}_k^T, \tag{7.58}$$

$$d\tilde{P}_{k+1} = \tilde{A}_k d\tilde{P}_k \tilde{A}_k^T + \Psi \tilde{K}_{k-1} \tilde{R}_{e,k-1}^{-1} \tilde{K}_{k-1}^T \Psi^T - \tilde{K}_k \tilde{R}_{e,k}^{-1} \tilde{K}_k^T. \tag{7.59}$$

Suppose $d\tilde{P}_k$ can be factored as

$$d\tilde{P}_k = \tilde{L}_{k-1} M_{k-1} \tilde{L}_{k-1}^T,$$

where $\tilde{L}_{k-1} \in \mathbb{R}^{n_w+1+n_s \times \alpha}$ and $M_{k-1} \in \mathbb{R}^{\alpha \times \alpha}$ for some $\alpha \ll n_w + n_s + 1$. Then it turns out that $\tilde{R}_{e,k-1}$, $\tilde{K}_{k-1}$ and $\tilde{L}_{k-1}$ can be updated to $\tilde{R}_{e,k}$, $\tilde{K}_k$ and $\tilde{L}_k$ as follows. Multiply the pre-array on the left-hand side below by a transformation Θ_{k-1}

$$
\begin{bmatrix}
\tilde{R}_{e,k-1}^{1/2} & \tilde{C}_k \tilde{L}_{k-1} \\
\Psi \tilde{K}_{k-1} \tilde{R}_{e,k-1}^{-T/2} & \tilde{A}_k \tilde{L}_{k-1}
\end{bmatrix} \Theta_{k-1} =
\begin{bmatrix}
\tilde{R}_{e,k}^{1/2} & 0_{1\times\alpha} \\
\tilde{K}_k \tilde{R}_{e,k}^{-T/2} & \tilde{L}_k
\end{bmatrix}, \qquad (7.60)
$$

so as to result in the $1 \times \alpha$ zero block in the post-array on the right hand side. The matrix Θ_{k-1} is required to be J_{k-1}-unitary, i.e., it should satisfy

$$
\Theta_{k-1} J_{k-1} \Theta_{k-1}^T = J_{k-1},
$$

where

$$
J_{k-1} =
\begin{bmatrix}
1 & 0_{1\times\alpha} \\
0_{\alpha\times 1} & M_{k-1}
\end{bmatrix}.
$$

This fact can be verified by 'squaring' the left- and the right hand sides of (7.60) and using relations (7.57)-(7.59). By 'squaring' the left hand side of (7.60) we obtain

$$
\begin{bmatrix}
\tilde{R}_{e,k-1}^{1/2} & \tilde{C}_k \tilde{L}_{k-1} \\
\Psi \tilde{K}_{k-1} \tilde{R}_{e,k-1}^{-T/2} & \tilde{A}_k \tilde{L}_{k-1}
\end{bmatrix} \underbrace{\Theta_{k-1} J_{k-1} \Theta_{k-1}^T}_{=J_{k-1}}
\begin{bmatrix}
\tilde{R}_{e,k-1}^{1/2} & \tilde{C}_k \tilde{L}_{k-1} \\
\Psi \tilde{K}_{k-1} \tilde{R}_{e,k-1}^{-T/2} & \tilde{A}_k \tilde{L}_{k-1}
\end{bmatrix}^T =
$$

$$
\begin{bmatrix}
\tilde{R}_{e,k-1} + \tilde{C}_k \tilde{L}_{k-1} M_{k-1} \tilde{L}_{k-1}^T \tilde{C}_k^T & \tilde{K}_{k-1}^T \Psi^T + \tilde{C}_k \tilde{L}_{k-1} M_{k-1} \tilde{L}_{k-1}^T \tilde{A}_k^T \\
\Psi \tilde{K}_{k-1} + \tilde{A}_k \tilde{L}_{k-1} M_{k-1} \tilde{L}_{k-1}^T \tilde{C}_k^T & \Psi \tilde{K}_{k-1} \tilde{R}_{e,k-1}^{-1} \tilde{K}_{k-1}^T \Psi^T + \tilde{A}_k \tilde{L}_{k-1} M_{k-1} \tilde{L}_{k-1}^T \tilde{A}_k^T
\end{bmatrix} =
$$

$$
\begin{bmatrix}
\tilde{R}_{e,k} & \tilde{K}_k^T \\
\tilde{K}_k & \tilde{A}_k d\tilde{P}_k \tilde{A}_k^T + \Psi \tilde{K}_{k-1} \tilde{R}_{e,k-1}^{-1} \tilde{K}_{k-1}^T \Psi^T
\end{bmatrix}.
$$

And, on the other hand, by 'squaring' the right hand side of (7.60) we obtain

$$
\begin{bmatrix}
\tilde{R}_{e,k}^{1/2} & 0_{1\times\alpha} \\
\tilde{K}_k \tilde{R}_{e,k}^{-T/2} & \tilde{L}_k
\end{bmatrix} J_{k-1}
\begin{bmatrix}
\tilde{R}_{e,k}^{1/2} & 0_{1\times\alpha} \\
\tilde{K}_k \tilde{R}_{e,k}^{-T/2} & \tilde{L}_k
\end{bmatrix}^T =
$$

$$
= \begin{bmatrix}
\tilde{R}_{e,k} & \tilde{K}_k^T \\
\tilde{K}_k & \tilde{K}_k \tilde{R}_{e,k}^{-1} \tilde{K}_k^T + \tilde{L}_k M_{k-1} \tilde{L}_k^T
\end{bmatrix} =
$$

$$
= \begin{bmatrix}
\tilde{R}_{e,k} & \tilde{K}_k^T \\
\tilde{K}_k & \tilde{K}_k \tilde{R}_{e,k}^{-1} \tilde{K}_k^T + d\tilde{P}_{k+1}
\end{bmatrix}.
$$

Equality (7.60) then holds once we make the identification

$$
d\tilde{P}_{k+1} = \tilde{L}_k M_k \tilde{L}_k^T, \quad \text{with } M_k = M_{k-1}. \qquad (7.61)
$$

We thus conclude that if $d\tilde{P}_k$ has (low) rank α, then $d\tilde{P}_{k+1}$ also has (low) rank α. Furthermore, the matrix M_k in the factorization of $d\tilde{P}_{k+1}$ is equal to M_{k-1} and thus we may set

$$M = M_k, \quad \text{and} \quad J = \begin{bmatrix} 1 & 0_{1 \times \alpha} \\ 0_{\alpha \times 1} & M \end{bmatrix} = J_k$$

for all $k \geq 0$.

Now, the problem is to determine an initial factorization

$$d\tilde{P}_0 = \tilde{L}_{-1} M \tilde{L}_{-1}^T,$$

with $M \in \mathbb{R}^{\alpha \times \alpha}$ and α as small as possible. This problem will be solved in the next subsection.

Note that the update equation (7.60) is independent of R, G and Q. These parameters enter into the initialization of the algorithm.

7.3.3 Initialization

We now seek a matrix Π_{-1} such that the difference

$$\tilde{P}_0 - \Psi \tilde{P}_{-1} \Psi^T = \tilde{L}_{-1} M \tilde{L}_{-1}$$

has low rank $\alpha \ll (n_w + 1 + n_s)$. Note that since we iterate beginning from $k = 0$, we only need to know $\tilde{L}_{-1}$ and M. In the following, we will assume the *prewindowed-data* case, i.e.,

$$r(k) = 0, \quad -n_w - 1 \leq k \leq -1, \tag{7.62}$$

and thus

$$\tilde{A}_{-1} = \begin{bmatrix} \lambda^{-1/2} I_{n_w+1} & 0_{n_w+1 \times n_s} \\ 0_{n_s \times n_w+1} & A \end{bmatrix}, \quad \text{and} \quad \tilde{C}_{-1} = \begin{bmatrix} 0_{1 \times n_w+1} & C_e \end{bmatrix}.$$

Then, according to (7.53), (7.54), (7.56) and

$$\tilde{P}_k = \begin{bmatrix} P_k^{ww} & 0_{n_w \times 1} & P_k^{w\theta} \\ 0_{1 \times n_w} & 0 & 0_{1 \times n_s} \\ P_k^{\theta w} & 0_{n_s \times 1} & P_k^{\theta\theta} \end{bmatrix},$$

we get

$$\tilde{P}_0 = \tilde{A}_{-1} \tilde{P}_{-1} \tilde{A}_{-1}^T - \tilde{K}_{-1} \tilde{R}_{e,-1}^{-1} \tilde{K}_{-1}^T + \tilde{G} Q \tilde{G}^T =$$

$$= \begin{bmatrix} \lambda^{-1} P_{-1}^{ww} & 0_{n_w \times 1} & \lambda^{-1/2} P_{-1}^{w\theta} A^T \\ 0_{1 \times n_w} & 0 & 0_{1 \times n_s} \\ \lambda^{-1/2} A P_{-1}^{\theta w} & 0_{n_s \times 1} & A P_{-1}^{\theta\theta} A^T \end{bmatrix} +$$

$$- \begin{bmatrix} \lambda^{-1/2} P_{-1}^{w\theta} C_e^T \\ 0 \\ A P_{-1}^{\theta\theta} C_e^T \end{bmatrix} (R + C_e P_{-1}^{\theta\theta} C_e^T)^{-1} \begin{bmatrix} \lambda^{-1/2} P_{-1}^{w\theta} C_e^T \\ 0 \\ A P_{-1}^{\theta\theta} C_e^T \end{bmatrix}^T +$$

$$+ \begin{bmatrix} 0_{n_w \times n_w} & 0_{n_w \times 1} & 0_{n_w \times n_s} \\ 0_{1 \times n_w} & 0 & 0_{1 \times n_s} \\ 0_{n_s \times n_w} & 0_{n_s \times 1} & G_s Q G_s^T \end{bmatrix}.$$

For simplicity, we set the 1-2 and 2-1 blocks in P_{-1} to zero, i.e.,

$$P_{-1} = \begin{bmatrix} \Pi_{-1}^{ww} & 0_{n_w \times n_s} \\ 0_{n_s \times n_w} & \Pi_{-1}^{\theta\theta} \end{bmatrix} = \begin{bmatrix} P_{-1}^{ww} & 0_{n_w \times n_s} \\ 0_{n_s \times n_w} & P_{-1}^{\theta\theta} \end{bmatrix}.$$

with $\Pi_{-1}^{ww} > 0$ and $\Pi_{-1}^{\theta\theta} > 0$ to be determined. Then, we get

$$\tilde{P}_0 - \Psi \tilde{P}_{-1} \Psi^T =$$
$$\left[\begin{array}{c|c} \lambda^{-1} \begin{bmatrix} \Pi_{-1}^{ww} & 0_{n_w \times 1} \\ 0_{1 \times n_w} & 0 \end{bmatrix} - Z_{n_w+1} \begin{bmatrix} \Pi_{-1}^{ww} & 0_{n_w \times 1} \\ 0_{1 \times n_w} & 0 \end{bmatrix} Z_{n_w+1}^T & 0_{n_w+1 \times n_s} \\ \hline 0_{n_s \times n_w+1} & A\Pi_{-1}^{\theta\theta} A^T - A\Pi_{-1}^{\theta\theta} C_e^T (R + C_e \Pi_{-1}^{\theta\theta} C_e^T)^{-1} C_e \Pi_{-1}^{\theta\theta} A^T + G_s Q G_s^T - \Pi_{-1}^{\theta\theta} \end{array} \right]$$

Let us choose

$$\Pi_{-1}^{ww} = \delta \cdot \text{diag}\{\lambda, \lambda^2, \cdots, \lambda^{n_w}\}, \tag{7.63}$$

which yields

$$\lambda^{-1} \begin{bmatrix} \Pi_{-1}^{ww} & 0_{n_w \times 1} \\ 0_{1 \times n_w} & 0 \end{bmatrix} - Z_{n_w+1} \begin{bmatrix} \Pi_{-1}^{ww} & 0_{n_w \times 1} \\ 0_{1 \times n_w} & 0 \end{bmatrix} Z_{n_w+1}^T =$$
$$\delta \cdot \begin{bmatrix} 1 & 0_{1 \times n_w - 1} & 0 \\ 0_{n_w-1 \times 1} & 0_{n_w-1 \times n_w -1} & 0_{n_w-1 \times 1} \\ 0 & 0_{1 \times n_w - 1} & -\lambda^{n_w} \end{bmatrix}.$$

Furthermore, if the pair (A, C_e) is detectable, and $(A, G_s Q^{1/2})$ is stabilizable, then there exists $\Pi_{-1}^{\theta\theta} > 0$, such that the discrete algebraic Riccati equation (DARE)

$$A\Pi_{-1}^{\theta\theta} A^T - A\Pi_{-1}^{\theta\theta} C_e^T (R + C_e \Pi_{-1}^{\theta\theta} C_e^T)^{-1} C_e \Pi_{-1}^{\theta\theta} A^T + G_s Q G_s^T - \Pi_{-1}^{\theta\theta} =$$
$$= 0_{n_s \times n_s} \tag{7.64}$$

holds [85, Theorem E.6.2, p.786]. Note that to ensure the pair $(A, G_s Q^{1/2})$ is unit-circle controllable Q should be positive definite, $Q > 0$. Let Π_{-1}^{ww} and $\Pi_{-1}^{\theta\theta}$ satisfy (7.63) and (7.64) respectively, then we have

$$\tilde{P}_0 - \Psi \tilde{P}_{-1} \Psi^T = \delta \begin{bmatrix} 1 & & & \\ & 0_{n_w-1 \times n_w-1} & & \\ & & -\lambda^{n_w} & \\ & & & 0_{n_s \times n_s} \end{bmatrix}, \tag{7.65}$$

$$= \tilde{L}_{-1} M \tilde{L}_{-1}^T, \tag{7.66}$$

where

$$\tilde{L}_{-1} \; = \; \sqrt{\delta} \begin{bmatrix} 1 & 0 \\ 0_{n_w-1\times 1} & 0_{n_w-1\times 1} \\ 0 & \lambda^{n_w/2} \\ 0_{n_s\times 1} & 0_{n_s\times 1} \end{bmatrix}, \qquad M \; = \; \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Hence, we have obtained a factorization for $d\tilde{P}_0$ with rank $\alpha = 2$. Note that now we have

$$J \; = \; (1 \oplus M) \; = \; \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Due to the -1 in the 3-3 element of J, the transformations Θ_{k-1} in (7.60) are *hyperbolic*. For comments on the implementation and numerical accuracy of these rotations, we refer to [149, Ch. 14], see, also [151, Section 2]. The resulting fast-array implementation of the Kalman algorithm, is listed in the second column of Table 7.1.

7.3.4   Comments on the extension to the MIMO case

In principle, the extension to the MIMO case is straightforward. Still we shall comment on a few steps in the extension that need some care. Again a state-space description similar to (7.16)-(7.17) can be derived. Let N_r be the number of reference channels, $r(k) \in \mathbb{R}^{m_r}$, N_u be the number of control channels, $u(k) \in \mathbb{R}^{N_u}$, and N_e be the number of residual channels, $e(k) \in \mathbb{R}^{N_e}$. Then, the adaptive feedforward controller $\widehat{W}_k(q^{-1})$ is an $N_u \times N_r$ matrix with FIR filters of length n_w; its i,j^{th} element is given by

$$\widehat{W}_k^{ij}(q^{-1}) \; = \; \widehat{w}_0^{ij}(k) + \widehat{w}_1^{ij}(k)q^{-1} + \cdots + \widehat{w}_{n_w-1}^{ij}(k)q^{-n_w+1},$$

with $\widehat{w}_l^{ij}(k) \in \mathbb{R}$. We define $\widehat{\mathbf{w}}_{n_w}^{ij}(k) = [\widehat{w}_0^{ij}(k) \cdots \widehat{w}_{n_w-1}^{ij}(k)]^T \in \mathbb{R}^{n_w}$, $\widehat{\mathbf{w}}_{n_w}^{i}(k) = [\widehat{\mathbf{w}}_{n_w}^{i1}(k)^T \cdots \widehat{\mathbf{w}}_{n_w}^{iN_r}(k)^T]^T \in \mathbb{R}^{n_w N_r}$, $\widehat{\mathbf{W}}_{n_w}(k) = [\widehat{\mathbf{w}}_{n_w}^{1}(k) \cdots \widehat{\mathbf{w}}_{n_w}^{N_u}(k)] \in \mathbb{R}^{n_w N_r \times N_u}$, and the vector stacking all controller coefficients $\widehat{\mathbf{w}}_{n_w}(k) = \text{vec}(\widehat{\mathbf{W}}_{n_w}(k)) \in \mathbb{R}^{n_w N_u N_r}$. Further, let $\mathbf{r}_{n_w}(k) = [r_1(k) \cdots r_1(k - n_w + 1) \cdots r_{N_r}(k) \cdots r_{N_r}(k - n_w + 1)]^T \in \mathbb{R}^{n_w N_r}$. Then, the control signal $u(k)$ is given by

$$u(k) = - \widehat{W}_k(q^{-1})r(k) \; = \; -\widehat{\mathbf{W}}_{n_w}^T(k)\mathbf{r}_{n_w}(k)$$
$$= - \text{vec}(\mathbf{r}_{n_w}^T(k)\widehat{\mathbf{W}}_{n_w}(k)I_{N_u}) \; = \; -\left(I_{N_u} \otimes \mathbf{r}_{n_w}^T(k)\right)\widehat{\mathbf{w}}_{n_w}(k), \quad (7.67)$$

where in the last equality we made use of the matrix rule $\text{vec}(ABC) = (C^T \otimes A)\text{vec}(B)$. Similarly, we define the vector stacking all *optimal* controller coefficients $\mathbf{w}_{n_w}^o$ so that

$$u^o(k) = \left(I_{N_u} \otimes \mathbf{r}_{n_w}^T(k)\right)\mathbf{w}_{n_w}^o. \tag{7.68}$$

Table 7.1: Kalman algorithm in covariance and fast-array forms

Kalman covariance form	Fast-array form
Assumptions:	
$0 \ll \lambda \leq 1$ $\begin{cases} R & > & 0 \\ Q & \geq & 0 \end{cases}$ $\Pi_0 = \begin{bmatrix} \Pi_0^{ww} & \Pi_0^{w\theta} \\ \Pi_0^{\theta w} & \Pi_0^{\theta\theta} \end{bmatrix} > 0$	idem $\begin{cases} \text{idem} \\ Q > 0 \text{ such that the pair } (A, G_s Q^{1/2}) \\ \qquad \text{is unit-circle controllable;} \end{cases}$ the pair (A, C_e) is detectable $\begin{cases} \Pi_{-1} = \begin{bmatrix} \Pi_{-1}^{ww} & 0_{n_w \times n_s} \\ 0_{n_s \times n_w} & \Pi_{-1}^{\theta\theta} \end{bmatrix} > 0 \\ \text{with} \\ \Pi_{-1}^{ww} = \sqrt{\delta} \cdot \text{diag}\{\lambda, \lambda^2, \cdots, \lambda^{n_w}\}, \quad \delta > 0 \\ \text{and } \Pi_{-1}^{\theta\theta} > 0 \text{ satisfies the DARE} \\ \Pi_{-1}^{\theta\theta} = A\Pi_{-1}^{\theta\theta}A^T + G_s Q G_s^T + \\ \qquad -A\Pi_1^{\theta\theta}C_e^T(R + C_e\Pi_{-1}^{\theta\theta}C_e^T)^{-1}C_e\Pi_{-1}^{\theta\theta}A^T \end{cases}$ $r(k) = 0, \quad \text{for } -n_w - 1 \leq k \leq -1$ (i.e., prewindowed data)
Initialization:	
$\begin{cases} \widehat{w}(0) & = & 0_{n_w \times 1} \\ \widehat{\theta}(0) & = & 0_{n_s \times 1} \end{cases}$ $r_{n_w}(-1) = \begin{bmatrix} r(-1) & \cdots & r(-n_w) \end{bmatrix}^T$ $P_0 = \Pi_0$	idem $r_{n_w+1}(-1) = 0_{n_w+1 \times 1}$ $\begin{cases} \tilde{L}_{-1} = \sqrt{\delta}\begin{bmatrix} 1 & 0 \\ 0_{n_w-1 \times 1} & 0_{n_w-1 \times 1} \\ 0 & \lambda^{n_w/2} \\ 0_{n_s \times 1} & 0_{n_s \times 1} \end{bmatrix} \\ R_{e,-1}^{1/2} = (R + C_e\Pi_{-1}^{\theta\theta}C_e^T)^{1/2} \\ \overline{K}_{-1}^w = 0_{n_w \times 1} \\ \overline{K}_{-1}^\theta = A\Pi_{-1}^{\theta\theta}C_e^T R_{e,-1}^{-1/2} \end{cases}$
Iterate for $k \geq 0$:	
$r_{n_w}(k) = \begin{bmatrix} r(k) & r_{n_w-1}^T(k-1) \end{bmatrix}^T$ $\begin{cases} A_k = \begin{bmatrix} \lambda^{-1/2}I_{n_w} & 0_{n_w \times n_s} \\ B_u r_{n_w}^T(k) & A \end{bmatrix} \\ C_k = \begin{bmatrix} D_{eu} r_{n_w}^T(k) & C_e \end{bmatrix} \end{cases}$ $\epsilon(k) = e(k) - C_e\widehat{\theta}(k)$ $\begin{cases} \begin{bmatrix} K_k^w \\ K_k^\theta \end{bmatrix} = K_k = A_k P_k C_k^T \\ R_{e,k} = R + C_k P_k C_k^T \\ P_{k+1} = A_k P_k A_k^T - K_k R_{e,k}^{-1}K_k^T + GQG^T \end{cases}$ $\begin{bmatrix} \widehat{w}(k+1) \\ \widehat{\theta}(k+1) \end{bmatrix} = \begin{bmatrix} \lambda^{-1/2}\widehat{w}(k) \\ A\widehat{\theta}(k) \end{bmatrix} + \begin{bmatrix} K_k^w \\ K_k^\theta \end{bmatrix} R_{e,k}^{-1}\epsilon(k)$	$r_{n_w+1}(k) = \begin{bmatrix} r(k) & r_{n_w}^T(k-1) \end{bmatrix}^T$ $\begin{cases} \tilde{A}_k = \begin{bmatrix} \lambda^{-1/2}I_{n_w+1} & 0_{n_w+1 \times n_s} \\ B_u r_{n_w+1}^T(k) & A \end{bmatrix} \\ \tilde{C}_k = \begin{bmatrix} D_{eu} r_{n_w+1}^T(k) & C_e \end{bmatrix} \end{cases}$ idem Perform J-unitary rotation to make 1-2 block of post-array zero, $J = (I_2 \oplus -1), \Theta_{k-1}J\Theta_{k-1}^T = J$ $\begin{bmatrix} R_{e,k-1}^{1/2} & \tilde{C}_k\tilde{L}_{k-1} \\ \begin{bmatrix} 0 \\ \overline{K}_{k-1}^w \\ \overline{K}_{k-1}^\theta \end{bmatrix} & \tilde{A}_k\tilde{L}_{k-1} \end{bmatrix}\Theta_{k-1} = \begin{bmatrix} R_{e,k}^{1/2} & 0_{1 \times 2} \\ \begin{bmatrix} \overline{K}_k^w \\ 0 \\ \overline{K}_k^\theta \end{bmatrix} & \tilde{L}_k \end{bmatrix}$ $\begin{bmatrix} \widehat{w}(k+1) \\ \widehat{\theta}(k+1) \end{bmatrix} = \begin{bmatrix} \lambda^{-1/2}\widehat{w}(k) \\ A\widehat{\theta}(k) \end{bmatrix} + \begin{bmatrix} \overline{K}_k^w \\ \overline{K}_k^\theta \end{bmatrix} R_{e,k}^{-1/2}\epsilon(k)$

Combining (7.67) and (7.68) with the (MIMO) state-space description of the secondary-path system (7.14)-(7.15) yields

$$
\left[\begin{array}{c} w(k+1) \\ \theta(k+1) \end{array}\right] = \left[\begin{array}{cc} \lambda^{-1/2}I_{n_w N_u N_r} & 0_{n_w N_u N_r \times n_s} \\ B_u\left(I_{N_u} \otimes \mathbf{r}_{n_w}^T(k)\right) & A \end{array}\right] \left[\begin{array}{c} w(k) \\ \theta(k) \end{array}\right] +
$$

$$
+ \left[\begin{array}{c} 0_{n_w N_u N_r \times n_w N_u N_r} \\ -B_u\left(I_{N_u} \otimes \mathbf{r}_{n_w}^T(k)\right) \end{array}\right] \widehat{\mathbf{w}}_{n_w}(k) + \left[\begin{array}{c} 0_{n_w N_u N_r \times n_v} \\ G_s \end{array}\right] v_s(k), \quad \left[\begin{array}{c} w(0) \\ \theta(0) \end{array}\right] = \left[\begin{array}{c} \mathbf{w}_{n_w}^o \\ \theta_0 \end{array}\right]
$$

$$
e(k) = \left[\begin{array}{cc} D_{eu}\left(I_{N_u} \otimes \mathbf{r}_{n_w}^T(k)\right) & C_e \end{array}\right] \left[\begin{array}{c} w(k) \\ \theta(k) \end{array}\right] - D_{eu}\left(I_{N_u} \otimes \mathbf{r}_{n_w}^T(k)\right)\widehat{\mathbf{w}}_{n_w}(k) + v_s(k)
$$

Like in Section 7.2.2, the Kalman filter algorithm can be derived for this state-space model.

The derivation of the fast implementation is similar to the derivation in Section 7.3.1–7.3.3. But now the state should be augmented with $N_u N_r$ additional zero-states as follows:

$$
\tilde{x}(0) = \left[\begin{array}{ccccccccc} \mathbf{w}_{n_w}^{o11\,T} & 0 & \mathbf{w}_{n_w}^{o12\,T} & 0 & \cdots & \mathbf{w}_{n_w}^{oN_u N_r\,T} & 0 & \theta_0^T \end{array}\right]^T \in \mathbb{R}^{(n_w+1)N_u N_r + n_s}
$$

(cf. (7.43)). The elements of the augmented state-error covariance matrix $\tilde{P}_{-1}$ referring to those zero-states are set to zero. If we choose

$$
P_{-1}^{ww} = \delta \cdot I_{N_u N_r} \otimes \text{diag}\{\lambda, \lambda^2, \cdots, \lambda^{n_w}\}
$$

in Section 7.3.3, we end up with the following MIMO extension for $\tilde{L}_{-1}$ and M

$$
\tilde{L}_{-1} = \sqrt{\delta}\left[\begin{array}{c} I_{N_u N_r} \otimes \left[\begin{array}{cc} 1 & 0 \\ 0_{n_w-1\times 1} & 0_{n_w-1\times 1} \\ 0 & \lambda^{n_w/2} \end{array}\right] \\ \\ 0_{n_s \times 2N_u N_r} \end{array}\right], \quad M = I_{N_u N_r} \otimes \left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right].
$$

Note that now $d\tilde{P}_0$ has rank $2N_u N_r$, and $\tilde{L}_{-1} \in \mathbb{R}^{((n_w+1)N_u N_r + n_s) \times 2N_u N_r}$.

7.4 Comparison with modified Filtered-RLS

Figures 7.2 and 7.3 show the block diagrams of the Filtered-RLS and the modified Filtered-RLS algorithm respectively. Because the adaptive filter $\widehat{W}_k(q^{-1})$ significantly varies in time, the adaptive filter and the secondary path system may not be interchanged as assumed in the Filtered-RLS algorithm. For this reason, the *modified* Filtered-RLS algorithm has been proposed [51], which shows better convergence. In the Filtered-RLS algorithm the reference signal $r(k)$ is replaced by a *filtered* reference signal $r'(k)$ that is generated by[1]

$$
\theta_r'(k+1) = \lambda^{1/2}A\theta_r'(k) + \lambda^{1/2}B_u r(k), \quad \theta_r'(0) = 0_{n_s \times 1} \tag{7.69}
$$

$$
r'(k) = C_e \theta_r'(k) + D_{eu} r(k) \tag{7.70}
$$

---

[1]Variables with subscript $_r$ or superscript ^{r} refer the variables from the modified Filtered-RLS algorithm (in order to prevent confusion with variables from the Kalman filter algorithm of Table 7.1).

PSfrag replacements
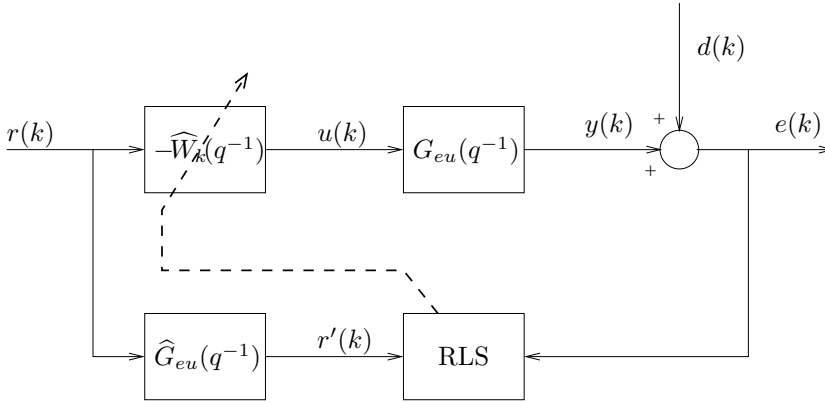
Figure 7.2: Block diagram of the Filtered-RLS algorithm.

PSfrag replacements
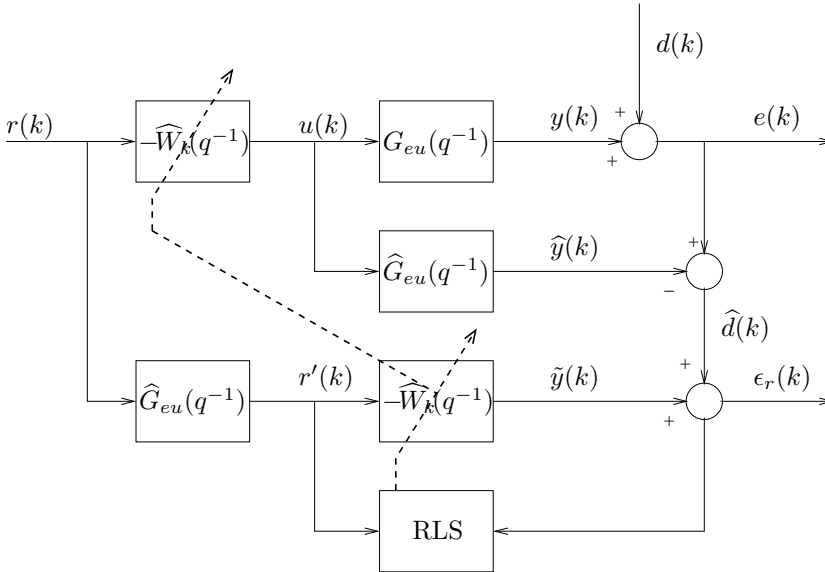
Figure 7.3: Block diagram of the modified Filtered-RLS algorithm.

Note that in the expression for $\theta'_r(k+1)$ we are using the exponential forgetting factor λ; this choice for generating $r'(k)$ is motivated by the proof of Theorem 7.1. The adaptive filter $\widehat{W}_k(q^{-1})$ is tuned by the RLS algorithm such that the error

$$\epsilon_r(k) \;=\; \widehat{d}(k) + \tilde{y}(k) \tag{7.71}$$

is minimized, where $\widehat{d}(k)$ is the estimated disturbance determined by

$$\widehat{\theta}_r(k+1) = A\widehat{\theta}_r(k) + B_u u(k), \quad \widehat{\theta}_r(0) \;=\; 0_{n_s \times 1}, \tag{7.72}$$

$$\widehat{y}(k) = C_e \widehat{\theta}_r(k) + D_{eu} u(k), \tag{7.73}$$

$$\widehat{d}(k) = e(k) - \widehat{y}(k) \tag{7.74}$$

and $\tilde{y}(k)$ is the output of the adative filter given by

$$\tilde{y}(k) = -\, r'^{T}_{n_w}(k)\widehat{w}_r(k) \tag{7.75}$$

where

$$\widehat{w}_r = \begin{bmatrix} \widehat{w}_0(k) & \widehat{w}_1(k) & \cdots & \widehat{w}_{n_w-1}(k) \end{bmatrix}^T, \tag{7.76}$$

$$r'_{n_w}(k) = \begin{bmatrix} r'(k) & r'(k-1) & \cdots & r'(k-n_w+1) \end{bmatrix}^T. \tag{7.77}$$

Table 7.2 lists the modified Filtered-RLS algorithm in its standard covariance and fast-array forms, which are derived according to [151]. The computational complexity of the modified Filtered-RLS algorithm can be reduced further by using the Fast Transversal Filter (FTF), see [149, Ch. 14], but often at the expense of numerical accuracy. The derivation of the modified RLS algorithm is quite ad hoc, and no systematic derivation of the modification and conditions for its optimality have been given yet.

In this section, we will compare the Kalman algorithm with the modified RLS algorithm. Our main result in this section is that the modified RLS algorithm is a special case of the Kalman algorithm of the previous section when there is no uncertainty on the secondary-path state (due to initial-state uncertainty and/or noise). By showing the equivalence, we have thus provided a systematic derivation of the modified filtered-RLS algorithm and conditions for its optimality.

Theorem 7.1 *The Kalman algorithm listed in Table 7.1 and the modified Filtered-RLS algorithm listed in Table 7.2 are equivalent, under the condition that*

$$\Pi^{\theta\theta}_0 \;=\; 0_{n_s \times n_s}, \quad \Pi^{w\theta}_0 \;=\; \Pi^{\theta w}_0{}^{T} \;=\; 0_{n_w \times n_s}, \quad Q \;=\; 0_{n_v \times n_v},$$

in the Kalman algorithm and $r(k) = 0$ for $-n_w \leq k \leq 0$.

Proof: First, let us define

$$\overline{S}_{k+1} = \lambda^{1/2} A\overline{S}_k + \lambda^{1/2} B_u r^T_{n_w}(k), \qquad \overline{S}_0 \;=\; 0_{n_s \times n_w} \tag{7.78}$$

Using expression (7.69) and the expression for $r_{n_w}(k)$ from Table 7.2 it can be verified that

$$\overline{S}_k = \begin{bmatrix} \theta'_r(k) & \theta'_r(k-1) & \cdots & \theta'_r(k-n_w+1) \end{bmatrix}. \tag{7.79}$$

Table 7.2: Modified Filtered-RLS algorithm in Kalman covariance and fast-array forms

KALMAN COVARIANCE FORM	FAST-ARRAY FORM
Assumptions:	
$0 \ll \lambda \le 1$ $\Pi_0^{ww} > 0$	idem $\Pi_{-1}^{ww} = \sqrt{\delta} \cdot \mathrm{diag}\{\lambda, \lambda^2, \cdots, \lambda^{n_w}\}, \quad \delta > 0$ $r'(k) = 0, \quad \text{for } -n_w-1 \le k \le -1$ (i.e. prewindowed data)
Initialization:	
$\begin{cases} \widehat{\theta}_r(0) = \theta'_r(0) = 0_{n_s \times 1} \\ \widehat{w}_r(0) = 0_{n_w \times 1} \end{cases}$ $\begin{cases} r_{n_w}(-1) = \begin{bmatrix} r(-1) & \cdots & r(-n_w) \end{bmatrix}^T \\ r'_{n_w}(-1) = 0_{n_w \times 1} \end{cases}$ $P_0^r = \Pi_0^{ww}$	idem $\begin{cases} \text{idem} \\ r'_{n_w+1}(-1) = 0_{n_w+1 \times 1} \\ \tilde{L}_{-1}^r = \sqrt{\delta}\begin{bmatrix} 1 & 0 \\ 0_{n_w-1 \times 1} & 0_{n_w-1 \times 1} \\ 0 & \lambda^{n_w/2} \end{bmatrix} \\ \overline{K}_{-1}^r = 0 \\ R_{e,-1}^r{}^{1/2} = R^{1/2} \end{cases}$
Iterate for $k \ge 0$:	
$\begin{cases} \theta'_r(k+1) = \begin{cases} A\theta'_r(k) + B_u r(k) & \text{(standard)} \\ \lambda^{1/2} A\theta'_r(k) + \lambda^{1/2} B_u r(k) & \text{(new)} \end{cases} \\ r'(k) = C_e \theta'_r(k) + D_{eu} r(k) \end{cases}$	idem
$\begin{cases} r_{n_w}(k) = \begin{bmatrix} r(k) & r_{n_w-1}^T(k-1) \end{bmatrix}^T \\ r'_{n_w}(k) = \begin{bmatrix} r'(k) & r'^T_{n_w-1}(k-1) \end{bmatrix}^T \end{cases}$	$\begin{cases} \text{idem} \\ r'_{n_w+1}(k) = \begin{bmatrix} r'(k) & r'^T_{n_w}(k-1) \end{bmatrix}^T \end{cases}$
$\begin{cases} \widehat{\theta}_r(k+1) = A\widehat{\theta}_r(k) - B_u r_{n_w}^T(k)\widehat{w}_r(k) \\ \widehat{y}(k) = C_e \widehat{\theta}_r(k) - D_{eu} r_{n_w}^T(k)\widehat{w}_r(k) \end{cases}$	idem
$\begin{cases} \tilde{y}(k) = -r'^T_{n_w}(k)\widehat{w}_r(k) \\ \epsilon_r(k) = e(k) - \widehat{y}(k) + \tilde{y}(k) \end{cases}$	idem
$\begin{cases} K_k^r = \lambda^{-1/2} P_k^r r'_{n_w}(k) \\ R_{e,k}^r = R^r + r'^T_{n_w}(k) P_k^r r'_{n_w}(k) \\ P_{k+1}^r = \lambda^{-1} P_k^r - K_k^r R_{e,k}^r{}^{-1} K_k^{rT} \end{cases}$	$\begin{cases} \text{Perform } J\text{-unitary rotation to make 1-2 block} \\ \text{of post-array zero,} \\ J = (I_2 \oplus -1), \Theta_{k-1} J \Theta_{k-1}^T = J \\ \begin{bmatrix} R_{e,k-1}^{r1/2} & r'^T_{n_w+1}(k)\tilde{L}_{k-1}^r \\ \begin{bmatrix} 0 \\ \overline{K}_{k-1}^r \end{bmatrix} & \lambda^{-1/2}\tilde{L}_{k-1}^r \end{bmatrix} \Theta_{k-1} = \\ \qquad = \begin{bmatrix} R_{e,k}^{r1/2} & 0_{1\times2} \\ \begin{bmatrix} \overline{K}_k^r \\ 0 \end{bmatrix} & \tilde{L}_k^r \end{bmatrix} \end{cases}$
$\widehat{w}_r(k+1) = \lambda^{-1/2}\widehat{w}_r(k) + K_k^r R_{e,k}^r{}^{-1}\epsilon_r(k)$	$\widehat{w}_r(k+1) = \lambda^{-1/2}\widehat{w}_r(k) + \overline{K}_k^r R_{e,k}^{r-1/2}\epsilon_r(k)$

Substituting this result into the expressions (7.70), (7.75) and (7.77) yields

$$r_{n_w}'^T(k) = C_e \overline{S}_k + D_{eu} r_{n_w}^T(k) \tag{7.80}$$

$$\tilde{y}(k) = -C_e \overline{S}_k \widehat{w}_r(k) - D_{eu} r_{n_w}^T(k) \widehat{w}_r(k) \tag{7.81}$$

and thus

$$\widehat{y}(k) - \tilde{y}(k) = C_e \widehat{\theta}_r(k) + C_e \overline{S}_k \widehat{w}_r(k). \tag{7.82}$$

The relations just derived will be used in the sequel in the proof.

By induction the following relations can be verified:

$$\text{KALMAN:} \quad \text{RLS:}$$

$$\widehat{w}(k) = \widehat{w}_r(k) \tag{7.83}$$

$$\widehat{\theta}(k) = \widehat{\theta}_r(k) + \overline{S}_k \widehat{w}_r(k) \tag{7.84}$$

$$R_{e,k} = R_{e,k}^r \tag{7.85}$$

$$K_k^w = K_k^r \tag{7.86}$$

$$K_k^\theta = \overline{S}_{k+1} K_k^r \tag{7.87}$$

$$P_k = \begin{bmatrix} P_k^{ww} & P_k^{w\theta} \\ P_k^{\theta w} & P_k^{\theta\theta} \end{bmatrix} = \begin{bmatrix} I_{n_w} \\ \overline{S}_k \end{bmatrix} P_k^r \begin{bmatrix} I_{n_w} & \overline{S}_k^T \end{bmatrix} \tag{7.88}$$

The first step is to show that (7.83)-(7.88) hold for $k = 0$, which can be verified readily from the initialization of the algorithm from Table 7.1 and 7.2 and the assumptions in the theorem. Note that if (7.84) holds, then also $\epsilon(k) = \epsilon_r(k)$ holds. Further, since $\overline{S}_0 = 0_{n_s \times n_w}$ the equivalence (7.88) yields

$$P_0 = \begin{bmatrix} \Pi_0^{ww} & \Pi_0^{w\theta} \\ \Pi_0^{\theta w} & \Pi_0^{\theta\theta} \end{bmatrix} = \begin{bmatrix} \Pi_0^{ww} & 0_{n_w \times n_s} \\ 0_{n_s \times n_w} & 0_{n_s \times n_s} \end{bmatrix},$$

which is the reason to assume $\Pi_0^{\theta\theta}$, $\Pi_0^{w\theta}$ and $\Pi_0^{\theta w}$ to be zero in the theorem.

The second step is to show that if (7.83)-(7.88) hold for k, then (7.83)-(7.88) also hold for $k + 1$. Assume (7.83)-(7.88) hold for k. That (7.83) holds for $k + 1$ directly follows from (7.84) (i.e., $\epsilon(k) = \epsilon_r(k)$), (7.85) and (7.86) and the update rules of $w(k)$ and $w_r(k)$.

To show that (7.84) holds for $k + 1$, we write

$$\widehat{\theta}(k+1) = A\widehat{\theta}(k) + \overline{S}_{k+1} K_k^r R_{e,k}^{r}{}^{-1} \epsilon_r(k)$$
$$= A(\widehat{\theta}_r(k) + \overline{S}_k \widehat{w}_r(k)) + \overline{S}_{k+1} K_k^r R_{e,k}^{r}{}^{-1} \epsilon_r(k)$$
$$= A(\widehat{\theta}_r(k) + \overline{S}_k \widehat{w}_r(k)) + \overline{S}_{k+1}(\widehat{w}_r(k+1) - \lambda^{-1/2} \widehat{w}_r(k))$$

On the other hand, we have

$$\widehat{\theta}_r(k+1) + \overline{S}_{k+1} \widehat{w}_r(k+1) =$$
$$= A\widehat{\theta}_r(k) - B_u r_{n_w}^T(k) \widehat{w}_r(k) + \lambda^{1/2} A \overline{S}_k \widehat{w}_r(k+1) + \lambda^{1/2} B_u r_{n_w}^T(k) \widehat{w}_r(k+1)$$
$$= A(\widehat{\theta}_r(k) + \overline{S}_k \widehat{w}_r(k)) + (\lambda^{1/2} A \overline{S}_k + \lambda^{1/2} r_{n_w}^T(k))(\widehat{w}_r(k+1) - \lambda^{-1/2} \widehat{w}_r(k))$$
$$= A(\widehat{\theta}_r(k) + \overline{S}_k \widehat{w}_r(k)) + \overline{S}_{k+1}(\widehat{w}_r(k+1) - \lambda^{-1/2} \widehat{w}_r(k))$$

and thus $\widehat{\theta}(k+1) = \widehat{\theta}_r(k+1) + \overline{S}_{k+1}\widehat{w}_r(k+1)$.

Before showing that (7.85)-(7.87) hold for $k+1$, we show that (7.88) holds. Using the fact that (7.85)-(7.88) hold for k and the assumption that $Q = 0$, we can write

$$
P_{k+1} = \left[\begin{array}{cc} \lambda^{-1/2}I_{n_w} & 0_{n_w \times n_s} \\ B_u r_{n_w}^T(k) & A \end{array}\right] \left[\begin{array}{c} I_{n_w} \\ \overline{S}_k \end{array}\right] P_k^r \left[\begin{array}{c} I_{n_w} \\ \overline{S}_k \end{array}\right]^T \left[\begin{array}{cc} \lambda^{-1/2}I_{n_w} & 0_{n_w \times n_s} \\ B_u r_{n_w}^T(k) & A \end{array}\right]^T +
$$

$$
- \left[\begin{array}{c} I_{n_w} \\ \overline{S}_{k+1} \end{array}\right] \lambda^{-1/2} P_k^r r_{n_w}^{\prime T}(k) \left(R + r_{n_w}^{\prime T}(k)P_k^r r_{n_w}^{\prime}(k)\right)^{-1} r_{n_w}^{\prime}(k) P_k^r \lambda^{-1/2} \left[\begin{array}{c} I_{n_w} \\ \overline{S}_{k+1} \end{array}\right]^T
$$

$$
= \left[\begin{array}{c} I_{n_w} \\ \overline{S}_{k+1} \end{array}\right] \lambda^{-1} P_k^r \left[\begin{array}{c} I_{n_w} \\ \overline{S}_{k+1} \end{array}\right]^T - \left[\begin{array}{c} I_{n_w} \\ \overline{S}_{k+1} \end{array}\right] K_k^r R_{e,k}^{-1} K_k^{rT} \left[\begin{array}{c} I_{n_w} \\ \overline{S}_{k+1} \end{array}\right]^T
$$

$$
= \left[\begin{array}{c} I_{n_w} \\ \overline{S}_{k+1} \end{array}\right] P_{k+1}^r \left[\begin{array}{c} I_{n_w} \\ \overline{S}_{k+1} \end{array}\right]^T
$$

Thus (7.88) holds for $k+1$.

Using this result, we can write

$$
\begin{aligned}
R_{e,k+1} &= R + (D_{eu} r_{n_w}^T(k+1) + C_e \overline{S}_{k+1}) P_{k+1}^r (D_{eu} r_{n_w}^T(k+1) + C_e \overline{S}_{k+1})^T \\
&= R + r_{n_w}^{\prime T}(k+1) P_{k+1}^r r_{n_w}^{\prime}(k+1) \\
&= R_{e,k+1}^r
\end{aligned}
$$

$$
\begin{aligned}
K_{k+1} &= \left[\begin{array}{c} K_{k+1}^w \\ K_{k+1}^\theta \end{array}\right] \\
&= \left[\begin{array}{c} I_{n_w} \\ \lambda^{1/2}(B_u r_{n_w}^T(k+1) + A\overline{S}_{k+1}) \end{array}\right] \lambda^{-1/2} P_{k+1}^r \left[\begin{array}{c} I_{n_w} \\ \overline{S}_{k+1} \end{array}\right]^T \left[\begin{array}{c} r_{n_w}(k)D_{eu}^T \\ C_e^T \end{array}\right] \\
&= \left[\begin{array}{c} I_{n_w} \\ \overline{S}_{k+2} \end{array}\right] \lambda^{-1/2} P_{k+1}^r (D_{eu} r_{n_w}^T(k+1) + C_e \overline{S}_{k+1})^T \\
&= \left[\begin{array}{c} I_{n_w} \\ \overline{S}_{k+2} \end{array}\right] K_{k+1}^r
\end{aligned}
$$

Thus (7.85)-(7.87) hold for $k+1$.

Hence, (7.83)-(7.88) hold for all $k \geq 0$ and we conclude that the Kalman algorithm of Table 7.1 and the RLS algorithm of Table 7.2 (with $\theta_r'(k+1)$ given by (7.69)) are equivalent under the conditions given in the theorem. $\square$

The effect of λ in the expression for $\theta'(k+1)$ in (7.69) (see also Table 7.2) yields

$$
\begin{aligned}
r'(k) = D_{eu}r(k) &+ \lambda^{1/2}C_e B_u r(k-1) + \lambda C_e A B_u r(k-2) + \\
&+ \lambda^{3/2}C_a A^2 B_u r(k-3) + \cdots
\end{aligned}
$$

and thus λ has the effect of *exponential forgetting* in the generation of the filtered reference signal $r'(k)$.

Table 7.3: Computational load of the fast-array RLS (Table 7.2) and fast-array Kalman (Table 7.1) algorithms, with the secondary-path in full state-space and FIR parameterization in number floating point additions or multiplications (neglecting terms not depending on dimensions n_w and n_s).

Action:	RLS		Kalman	
	state-space	FIR	state-space	FIR
Filtered-reference:	$2n_s^2 + 3n_s$	$2n_s$	$-$	$-$
Disturbance estimate:	$2n_s^2 + 3n_s$	$2n_s$	$-$	$-$
Calculation innovation:	$2n_w$	$2n_w$	$2n_s$	0
Construction pre-array:	$6n_w$	$6n_w$	$6n_w + 4n_s^2$	$6n_w + 2n_s$
Performing rotations:	$12n_w$	$12n_w$	$12n_w + 12n_s$	$12n_w + 12n_s$
Updating coefficients/state:	$3n_w$	$3n_w$	$3n_w + n_s^2 + n_s$	$3n_w + 2n_s$
Calculating control:	$2n_w$	$2n_w$	$2n_w$	$2n_w$
Total:	$25n_w + 4n_s^2 + 6n_s$	$25n_w + 4n_s$	$23n_w + 5n_s^2 + 15n_s$	$23n_w + 16n_s$

Finally, Table 7.3 compares the computational complexity of the fast-array implementations of the modified Filtered-RLS and the Kalman algorithm proposed in this chapter. From this table, we infer that the number of floating point operations are linearly increasing with n_w. The main computational step is the evaluation of the rotations. Each elementary rotation is of the form

$$x_{new} \leftarrow \alpha(x + \rho y)$$
$$y_{new} \leftarrow \rho x_{new} - \beta y$$

which takes 6 floating point operations. The rotations need to be evaluated for all rows in the pre-array and by operating on the elements in the column pair 1-2 and the column pair 1-3. Alternative implementations are also possible, see, e.g., [149, Ch. 14].

7.5 Simulation results

To illustrate the method, simulations are performed on an $n_s = 19^{th}$ order discrete acoustic duct system with signal-to-noise ratio of 30dB and $r(k)$ is a zero-mean white-noise signal with unit variance. The number of filter coefficients was chosen to be $n_w = 150$. Only the *fast-array* implementations contained in Tables 7.1 and 7.2 are used, with $\lambda = 1$ (no exponential forgetting). The measurement noise variance was $R = 2.1 \cdot 10^{-5}$. The value of δ, which determines the magnitude of the initial state covariance P_{-1}, was set to $\delta = 10^{-3}$. In the (fast-array) Kalman filter algorithm Q has chosen to be $Q = 2 \cdot 10^{-3}$. For comparison, also the FxLMS and the preconditioned FxLMS (c.f. Chapter 6) have been used with the normalized stepsize chosen to be 0.05, optimized by trial and error. All algorithms are turned on after 1000 samples.

Figure 7.4 shows the learning curves obtained by the algorithms, averaged over 50 experiments. From this figure, it can be concluded that both the RLS and the Kalman filter algorithm converge to (approximately) the same performance level. However, the RLS algorithm shows a significant overshoot directly after turning on the algorithm. This overshoot can be explained by the fact that the uncertainty in
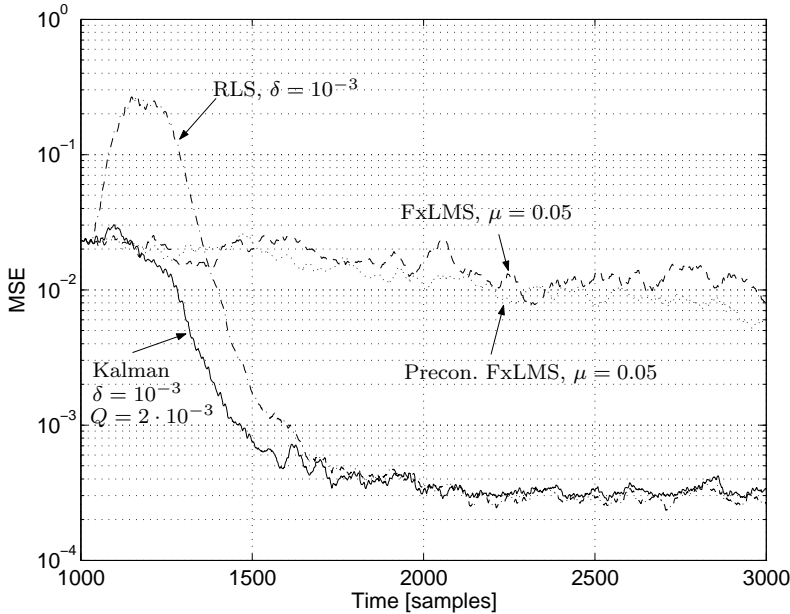
Figure 7.4: Learning curves obtained by the Kalman, RLS, FxLMS and the Preconditioned FxLMS algorithm, averaged over 50 experiments.

the secondary path state is not taken into account, contrary to the Kalman filter algorithm for $Q \neq 0$. The FxLMS and the preconditioned FxLMS are converging much slower, as is expected since they are based on an LMS estimated gradient update. But, note that the computational complexity per iteration of FxLMS and preconditioned FxLMS is still lower than the computational complexity of the fast-array implementations of the RLS and Kalman filter algorithms.

The same experiment was performed by chosing δ in the RLS and the Kalman filter algorithm to be $\delta = 10^{-4}$, see Figure 7.5. From this figure, it is clear that the overshoot of the RLS algorithm with $\delta = 10^{-3}$ can be considerably reduced by lowering δ to $\delta = 10^{-4}$, but at the expense of convergence rate. Using $\delta = 10^{-4}$ in the Kalman algorithm, shows fast convergence at the first few hundred samples, but then it convergence rate slows down to the convergence of the RLS algorithm.

From these observations, we conclude that the overshoot or bad convergence of the RLS algorithm at startup can be prevented by the Kalman filter algorithm, since uncertainty in the secondary path state is accounted for.

The same experiment has been repeated for an uncertain secondary path model, which contains 1 sample in addition to the secondary path system. Figure 7.6 shows the learning curves obtained by the Kalman and the RLS algorithm for $\delta = 10^{-3}$ and $\delta = 10^{-4}$. From this figure, we observe that the algorithms are, within some extend, robust for the model uncertainty in the secondary path model, at least no divergence is obtained. However, the all algorithms converge to a suboptimal solution (c.f. with the performance obtained in Figure 7.4 and 7.5), which is lower $\delta = 10^{-4}$ than for $\delta = 10^{-3}$ for both the Kalman and the RLS algorithms. From
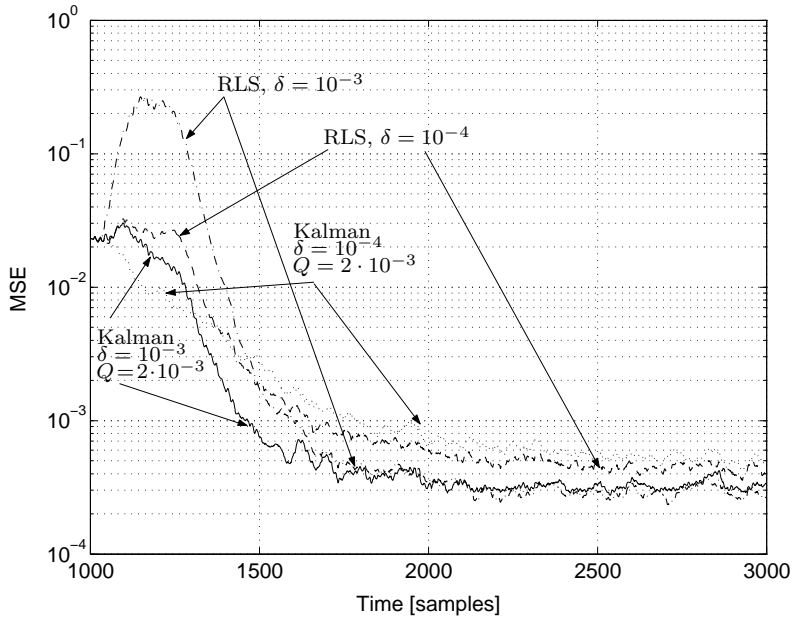
Figure 7.5: Learning curves obtained by the Kalman and the RLS algorithm for $\delta = 10^{-3}$ and for $\delta = 10^{-4}$, averaged over 50 experiments.
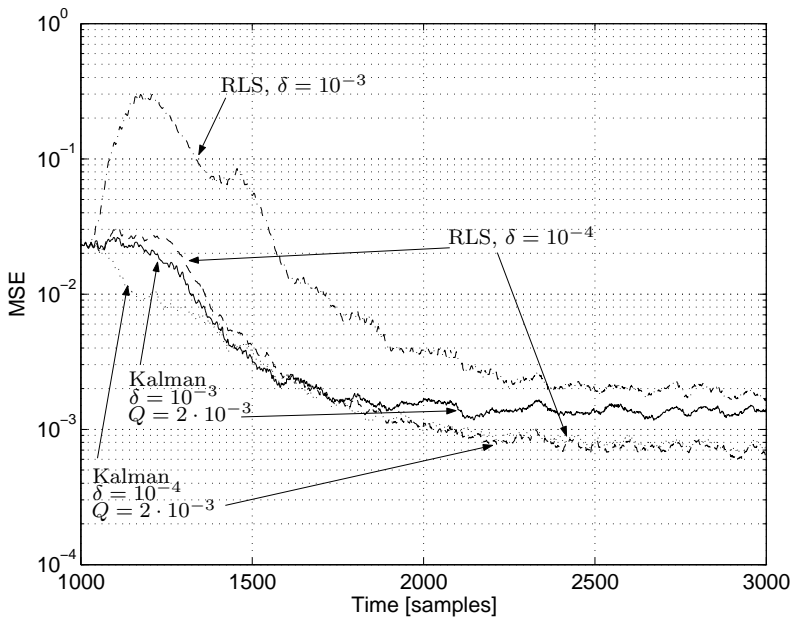
Figure 7.6: Learning curves obtained by the Kalman and the RLS algorithm for $\delta = 10^{-3}$ and for $\delta = 10^{-4}$ with 1 sample delay uncertainty in the secondary path model, averaged over 50 experiments.

this observation, we conclude, that the robustness w.r.t. model uncertainty (for this particular case) can be improved by lowering δ to $\delta = 10^{-4}$. But both, the Kalman and the RLS algorithm, converge to suboptimal solutions, and do not show significant different convergence behavior for this model uncertainty (apart from the uncertainty in the secondary-path state).

7.6   Conclusions

The active control problem can be reformulated in state-space form, which overcomes formulating the control problem in terms of interchanging the adaptive filter and the secondary path. In this way, uncertainty due to initial-state and time-variations are taken into account explicitly. The state-estimation problem was solved by the Kalman filter and the structure in the state-space matrices was exploited to develop a fast-array implementation of the algorithm. Under the theoretical condition that there is no uncertainty in the secondary path state, it is proven that the Kalman algorithm is equivalent to the modified Filtered-RLS algorithm. Hence, the Kalman algorithm can be seen as a generalization of the modified Filtered-RLS algorithm. At the same time, conditions for optimality of the modified Filtered-RLS algorithm are derived. When using exponential forgetting in the modified Filtered-RLS algorithm, the forgetting factor should also be applied to the reference signal.

Chapter 8

Conclusions, evaluation and further research

8.1 Conclusions

In this thesis various algorithms for fast and robust active control of broadband noise and vibration are proposed. In the following, we summarize the main conclusions and the more detailed conclusions for each chapter. Section 8.2 evaluates the practical relevance of the contributions and Section 8.3 comments on further research.

The main conclusions of this thesis are:

- The nominal broadband feedforward controller can be estimated by solving a control-relevant identification problem, which compensates for model errors in the spectral factor of the reference signal and enables controller order reduction.

- The nominal design method has been extended to a probabilistic robust design method, in which the model uncertainty is considered as a stochastic variable (in the frequency domain) with zero-mean and known covariance and optimizes the performance averaged over the stochastic model uncertainty.

- Both, the nominal and the robust, design methods are applied to feedback problems using IMC. It is shown that the robust design method increases stability robustness.

- Conditions for global convergence of the FuLMS algorithm are derived which hold also in case perfect cancellation is not achievable.

- The robustness of FxLMS and preconditioned FxLMS w.r.t. model uncertainty in the secondary path has been increased using the probabilistic robust approach.

- A Kalman filter solution has been proposed which optimally (in the minimum variance sense) estimates the filter coefficients. The derivation overcomes formulating the control problem in terms of *interchanging* the adaptive filter

and the secondary path as in the modified Filtered-RLS algorithm. In case there is no uncertainty on the state of the secondary path (due to initial state uncertainty or noise), it is shown that the Kalman filter solution reduces to the modified Filtered-RLS algorithm.

- A fast-array implementation of the Kalman filter solution is derived with complexity $O(23n_w + 16n_s)$ for the SISO case, with n_w the number of coefficients of the adaptive controller and n_s the number of coefficients of the secondary path model (in FIR realization).

- A fast implementation of state-space iterations via input- and output-normal forms has been obtained. An application is the use of the input-normal parameterization to efficiently calculate the regression vector of the FxLMS algorithm for multi-channel secondary path systems, see Appendix C.

The more detailed conclusions per chapter are:

Chapter 2 Optimal feedforward and feedback control. The solution to the optimal feedforward control problem is provided by the Causal Wiener filter. When the detector and the secondary path contain non-minimum phase zeros that are not canceled by non-minimum phase zeros in the primary path, there will be still correlation between the optimal residual disturbance and the disturbance source signal, which means that perfect cancellation is not achievable. Measurement noise on the reference signal uncorrelated with the disturbance source signal, will reduce the magnitude of the Causal Wiener filter. Dually, control effort weighting also will reduce the magnitude of the Causal Wiener filter. For systems with feedback, where the feedback is determined by a *stable* system, the feedback can be compensated for by means of Internal Model Control. Since in practice, the model of the feedback is distorted by model errors, stability robustness of the closed loop has to be taken into account in the controller design.

Chapter 3 Nominal controller estimation. Subspace Model Identification (SMI) provides a compact state-space model which models the common dynamics at several outputs only once, contrary to Prediction Error Modeling (PEM). Simulation experiments on a 4×4 vibrating plate of order 20, show that with the SMI (PO-MOESP) implementation contained in the SLICOT library [159] an accurate model could be obtained within less then one minute computation time. In comparison, the PEM algorithm implemented in the MATLAB Identification toolbox [105] provided a less accurate (output-error) model within more than 5 minutes computation time. Therefore, in this thesis, the choice of SMI to estimate models of the systems (which are typically infinite dimensional) is preferred over the PEM approach.

For the identification of the secondary path G_{eu} and the feedback path G_{ru} it is preferred that the disturbance source is turned off (i.e., $s(k) = 0$, which is possible in many practical applications). In this way, $s(k)$ does not act as a disturbance on the outputs of G_{eu} and G_{ru}, and thus can be identified more accurately.

Assuming perfect modeling of G_{eu} (and for systems with feedback G_{ru} it is assumed that the feedback is perfectly compensated by IMC), the Causal Wiener filter has been estimated by solving a control-relevant identification problem. This means that during identification the same cost-function has been minimized as obtained for control.

Chapter 4 Robust controller estimation. Model uncertainty has been explicitly taken into account by using a probabilistic uncertainty model, cf. [164]. The uncertainty is supposed to be a random variable in the frequency domain, with zero-mean and known covariance. The (frequency-dependent) covariance is given by a state-space realization of a (co-)spectral factor, and can be determined e.g. by identifying the system under various realistic conditions.

The feedforward controller, which minimizes the variance of the residual signal summed over all channels and *averaged over the distribution of the model uncertainty*, is given by the Cautious Wiener filter. It is shown, that the uncertainty in the detector and the secondary path act similarly as measurement noise on the reference signal and control effort weighting respectively. Furthermore, at the frequency bands where the uncertainty in the detector and/or secondary path is large, the gain of the Cautious Wiener filter is reduced. The Cautious Wiener solution is independent of the uncertainty in the primary path.

The nominal control-relevant estimation method of Chapter 3 is extended to the control-relevant estimation of the Cautious Wiener filter. In case of uncertainty in the secondary path G_{eu}, the inner-outer factorization of the secondary path model is replaced by the inner-outer factorization of the secondary path model augmented with the co-spectral factor $\widetilde{\Delta G_{eu}}$ which quantifies the uncertainty. In case of uncertainty in the detector path G_{rs}, an additional disturbance signal, uncorrelated with the primary disturbance and shaped by the spectral factor $\widetilde{\Delta G_{rs}}$ which quantifies the uncertainty, is added on the reference signal.

For feedback systems, the Cautious Wiener filter increases stability robustness, since the loop-gain is reduced at the frequencies where the uncertainty is large. Because closed-loop stability is not *guaranteed*, a scaling of the uncertainty is introduced which allows further reduction of the loop-gain until stability is obtained.

Chapter 5 Convergence Analysis of Filtered-U LMS. The convergence analysis of the FuLMS algorithm, which adapts an IIR filter, as given in [187] was based on Ljung's Ordinary Differential Equation (ODE) approach to the analysis of recursive identification schemes. In the ODE approach, it is assumed that the system is contained in the model set, such that when the recursive identification algorithm has been converged to its global optimum, the prediction error is only determined by measurement noise. To adopt the ODE approach to study the convergence of the FuLMS algorithm, it was assumed in [187] that perfect cancellation is achievable, such that at the global optimum, the residual signal is determined by measurement noise only. In Chapter 5 it has been shown that this assumption is not necessary and the ODE approach can be used, even for the case perfect cancellation is not achievable. The reason is, that it could be shown that the *optimal* residual signal (which may still contain contributions from the disturbance source) is uncorrelated with the FuLMS regression vector.

The convergence rate of the FuLMS algorithm can be significantly improved by a preconditioning similar as proposed for the FxLMS algorithm in [47]. The preconditioning consists of prewhitening the reference signal and filtering the output of the adaptive filter with the inverse secondary path outer-factor. Due to the nonlinearity in the adaptive IIR filter, the analysis of the convergence rate is highly complicated and no proof could yet be given that preconditioning increases the convergence rate of the FuLMS algorithm. But, significant increase of the convergence rate has been demonstrated by simulation experiments on an acoustical duct system.

Chapter 6 Robust preconditioned Filtered-X LMS. The robustness of the preconditioned FxLMS algorithm proposed in [47] has been increased using the probabilistic robust filtering approach considered in Chapter 4. The robust FxLMS and robust preconditioned FxLMS algorithms have been derived by minimizing the variance of the residual signal (summed over all channels) averaged over the distribution of the model uncertainty. It has been shown that, accounting for uncertainty in the secondary path model, the SPR condition which guarantees global convergence is relaxed. Furthermore, the gain of the preconditioning filter is reduced, which may prevent over steering problems.

The robust (preconditioned) FxLMS algorithm is equivalent to the FxLMS algorithm for the system with residual signal channels augmented with the control signal filtered by the uncertainty spectral factor $\widetilde{\Delta G}_{eu}$. In case the uncertainty is frequency independent, i.e. $\widetilde{\Delta G}_{eu}$ is a scalar, the robust (preconditioned) FxLMS algorithm reduces to the *Leaky* (preconditioned) FxLMS algorithm.

Chapter 7 A Fast-array Kalman filter solution. The ANVC adaptive feedforward control problem has been formulated as a state-estimation problem. This approach was also taken in [154]. However, [154] still uses an estimation of the disturbance signal obtained by means of a Internal Model Control approach, which is not necessary in the approach proposed in Chapter 7. The state estimation problem has been solved optimally (in the minimum variance sense) by using the non-stationary Kalman filter solution.

It has been shown that the state-space model is structured, in the sense considered in [150], such that a fast-array implementation could be derived. In fast-array algorithms, the initialization of the state-error covariance matrix is important, because for all following iterations, it determines the rank of the 'difference' quantity of the state-error covariance matrix between two iterations. The rank of this matrix determines the size of the array which will be rotated. An efficient initialization has been found, such that the rank of the 'difference' is just 2 for all iterations (for the single channel case).

Furthermore, it has been shown that the modified Filtered-RLS algorithm is a special case of the Kalman filter solution. Under the theoretical condition that there is no uncertainty in the secondary path state, it is proven that the Kalman algorithm is equivalent to the modified Filtered-RLS algorithm. In this way, a theoretical sound derivation of the modified Filtered-RLS algorithm has been obtained, which overcomes formulating the control problem in terms of *interchanging* the adaptive filter and the secondary path. It also shows, that when using exponential

forgetting in the modified Filtered-RLS algorithm, the forgetting factor should also be applied to the reference signal.

8.2   Evaluation of practical relevance

The thesis started in Section 1.2 on page 5 by distinguishing, besides the necessity of electronic equipment, the following problems with *practical* ANVC control systems:

- Lack of robustness;

- Slow convergence for broadband applications;

- Computationally complex algorithms;

- Lack of tracking performance.

At the end of the thesis, do we do better? Though further improvements are necessary, we can answer this question positively, since the proposed algorithms show promising results in laboratory experiments. We may conclude, that all aspects —robustness, convergence time, computational complexity and in less extend also tracking— are addressed, better understanding of the problems and existing algorithms has been obtained and improved algorithms are proposed.

In the following, we briefly discuss the practical relevance the contributions.

Control-relevant controller estimation. In case the spectral factor of the reference signal cannot be accurately identified, the control-relevant nominal and robust design methods, of Chapter 3 and 4 respectively, are preferred to obtain better performance, instead of the fully model based Causal/Cautious Wiener or LQG approach.

The drawback is that the causal factor $[G^*_{eu,i}G_{es}G^*_{rs,ci}]_+$ of the Causal Wiener filter needs to be identified, and the user needs to determine system identification parameters (the number of samples for identification, the number of blockrows for subspace identification, and the system order). However, these parameters are determined by the dynamics of the system and hence can often be chosen the same as those during identification of G_{eu} and G_{ru}. In addition, subspace identification methods provide information on the system order and in general the number of blockrows can be chosen to be two times the expected system order.

Furthermore, the nominal and robust design methods, of Chapter 3 and 4, allow the estimation of a reduced order Causal or Cautious Wiener filter which reduces the number of real-time computations to calculate the control signal.

Nominal and robust controller estimation. In general robustness has to be taken into account in every controller design for practical application. Usually, the robustness (performance and for feedback applications also stability robustness) of the nominal controller is very poor, but can be easily improved by control-effort weighting and/or increasing the variance of the measurement noise on the reference signal, cf. Section 2.2.3 and 2.2.4 on page 41. Which is the same as setting the

uncertainty models $\widetilde{\Delta G}_{rs}$ and $\widetilde{\Delta G}_{eu}$ to a scalar times the identity matrix in the robust design approach (see the end of Section 4.3 on page 86).

In applications where the uncertainty in the model is dominant *only in particular frequency bands*, robustness can be improved without paying too much performance, by using the probabilistic robust controller design approach discussed in Chapter 4. Uncertainty models $\widetilde{\Delta G}_{rs}$ and $\widetilde{\Delta G}_{eu}$ of the uncertainty in the detector path model and the secondary path model respectively, needs to be determined by estimating them as discussed in Section 4.4 on page 87 or by tuning.

Blockwise/offline controller estimation versus FxLMS/FuLMS. For the experiments considered in this thesis (systems with order 30–80 and up to 4 inputs and 4 outputs), the computation time to calculate the nominal and robust controllers by the estimation approach of Chapter 3 and 4 (about 40 seconds, on a 2.6GHz PC) is lower than the convergence of the FxLMS or FuLMS algorithm (about 8 minutes). With increasing computational performance of PC's the computation time to calculate the nominal and robust controllers is expected to be further reduced in the future, whereas the convergence time of adaptive algorithms is related to the number of samples and the sampling rate.

Furthermore, since the controllers obtained by the nominal and robust estimation methods are not adaptive, the number of real-time computations is significantly less (especially when using the input- or output-normal form parameterization of Appendix C) then the number of real-time computations of FxLMS and FuLMS.

On the other side, when the model of the secondary path G_{eu} is very *coarse*, the performance of the off-line/blockwise nominal and also robust design methods is poor, whereas FxLMS and FuLMS yield still good performance in case the conditions (such as the SPR condition) for global convergence are satisfied.

Global convergence and preconditioning of FuLMS. Though the conditions for global convergence of FuLMS are relaxed to the more realistic case where no perfect cancellation is achievable due to non-minimum phase zeros, the convergence analysis of the FuLMS algorithm is still only of theoretical interest. This is because to check the conditions for global convergence knowledge of the optimal denominator polynomial of the adaptive filter should be known as well as the orders of the optimal numerator and denominator polynomials. Only in case accurate models of the system are available, the optimal controller can be calculated and the conditions can be analyzed.

The preconditioning of the FuLMS algorithm considerably increases the convergence rate, as was illustrated by simulation examples. The computational costs are only slightly increased since just 2 additional filtering actions (of the reference signal and the control signal) are necessary and the regression vector is constructed by filtering the pre-withened reference signal with the inner-factor of the secondary path model which is usually of much smaller order —the number of non-minimum phase zeros— than the complete secondary path model. In multiple channel applications the computational complexity may be even *reduced*, since each reference signal and each control signal need to be filtered only with the inner-factor instead of the complete secondary path model, which may yield a larger reduction of

real-time computations than the increase due to the 2 additional filtering actions. This also holds for the preconditioned FxLMS algorithm, except that the (filtered) control signal is not contained in the regression vector.

Robust preconditioned FxLMS. The robust preconditioned FxLMS algorithm increases the robustness of preconditioned FxLMS in specific frequency bands. In this way, the robustness of the preconditioned FxLMS algorithm was considerably increased without paying too much performance in a simulation example on a duct system with uncertainty in the secondary path delay. It was shown, that the performance versus robustness trade-off was much better than by just using a scalar weighting of the control-effort signal (equivalent to leakage). However, the number of columns of the regression vector is increased by the number of control-signals, which rather increases computational complexity. Hence, in case computational complexity is more a key issue than the performance-robustness trade-off, the robustness of preconditioned FxLMS should be increased by using a scalar leakage factor in the preconditioned FxLMS algorithm.

Fast-array Kalman filter. The advantage of the Kalman filter solution over the modified Filtered-RLS algorithm is that uncertainty in the secondary path state is taken into account. In practice the secondary path state is not known exactly due to noise and the unknown initial secondary path state. Because of the uncertainty in the secondary path state, the modified Filtered-RLS algorithm shows a large overshoot, which is prevented using the Kalman filter solution.

The Kalman filter solution yields optimal convergence rate of the controller coefficients, e.g., in the broadband simulation example with an acoustical duct in Section 7.5 on page 145 the 150 filter coefficients converged within about 1000 samples. This makes the Kalman filter solution very suitable in tracking of non-stationary disturbances.

Because a fast-array implementation of the Kalman filter solution could be derived, the obtained algorithm can be applied in single channel ANVC systems, where the complexity is $O(23n_w + 16n_s)$, with n_w the number of coefficients of the adaptive controller and n_s the number of coefficients of the (FIR) secondary path model. For multiple channel applications, further research is necessary to reduce the computational complexity. In the following section, we provide some recommendations for further research on the reduction of computational complexity.

8.3 Further research

It would not be remarkable, that the contributed algorithms do not provide improvement of all criteria at the same time. For example, the robust version of the preconditioned FxLMS algorithm improves the robustness, but increases computational complexity. This is one motivation for further improvement of the contributions and the search for new algorithms for ANVC applications. Furthermore, the potential of the algorithms have been demonstrated by laboratory experiments, but no extensive comparisons have been made between various algorithms such as the (robust) preconditioned FxLMS algorithm and the nominal (robust) controller

estimation method. The results of these comparison studies will provide useful additional information about the choice of the control algorithm for different practical ANVC problems.

In the following, a list of further research recommendations are provided. First, we give the theoretical research recommendations on the further development of the algorithms. Second, we give the practical research recommendations on the choice and the use of the algorithms in practice.

Theoretical research recommendations:

- **Automatic iterative identification and robust controller design:** In practice, it may happen that the model uncertainty is so large, that also the probabilistic robust controller of Chapter 4 does not yield good performance anymore. In this case, it is desirable to update the model during control, such that the model accuracy is improved in the frequency band of interest. Using this improved model a new (robust) controller can be computed which is expected to yield better performance. This iterative identification and controller design approach has been a quite active research field, cf. e.g., [69, 71, 175]. However, further research is necessary to develop reliable fully automatic procedures for high order systems *without intervention of the user*.

  In this line, also the iterative design proposed in [115, 116] which has been based on the iterative feedback tuning approach of [82]. Here, no model is necessary at all but during control in each iteration the disturbance source needs to be turned off for some time, which is not always possible during operation.

- **Stable predictor estimation with SMI:** In the control-relevant estimation approach of Chapter 3 and 4, the algorithm based on prediction error model (PEM) identification differs from the one based on subspace model identification (SMI). This is due to the fact that SMI cannot estimate a predictor with explicitly taking into account the stability constraint. In PEM identification, this stability constraint is accounted for every iteration by projection on the set of stable systems. Further research should give insight, whether the stability constraint can be incorporated in SMI algorithms, e.g., by means of an iterative search algorithm with projection on the stable set of system (e.g., by means of the methods proposed in [27, 111].

- **Control-relevant estimation of probabilistic robust feedback controller proposed in [73]:** Given a stabilizing controller which yields a desired sensitivity function for the nominal model, Goodwin et al. [73] propose a control design method based on the IMC principle to minimize the H_2-norm of the difference between the actual and the desired sensitivity function, averaged over the stochastic model uncertainty. As in Chapter 4, where the *Cautious Wiener* has been used in an IMC framework for feedback applications, a linearizing approximation had to be made to reformulate the non-linear closed-loop problem in a linear open-loop problem. Further research is necessary to compare both approximations in more detail and to determine which approach is better (i.e., less conservative). Further research

is also necessary to develop a control-relevant estimation approach to estimate the controller proposed in [73] using measured data directly.

- **Convergence of Filtered-U LMS:**

  1. **Convergence analysis of FuLMS for problems with feedback:** Conditions for global convergence of the FuLMS algorithm were proposed in [187] assuming perfect cancellation is achievable. Chapter 5 shows this assumption of perfect cancellation is achievable is not necessary. But, in both convergence results it has been assumed that the reference signal is not distorted by feedback of the control signal. In [187], and by reference also in [46, 58, 125], it was erroneously claimed that the convergence result could be extended to systems with feedback, as was pointed out in [168]. Hence, the problem of the convergence of the FuLMS algorithm in case there is feedback of the control signal to the reference signal is still an interesting research question.

     Note, that active control problems with feedback were the original motivation of Eriksson et al. [49] to propose the FuLMS algorithm.

  2. **Proof that preconditioning increases the convergence rate of the FuLMS algorithm:** Due to the non-linear structure of the IIR adaptive filter in the FuLMS algorithm, the analysis of the convergence rate of the FuLMS algorithm is very difficult. Because of this reason, no theoretical analysis of convergence of the preconditioned FuLMS algorithm had been given in Chapter 5. It would be interesting to investigate whether bounds on the convergence rate of the FuLMS and the preconditioned FuLMS algorithm can be obtained. For example by using a linear approximation of the non-linear ODE which describes the convergence of filter-coefficients obtained by using Ljungs' ODE approach.

- **Fast-array Kalman solution:**

  1. **Block processing for further reduction of computations:** In [110] it is shown for the fast implementation of the RLS algorithm via the Fast Transversal Filter (FTF) a block processing approach can be used to obtain a further reduction of the computational complexity. The block processing approach exploits the (Fast Fourier Transform) FFT, such that the total number of computations for one block of samples is less than the number of computations of the sample-by-sample algorithm using the same number of samples. It would be interesting to investigate whether such a block processing algorithm exists for the fast-array Kalman solution[1].

  2. **Fast-array robust filter solutions:** Recently, the fast-array H_∞-filter has been developed in [74] (see also [75]). We believe that the fast-array H_∞-filter can also be used for the time-varying state-space equations

---

[1]Note, that it is not trivial —if possible at all— to derive a FTF implementation of the fast-array filter. This is because, in the Kalman filter solution the state-transition matrix A_k is not just a scalar times the identity matrix as in the RLS algorithm (where the scalar is the forgetting parameter), cf. [151, Sec. 7.5.3].

in the state-space approach to active noise control of Chapter 7. The H_∞-filter appears to be more robust w.r.t. uncertainty in the statistics of the measurement noise than the Kalman filter. And therefore better robustness properties are expected by using the H_∞-filter, as was also discussed in [154].

However, in the H_∞-filter every iteration a particular matrix should be checked to be positive definite. If this matrix becomes negative definite the H_∞ filter does not yield reliable results anymore.

This 'deregularization' of the H_∞-filter is prevented by the robust framework to state-space estimation proposed in [148]. Here a robust least squares approach [152] to state-estimation has been taken, where robustness has been obtained by *regularization* of the least-squares problem. Also see the robust solutions in [167], which guarantees a minimum state-error covariance. However, the robust solutions of [148,167] breaks down the structure in the state-space model to derive a fast-array implementation. It is interesting to investigate fast-array robust solutions in line of the regularization approach taken in [148].

3. **Generalization to the case where perfect cancellation cannot be obtained:** In the state-space approach to adaptive active noise control of Chapter 7, it is assumed that the FIR adaptive filter is such that there exist optimal values of the coefficients such that the residual disturbance only consists of measurement noise. It is worth to investigate whether this assumption can be relaxed. For example, to enable to choose the FIR adaptive filters of reduced length which does not yield perfect cancellation but some desired level of cancellation. Or another application would be to ANVC systems in which perfect cancellation cannot be obtained due to non-minimum phase zeros.

4. **Analyze the convergence of the Kalman filter solution for systems with feedback:** In case the reference signal is 'distorted' by feedback from the control signal, the state-transition matrix A_k and the output matrix C_k are depending on past samples of the control signal. The question which needs to be investigated, is whether the Kalman filter still provides an *unbiased* minimum-variance estimate of the filter-coefficients and the secondary path state.

- **Decentralized and distributed control:** For systems with a large number of inputs and outputs (e.g., a few hundreds till thousands in future expected smart materials) the centralized algorithms, as considered in this thesis, are not feasible anymore. For these systems, the complexity of the control problem is too large and decentralized or distributed control methods need to be used. Currently the development of these algorithms is an active research field, and interesting approaches are already investigated in literature. For example the decentralized control approach for smart panels proposed in the series of papers [11, 66, 67], which exploits a passivity relation between (perfectly) collocated sensor/actuator pairs. In this decentralized control method, there is no connection between the controller units. Another approach is distributed control, see, e.g., [5, 31], which exploits (spatial) structure in the sys-

tem and allows some degree of communication between the controller units. Some important research questions are: 1) how large is the loss of performance compared with centralized control design methods, 2) how large is the loss of stability robustness, and 3) how to identify (spatially) structured models for use in distributed controller design?

Practical research recommendations:

1. **Validation of the robust controller estimation method using a large number of conditions:** In the vibrating plate experiments in Section 4.6 on page 93 the robust controller estimation method has been demonstrated under two conditions: without and with additional mass mounted on the plate. It would be interesting to validate the robust controller estimation method for multiple conditions.

2. **Relation between performance and accuracy of nominal secondary path model:** In some cases, the accuracy of the secondary path model is bad, such that no good performance can be obtained by the (estimated) Causal or Cautious Wiener filter (with IMC in feedback applications). However, still the SPR condition for convergence of the FxLMS algorithm is satisfied, such that the FxLMS algorithm converges and better performance can be obtained than by the Causal or Cautious Wiener filter. Further research is necessary to obtain better insight in the relation between the model accuracy and the performance obtained by Causal or Cautious Wiener on one side and the FxLMS algorithm on the other side. For example, how important is the accurate modeling of delays in secondary path system?

3. **Frequency domain implementation of the (robust) estimation method:** The estimation of the nominal and robust controller in Chapter 3 and 4 respectively, have been performed by subspace identification using time domain samples. In most ANVC systems, the acoustical or mechanical system contains many resonance frequencies, such that a huge amount of samples is necessary to model the system. Since the number of data-samples determines the computational complexity, it is worth investigating whether the number of data samples can be reduced using *frequency* domain subspace identification methods [112]. By careful experiment design accurate Frequency Response Functions (FRF's) of systems can be estimated [137], which are used as the input to the frequency domain subspace identification method. The identification of the causal factor $[G^*_{eu,i} G_{es} G^*_{rs,ci}]_+$ of the Causal Wiener filter (see (2.20) on page 36) can be estimated using the frequency domain causal/anti-causal state-space identification method developed in Appendix D, which was published before in [63].

Appendix A

Proof of the Causal Wiener Theorem

Theorem A.1 (Causal Wiener filter) *Given* $G_{es}(q^{-1}) \in \mathcal{RH}_\infty^{m_e \times m_s}$, $G_{eu}(q^{-1}) \in \mathcal{RH}_\infty^{m_e \times m_u}$, $G_{rs}(q^{-1}) \in \mathcal{RH}_\infty^{m_r \times m_s}$. *Assume that* $G_{eu}(q^{-1})$ *and* $G_{rs}(q^{-1})$ *do not loose rank* $\forall |q| = 1$. *Then, let*

$$G_{eu} = G_{eu,i} G_{eu,o} \tag{A.1}$$

$$G_{rs} = G_{rs,co} G_{rs,ci} \tag{A.2}$$

be the inner-outer and outer-inner factorization of G_{eu} *and* G_{rs} *respectively and* $G_{eu,o}^\dagger$ *a right-inverse of* $G_{eu,o}$ *and* $G_{rs,co}^\dagger$ *a left-inverse of* $G_{rs,co}$. *Let* $G_{eu,i}^\perp$ *and* $G_{rs,ci}^\perp$ *be such that* $[G_{eu,i} \ G_{eu,i}^\perp]$ *and* $[G_{rs,ci}^* \ G_{rs,ci}^{\perp *}]^*$ *are unitary. Then*

$$W = -G_{eu,o}^\dagger \left[G_{eu,i}^* G_{es} G_{rs,ci}^* \right]_+ G_{rs,co}^\dagger \tag{A.3}$$

minimizes

$$||G_{es} + G_{eu} \mathcal{W} G_{rs}||_2, \qquad subject \ to \ \mathcal{W} \in \mathcal{RH}_\infty^{m_u \times m_r} \tag{A.4}$$

and its minimum value is given by

$$\begin{aligned} ||G_{es} + G_{eu} W G_{rs}||_2 = \\ \sqrt{||G_{es} G_{rs,ci}^{\perp *}||_2^2 + ||G_{eu,i}^{\perp *} G_{es} G_{rs,ci}^*||_2^2 + ||\left[G_{eu,i}^* G_{es} G_{rs,ci}^* \right]_-||_2^2} \end{aligned} \tag{A.5}$$

Proof: The proof is given by [186, Section 6.2], and the strategy of the proof is by completing the squares. Here, we give the proof in a general and slightly different presentation.

First, note that by the inner-outer and outer-inner factorization lemmas, Lemma 2.1 and 2.2 on page 35, the inner-outer and outer inner factorization of G_{eu} and G_{rs} exists. Furthermore, since G_{eu} and G_{rs} do not loose rank $\forall |q| = 1$, $G_{eu,o}^\dagger$ and $G_{rs,co}^\dagger$ are stable.

Second, it will be usefull to write the (squared) cost function in the frequency domain, as follows

$$J = ||G_{es} + G_{eu}WG_{rs}||_2^2 = \frac{1}{2\pi}\text{tr}\int\limits_{\omega=-\pi}^{\pi}\Big(G_{es} + G_{eu}WG_{rs}\Big)\Big(G_{es} + G_{eu}WG_{rs}\Big)^*d\omega$$

using Parseval, where q is replaced by $e^{j\omega}$. Because $[G_{rs,ci}^* \quad G_{rs,ci}^{\perp*}]^*$ is unitary, we can write

$$G_{es} \;=\; G_{es}G_{rs,ci}^*G_{rs,ci} + G_{es}G_{rs,ci}^{\perp*}G_{rs,ci}^{\perp} \;=\; G_{es,1}G_{rs,ci} + G_{es,2}G_{rs,ci}^{\perp}$$

where $G_{es,1} = G_{es}G_{rs,ci}^*$ and $G_{es,2} = G_{es}G_{rs,ci}^{\perp*}$. Hence, we can write

$$J = \frac{1}{2\pi}\text{tr}\int\limits_{\omega=-\pi}^{\pi}\Big(G_{es,1}G_{rs,ci} + G_{es,2}G_{rs,ci}^{\perp} + G_{eu}WG_{rs,co}G_{rs,ci}\Big)\big(.\big)^*d\omega$$

$$= \frac{1}{2\pi}\text{tr}\int\limits_{\omega=-\pi}^{\pi}G_{es,2}G_{es,2}^*d\omega + \frac{1}{2\pi}\text{tr}\int\limits_{\omega=-\pi}^{\pi}\Big(G_{es,1} + G_{eu}WG_{rs,co}\Big)\big(.\big)^*d\omega$$

$$= \frac{1}{2\pi}\text{tr}\int\limits_{\omega=-\pi}^{\pi}G_{es}G_{rs,ci}^{\perp*}G_{rs,ci}^{\perp}G_{es}^*d\omega + \frac{1}{2\pi}\text{tr}\int\limits_{\omega=-\pi}^{\pi}\Big(G_{es}G_{rs,ci}^* + G_{eu}WG_{rs,co}\Big)\big(.\big)^*d\omega$$

$$= \frac{1}{2\pi}\text{tr}\int\limits_{\omega=-\pi}^{\pi}G_{rs,ci}^{\perp}G_{es}^*G_{es}G_{rs,ci}^{\perp*}d\omega + \frac{1}{2\pi}\text{tr}\int\limits_{\omega=-\pi}^{\pi}\big(.\big)^*\Big(G_{es}G_{rs,ci}^* + G_{es}WG_{rs,co}\Big)d\omega$$

where we made use of the identity $\text{tr}(AB) = \text{tr}(BA)$ in the last step.

Now, because $[G_{eu,i} \quad G_{eu,i}^{\perp}]$ is also unitary, we can write

$$G_{es} \;=\; G_{eu,i}G_{eu,i}^*G_{es} + G_{eu,i}^{\perp}G_{eu,i}^{\perp*}G_{es} \;=\; G_{eu,i}\widetilde{G}_{es,1} + G_{eu,i}^{\perp}\widetilde{G}_{es,2}$$

where $\widetilde{G}_{es,1} = G_{eu,i}^*G_{es}$ and $\widetilde{G}_{es,2} = G_{eu,i}^{\perp*}G_{es}$. Hence, we can write

$$J = \frac{1}{2\pi}\text{tr}\int\limits_{\omega=-\pi}^{\pi}G_{rs,ci}^{\perp}G_{es}^*G_{es}G_{rs,ci}^{\perp*}d\omega+$$

$$+ \frac{1}{2\pi}\text{tr}\int\limits_{\omega=-\pi}^{\pi}\big(.\big)^*\Big(G_{eu,i}\widetilde{G}_{es,1}G_{rs,ci}^* + G_{eu,i}^{\perp}\widetilde{G}_{es,2}G_{rs,ci}^* + G_{eu,i}G_{eu,o}WG_{rs,co}\Big)d\omega$$

$$= \frac{1}{2\pi}\text{tr}\int\limits_{\omega=-\pi}^{\pi}\Big(G_{rs,ci}^{\perp}G_{es}^*G_{es}G_{rs,ci}^{\perp*} + G_{rs,ci}\widetilde{G}_{es,2}^*\widetilde{G}_{es,2}G_{rs,ci}^*\Big)d\omega+$$

$$+ \frac{1}{2\pi}\text{tr}\int\limits_{\omega=-\pi}^{\pi}\big(.\big)^*\Big(\widetilde{G}_{es,1}G_{rs,ci}^* + G_{eu,o}WG_{rs,co}\Big)d\omega$$

$$= \frac{1}{2\pi} \text{tr} \int\limits_{\omega=-\pi}^{\pi} \Big(G_{rs,ci}^{\perp} G_{es}^* G_{es} G_{rs,ci}^{\perp *} + G_{rs,ci} G_{es}^* G_{eu,i}^{\perp} G_{eu,i}^{\perp *} G_{es} G_{rs,ci}^* \Big) d\omega +$$

$$+ \frac{1}{2\pi} \text{tr} \int\limits_{\omega=-\pi}^{\pi} (.)^* \Big(G_{eu,i}^* G_{es} G_{rs,ci}^* + G_{eu,o} W G_{rs,co} \Big) d\omega$$

Because W is constrained to be stable, $G_{eu,o} W G_{rs,co}$ is stable, and thus $G_{eu,o} W G_{rs,co} = [G_{eu,o} W G_{rs,co}]_+, \forall W \in \mathcal{RH}_\infty^{m_u \times m_r}$, we can write

$$J = \frac{1}{2\pi} \text{tr} \int\limits_{\omega=-\pi}^{\pi} \Big(G_{rs,ci}^{\perp} G_{es}^* G_{es} G_{rs,ci}^{\perp *} + G_{rs,ci} G_{es}^* G_{eu,i}^{\perp} G_{eu,i}^{\perp *} G_{es} G_{rs,ci}^* \Big) d\omega +$$

$$+ \frac{1}{2\pi} \text{tr} \int\limits_{\omega=-\pi}^{\pi} (.)^* \Big([G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ + [G_{eu,i}^* G_{es} G_{rs,ci}^*]_- + G_{eu,o} W G_{rs,co} \Big) d\omega$$

$$= \frac{1}{2\pi} \text{tr} \int\limits_{\omega=-\pi}^{\pi} \Big(G_{rs,ci}^{\perp} G_{es}^* G_{es} G_{rs,ci}^{\perp *} + G_{rs,ci} G_{es}^* G_{eu,i}^{\perp} G_{eu,i}^{\perp *} G_{es} G_{rs,ci}^* +$$

$$+ [G_{eu,i}^* G_{es} G_{rs,ci}^*]_-^* [G_{eu,i}^* G_{es} G_{rs,ci}^*]_- \Big) d\omega +$$

$$+ \frac{1}{2\pi} \text{tr} \int\limits_{\omega=-\pi}^{\pi} (.)^* \Big([G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ + G_{eu,o} W G_{rs,co} \Big) d\omega \tag{A.6}$$

where we made use of the fact, that for X, Y transfer function matrices of equal dimensions and $[X]_+ [Y]_-$ has no poles on the unit circle, we have

$$\frac{1}{2\pi} \text{tr} \int\limits_{\omega=-\pi}^{\pi} \Big([X]_+ + [Y]_- \Big)^* \Big([X]_+ + [Y]_- \Big) d\omega = \frac{1}{2\pi} \text{tr} \int\limits_{\omega=-\pi}^{\pi} \Big([X]_+^* [X]_+ + [Y]_-^* [Y]_- \Big) d\omega$$

because $\frac{1}{2\pi} \int\limits_{\omega=-\pi}^{\pi} [X]_+^* [Y]_- d\omega = 0$. This can be seen directly from the Laurent series of $[X]_+^* [Y]_- z^{-1}$ (which exists because $[X]_+ [Y]_-$ has no poles on the unit circle):

$$[X]_+^* [Y]_- z^{-1} = \sum_{i=-\infty}^{\infty} \gamma_i z^i$$

Because $[X]_+$ only consists of a direct feed through term and anti-causal terms and $[Y]_-$ is anti-causal, we have $\gamma_i = 0$ for all $i \leq -1$. Using the residual theorem (see, e.g., [143]), we have

$$\frac{1}{2\pi} \int\limits_{\omega=-\pi}^{\pi} [X]_+^* [Y]_- d\omega = \oint_{|z|=1} [X]_+^* [Y]_- z^{-1} dz = \gamma_{-1} = 0$$

Hence (A.6) is minimized subject to $W \in \mathcal{RH}_\infty^{m_u \times m_r}$ if and only if W satisfies

$$G_{eu,o} W G_{rs,co} = -[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+$$

Because $G_{eu,o}^\dagger$ and $G_{rs,co}^\dagger$ are stable and $G_{eu,o} G_{eu,o}^\dagger = I$ and $G_{rs,co}^\dagger G_{rs,co} = I$, a stable transfer function W which minimizes (A.6) —and thus (A.4)— is given by

$$W = -G_{eu,o}^\dagger [G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ G_{rs,co}^\dagger$$

Filling this W in (A.6) directly yields the minimal value of the cost-function given by (A.5), which completes the proof. $\qquad\square$

Appendix B

The Causal Wiener state-space filter and relation with LQG

B.1 Introduction

The goal of this report is to twofold:

1. to derive a state-space expression of the Causal Wiener filter;

2. to proof that the internal model controller with the Causal Wiener filter is equivalent to the LQG controller.

To achieve these goals, we start in Section B.2 with a brief review on calculus with state-space realizations. We refer to [24, 36, 85, 186] for more details.

B.2 Calculus with state-space realizations

Relation between transfer function and state-space models. Let

$$y(k) \ = \ P(z)u(k)$$

with P given by the following state-space realization

$$P \sim \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right]$$

Then, the input-output relation between $u(k)$ and $y(k)$ is given by

$$x(k+1) = Ax(k) + Bu(k) \qquad (B.1)$$
$$y(k) = Cx(k) + Du(k) \qquad (B.2)$$

and the transfer function matrix of P is

$$P = D + C(zI - A)^{-1}B$$
$$= D + CBz^{-1} + CABz^{-2} + CA^2Bz^{-3} + \cdots$$

with z the unit shift-forward operator in time-domain and a complex-valued number in frequency domain.

The input-output relation (B.1) is given in a causal-form. If A is nonsingular, the same input-output relation can be given in the anti-causal form. Therefore, let us rewrite (B.1) as

$$x(k) = A^{-1}x(k+1) - A^{-1}Bu(k).$$

Now, let us define $\tilde{x}(k) = x(k+1)$, such that we have

$$\tilde{x}(k-1) = A^{-1}\tilde{x}(k) - A^{-1}Bu(k)$$

and (B.2) rewrites as

$$y(k) = C\tilde{x}(k-1) + Du(k)$$
$$= CA^{-1}\tilde{x}(k) + (D - CA^{-1}B)u(k).$$

Finally, writing $\tilde{x}(k)$ as $x(k)$ yields the following *anti-causal* state-space description of (B.1),(B.2)

$$x(k-1) = A^{-1}x(k) - A^{-1}Bu(k)$$
$$y(k) = CA^{-1}x(k) + (D - CA^{-1}B)u(k).$$

For ease of notation, this anti-causal state-space description is written as

$$P \sim \left[\begin{array}{c|c} A^{-1} & -A^{-1}B \\ \hline CA^{-1} & D - CA^{-1}B \end{array} \right]_{ac}$$

And thus, if A is nonsingular P can be written as

$$P = -CA^{-1}(z^{-1}I - A^{-1})^{-1}A^{-1}B + (D - CA^{-1}B)$$
$$= \cdots - CA^{-3}Bz^2 - CA^{-2}Bz + (D - CA^{-1}B)$$

The adjoint operator. Let

$$P \sim \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right]$$

then the adjoint of P is given by

$$P^* = B^T(z^{-1}I - A^T)^{-1}C^T + D^T$$

thus

$$P^* \sim \left[\begin{array}{c|c} A^T & C^T \\ \hline B^T & D^T \end{array} \right]_{ac}$$

If, A is nonsingular, then a causal state-space realization of P^* is given by

$$P^* \sim \left[\begin{array}{c|c} A^{-T} & -A^{-T}C^T \\ \hline B^TA^{-T} & D^T - B^TA^{-T}C^T \end{array} \right]$$

Similarity transformation. Let

$$P \sim \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right]$$

$$P' \sim \left[\begin{array}{c|c} TAT^{-1} & TB \\ \hline CT^{-1} & D \end{array} \right]$$

with T any nonsingular matrix, then

$$P' = P$$

and P' is said to be equal to P up to a similarity transformation T.

Parallel and series connections. Let

$$P_1 \sim \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array} \right]$$

$$P_2 \sim \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 & D_2 \end{array} \right]$$

then

$$P_1 P_2 \sim \left[\begin{array}{cc|c} A_2 & 0 & B_2 \\ B_1 C_2 & A_1 & B_1 D_2 \\ \hline D_1 C_2 & C_1 & D_1 D_2 \end{array} \right]$$

$$P_1 + P_2 \sim \left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \\ \hline C_1 & C_2 & D_1 + D_2 \end{array} \right]$$

Linear Fractional Transformation — feedback connection. Let

$$P \sim \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & 0 \end{array} \right]$$

$$C \sim \left[\begin{array}{c|c} A_c & B_c \\ \hline C_c & D_c \end{array} \right]$$

then the lower LFT (linear fractional transformation) is given by

$$P_{11} + P_{12}(I - CP_{22})^{-1} CP_{21} \sim \left[\begin{array}{cc|c} A + B_2 D_c C_2 & B_2 C_c & B_1 + B_2 D_c D_{21} \\ B_c C_2 & A_c & B_c D_{21} \\ \hline C_1 + D_{12} D_c C_2 & D_{12} C_c & D_{11} + D_{12} D_c D_{21} \end{array} \right]$$

which is the feedback connection of P and C as depicted in Figure B.1.
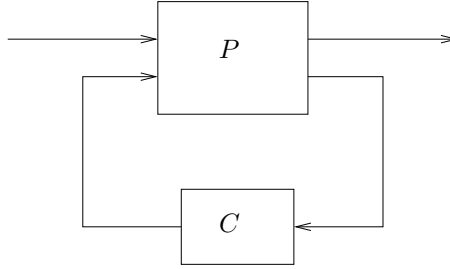
Figure B.1: Lower LFT connection of P and C.

(Pseudo) inverse model.   Let

$$P \sim \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right]$$

and

$$P^\dagger \sim \left[\begin{array}{c|c} A - BD^\dagger C & BD^\dagger \\ \hline -D^\dagger C & D^\dagger \end{array} \right]$$

then

- if D is nonsingular, $D^\dagger = D^{-1}$ and $P^\dagger = P^{-1}$ is the inverse of P, i.e. $PP^{-1} = P^{-1}P = I$;

- if $D^T D$ is nonsingular, $D^\dagger = (D^T D)^{-1} D^T$ and $P^\dagger$ is a left-inverse of P, i.e. $P^\dagger P = I$;

- if DD^T is nonsingular, $D^\dagger = D^T (DD^T)^{-1}$ and $P^\dagger$ is a right-inverse of P, i.e. $PP^\dagger = I$.

The (anti-)causality operator.   Let P_1 and P_2 be strictly stable, i.e. A_1 and A_2 has all eigenvalues inside the unit-disk. Then

$$P_1 P_2^* = [P_1 P_2^*]_+ + [P_1 P_2^*]_-$$

with

$$[P_1 P_2^*]_+ \sim \left[\begin{array}{c|c} A_1 & B_1 D_2^T + A_1 X_{12} C_2^T \\ \hline C_1 & D_1 D_2^T + C_1 X_{12} C_2^T \end{array} \right]$$

$$[P_1 P_2^*]_- \sim \left[\begin{array}{c|c} A_2^T & C_2^T \\ \hline D_1 B_2^T + C_1 X_{12} A_2^T & 0 \end{array} \right]_{ac}$$

with X_{12} the solution of the Sylvester equation

$$X_{12} - A_1 X_{12} A_2^T = B_1 B_2^T$$

Analogue, the following holds too

$$P_2^* P_1 = [P_2^* P_1]_+ + [P_2^* P_1]_-$$

with

$$[P_2^* P_1]_+ \sim \left[\begin{array}{c|c} A_1 & B_1 \\ \hline D_2^T C_1 + B_2^T X_{21} A_1 & D_2^T D_1 + B_2^T X_{21} B_1 \end{array} \right]$$

$$[P_2^* P_1]_- \sim \left[\begin{array}{c|c} A_2^T & C_2^T D_1 + A_2^T X_{21} B_1 \\ \hline B_2^T & 0 \end{array} \right]_{ac}$$

with X_{21} the solution of the Sylvester equation

$$X_{21} - A_2^T X_{21} A_1 = C_2^T C_1$$

These results are obtained by a partial fraction expansion of $P_1 P_2^*$ and $P_2^* P_1$ respectively.

Inner-outer (outer-inner) factorization. The inner-outer factorization of P is given by

$$P = P_i P_o$$

with $P_i^* P_i = I$ and P_o has a stable right inverse. Let $X = X^T > 0$ be the stabilizing solution to the Riccati equation

$$X = A^T X A - (A^T X B + C^T D)(B^T X B + D^T D)^{-1}(A^T X B + C^T D) + C^T C$$

and Y be a square-root of X, such that $Y^T Y = X$. Let the QR factorization be given

$$\left[\begin{array}{cc} D & C \\ YB & YA \end{array} \right] = \left[\begin{array}{cc} D^i & C^i \\ B^i & A^i \end{array} \right] \left[\begin{array}{cc} D^o & C^o \\ 0 & Y \end{array} \right]$$

Then

$$P_i \sim \left[\begin{array}{c|c} A^i & B^i \\ \hline C^i & D^i \end{array} \right]$$

$$P_o \sim \left[\begin{array}{c|c} A & B \\ \hline C^o & D^o \end{array} \right]$$

Analogue to the inner-outer factorization, the outer-inner factorization of P is given by

$$P = P_o P_i$$

with P_o has a stable left inverse and $P_i P_i^* = I$. Let $X = X^T > 0$ be the stabilizing solution to the Riccati equation

$$X = A X A^T - (A X C^T + B D^T)(C X C^T + D D^T)^{-1}(A X C^T + B D^T)^T + B B^T$$

and Y be a square-root of X, such that $YY^T = X$. Let the RQ-factorization be given

$$\left[\begin{array}{cc} D & CY \\ B & AY \end{array} \right] = \left[\begin{array}{cc} D^o & 0 \\ B^o & Y \end{array} \right] \left[\begin{array}{cc} D^i & C^i \\ B^i & A^i \end{array} \right]$$

Then

$$P_o \sim \left[\begin{array}{c|c} A & B^o \\ \hline C & D^o \end{array} \right]$$

$$P_i \sim \left[\begin{array}{c|c} A^i & B^i \\ \hline C^i & D^i \end{array} \right]$$

B.3   State-space realization of the Causal Wiener filter

Using these results, a state-space realization of the Causal Wiener filter can be derived.

The plant is as in Chapter 2 and given by

$$G(q^{-1}) = \left[\begin{array}{cc} G_{es}(q^{-1}) & G_{eu}(q^{-1}) \\ G_{rs}(q^{-1}) & G_{ru}(q^{-1}) \end{array} \right] \sim \left[\begin{array}{c|cc} A & B_s & B_u \\ \hline C_e & D_{es} & D_{eu} \\ C_r & D_{rs} & 0 \end{array} \right] \tag{B.3}$$

with G_{es} the transfer from the disturbance input to the performance output (primary path), G_{eu} the transfer from the control input to the performance output (secondary path), G_{rs} the transfer from the disturbance input to the measured output (detector path) and G_{ru} the transfer from the control input to the measured output (feedback path). The assumption that $D_{ru} = 0$ is standard and satisfied in almost all application due to discretization delay.

It has been shown in Appendix A that the Causal Wiener filter is given by

$$W = -G_{eu,o}^{\dagger}[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ G_{rs,co}^{\dagger} \tag{B.4}$$

with $(.)^*$ and $[.]_+$ are the complex conjugate transpose and the causality operator respectively. It has been shown by means of the Youla parameterization that the IMC controller

$$C = (I + WG_{eu})^{-1}W$$

with W the Causal Wiener filter optimally minimizes the H_2-norm of the LFT closed-loop connection between G and C.

First, we will derive the state-space realizations for the inner and outer factors. Then, we will derive the state-space realization of the Causal Wiener filter W and the IMC controller

$$C = (I + WG_{22})^{-1}W$$

The following section, shows that this controller is equivalent to the LQG controller.

B.3.1   IO-factorization of G_{eu}

Let IO-factorization of G_{eu} be given by

$$G_{eu} = G_{eu,i}G_{eu,o}$$

with

$$G_{eu,i} \sim \left[\begin{array}{c|c} A_{eu}^i & B_{eu}^i \\ \hline C_{eu}^i & D_{eu}^i \end{array} \right]$$

$$G_{eu,o} \sim \left[\begin{array}{c|c} A & B_u \\ \hline C_{eu}^o & D_{eu}^o \end{array} \right]$$

Then $A_{eu}^i, B_{eu}^i, C_{eu}^i, D_{eu}^i, C_{eu}^o, D_{eu}^o$ can be calculated as follows. Let $P_{eu} = P_{eu}^T \geq 0$ be a solution of the discrete algebraic Riccati equation

$$
\begin{aligned}
P_{eu} = A^T P_{eu} A - (A^T P_{eu} B_u + C_e^T D_{eu})(B_u^T P_{eu} B_u + D_{eu}^T D_{eu})^{-1} \cdot \\
(A^T P_{eu} B_u + C_e^T D_{eu})^T + C_e^T C_e
\end{aligned}
\tag{B.5}
$$

Further let Y_{eu} be an upper triangular matrix such that

$$Y_{eu}^T Y_{eu} = P_{eu} \tag{B.6}$$

Then, the following QR-decomposition can be calculated from which the matrices to be calculated can be read off.

$$
\left[\begin{array}{cc} D_{eu} & C_e \\ Y_{eu}B_u & Y_{eu}A \end{array} \right] = \left[\begin{array}{cc} D_{eu}^i & C_{eu}^i \\ B_{eu}^i & A_{eu}^i \end{array} \right] \left[\begin{array}{cc} D_{eu}^o & C_{eu}^o \\ 0 & Y_{eu} \end{array} \right]
\tag{B.7}
$$

Because $\left[\begin{array}{cc} D_{eu}^i & C_{eu}^i \\ B_{eu}^i & A_{eu}^i \end{array} \right]$ is orthogonal, we have the following relations

$$D_{eu}^{iT} D_{eu} + B_{eu}^{iT} Y_{eu} B_u = D_{eu}^o \tag{B.8}$$

$$D_{eu}^{iT} C_e + B_{eu}^{iT} Y_{eu} A = C_{eu}^o \tag{B.9}$$

$$C_{eu}^{iT} D_{eu} + A_{eu}^{iT} Y_{eu} B_u = 0 \tag{B.10}$$

$$C_{eu}^{iT} C_e + A_{eu}^{iT} Y_{eu} A = Y_{eu} \tag{B.11}$$

and by 'squaring' (B.7) we get

$$
\left[\begin{array}{cc} D_{eu}^T D_{eu} + B_u^T Y_{eu}^T Y_{eu} B_u & D_{eu}^T C_e + B_u^T Y_{eu}^T Y_{eu} A \\ C_e^T D_{eu} + A^T Y_{eu}^T Y_{eu} B_u & C_e^T C_e + A^T Y_{eu}^T Y_{eu} A \end{array} \right] =
$$

$$
= \left[\begin{array}{cc} D_{eu}^{oT} D_{eu}^o & D_{eu}^{oT} C_{eu}^o \\ C_{eu}^{oT} D_{eu}^o & C_{eu}^{oT} C_{eu}^o + Y_{eu}^T Y_{eu} \end{array} \right]
\tag{B.12}
$$

B.3.2   OI-factorization of G_{rs}

Let OI-factorization of G_{rs} be given by

$$G_{rs} = G_{rs,co}G_{rs,ci}$$

with

$$G_{rs,ci} \sim \left[\begin{array}{c|c} A^i_{rs} & B^i_{rs} \\ \hline C^i_{rs} & D^i_{rs} \end{array} \right]$$

$$G_{rs,co} \sim \left[\begin{array}{c|c} A & B^o_{rs} \\ \hline C_r & D^o_{rs} \end{array} \right]$$

Then $A^i_{rs}, B^i_{rs}, C^i_{rs}, D^i_{rs}, B^o_{rs}, D^o_{rs}$ can be calculated as follows. Let $P_{rs} = P^T_{rs} \geq 0$ be a solution of the discrete algebraic Riccati equation

$$P_{rs} = AP_{rs}A^T - (AP_{rs}C^T_r + B_sD^T_{rs})(C_rP_{rs}C^T_r + D_{rs}D^T_{rs})^{-1}(AP_{rs}C^T_r + B_sD^T_{rs})^T + \\ + B_sB^T_s \tag{B.13}$$

Further let Y_{rs} be a lower triangular matrix such that

$$Y_{rs}Y^T_{rs} = P_{rs} \tag{B.14}$$

Then, the following LQ-decomposition can be calculated from which the matrices to be calculated can be read off.

$$\left[\begin{array}{cc} D_{rs} & C_rY_{rs} \\ B_s & AY_{rs} \end{array} \right] = \left[\begin{array}{cc} D^o_{rs} & 0 \\ B^o_{rs} & Y_{rs} \end{array} \right] \left[\begin{array}{cc} D^i_{rs} & C^i_{rs} \\ B^i_{rs} & A^i_{rs} \end{array} \right] \tag{B.15}$$

Because $\left[\begin{array}{cc} D^i_{rs} & C^i_{rs} \\ B^i_{rs} & A^i_{rs} \end{array} \right]$ is orthogonal, we have the following relations

$$D_{rs}D^{iT}_{rs} + C_rY_{rs}C^{iT}_{rs} = D^o_{rs} \tag{B.16}$$

$$D_{rs}B^{iT}_{rs} + C_rY_{rs}A^{iT}_{rs} = 0 \tag{B.17}$$

$$B_sD^{iT}_{rs} + AY_{rs}C^{iT}_{rs} = B^o_{rs} \tag{B.18}$$

$$B_sB^{iT}_{rs} + AY_{rs}A^{iT}_{rs} = Y_{rs} \tag{B.19}$$

and by 'squaring' (B.15) we get

$$\left[\begin{array}{cc} D_{rs}D^T_{rs} + C_rY_{rs}Y^T_{rs}C^T_r & D_{rs}B^T_s + C_rY_{rs}Y^T_{rs}A^T \\ B_sD^T_{rs} + AY_{rs}Y^T_{rs}C^T_r & B_sB^T_s + AY_{rs}Y^T_{rs}A^T \end{array} \right] = \\ = \left[\begin{array}{cc} D^o_{rs}D^{oT}_{rs} & D^o_{rs}B^{oT}_{rs} \\ B^o_{rs}D^{oT}_{rs} & B^o_{rs}B^{oT}_{rs} + Y_{rs}Y^T_{rs} \end{array} \right] \tag{B.20}$$

B.3.3 Causal factor $[G^*_{eu,i}G_{es}G^*_{rs,ci}]_+$

First note, that the following holds

$$G^*_{eu,i}G_{es}G^*_{rs,ci} = [G^*_{eu,i}G_{es}]_+ G^*_{rs,ci} + [G^*_{eu,i}G_{es}]_- G^*_{rs,ci}$$

and because $G^*_{rs,ci}$ is anti-causal only up to a direct feedthrough term, we have

$$[G^*_{eu,i}G_{es}G^*_{rs,ci}]_+ = [[G^*_{eu,i}G_{es}]_+ G^*_{rs,ci}]_+$$

Therefore, we first calculate $[G^*_{eu,i}G_{es}]_+$.

It can be shown that

$$[G^{i*}_{eu}G_{es}]_+ \sim \left[\begin{array}{c|c} A & B_s \\ \hline D^{iT}_{eu}C_e + B^{iT}_{eu}X_{eu}A & D^{iT}_{eu}D_{es} + B^{iT}_{eu}X_{eu}B_s \end{array}\right]$$

with X_{eu} such that the following Lyapunov equation holds

$$C^{iT}_{eu}C_e + A^{iT}_{eu}X_{eu}A = X_{eu}$$

Note, that this expression is the same as for Y_{eu} in (B.11), such that we have

$$X_{eu} = Y_{eu}$$

Using this result and (B.9), we have

$$[G^*_{eu,i}G_{es}]_+ \sim \left[\begin{array}{c|c} A & B_s \\ \hline C^o_{eu} & D^{iT}_{eu}D_{es} + B^{iT}_{eu}Y_{eu}B_s \end{array}\right]$$

Now, we are ready to calculate $[G^*_{eu,i}G_{es}G^*_{rs,ci}]_+$. It can be shown that

$$[[G^*_{eu,i}G_{es}]_+ G^*_{rs,ci}]_+ \sim \left[\begin{array}{c|c} A & B_s D^{iT}_{rs} + AX_{rs}C^{iT}_{rs} \\ \hline C^o_{eu} & (D^{iT}_{eu}D_{es} + B^{iT}_{eu}Y_{eu}B_1)D^{iT}_{rs} + C^o_{eu}X_{rs}C^{iT}_{rs} \end{array}\right]$$

with X_{rs} such that the following Sylvester equation holds

$$B_1 B^{iT}_{rs} + AX_{rs}A^{iT}_{rs} = X_{rs}$$

Note, that this expression is the same as for Y_{rs} in (B.19), such that we have

$$X_{rs} = Y_{rs}$$

Using this result and (B.18), we have

$$[G^*_{eu,i}G_{es}G^*_{rs,ci}]_+ \sim \left[\begin{array}{c|c} A & B^o_{rs} \\ \hline C^o_{eu} & D^o_{es} \end{array}\right]$$

with

$$D^o_{es} = D^{iT}_{eu}D_{es}D^{iT}_{rs} + B^{iT}_{eu}Y_{eu}B_s D^{iT}_{rs} + C^o_{eu}Y_{rs}C^{iT}_{rs} \tag{B.21}$$

B.3.4   Calculation of $G_{eu,o}^{\dagger}[G_{eu,i}^{*}G_{es}G_{rs,ci}^{*}]_{+}$

First, note that the right inverse of $G_{eu,o}^{\dagger}$ is given by

$$
G_{eu,o}^{\dagger} \sim \left[\begin{array}{c|c} A - B_u D_{eu}^{o\dagger} C_{eu} & B_u D_{eu}^{o\dagger} \\ \hline -D_{eu}^{o\dagger} C_{eu} & D_{eu}^{o\dagger} \end{array}\right]
$$

with $D_{eu}^{o\dagger}$ a right inverse of D_{eu}^{o}.

Then using the rule for series-connection, we obtain

$$
G_{eu,o}^{\dagger}[G_{eu,i}^{*}G_{es}G_{rs,ci}^{*}]_{+} \sim \left[\begin{array}{cc|c} A - B_u D_{eu}^{o\dagger} C_{eu}^{o} & B_u D_{eu}^{o\dagger} C_{eu}^{o} & B_u D_{eu}^{o\dagger} D_{es}^{o} \\ 0 & A & B_{rs}^{o} \\ \hline -D_{eu}^{o\dagger} C_{eu}^{o} & D_{eu}^{o\dagger} C_{eu}^{o} & D_{eu}^{o\dagger} D_{es}^{o} \end{array}\right]
$$

By evaluating a similarity transformation with $T = \left[\begin{array}{cc} I & -I \\ 0 & I \end{array}\right]$, we obtain

$$
G_{eu}^{o\dagger}[G_{eu,i}^{*}G_{es}G_{rs,ci}^{*}]_{+} \sim \left[\begin{array}{cc|c} A - B_u D_{eu}^{o\dagger} C_{eu}^{o} & 0 & B_u D_{eu}^{o\dagger} D_{es}^{o} - B_{rs}^{o} \\ 0 & A & B_{rs}^{o} \\ \hline -D_{eu}^{o\dagger} C_{eu}^{o} & 0 & D_{eu}^{o\dagger} D_{es}^{o} \end{array}\right]
$$

From which is clear, that the second part of the state is unobservable. Hence, we may write

$$
G_{eu}^{o\dagger}[G_{eu,i}^{*}G_{es}G_{rs,ci}^{*}]_{+} \sim \left[\begin{array}{c|c} A - B_u D_{eu}^{o\dagger} C_{eu}^{o} & B_u D_{eu}^{o\dagger} D_{es}^{o} - B_{rs}^{o} \\ \hline -D_{eu}^{o\dagger} C_{eu}^{o} & D_{eu}^{o\dagger} D_{es}^{o} \end{array}\right]
$$

B.3.5   Calculation of $G_{eu,o}^{\dagger}[G_{eu,i}^{*}G_{es}G_{rs,ci}^{*}]_{+}G_{rs}^{o\dagger}$

First, note that the left inverse of $G_{rs,co}^{\dagger}$ is given by

$$
G_{rs,co}^{\dagger} \sim \left[\begin{array}{c|c} A - B_{rs}^{o} D_{rs}^{o\dagger} C_r & B_{rs}^{o} D_{rs}^{o\dagger} \\ \hline -D_{rs}^{o\dagger} C_r & D_{rs}^{o\dagger} \end{array}\right]
$$

with $D_{rs}^{o\dagger}$ a left inverse of D_{rs}^{o}.

Using the rule for series connection, we obtain

$$
G_{eu}^{o\dagger}[G_{eu,i}^{*}G_{es}G_{rs,ci}^{*}]_{+}G_{rs,co}^{\dagger} \sim
$$

$$
\left[\begin{array}{cc|c} A - B_u D_{eu}^{o\dagger} C_{eu}^{o} & -(B_u D_{eu}^{o\dagger} D_{es}^{o} - B_{rs}^{o}) D_{rs}^{o\dagger} C_r & (B_u D_{eu}^{o\dagger} D_{es}^{o} - B_{rs}^{o}) D_{rs}^{o\dagger} \\ 0 & A - B_{rs}^{o} D_{rs}^{o\dagger} C_r & B_{rs}^{o} D_{rs}^{o\dagger} \\ \hline -D_{eu}^{o\dagger} C_{eu}^{o} & -D_{eu}^{o\dagger} D_{es}^{o} D_{rs}^{o\dagger} C_r & D_{eu}^{o\dagger} D_{es}^{o} D_{rs}^{o\dagger} \end{array}\right]
$$

Theorem B.1 *A state-space realization of the Causal Wiener filter given by (B.4) is*

$$
W \sim \left[\begin{array}{cc|c} A - B_u D_{eu}^{o\dagger} C_{eu}^{o} & -(B_u D_{eu}^{o\dagger} D_{es}^{o} - B_{rs}^{o}) D_{rs}^{o\dagger} C_r & (B_u D_{eu}^{o\dagger} D_{es}^{o} - B_{rs}^{o}) D_{rs}^{o\dagger} \\ 0 & A - B_{rs}^{o} D_{rs}^{o\dagger} C_r & B_{rs}^{o} D_{rs}^{o\dagger} \\ \hline D_{eu}^{o\dagger} C_{eu}^{o} & D_{eu}^{o\dagger} D_{es}^{o} D_{rs}^{o\dagger} C_r & -D_{eu}^{o\dagger} D_{es}^{o} D_{rs}^{o\dagger} \end{array}\right]
$$

$$(B.22)$$

with C_{eu}^o, D_{eu}^o given by (B.7), B_{rs}^o, D_{rs}^o given by (B.15) and D_{es}^o given by (B.21).

Proof: The proof is obtained by following the reasoning in this section. $\qquad\square$

Note, that the Causal Wiener filter is of order 2 times the system order.

B.3.6 Feedback connection of $G_{eu,o}^\dagger[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ G_{rs,co}^\dagger$ and G_{ru}

First, note that

$$G_{ru} \sim \left[\begin{array}{c|c} A & B_u \\ \hline C_r & 0 \end{array}\right]$$

Then using, the rule for a negative feedback connection of $W = -G_{eu,o}^\dagger[G_{eu,i}^* G_{es} G_{rs,ci}^*]_+ G_{rs,co}^\dagger$ with G_{ru}, we obtain

$$C = (I + WG_{ru})^{-1}W \sim$$

$$\left[\begin{array}{ccc|}
A - B_u D_{eu}^{o\dagger} C_{eu}^o & -(B_u D_{eu}^{o\dagger} D_{es}^o - B_{rs}^o)D_{rs}^{o\dagger}C_r & (B_u D_{eu}^{o\dagger} D_{es}^o - B_{rs}^o)D_{rs}^{o\dagger}C_r \\
0 & A - B_{rs}^o D_{rs}^{o\dagger}C_r & B_{rs}^o D_{rs}^{o\dagger}C_r \\
-B_u D_{eu}^{o\dagger} C_{eu}^o & -B_u D_{eu}^{o\dagger} D_{es}^o D_{rs}^{o\dagger}C_r & A + B_u D_{eu}^{o\dagger} D_{es}^o D_{rs}^{o\dagger}C_r \\
\hline
-D_{eu}^{o\dagger} C_{eu}^o & -D_{eu}^{o\dagger} D_{es}^o D_{rs}^{o\dagger}C_r & D_{eu}^{o\dagger} D_{es}^o D_{rs}^{o\dagger}C_r
\end{array}\right.$$

$$\left.\begin{array}{c}
-(B_u D_{eu}^{o\dagger} D_{es}^o - B_{rs}^o)D_{rs}^{o\dagger} \\
-B_{rs}^o D_{rs}^{o\dagger} \\
-B_u D_{eu}^{o\dagger} D_{es}^o D_{rs}^{o\dagger} \\
\hline
-D_{eu}^{o\dagger} D_{es}^o D_{rs}^{o\dagger}
\end{array}\right]$$

By performing a similarity transformation by $T = \left[\begin{array}{ccc} I & & \\ & I & -I \\ & & I \end{array}\right]$, we obtain

$$C = (I + WG_{ru})^{-1}W \sim$$

$$\left[\begin{array}{ccc|c}
A - B_u D_{eu}^{o\dagger} C_{eu}^o & -(B_u D_{eu}^{o\dagger} D_{es}^o - B_{rs}^o)D_{rs}^{o\dagger}C_r & 0 & -(B_u D_{eu}^{o\dagger} D_{es}^o - B_{rs}^o)D_{rs}^{o\dagger} \\
B_u D_{eu}^{o\dagger} C_{eu}^o & A + (B_u D_{eu}^{o\dagger} D_{es}^o - B_{rs}^o)D_{rs}^{o\dagger}C_r & 0 & (B_u D_{eu}^{o\dagger} D_{es}^o - B_{rs}^o)D_{rs}^{o\dagger} \\
-B_u D_{eu}^{o\dagger} C_{eu}^o & -B_u D_{eu}^{o\dagger} D_{es}^o D_{rs}^{o\dagger}C_r & A & -B_u D_{eu}^{o\dagger} D_{es}^o D_{rs}^{o\dagger} \\
\hline
-D_{eu}^{o\dagger} C_{eu}^o & -D_{eu}^{o\dagger} D_{es}^o D_{rs}^{o\dagger}C_r & 0 & -D_{eu}^{o\dagger} D_{es}^o D_{rs}^{o\dagger}
\end{array}\right]$$

from which is clear, that the last part of the state is unobservable. Hence, we may write

$$C = (I + WG_{ru})^{-1}W \sim$$

$$\left[\begin{array}{cc|c}
A - B_u D_{eu}^{o\dagger} C_{eu}^o & -(B_u D_{eu}^{o\dagger} D_{es}^o - B_{rs}^o)D_{rs}^{o\dagger}C_r & -(B_u D_{eu}^{o\dagger} D_{es}^o - B_{rs}^o)D_{rs}^{o\dagger} \\
B_u D_{eu}^{o\dagger} C_{eu}^o & A + (B_u D_{eu}^{o\dagger} D_{es}^o - B_{rs}^o)D_{rs}^{o\dagger}C_r & (B_u D_{eu}^{o\dagger} D_{es}^o - B_{rs}^o)D_{rs}^{o\dagger} \\
\hline
-D_{eu}^{o\dagger} C_{eu}^o & -D_{eu}^{o\dagger} D_{es}^o D_{rs}^{o\dagger}C_r & -D_{eu}^{o\dagger} D_{es}^o D_{rs}^{o\dagger}
\end{array}\right]$$

Again, perform a similarity transformation with $T = \begin{bmatrix} I & I \\ 0 & I \end{bmatrix}$, which yields

$$C = (I + WG_{ru})^{-1}W \sim$$

$$\left[\begin{array}{cc|c} A & 0 & 0 \\ B_u D_{eu}^{o\dagger} C_{eu}^o & A + (B_u D_{eu}^{o\dagger} D_{es}^o - B_{rs}^o) D_{rs}^{o\dagger} C_r - B_u D_{eu}^{o\dagger} C_{eu}^o & (B_u D_{eu}^{o\dagger} D_{es}^o - B_{rs}^o) D_{rs}^{o\dagger} \\ \hline -D_{eu}^{o\dagger} C_{eu}^o & D_{eu}^{o\dagger}(C_{eu}^o - D_{es}^o D_{rs}^{o\dagger} C_r) & -D_{eu}^{o\dagger} D_{es}^o D_{rs}^{o\dagger} \end{array} \right]$$

from which is clear that the first part of the state is uncontrollable (but stable). Hence, we may write

$$C = (I + WG_{ru})^{-1}W \sim$$

$$\left[\begin{array}{c|c} A + (B_u D_{eu}^{o\dagger} D_{es}^o - B_{rs}^o) D_{rs}^{o\dagger} C_r - B_u D_{eu}^{o\dagger} C_{eu}^o & (B_u D_{eu}^{o\dagger} D_{es}^o - B_{rs}^o) D_{rs}^{o\dagger} \\ \hline D_{eu}^{o\dagger}(C_{eu}^o - D_{es}^o D_{rs}^{o\dagger} C_r) & -D_{eu}^{o\dagger} D_{es}^o D_{rs}^{o\dagger} \end{array} \right] \quad \text{(B.23)}$$

B.4   Relation between Causal Wiener and LQG

Recall from Theorem 2.2 on page 47 that, given the assumptions in the theorem, the LQG controller is given the following relations:

Let $P_{rs} = P_{rs}^T > 0$ and $P_{eu} = P_{eu}^T > 0$ be the stabilizing solutions to the following Riccati equations

$$P_{rs} = AP_{rs}A^T - (AP_{rs}C_r^T + S_{rs})(C_r P_{rs} C_r^T + R_{rs})^{-1}(AP_{rs}C_r^T + S_{rs})^T + Q_s, \text{(B.24)}$$
$$P_{eu} = A^T P_{eu} A - (A^T P_{eu} B_u + S_{eu})(B_u^T P_{eu} B_u + R_{eu})^{-1}(A^T P_{eu} B_u + S_{eu})^T + Q_e \text{(B.25)}$$

Let

$$F_{eu} = (B_u^T P_{eu} B_u + R_{eu})^{-1}(B_u^T P_{eu} A + S_{eu}^T), \quad \text{(B.26)}$$
$$F_{eu}^o = (B_u^T P_{eu} B_u + R_{eu})^{-1}(B_u^T P_{eu} B_s + D_{eu}^T D_{es}) \quad \text{(B.27)}$$

and

$$K_{rs} = (AP_{rs}C_r^T + S_{rs})(C_r P_{rs} C_r^T + R_{rs})^{-1}, \quad \text{(B.28)}$$
$$K_{rs}^o = (F_{eu} P_{rs} C_r^T + F_{eu}^o D_{rs}^T)(C_r P_{rs} C_r^T + R_{rs})^{-1}, \quad \text{(B.29)}$$

Then, the LQG controller is given by

$$\left[\begin{array}{c} \widehat{x}(k+1|k) \\ \hline u(k) \end{array} \right] = \left[\begin{array}{c|c} A + B_u K_{rs}^o C_r - B_u F_{eu} - K_{rs} C_r & B_u K_{rs}^o - K_{rs} \\ \hline F_{eu} - K_{rs}^o C_r & -K_{rs}^o \end{array} \right]$$

$$\cdot \left[\begin{array}{c} \widehat{x}(k|k-1) \\ \hline y(k) \end{array} \right] \quad \text{(B.30)}$$

Comparing the state-space matrices in (B.30) with (B.23), we conclude that if the following equalities hold

$$F_{eu} = D_{eu}^{o\dagger} C_{eu}^o \quad \text{(B.31)}$$
$$K_{rs} = B_{rs}^o D_{rs}^{o\dagger} \quad \text{(B.32)}$$
$$K_{rs}^o = D_{eu}^{o\dagger} D_{es}^o D_{rs}^{o\dagger} \quad \text{(B.33)}$$

the LQG controller given by (B.30) is equivalent to the IMC controller with the Causal Wiener filter, given by (B.23).

First, note that the Riccati equations (B.24) and (B.25) are the same as the Riccati equations (B.5) and (B.13) respectively, to compute the inner-outer factorization of G_{eu} and outer-inner factorization of G_{rs}.

Since, $P_{eu} = Y_{eu}^T Y_{eu}$ we can write

$$F_{eu} = (B_u^T P_{eu} B_u + R_{eu})^{-1} (B_u^T P_{eu} A + S_{eu}^T)$$
$$= (B_u^T Y_{eu}^T Y_{eu} B_u + D_{eu}^T D_{eu})^{-1} (B_u^T Y_{eu}^T Y_{eu} A + D_{eu}^T C_e)$$

Using the relations in the (1,1) and (1,2) block of (B.12) we obtain

$$F_{eu} = (D_{eu}^{oT} D_{eu}^o)^{-1} D_{eu}^{oT} C_{eu}^o$$
$$= D_{eu}^{o\dagger} C_{eu}^o$$

Similarly, since $P_{rs} = Y_{rs} Y_{rs}^T$ and using the relations in the (1,1) and (2,1) blocks of (B.20) we obtain

$$K_{rs} = (A P_{rs} C_r^T + S_{rs})(C_r P_{rs} C_r^T + R_{rs})^{-1},$$
$$= (A Y_{rs} Y_{rs}^T C_r^T + B_s D_{rs}^T)(C_r Y_{rs} Y_{rs}^T C_r^T + D_{rs} D_{rs}^T)^{-1},$$
$$= B_{rs}^o D_{rs}^{oT} (D_{rs}^o D_{rs}^{oT})^{-1},$$
$$= B_{rs}^o D_{rs}^{o\dagger}.$$

Since K_{rs}^o is depending on F_{eu}^o, let us first rewrite F_{eu}^o as follows

$$F_{eu}^o = (B_u^T P_{eu} B_u^T + R_{eu})^{-1} (B_u^T P_{eu} B_s + D_{eu}^T D_{es}),$$
$$= (D_{eu}^{oT} D_{eu}^o)^{-1} (B_u^T Y_{eu}^T Y_{eu} B_s + D_{eu}^T D_{es}).$$

Then, filling in the expressions for F_{eu} and F_{eu}^o we obtain for K_{rs}^o

$$K_{rs}^o = (F_{eu} P_{rs} C_r^T + F_{eu}^o D_{rs}^T)(C_r P_{rs} C_r^T + R_{rs})^{-1},$$
$$= (D_{eu}^{o\dagger} C_{eu}^o Y_{rs} Y_{rs}^T C_r^T + (D_{eu}^{oT} D_{eu}^o)^{-1} (B_u^T Y_{eu}^T Y_{eu} B_s + D_{eu}^T D_{es}) D_{rs}^T)(D_{rs}^o D_{rs}^{oT})^{-1}$$

Using the relations of D_{eu} and $Y_{eu} B_u$ given by (B.7) we can write

$$K_{rs}^o = D_{eu}^{o\dagger}(C_{eu}^o Y_{rs} Y_{rs}^T C_r^T (D_{rs}^o D_{rs}^{oT})^{-1} + D_{eu}^{iT} D_{es} D_{rs}^T (D_{rs}^o D_{rs}^{oT})^{-1} +$$
$$+ B_{eu}^{iT} Y_{eu} B_s D_{rs}^T (D_{rs}^o D_{rs}^{oT})^{-1})$$

Furthermore, using the relations of D_{rs} and $C_r Y_{rs}$ given by (B.15) we can write

$$K_{rs}^o = D_{eu}^{o\dagger}(C_{eu}^o Y_{rs} C_{rs}^{iT} + D_{eu}^{iT} D_{es} D_{rs}^{iT} + B_{eu}^{iT} Y_{eu} B_s D_{rs}^{iT}) D_{rs}^{o\dagger}$$
$$= D_{eu}^{o\dagger} D_{es}^o D_{rs}^{o\dagger}$$

where the latter equality follows from (B.21). Hence, we may conclude, that indeed (B.31)–(B.33) hold and thus the state-space expression of the LQG controller (B.30) is equivalent to (B.23).

B.5   Conclusions

A state-space representation of the Causal Wiener filter was derived, which is of order 2 times the system order. Further it is proven that the IMC controller using the Causal Wiener filter, is similarly equivalent to the LQG controller.

Appendix C

Efficient state-space filtering using input- and output normal forms and application to FxLMS

C.1   Introduction

Consider the following state-space controller equations

$$x(k+1) = Ax(k) + Bu(k) \tag{C.1}$$
$$y(k) = Cx(k) + Du(k) \tag{C.2}$$

with $u(k) \in \mathbb{R}^m$ the input to the controller , $x(k) \in \mathbb{R}^n$ the state of the controller, $y(k) \in \mathbb{R}^l$ the output of the controller (the control signal) and A,B,C and D of appropriate dimensions. To calculate the control signal $y(k)$ and the state-update $x(k+1)$ from $u(k)$ and $x(k)$, one can use (C.1) and (C.2) directly. However, the matrix-vector multiplication $Ax(k)$ needs $(2n-1)n$ floating point operations (flops[1]). This can be computationally complex if n is large, as in noise and vibration control.

The computational complexity can be reduced by using another parameterizations of the controller. In this report we show how the computational complexity can be reduced via the *output-normal* (Section C.2) and *input-normal* parameterizations (Section C.3). Section C.4 compares the computational complexity of the filter iteration. Section C.5 shows how the input-normal parameterization can be used to reduce the computational complexity of the Filtered-X LMS algorithm. Section C.6 concludes the report.

---

[1]One flop is one multiplication or one addition of two floating point numbers.

C.2 Efficient filtering via output-normal parameterization

To calculate the output-normal parameterization, the state-space model (C.1),(C.2) should be in output-normal form, i.e. the Observability grammian should be the identity matrix:

$$A^T A + C^T C = I$$

In general the state-space model (e.g. obtained from subspace identification) is not in output-normal form. But, if the state-space model is observable, then the state-space model can be transformed to the output-normal form via a similarity transformation:

$$x_T(k + 1) = A_T x_T(k) + B_T u(k), \quad x_T(0) = T x_0 \qquad (C.3)$$
$$y(k) = C_T x_T(k) + D_T u(k) \qquad (C.4)$$

with $x_T(k) = Tx(k)$, $A_T = TAT^{-1}$, $B_T = TB$, $C_T = CT^{-1}$, $D_T = D$ and $T \in \mathbb{R}^{n \times n}$ a non-singular matrix (determined by the observability grammian of the state-space model which is not in output-normal form).

Now, suppose the state-space model (C.1),(C.2) is in output-normal form. Then, the pair (A, C) can be decomposed in a series of rotations. Let us define the 2×2 matrix $U(\alpha)$

$$U(\alpha) = \begin{bmatrix} -\alpha & \sqrt{1 - \alpha^2} \\ \sqrt{1 - \alpha^2} & \alpha \end{bmatrix}$$

with the scalar $\alpha \in [-1, 1]$. Then, (A, C) is decomposed in the following way

$$\begin{bmatrix} C \\ A \end{bmatrix} = T_1(\theta(1)) T_2(\theta(2)) \cdots T_{nl}(\theta(nl)) \begin{bmatrix} 0 \\ I_n \end{bmatrix} \qquad (C.5)$$

with $T_i(\theta(i))$ given as:

$$T_1(\theta(1)) = \begin{bmatrix} I_{n-1} & 0 & 0 \\ 0 & U(\theta(1)) & 0 \\ 0 & 0 & I_{l-1} \end{bmatrix} \in \mathbb{R}^{(n+l) \times (n+l)}$$

$$\vdots$$

$$T_l(\theta(l)) = \begin{bmatrix} I_{n+l-2} & 0 \\ 0 & U(\theta(l)) \end{bmatrix}$$

$$T_{l+1}(\theta(l+1)) = \begin{bmatrix} I_{n-2} & 0 & 0 \\ 0 & U(\theta(l+1)) & 0 \\ 0 & 0 & I_l \end{bmatrix}$$

$$\vdots$$

$$T_{2l}(\theta(2l)) = \begin{bmatrix} I_{n+l-1} & 0 & 0 \\ 0 & U(\theta(2l)) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\vdots$$

$$T_{(n-1)l+1}(\theta((n-1)l+1)) = \begin{bmatrix} U(\theta((n-1)l+1)) & 0 \\ 0 & I_{n+l-2} \end{bmatrix}$$

$$\vdots$$

$$T_{nl}(\theta(nl)) = \begin{bmatrix} I_{l-1} & 0 & 0 \\ 0 & U(\theta(nl)) & 0 \\ 0 & 0 & I_{n-1} \end{bmatrix}$$

with $\theta \in \mathbb{R}^{nl}$ a vector with reflection coefficient. The parameterization (θ, B, D) is called the output-normal parameterization. Using the output-normal parameterization y_k and x_{k+1} can be calculated in the following way:

$$\begin{bmatrix} y_k \\ x_{k+1} \end{bmatrix} = T_1(\theta(1))T_2(\theta(2)) \cdots T_{nl}(\theta(nl)) \begin{bmatrix} 0 \\ x_k \end{bmatrix} + \begin{bmatrix} D \\ B \end{bmatrix} u_k \qquad (C.6)$$

Because a multiplication with $T_i(\theta(i))$ is just a 2-dimensional rotation, it can be performed by 2 additions and 4 multiplications. Hence, the computational complexity of the nl multiplications with T_i is $2nl$ additions and $4nl$ multiplications.

The computational complexity of the multiplications Ax_k and Cx_k in (C.1) and (C.2) are $(n + l)(n - 1)$ additions and $(n + l)n$ multiplications.

So, in the case $6nl < (n + l)(2n - 1)$, the number of floating point operations for the output-normal form controller is less than for the state-space controller of (C.1),(C.2). In this case, we have

$$\begin{aligned} 6nl \quad &< \quad (n + l)(2n - 1) \;=\; 2n^2 + 2nl - n - l \\ &\Longleftrightarrow \\ 4nl + l \quad &< \quad 2n^2 - n \\ &\Longleftrightarrow \\ l \quad &< \quad \frac{n(2n - 1)}{4n + 1} \end{aligned}$$

In case of high orders n, the expression can be simplified by $l < \dfrac{n}{2}$.

C.3 Efficient filtering via input-normal parameterization

The input-normal parameterization is analogue to the output-normal parameterization. Now, the state-space model (C.1),(C.2) should be in input-normal form, which means the Controllability grammian should be the identitiy matrix:

$$AA^T + BB^T = I$$

Every *controllable* state-space model can be transformed to the input-normal form via a similarity transformation.

Now, suppose the state-space model (C.1),(C.2) is in input-normal form. Then, the pair (A, B) can be decomposed in a series of rotations in the following way

$$\begin{bmatrix} B & A \end{bmatrix} = \begin{bmatrix} 0 & I_n \end{bmatrix} T_{mn}(\theta(mn))T_{mn-1}(\theta(mn-1))\cdots T_1(\theta(1)) \qquad \text{(C.7)}$$

with $T_i(\theta(i))$ given as:

$$T_1(\theta(1)) = \begin{bmatrix} I_{n-1} & 0 & 0 \\ 0 & U(\theta(1)) & 0 \\ 0 & 0 & I_{m-1} \end{bmatrix} \in \mathbb{R}^{(n+m)\times(n+m)}$$

$$\vdots$$

$$T_m(\theta(m)) = \begin{bmatrix} I_{n+m-2} & 0 \\ 0 & U(\theta(m)) \end{bmatrix}$$

$$T_{m+1}(\theta(m+1)) = \begin{bmatrix} I_{n-2} & 0 & 0 \\ 0 & U(\theta(m+1)) & 0 \\ 0 & 0 & I_m \end{bmatrix}$$

$$\vdots$$

$$T_{2m}(\theta(2m)) = \begin{bmatrix} I_{n+m-1} & 0 & 0 \\ 0 & U(\theta(2m)) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\vdots$$

$$T_{(n-1)m+1}(\theta((n-1)m+1)) = \begin{bmatrix} U(\theta((n-1)m+1)) & 0 \\ 0 & I_{n+m-2} \end{bmatrix}$$

$$\vdots$$

$$T_{nm}(\theta(nm)) = \begin{bmatrix} I_{m-1} & 0 & 0 \\ 0 & U(\theta(nm)) & 0 \\ 0 & 0 & I_{n-1} \end{bmatrix}$$

with $\theta \in \mathbb{R}^{nm}$ a vector with reflection coefficient. The parameterization (θ, C, D) is called the input-normal parameterization. Using the input-normal parameterization $y(k)$ and $x(k+1)$ can be calculated in the following way:

$$y(k) = \begin{bmatrix} D & C \end{bmatrix} \begin{bmatrix} u(k) \\ x(k) \end{bmatrix}$$

$$x(k+1) = \begin{bmatrix} 0 & I_n \end{bmatrix} T_{mn}T_{mn-1}\cdots T_1 \begin{bmatrix} u(k) \\ x(k) \end{bmatrix}$$

The computational complexity of the state update via the rotations is $6nm$ flops, whereas using (C.1) explicitly it is $n(2n+2m-1)$ flops. So, in the case $6nm < n(2n+2m-1)$, the number of floating point operations for the input-normal form

Table C.1: Computational complexity of one iteration of the state-space equations, the output-normal and the input-normal parameterization.

	state-space	output-normal	input-normal
# additions	$(n+l)(n+m-1)$	$2nl+(n+l)m$	$2nm+l(m+n-1)$
# multiplications	$(n+l)(n+m)$	$4nl+(n+l)m$	$4nm+l(m+n)$
total # of flops	$(n+l)(2n+2m-1)$	$6nl+2(n+l)m$	$6nm+l(2n+2m-1)$

Table C.2: Complexity of one filter iteration for different parameterizations (l outputs, m inputs, n is the state-dimension and I the FIR filter length).

Method	# coefficients	# memory (×8 Bytes)	# flops
FIR	lmI	l+mI	lm(2I-1)
Full matrix transferf.	lm(2n+1)	(m+l)(n+1)	lm(6n+1)
Full state-space	n(n+m+l)+lm	l+m+2n	(n+l)(2n+2m-1)
Output normal	2nl+(l+n)m	l+m+n+1	6nl+2(n+l)m
Input normal	2nm+l(n+m)	l+m+n+1	6nm+l(2n+2m-1)

controller is less than for the state-space controller of (C.1). In this case, we have

$$
\begin{aligned}
6nm \quad &< \quad n(2n+2m-1) \\
&< \quad 2n^2 + 2nm - n \\
&\Longleftrightarrow \\
4nm \quad &< \quad 2n^2 - n \\
&\Longleftrightarrow \\
m \quad &< \quad \frac{(2n-1)}{4}
\end{aligned}
$$

In case of high orders n, the expression can be simplified by $m < \dfrac{n}{2}$.

C.4   Comparison

The computational complexity of the state-space equations and the output-normal and input-normal parameterizations is summarized in Table C.1.

Table C.2 compares different filter parameterizations, w.r.t. number of coefficients, necessary memory and computational complexity.

Table C.1–C.3 illustrate that state-space filtering can be evaluated at significant lower computational complexity using input-normal and output-normal parameter-

Table C.3: Number of flops of a state-space (SS), output-normal (ON) and input-normal (IN) iteration for different dimensions and orders.

$m = 1, l = 1$

order n	SS	ON	IN
2	15	18	17
4	45	34	33
8	153	66	65
16	561	130	129
32	2145	258	257
64	8385	514	513

$m = 1, l = 4$

order n	SS	ON	IN
2	30	60	32
4	72	112	60
8	204	216	116
16	660	424	228
32	2340	840	452
64	8772	1672	900

$m = 4, l = 1$

order n	SS	ON	IN
2	33	36	59
4	75	64	111
8	207	120	215
16	663	232	423
32	2343	456	839
64	8775	904	1671

$m = 4, l = 4$

order n	SS	ON	IN
2	66	96	92
4	120	160	156
8	276	288	284
16	780	544	540
32	2556	1056	1052
64	9180	2080	2076

izations. In general, we can conclude

$$
\begin{aligned}
&\text{if} \quad m < l \ \ \& \ \ m < \tfrac{n}{2} \quad \rightarrow \quad \text{input-normal} \\
&\text{if} \quad l < m \ \ \& \ \ l < \tfrac{n}{2} \quad \rightarrow \quad \text{output-normal} \\
&\text{else} \quad\quad\quad\quad\quad\quad\quad\ \ \rightarrow \quad \text{state-space}
\end{aligned}
$$

C.5 Application to Filtered-X LMS

This section shows how the input-normal parameterization can be used to efficiently evaluate the filtering of the reference signal in the multi-reference Filtered-X LMS algorithm. Let

m_r: number of reference signals
m_u: number of control signals
m_e: number of error signals
$\widehat{G}_{eu}$: $m_e \times m_u$ model of secondary path
n_s: number of states sec. path model $\widehat{G}_{eu}$
I_s: number of coefficients of FIR realization of sec. path model
I: number of coefficients of adaptive filter
$\theta(k)$: $m_u m_r I$ dimensional vector with controller coefficients
$r(k)$: m_r dimensional vector with reference signal
$\phi(k)$: $m_r I$ dimensional vector with samples from reference signal $r(k)$
$e(k)$: m_e dimensional vector with measured error signal

Then, the update equation of the Filtered-X LMS algorithm is given by

$$\theta(n+1) = \theta(k) - \mu[\widehat{G}_{eu}^T(q^{-1}) \otimes \phi(k)]e(k)$$

with $\mu > 0$ a scalar stepsize parameter.

To evaluate $\widehat{G}_{eu}^T(q^{-1}) \otimes \phi(k)$ every sampling instant n we have to evaluate

$$\begin{bmatrix} \widehat{G}_{eu,11} \\ \vdots \\ \widehat{G}_{eu,m_e 1} \\ \vdots \\ \widehat{G}_{eu,m_e m_u} \end{bmatrix} r_i(k), \qquad i = 1, \cdots, m_r$$

This filter, i.e. the vector stacking of the secondary path model, can be realized as a state-space filter of order n_s (so the state has not to be increased as implicitly done in [128]). This single-input multi-output state-space filter can be transformed to an input-normal parameterization, which will reduce the computational complexity.

Table C.4 shows the computational complexity of the Filtered-X LMS algorithm using the FIR and the input-normal parameterization. Here we assumed that the FIR filter length of the controller is the same as the FIR length of the secondary path model.

Note, that for both cases the computational complexity is linear in the number of reference signals m_r. Further note, that (also for both cases), the computational complexity (approximately) depends on the product of m_e and m_u.

The input-normal filter version is computationally more efficient than the FIR filter version of the Filtered-X LMS algorithm if

$$n_s(6m_r + 2m_r m_e m_u) + m_u m_r I(2m_e + 3) + m_u(m_r m_e - 1) <$$
$$< m_u m_r I(4m_e + 3) - m_u(m_r m_e + 1) \iff$$
$$n_s(6m_r + 2m_r m_e m_u) < 2m_r m_e m_u I - 2m_r m_e m_u \iff$$
$$n_s(3 + m_e m_u) < m_e m_u(I - 1) \iff$$
$$n_s < \frac{3 + m_e m_u}{m_e m_u}(I - 1)$$

Table C.4: Number of flops of the Filtered-X LMS algorithm using an FIR and an input-normal filter to generate the regression vector.

	FIR	input-normal
Filter regression	$m_r((2I-1)m_e m_u)$	$m_r(n_s(6 + 2m_e m_u) + m_e m_u)$
Update controller	$(2m_e + 1)m_u m_r I$	$(2m_e + 1)m_u m_r I$
Calc. control	$2m_u m_r I - m_u$	$2m_u m_r I - m_u$
Total	$m_u m_r I(4m_e + 3)+$	$n_s m_r(6 + 2m_e m_u)+$
	$-m_u(m_r m_e + 1)$	$+m_u m_r I(2m_e + 3) + m_u(m_r m_e - 1)$
Total (approx.)	$4m_r m_e m_u I$	$n_s(6m_r + 2m_r m_e m_u) + 2m_r m_e m_u I$

To get insight how much reduction is obtained by the input-normal parameterization, the ratio of the flops necessary for the input-normal parameterization and the FIR filter are calculated for different values of m_r, m_e and m_u, i.e.

$$\frac{m_r(n_s(6 + 2m_e m_u) + m_e m_u)}{m_r((2I-1)m_e m_u)} \times 100\%$$

To get insight in the reduction of the complete Filtered-X LMS algorithm, we also calculated the ratio between the flops necessary for the Filtered-X LMS algorithm using the input-normal parameterization and for using the FIR filter, i.e.

$$\frac{n_s m_r(6 + 2m_e m_u) + m_u m_r I(2m_e + 3) + m_u(m_r m_e - 1)}{m_u m_r I(4m_e + 3) - m_u(m_r m_e + 1)} \times 100\%$$

The results are given in Table C.5.

From Table C.5 it can be concluded, that indeed reduction of the computational complexity can be obtained by using the input-normal parameterization, especially for low orders n_s and large number of FIR coefficients I_s. Table C.6 shows the number of kiloflops per sample for the Filtered-X LMS algorihm using the input-normal and FIR filter implementations. Note, that this is obtained for the standard Filtered-X LMS algorithm without normalization of the stepsize.

C.6 Conclusions

Output-normal and input-normal parameterizations were proposed to efficiently evaluate state-space filter iterations. It was derived that for filters of order n with m inputs and l outputs, the following choice leads to most efficient implementation

$$
\begin{aligned}
&\text{if} \quad m < l \ \ \& \ \ m < \tfrac{n}{2} \quad \rightarrow \quad \text{input-normal} \\
&\text{if} \quad l < m \ \ \& \ \ l < \tfrac{n}{2} \quad \rightarrow \quad \text{output-normal} \\
&\text{else} \qquad\qquad\qquad\qquad\ \rightarrow \quad \text{state-space}
\end{aligned}
$$

It was also shown, that the input-normal parameterization can be used to reduce the computational complexity of filtering the reference signal in the Filtered-X LMS algorithm. In case, the length of the FIR controller and the FIR filter model of

Table C.5: For every entry, the first number is the ratio (in %) between the flops for input-normal and FIR filtering, the second number is the ratio between the flops for Filtered-X LMS using input-normal and FIR filtering.

$m_u = 1, m_e = 1$

$n_s =$	$I_s =$ 100	200	300	400	500
20	81/95	40/83	27/79	20/77	16/76
40	161/118	81/94	54/87	40/83	32/81
60	242/140	121/106	80/94	60/89	48/85
80	322/163	161/117	107/102	80/94	64/90
100	403/186	201/129	134/110	100/100	80/94

$m_u = 1, m_e = 8$

$n_s =$	$I_s =$ 100	200	300	400	500
20	28/67	14/61	9.3/59	7.0/58	5.6/57
40	56/80	28/67	19/63	14/61	11/59
60	83/92	42/73	28/67	21/64	17/62
80	111/105	55/80	37/71	28/67	22/64
100	139/118	69/86	46/75	35/70	28/67

$m_u = 4, m_e = 4$

$n_s =$	$I_s =$ 100	200	300	400	500
20	24/68	12/63	8.1/61	6.1/61	4.9/60
40	48/78	24/68	16/65	12/63	9.6/62
60	72/88	36/73	24/68	18/66	14/64
80	96/98	48/78	32/71	24/68	19/66
100	120/108	60/83	40/75	30/71	24/68

$m_u = 6, m_e = 8$

$n_s =$	$I_s =$ 100	200	300	400	500
20	22/64	11/59	7.3/58	5.4/57	4.4/56
40	43/74	22/64	14/61	11/59	8.6/58
60	65/84	32/69	22/64	16/62	13/60
80	86/94	43/74	29/67	21/64	17/62
100	107/103	54/79	36/71	27/67	21/64

Table C.6: Number of kiloflops per sample of Filtered-X LMS using input-normal and FIR filtering.

$m_u = 6, m_e = 8, m_r = 8$

$n_s =$	$I_s =$ 100	200	300	400	500
20	108/168	199/336	290/504	381/672	473/840
40	124/168	215/336	307/504	398/672	489/840
60	141/168	232/336	323/504	414/672	505/840
80	157/168	248/336	339/504	430/672	522/840
100	173/168	264/336	356/504	447/672	538/840

the secondary path are both of length I, then an input-normal parameterization of order n_s with

$$n_s < \frac{3 + m_e m_u}{m_e m_u}(I - 1)$$

will reduce the computational complexity (where m_e and m_u are the number of error and control signals respectively, no matter the number of reference signals).

Appendix D

A frequency domain subspace algorithm for mixed causal, anti-causal LTI systems

Abstract

The paper extends the subspace identification method to estimate state-space models from frequency response function (FRF) samples, proposed by [113] for mixed causal/anti-causal systems, and shows that other frequency domain subspace algorithms can be extended similarly. The method is demonstrated by simulation experiments.

keywords: frequency domain identification, subspace method, descriptor system, Kronecker canonical form, state-space model

This paper has been published before in [63].

D.1   Introduction

Subspace identification methods are powerful methods in identifying linear multi variable systems. This is because these methods are based on numerically reliable algorithms as SVD and QR-decomposition and directly yield a state-space model. Advantages of a state-space model over a transfer function model are e.g. that a resonance mode of the system, which is observed at multiple outputs, is modeled only once and its numerical sensitivity to round-off errors is in general significantly smaller (see e.g. [68]). Further, most modern control methods are based on state-space models.

Originally, subspace identification methods were based on time-domain measured input/output data [177, 181]. Not much time later, subspace identification methods were proposed which are based on frequency-domain data: Fourier transformed input/output data [112] or FRF samples [113]; see [137] for new results on consistency and convergence. Using frequency domain data the number of samples may be significantly reduced, especially in case of systems with (many) widely separated resonance modes (e.g. stiff systems and systems with a high number of resonances like acoustical systems) and a non-uniform frequency grid can be exploited to accurately model the system at specific frequencies (see [138] for more details on system identification in the frequency domain).

Using frequency domain subspace identification methods systems with anti-causal/unstable modes can be identified too, which is an advantage over time-domain subspace algorithms. However, to study the causal and/or anti-causal behavior, additional post-processing has to be a applied to separate the causal and anti-causal modes. This separation is also necessary, when only a model of the causal part of the system has to be used for control/filter design, such as in the Causal Wiener filter, see e.g. [57].

This paper proposes a frequency domain subspace method to identify a state-space model of the causal/stable part and the anti-causal/unstable part of a system *directly*. The method is obtained by adjusting *Algorithm 2* from [112] using techniques from [182], who proposed a *time-domain* algorithm to identify mixed causal, anti-causal systems. We will base our method for mixed causal, anti-causal systems on FRF samples, however the same reasoning can be used to extend other subspace methods based on e.g. discrete Fourier transforms (DFT's) of input and output data as discussed by [112].

Problems with identification of anti-causal (or unstable) systems do arise e.g. in closed-loop identification of a stabilized unstable system, in discretized systems with fractional delay [94], in estimating the inverse of non-minimum phase systems (usually due to time delay) or direct estimation of a deconvolution filter. Furthermore, this problem arises in the frequency domain implementation of a method we proposed to estimate the Causal Wiener filter [57].

The paper is organized as follows. Section D.2 describes the problem of frequency domain subspace algorithm for mixed causal, anti-causal systems in more detail. Section D.3 derives a solution of the problem based on FRF samples. Section D.4 illustrates the method by simulation examples.

D.2   Problem description

Consider the following discrete time mixed causal, anti-causal state-space system

$$x^c(k+1) = A^c\ x^c(k)\ + B^c\ u(k) \quad \text{(causal)} \tag{D.1}$$

$$x^{ac}(k-1) = A^{ac}x^{ac}(k) + B^{ac}u(k) \quad \text{(anti-causal)} \tag{D.2}$$

$$y(k) = C^c\ x^c(k)\ + C^{ac}x^{ac}(k) + Du(k) \tag{D.3}$$

with $u(k) \in \mathbb{R}^m$, $y(k) \in \mathbb{R}^l$, $x^c(k) \in \mathbb{R}^{n_c}$ and $x^{ac}(k) \in \mathbb{R}^{n_{ac}}$ and A^c, B^c, C^c, A^{ac}, B^{ac}, C^{ac} and D of appropriate dimensions. Equation (D.1) models the causal and (D.2) the anti-causal dynamics. The order of the system is given by $n = n_c +$

n_{ac}. Furthermore, the eigenvalues of A^c and A^{ac} are inside the unit-circle and we assume that the state-space description is minimal, i.e. there are no unobservable or uncontrollable modes. This class of systems is a special case of descriptor systems, described in the so-called Kronecker canonical form [182]. The transfer function of the system is given by

$$G(z) = \underbrace{\sum_{i=-\infty}^{-1} C^{ac} A^{ac(-i-1)} B^{ac} z^{-i}}_{anti-causal} + \underbrace{D + \sum_{i=1}^{\infty} C^c A^{c(i-1)} B^c z^{-i}}_{causal}$$

and clearly consists of a causal and an anti-causal part. The FRF of the system is given by

$$G(e^{j\omega}) = D + C^c(e^{j\omega} I_{n_c} - A^c)^{-1} B^c + C^{ac}(e^{-j\omega} I_{n_{ac}} - A^{ac})^{-1} B^{ac} \quad \text{(D.4)}$$

The problem is to estimate the matrices A^c, B^c, C^c (up to a similarity transformation T_c), A^{ac}, B^{ac}, C^{ac} (up to a similarity transformation T_{ac}) and D using M noise corrupted estimates of the frequency response

$$G_k = G(e^{j\omega_k}) + n_k, \quad k = 1, 2, \cdots, M \quad \text{(D.5)}$$

at a given but arbitrary number of distinct frequencies ω_k (cf. [113]).

Note, that if $\omega_k = \pi k/M$, $k = 1, \cdots, M$ (uniformly spaced frequencies) the impulse response coefficients g_i of $G(z) = \sum_{i=-\infty}^{\infty} g_i z^{-i}$ can be calculated by the two-sided inverse discrete Fourier transform

$$g_i = \frac{1}{2M} \sum_{k=1-M}^{M} G_k e^{j2\pi ik/2M}, \quad i = 1 - M, \cdots, M$$

with $G_{-k} = G_k^*$, $(k = 1, \cdots, M-1)$. Then, D can be set to $D = g_0$. A_T^c, B_T^c and C_T^c can be calculated from g_i ($i = 1, \cdots, M$) by well known realization algorithms (e.g. [89]). Dually A_T^{ac}, B_T^{ac} and C_T^{ac} can be calculated from g_{-i} ($i = 1, \cdots, M$) by these same realization algorithms. This is basically an extension of *Algorithm 1* of [113] for mixed causal, anti-causal systems. The following Section derives an alternative algorithm, which can also be used for non-uniformly spaced frequencies, which is basically an extension of *Algorithm 2* of [113].

D.3 Derivation of the algorithm

First consider the DFT of (D.1)–(D.3), where we shifted equation (D.2) $i - 1$ samples forward in time

$$e^{j\omega} X^c(\omega) = A^c X^c(\omega) + B^c U(\omega) \quad \text{(D.6)}$$
$$e^{-j\omega} X^{ac,i}(\omega) = A^{ac} X^{ac,i}(\omega) + B^{ac} e^{j(i-1)\omega} U(\omega) \quad \text{(D.7)}$$
$$Y(\omega) = C^c X^c(\omega) + C^{ac} e^{-j(i-1)\omega} X^{ac,i}(\omega) + D U(\omega) \quad \text{(D.8)}$$

where $X^c(\omega)$, $X^{ac,i}(\omega)$, $U(\omega)$ and $Y(\omega)$ denoted the DFT of $x^c(k)$, $x^{ac}(k+i-1)$, $u(k)$ and $y(k)$ respectively and let $i > n$.

As in [113], let $X_i^c(\omega)$ the resulting state transform when $U(\omega) = e_r$, with e_r the unit-vector with 1 on the rth position, $X_r^{ac,i}(\omega)$ is defined similarly. By defining the compound state matrix

$$X_C^c(\omega) \;=\; [X_1^c(\omega)\; X_2^c(\omega)\; \cdots\; X_m^c(\omega)]$$

and $X_C^{ac,i}$ similarly, the transfer function can be implicitly described as

$$G(e^{j\omega}) = C^c X_C^c(\omega) + C^{ac} e^{-j(i-1)\omega} X_C^{ac,i}(\omega) + D \tag{D.9}$$

with

$$e^{j\omega} X_C^c(\omega) = A^c X_C^c(\omega) + B^c \tag{D.10}$$
$$e^{-j\omega} X_C^{ac,i}(\omega) = A^{ac} X_C^{ac,i}(\omega) + B^{ac} e^{j(i-1)\omega} \tag{D.11}$$

By iterative substituting the state-equations we obtain the relation

$$\begin{bmatrix} G(e^{j\omega}) \\ e^{j\omega} G(e^{j\omega}) \\ \vdots \\ e^{j(i-2)\omega} G(e^{j\omega}) \\ e^{j(i-1)\omega} G(e^{j\omega}) \end{bmatrix} = \mathcal{O}_i \begin{bmatrix} X_C^c(\omega) \\ X_C^{ac,i}(\omega) \end{bmatrix} + \Gamma_i \begin{bmatrix} I_m \\ e^{j\omega} I_m \\ \vdots \\ e^{j(i-2)\omega} I_m \\ e^{j(i-1)\omega} I_m \end{bmatrix} \tag{D.12}$$

with

$$\mathcal{O}_i = \begin{bmatrix} C^c & C^{ac} A^{ac(i-1)} \\ C^c A^c & C^{ac} A^{ac(i-2)} \\ \vdots & \vdots \\ C^c A^{c(i-2)} & C^{ac} A^{ac} \\ C^c A^{c(i-1)} & C^{ac} \end{bmatrix} \tag{D.13}$$
$$= \begin{bmatrix} \mathcal{O}_i^c & \mathcal{O}_i^{ac} \end{bmatrix} \tag{D.14}$$

and the following Toeplitz matrix filled with impulse response coefficients

$$\Gamma_i = \begin{bmatrix} D & C^{ac} B^{ac} & \cdots & C^{ac} A^{ac(i-2)} B^{ac} \\ C^c B^c & D & \ddots & \vdots \\ \vdots & \ddots & \ddots & C^{ac} B^{ac} \\ C^c A^{c(i-2)} B^c & \cdots & C^c B^c & D \end{bmatrix}$$

By repeating (D.12) for all ω_k $(k = 1, \cdots, M)$ and using the frequency response estimates G_k, the range space of $\mathcal{O}_i$ (and thus also the order n of the system) can be determined by a QR factorization and an SVD as explained by [113]. If the covariance function $E[n_k n_s^H] = R_k \delta_{ks}$ of the noise n_k is known, a weighting can be performed to compensate for n_k.

Let U_n be such that its columns span the range space of $\mathcal{O}_i$. Then, we look for an invertible $n \times n$ transformation matrix P such that

$$U_n P = \left[\begin{array}{c|c} \mathcal{O}_i^c & \mathcal{O}_i^{ac} \end{array} \right] \left[\begin{array}{c|c} T_c & 0 \\ \hline 0 & T_{ac} \end{array} \right] \qquad \text{(D.15)}$$

This problem to separate U_n in a causal part fully determined by the pair (A^c, C^c) and an anti-causal part fully determined by the pair (A^{ac}, C^{ac}) is exactly the problem which also appears in mixed causal, anti-causal subspace identification using time-domain data, and is solved e.g. by [182].

The matrix P and n_c, n_{ac} are calculated according to [182]. Then from the first l rows in $U_n' = U_n P$, C_T^c and C_T^{ac} can be picked up:

$$C_T^c = U_n'(1:l, 1:n_c), \qquad \text{(D.16)}$$
$$C_T^{ac} = U_n'(1:l, n_c+1:n_c+n_{ac}) \qquad \text{(D.17)}$$

and A_T^c, A_T^{ac} can be calculated by solving

$$U_n'(1:(i-1)l, 1:n_c)A_T^c = U_n'(l+1:il, 1:n_c) \qquad \text{(D.18)}$$

and

$$U_n'(1:(i-1)l, n_c+1:n_c+n_{ac})A_T^{ac} = U_n'(l+1:il, n_c+1:n_c+n_{ac}) \quad \text{(D.19)}$$

If A_T^c, C_T^c, A_T^{ac}, C_T^{ac} are calculated, B_T^c, B_T^{ac} and D can be calculated by solving a least squares problem using the samples (D.5), because B_T^c, B_T^{ac} and D appear linearly in $G(e^{j\omega})$, cf. (D.4). Again a weighting using the noise covariance matrices R_k can be used to compensate for n_k [113].

Let us summarize the method in the following *Algorithm 2 (Causal/Anti-causal)*.

Algorithm 2 (C/AC):

1. Given: Samples G_k of the frequency response, and the covariance R_k at frequency ω_k for $k = 1, \cdots, M$.

2. Calculate the estimate U_n of the extended observability matrix $\mathcal{O}_i$ as in *Algorithm 2* in [113].

3. Calculate the matrix P, which separates the causal and anti-causal part in U_n, and the orders of the causal and anti-causal part n_c and n_{ac} respectively as in Section 3.3 and 3.5 in [182].

4. Calculate $U_n' = U_n P$ and select C_T^c and C_T^{ac} according to (D.16) and (D.17) respectively. Further solve A_T^c and A_T^{ac} from (D.18) and (D.19) respectively.

5. Solve B_T^c, B_T^{ac} and D from:

$$(B_T^c, B_T^{ac}, D) = \arg\min \sum_{k=1}^{M} \|R_k^{-1/2}\Big(G_k - D - C_T^c(e^{j\omega_k}I_{n_c} - A_T^c)^{-1}B_T^c +$$
$$- C_T^{ac}(e^{-j\omega_k}I_{n_{ac}} - A_T^{ac})^{-1}B_T^{ac}\Big)\|_F^2$$

We remark, that solving the least squares problem for B_T^c, B_T^{ac} and D might be ill-conditioned, especially if the system has poles close to the unit circle [113]. Regularization with a small $\epsilon > 0$ parameter can improve the conditioning of the least squares problem, at the expense of a small bias, see e.g. [70].

D.4   Simulation examples

We will consider two simulation examples: the identification of an academic system and of an acoustic system, both with a stable and an unstable resonance modes.

D.4.1   Academic example

The discrete time academic system has a stable mode at 100Hz and an unstable mode at 300Hz, the sampling rate is $F_s = 1000$Hz, and its transfer function is given by

$$G(z) \;=\; \frac{(z-0.9e^{-j2\pi 0.15})(z-0.9e^{+j2\pi 0.15})}{(z-0.95e^{-j2\pi 0.1})(z-0.95e^{+j2\pi 0.1})} \cdot \frac{(z-0.5e^{-j2\pi 0.4})(z-0.5e^{+j2\pi 0.4})}{(z-1.1e^{-j2\pi 0.3})(z-1.1e^{+j2\pi 0.3})}$$

The system was excited with a Schroeder multi sine, with frequencies ranging from .5Hz to 500Hz in steps of $\Delta F = .5Hz$. The number of samples in one block to estimate the frequency response was chosen to be 2000 ($= F_s/\Delta F$) such that leakage due to Fourier transforming a finite block of samples is prevented [138]. The measured output was corrupted with unit variance Gaussian white noise $v(k)$ filtered by $H(z)$

$$y(k) = G(z)u(k) + H(z)v(k)$$

with $H(z)$ given by

$$H(z) \;=\; \frac{0.3z^2}{(z-0.95e^{-j2\pi 0.1})(z-0.95e^{+j2\pi 0.1})} \cdot \frac{z^2}{(z-0.91e^{-j2\pi 0.3})(z-0.91e^{+j2\pi 0.3})}$$

The output data was generated by splitting G in a stable/causal part and an unstable/anti-causal part. The latter was simulated by filtering anti-causally, to prevent the output from exploding.

Based on 100 blocks, three methods were used to calculate the state-space model of G:

PO-MOESP: Time-domain PO-MOESP [182];

A2: *Algorithm 2 (C/AC) without knowledge of R_k*;

A2wi: *Algorithm 2(C/AC) with estimated R_k.*

The i parameter of (D.12) was chosen to be $i = 10$. To make a fair comparison, for PO-MOESP the 100 blocks each of 2000 samples were used to average the output to reduce. In A2 and A2wi, the 100 blocks we used to average the estimated frequency responses, and in A2wi also to calculated the variance R_k ($k = 1, \cdots, 2000$).

Each experiment of 100 blocks was repeated 1000 times. Figure D.4.1 shows the real frequency response of G, the average of the frequency response error made by PO-MOESP, A2 and A2wi. From this Figure, we clearly see that the stable as well as the unstable mode are accurately modeled by all three methods. Further,
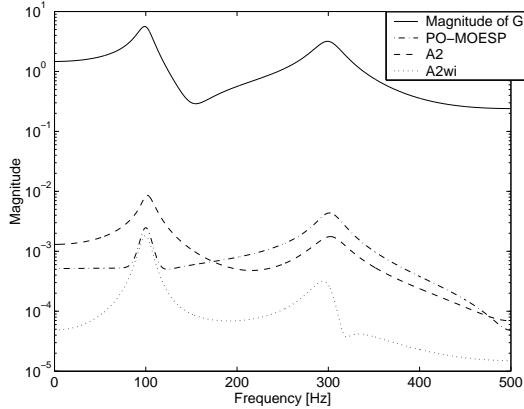
Figure D.1: Magnitude of G, and the frequency response estimation errors obtained by using 100 blocks of 2000 samples, which were averaged over 1000 experiments.

Table D.1: Frequency and magnitude of unstable poles.

Frequency	Magnitude	Frequency	Magnitude
0	1.64	± 171	1.31
0	3.65	± 278	1.41
0	8.21	± 388	1.46
± 32.1	1.01	500	1.59
± 33.5	1.08	500	2.44
± 76.0	1.04	500	12.6
± 91.7	1.00		

we infer that by taking the covariance information R_k of the noise into account, the model is more accurately estimated, as was also concluded in [113] for the causal method. Finally, we infer that on the average using the causal/anti-causal frequency domain with covariance information, a more accurate model was estimated then by using the causal/anti-causal time domain PO-MOESP method.

D.4.2 Acoustic system

The acoustic system to be identified is a transfer function in an acoustical duct, which contains unstable modes due to the inversion of delays between actuators and sensors which are not collocated, for more details we refer to [57]. The sampling rate is again 1000Hz and the frequency response of the real system G and the noise coloring H is shown in Figure D.2. The unstable poles of the system are given in Table D.1.

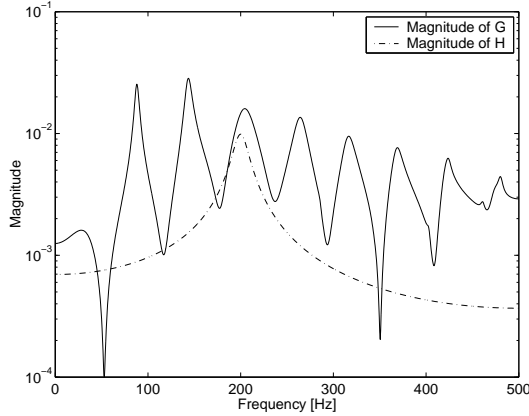The excitation signal was chosen the same as in the previous example, a

Figure D.2: Magnitude of the real system G and the noise model H.

Schroeder multi sine with frequencies ranging from 0.5Hz to 500Hz with steps of 0.5Hz. Each of the 100 blocks consists of 2000 samples. The results of the three methods, PO-MOESP, A2 and A2wi, were averaged over 200 experiments and the i parameter of (D.12) was chosen to be $i = 100$. We note, that the least squares problem to solve B_T^c, B_T^{ac} and D in A2 and A2wi was ill conditioned, due to poles close to the unit circle.

Figure D.3 shows the average frequency response estimation error when the order of the state-space model was chosen to be 44, which is the order of the real system G. We observe, that again A2wi gives the smallest estimation error, but the difference with A2 is not that large as in the previous academic example. The causal, anti-causal PO-MOESP method gives less accurate models, which is currently under study.   Finally, Figure D.4 gives the average estimation error, which is defined as [113]:

$$||\widehat{G} - G||_2 = \sqrt{\frac{1}{M} \sum_{k=1}^{M} |\widehat{G}(e^{j\omega_k}) - G(e^{j\omega_k})|^2}$$

for different model orders. From the Figure, we infer that for orders above 32 A2wi yields the best result, closely followed by A2. It is remarkable that for orders between 20 and 30, A2wi yields significantly less accurate results, whereas A2 gives reasonable good results for these orders.

Though, some questions remain on the precise interpretation of the simulation results, the simulation experiments show that the extension of frequency domain subspace identification methods for mixed causal, anti-causal systems was successful.
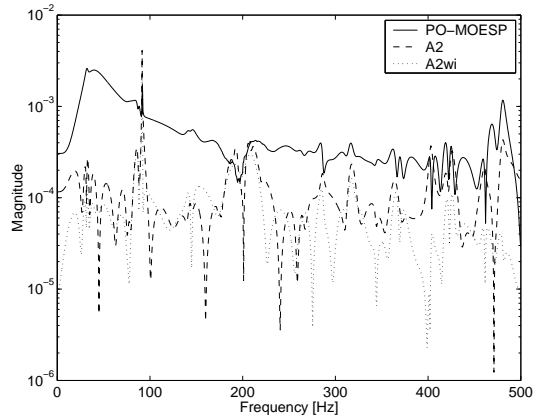
Figure D.3: Magnitude of the frequency response estimation errors obtained by the 44th order models identified by using 100 blocks of 2000 samples, which were averaged over 200 experiments.
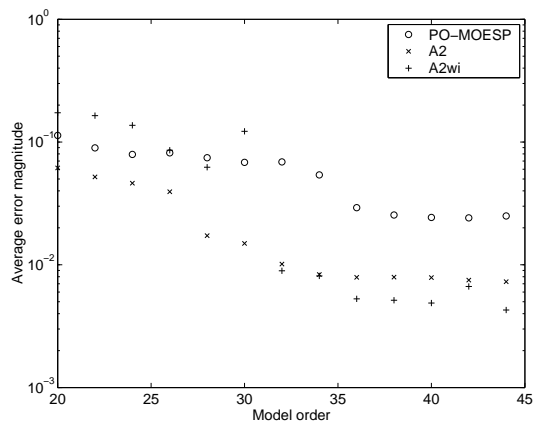

Figure D.4: Estimation error obtained by using 100 blocks of 2000 samples, averaged over 200 experiments.

D.5   Conclusions

It has been shown how subspace identification methods based on frequency domain data, can be adjusted to estimate a state-space model which models the causal and anti-causal part of the system separately. The crucial step in extending the frequency domain methods is to split up the extended observability matrix in a part due to causal modes and a part due to anti-causal modes. The two simulation experiments demonstrated that with the derived mixed causal, anti-causal subspace identification algorithm using FRF samples the causal and anti-causal modes of the systems could be accurately identified. We also observed, that including noise weighting to compensate for noise on the FRF samples, the obtained model error was better than using time-domain PO-MOESP for mixed causal, anti-causal systems.

Appendix E

Robust Decision Feedback Equalizer design via the solution of a regularized least squares problem

Abstract

This paper[1] presents a method to estimate a Decision Feedback Equalizer (DFE) directly from training data, which is robust w.r.t. time-variations in the communication channel. It is based on the indirect method proposed in [165], where the time variations in the channel are modeled as a probabilistic uncertainty. The robust DFE optimizes the performance by minimizing the mean squared error averaged over the distribution of the uncertainty in the channel. We show, that the robust DFE design problem can be solved by a regularized least squares problem. The main advantage of this direct method over [165] is, that no longer a spectral factorization in addition to a least squares problem is necessary. Another advantage is, that a model of the (average) channel and the noise color are not necessary anymore.

Keywords: robust filtering, estimation, signal processing, robust control, adaptive control.

This paper has been published before in [56].

E.1   Introduction

In a lot of applications, digital signals are transmitted over noisy and time disper-
sive channels, which cause Inter Symbol Interference (ISI). For example, in digital
mobile radio like GSM, in modem connections and digital magnetic recording sys-
tems. The distorted digital signal should be recovered. If the channel is constant
in time, this can be done in an optimal way by the Viterby algorithm [52], how-
ever this algorithm is computationally complex. A suboptimal method, which
approximates the performance of the Viterby algorithm with much lower compu-
tational complexity, is provided by the Decision Feedback Equalizer (DFE), see
e.g. [119, 162]. A DFE is a nonlinear filter, which consists of two linear filters and
a nonlinear decision circuit. The linear filters give an estimate of the transmit-
ted digital symbol by feed forward of the received samples and feedback of past
decisioned data. The decisioned data is the output of the DFE and obtained by
rounding the estimate to the closest symbol value. In most applications the channel
dynamics are not constant in time. Therefore, the linear filters should be adjusted
in time. This is partly solved by splitting up the data to be transmitted in a series
of bursts which contain training bits. The training bits can be used to estimate the
channel or the DFE filters directly. The remaining time variation in a burst can
be compensated by adaptive design methods (e.g. [98]), which are especially used
for fast time variations, or by robust design methods for slowly varying channels.
In [100, 165] design equations for a DFE which is robust for variations in the chan-
nel dynamics are given, and applied to the design of a DFE which is robust for
variations in mobile radio channels due to variations in the multi path propagation
of the radio waves. The design of the robust DFE was based on a robust open loop
filtering method, which uses probabilistic descriptions of model errors (related to
the stochastic embedding method [72]), and is referred to as the Cautious Wiener
(CW) approach [129, 164]. The advantage of this method is, that the filters are
optimized by taking the likelihood of the variations into account, rather than opti-
mizing for the worst case variation which might seldom happen. However, besides
the knowledge of the second order statistics of the uncertainty, this indirect method
requires a model of the average channel and the noise color. Using this information
the robust DFE is calculated by solving two coupled polynomial (Diophantine)
equations and a spectral factorization.

   Our contribution is a method to estimate a robust DFE directly from data, us-
ing second order statistics of the uncertainty in the channel too. The uncertainty
in the channel is modeled in the same way as in the CW method, such that the av-
eraged performance is optimized. The robust DFE is solved by a regularized least
squares problem, which is determined by the data and the second order statistics
of the uncertainty. So, we do not need a model of the channel and the noise color,
neither solving the coupled Diophantine equations and the spectral factorization.
We assume, that the training data used to design the DFE is coming from (approx-
imately) the average channel. Simulation experiments with a fading mobile radio
channel, showed that the obtained performance of this directly estimated robust
DFE was comparable to the indirectly estimated robust DFE of [165].

   The paper is organized as follows. Section E.2 describes the modeling of the
uncertainty in the channel and derives the robust performance criterion for the

robust DFE. Section E.3 describes the regularized least squares solution to the robust performance criterion. Section E.4 compares the robust DFE for a mobile radio application by simulation experiments and makes a comparison with the CW robust DFE. Section E.5 concludes the paper.

E.2 Problem setup

E.2.1 Remarks on notation

In this paper we assume, without loss of generality, the signals to be one-dimensional. A polynomial $P(q^{-1})$ in the unit shift backward operator q^{-1} of order n_p is denoted by $P(q^{-1}) = p_0 + p_1 q^{-1} + \cdots + p_{n_p} q^{-n_p}$, its complex conjugate is denoted by $P_*(q) = p_0^* + p_1^* q + \cdots + p_{n_p}^* q^{n_p}$. Transfer functions are denoted in calligraphic, e.g. $\mathcal{P}(q^{-1}) = P(q^{-1})/Q(q^{-1})$ with complex conjugate denoted as $\mathcal{P}_*(q) = P_*(q)/Q_*(q)$. Often, we omit the arguments q^{-1} and q. Further, we define the vector $\varphi_{n_p}(q^{-1}) := \begin{bmatrix} 1 & q^{-1} & \cdots & q^{-n_p} \end{bmatrix}^T$.

E.2.2 Modeling of time variations in the channel

Figure E.1 illustrates the principle of decision feedback equalization. A sequence of digital symbols $d(k)$ (possibly complex valued) is transmitted through a time dispersive channel $\mathcal{C} := C/A$ and corrupted by noise, which color is determined by $\mathcal{M} := M/N$. The received signal $y(k)$ can be expressed as

$$y(k) = \mathcal{C}(q^{-1})d(k) + \mathcal{M}(q^{-1})v(k) \qquad (E.1)$$

with $v(k)$ zero mean white noise with variance $\mathrm{E}[|v(k)|^2] = \sigma_v^2$ and uncorrelated with $d(k-i)$ for all i. We assume, that the transmitted symbols $d(k)$ are zero mean and white, with variance $\mathrm{E}[|d(k)|^2] = \sigma_d^2$. We want to reconstruct $d(k)$ using the received signal $y(k)$. This is done by a DFE, which is tuned during a training period in which a sequence of N_{tr} a priori known symbols $d_{tr}(k)$ is transmitted
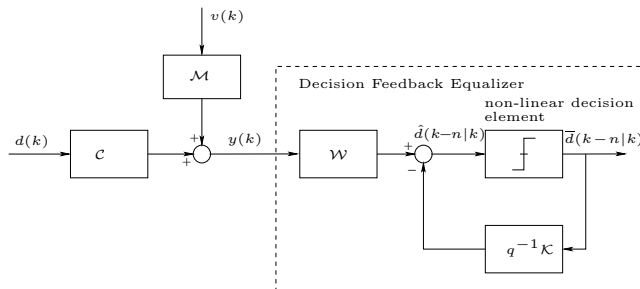

Figure E.1: The DFE, $\mathcal{W}$, $\mathcal{K}$ with the non-linear decision element, provide the estimate $\overline{d}(k-n|k)$ of the transmitted message $d(k)$ corrupted by the channel $\mathcal{C}$ and noise $\mathcal{M}v(k)$.

and $y_{tr}(k)$ is received. The training data is indicated with $\{d_{tr}(k), y_{tr}(k)\}_{k=1}^{N_{tr}}$. The coefficients of the transfer function $\mathcal{C}$ are (slowly) time varying due to e.g. environmental changes. Hence, the DFE which is designed to be optimal for the training period, might be suboptimal outside this time period. Because the time variation of the channel outside the training period is unknown, we cannot design a DFE which is optimal during the normal data transmission. However, we can model the time variation as uncertainty on the channel during the training and design a DFE which is robust for this uncertainty. This means, that the average performance of the robust DFE should be better than the DFE which does not account for the uncertainty due to time variations. Therefore, the time variation in the channel during data transmission is modeled in a probabilistic manner as in the CW approach [129, 164], such that the likelihood of the variations in the channel is also taken into account. During the training period and during the normal data transmission the transfer function of the channel is given by

$$\mathcal{C}(q^{-1}) := \mathcal{C}_0(q^{-1}) \qquad \text{during training} \tag{E.2}$$

$$\mathcal{C}(q^{-1}) := \mathcal{C}_0(q^{-1}) + \Delta\mathcal{C}(q^{-1}) \quad \text{during data transmission} \tag{E.3}$$

with $\Delta\mathcal{C}$ an uncertain transfer function $\Delta\mathcal{C}$ is modeled as a complex stochastic variable in the frequency domain, with zero mean and known covariance function $\Phi_{\Delta\mathcal{C}}(\omega)$:

$$\mathrm{E}[\Delta\mathcal{C}(e^{-j\omega})] := 0, \qquad -\pi \le \omega < \pi \tag{E.4}$$

$$\mathrm{E}[\Delta\mathcal{C}(e^{-j\omega})\Delta\mathcal{C}_*(e^{j\omega})] := \Phi_{\Delta\mathcal{C}}(\omega), \quad -\pi \le \omega < \pi \tag{E.5}$$

The modeling of the uncertainty $\Delta\mathcal{C}$ is similar to the modeling of uncertainty in the CW approach. In the CW approach $\Delta\mathcal{C}$ is of the form [164]

$$\Delta\mathcal{C}(q^{-1}) := \frac{C_1(q^{-1})\Delta C(q^{-1})}{A_1(q^{-1})}$$

with C_1 and A_1 known polynomials and $\Delta C(q^{-1}) = \Delta c_0 + \Delta c_1 q^{-1} + \cdots + \Delta c_{n_c} q^{-n_c}$ a polynomial which coefficients are stochastic variables with zero mean and known covariances $\mathrm{E}[\Delta c_i \Delta c_j^*]$ collected in the matrix $R_{\Delta C}$. Then the covariance function $\Phi_{\Delta\mathcal{C}}(\omega)$ is given by

$$\Phi_{\Delta\mathcal{C}}(\omega) = \frac{C_1(e^{-j\omega})C_{1*}(e^{j\omega})\varphi_{n_c}^T(e^{-j\omega})R_{\Delta C}\varphi_{n_c}(e^{j\omega})}{A_1(e^{-j\omega})A_{1*}(e^{j\omega})}$$

In the method proposed in this paper, we only need the inverse Fourier transform of $\Phi_{\Delta\mathcal{C}}(\omega)$, which can be determined e.g. by using the polynomials used in the CW method, i.e. A_1, C_1 and $\varphi_{n_c}^T(q^{-1})R_{\Delta C}\varphi_{n_c}(q)$.

E.2.3   Formulation of the robust DFE criterion

The DFE $d(k-n)$ by $\overline{d}(k-n|k)$ (with n a user chosen smoothing lag) using the received samples $y(i)$ ($i \le k$). The estimate $\overline{d}(k-n|k)$ is given by the digital symbol closest to the estimate $\hat{d}(k-n|k)$ made by

$$\hat{d}(k-n|k) = \mathcal{W}(q^{-1})y(k) + \mathcal{K}(q^{-1})\overline{d}(k-n-1|k-1) \tag{E.6}$$

with $\mathcal{W} = W/R$ and $\mathcal{K} = K/P$ the feed forward and feedback filters that need to be designed respectively. Due to the nonlinear decision element in the DFE, the design of an optimal DFE becomes very complicated. Therefore, errors in the decisioned data $\overline{d}(k - n|k)$ are modeled as zero mean white noise $\kappa(k - n)$ with variance $\sigma_\kappa^2 = \eta\sigma_d^2$, as proposed in [165]

$$\overline{d}(k - n|k) \;=\; d(k - n) + \kappa(k - n) \tag{E.7}$$

with $\kappa(k)$ independent of $d(k - i)$ and $v(k - i)$ for all i. Modeling the decision errors in this way, the feedback of the decisioned data can be interpreted as feed forward of the symbol $d(k-n)$ corrupted by noise $\kappa(k-n)$, see Figure E.2. Now, by removing the nonlinear decision element, the equalization problem can be solved by minimizing the Mean Squared Error (MSE) $\mathrm{E}[|d(k-n) - \hat{d}(k-n|k)|^2]$. However, we want to design a DFE which is robust for the uncertainty $\Delta\mathcal{C}$, hence the DFE will be designed by minimizing the MSE averaged over the distribution of $\Delta\mathcal{C}$. So we want to minimize the criterion used in the CW approach

$$J \;:=\; \overline{\mathrm{E}}\big[\mathrm{E}\big[|d(k-n) - \hat{d}(k-n|k)|^2\big]\big] \tag{E.8}$$

with $\overline{\mathrm{E}}[.]$ and $\mathrm{E}[.]$ denote expectation over $\Delta\mathcal{C}$ and over the noise quantities d, v and κ respectively.

E.3 Estimation of the robust DFE

E.3.1 Minimization of the criterion

Making use of the definition of $\hat{d}(k - n|k)$ by respectively (E.6) and (E.7), the definition of $y(k)$ by (E.1) and the channel $\mathcal{C}$ by (E.3), we can formulate the criterion (E.8) via Parseval's identity [79] in the frequency domain

$$J = \frac{1}{2\pi}\int_{-\pi}^{\pi} \overline{\mathrm{E}}\Big(|e^{-j\omega n} - \mathcal{W}(\mathcal{C}_0 + \Delta\mathcal{C}) + \mathcal{K}e^{-j\omega(n+1)}|^2\sigma_d^2 + |\mathcal{W}|^2|\mathcal{M}|^2\sigma_v^2 + |\mathcal{K}|^2\sigma_\kappa^2\Big)d\omega$$
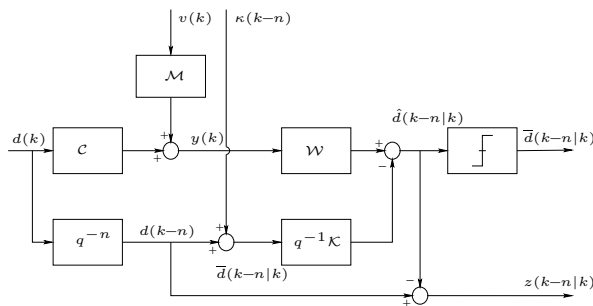

Figure E.2: Design scheme for the DFE assuming the errors in the decisioned data to be the white noise process $\kappa(k)$.

Because $\Delta\mathcal{C}$ is zero mean and its covariance is given by (E.5), this criterion J can be rewritten as

$$J = \frac{1}{2\pi}\int_{-\pi}^{\pi}|e^{-j\omega n} - \mathcal{W}\mathcal{C}_0 + \mathcal{K}e^{-j\omega(n+1)}|^2\sigma_d^2 + |\mathcal{W}|^2|\mathcal{M}|^2\sigma_v^2 + |\mathcal{W}|^2\Phi_{\Delta\mathcal{C}}\sigma_d^2 +$$
$$+ |\mathcal{K}|^2\sigma_\kappa^2 \; d\omega \tag{E.9}$$

The first two terms in the integrand determine the MSE when there is no uncertainty in the channel ($\Phi_{\Delta\mathcal{C}}(\omega) = 0$) and the decisioned data $\overline{d}(k{-}n|k)$ is assumed to be error free ($\sigma_\kappa = 0$). From the third term, we infer that by minimizing the criterion the gain of the feed forward filter $\mathcal{W}$ is decreased in the frequency region where the uncertainty in the channel is large. Note, that the same effect would have been obtained by measurement noise (uncorrelated with d and κ) with spectrum $|\mathcal{M}|^2\sigma_v^2 + \Phi_{\Delta\mathcal{C}}\sigma_d^2$ in the case the channel is the nominal channel $\mathcal{C}_0$. From the fourth term, we infer that by minimizing the criterion the gain of $\mathcal{K}$ is decreased over all frequencies, because we assumed $\kappa(k)$ to be white. This was also concluded in [165] where a minimization of (E.9) resulted in a spectral factorization and two coupled Diophantine equations for the filter coefficients of $\mathcal{W}$ and $\mathcal{K}$, which require a model of the average channel $\mathcal{C}_0$, the second order statistics $\Phi_{\Delta\mathcal{C}}$ of the uncertainty in the channel and the noise color $\mathcal{M}$. In the following, we will indicate how to minimize (E.9) in the time domain using the training data and the second order statistics $\Phi_{\Delta\mathcal{C}}$ of the uncertainty of the channel. Let us introduce the signal $d_f(k)$, with power spectrum $\Phi_{d_f}(\omega)$ defined as

$$\Phi_{d_f}(\omega) \; := \; \Phi_{\Delta\mathcal{C}}(\omega)\sigma_d^2 \tag{E.10}$$

Further, let us distinguish the individual contributions to J:

$$J_0 := \tfrac{1}{2\pi}\int_{-\pi}^{\pi}|e^{-j\omega n} - \mathcal{W}\mathcal{C}_0 + \mathcal{K}e^{-j\omega(n+1)}|^2\sigma_d^2 \; d\omega$$

$$J_v := \tfrac{1}{2\pi}\int_{-\pi}^{\pi}|\mathcal{W}|^2|\mathcal{M}|^2\sigma_v^2 \; d\omega$$

$$J_{d_f} := \tfrac{1}{2\pi}\int_{-\pi}^{\pi}|\mathcal{W}|^2\Phi_{d_f}(\omega) \; d\omega$$

$$J_\kappa := \tfrac{1}{2\pi}\int_{-\pi}^{\pi}|\mathcal{W}|^2\sigma_\kappa^2 \; d\omega$$

so $J \;=\; J_0 + J_v + J_{d_f} + J_\kappa$. By Parseval's rule we have:

$$J_0 = \mathrm{E}\big[\big(d(k - n) - \mathcal{W}\mathcal{C}_0 d(k) + \mathcal{K}d(k - n - 1)\big)^2\big]$$

$$J_v = \mathrm{E}\big[\big(\mathcal{W}\mathcal{M}v(k)\big)^2\big]$$

$$J_{d_f} = \mathrm{E}\big[\big(\mathcal{W}d_f(k)\big)^2\big]$$

$$J_\kappa = \mathrm{E}\big[\big(\mathcal{K}\kappa(k)\big)^2\big]$$

Because $\mathrm{E}[d(k)v^*(k{-}i)] = 0$ for all i, we can write

$$J_0 + J_v = \mathrm{E}\Big[\big(d(k - n) - \mathcal{W}\big(\mathcal{C}_0 d(k) + \mathcal{M}v(k)\big) + \mathcal{K}d(k - n - 1)\big)^2\Big]$$

Hence, (E.9) can be written as

$$J = \mathrm{E}\left[\left|d(k-n) - \mathcal{W}\big(\mathcal{C}_0 d(k) + \mathcal{M}v(k)\big) + \mathcal{K}d(k-n-1)\right|^2\right] + \mathrm{E}\left[\left|\mathcal{W}d_f(k)\right|^2\right] + \\ + \mathrm{E}\left[\left|\mathcal{K}\kappa(k)\right|^2\right]$$

Using the training data $d_{tr}(k)$, $y_{tr}(k) = \mathcal{C}_0 d_{tr}(k) + \mathcal{M}v(k)$ $(k = 1, ..., N_{tr})$ and replacing the ensemble average with the time average over the available training data in the first term of J, the criterion can be approximated with

$$\hat{J} = \frac{1}{N_{tr}-m} \sum_{k=m+1}^{N_{tr}} \left|d_{tr}(k-n) - \hat{\mathcal{W}}y_{tr}(k) + \hat{\mathcal{K}}d_{tr}(k-n-1)\right|^2 + \mathrm{E}\left[\left|\hat{\mathcal{W}}d_f(k)\right|^2\right] \\ + \mathrm{E}\left[\left|\hat{\mathcal{K}}\kappa(k)\right|^2\right] \tag{E.11}$$

with m determined by the smoothing lag n and the degrees of the numerator and denominator polynomials of $\hat{\mathcal{W}}$ and $\hat{\mathcal{K}}$. We can solve for the IIR filters $\hat{\mathcal{W}} = \hat{W}/\hat{R}$ and $\hat{\mathcal{K}} = \hat{K}/\hat{P}$ by minimizing (E.11) if $\sigma_\kappa \neq 0$. Note, that the criterion is not convex in the parameters, so local minima may exist. The problem of non-convexity can be solved by using subspace identification methods [181], which give unique state-space models of $\hat{\mathcal{W}}$ and $\hat{\mathcal{K}}$ (up to a similarity transformation) which are close to the optimal transfer functions, but do not explicitly minimize the MSE. In the case $\sigma_\kappa = 0$, the input data is not informative enough to estimate the polynomials $\hat{W}$, $\hat{R}$, $\hat{K}$ and $\hat{P}$, due to feedback of $d_{tr}(k-n-1)$ [104]. In this case, we have to restrict to estimate FIR filters, so $\hat{\mathcal{W}} = \hat{W}$ and $\hat{\mathcal{K}} = \hat{K}$.

E.3.2 Estimation of a robust DFE containing FIR filters

We will work out the estimation of the time domain robust DFE in the case that $\hat{\mathcal{W}}$ and $\hat{\mathcal{K}}$ are FIR filters. Let us define

$$\hat{\mathbf{w}} := \begin{bmatrix} \hat{w}_0 & \cdots & \hat{w}_{n_w} \end{bmatrix}^T, \quad \hat{\mathbf{k}} := \begin{bmatrix} \hat{k}_0 & \cdots & \hat{k}_{n_k} \end{bmatrix}^T$$
$$\mathbf{y}_{n_w}(k) := \begin{bmatrix} y_{tr}(k) & \cdots & y_{tr}(k-n_w) \end{bmatrix}^T$$
$$\mathbf{d_f}_{n_w}(k) := \begin{bmatrix} d_f(k) & \cdots & d_f(k-n_w) \end{bmatrix}^T$$
$$\mathbf{d}_{n_k}(k) := \begin{bmatrix} d_{tr}(k) & \cdots & d_{tr}(k-n_k) \end{bmatrix}^T$$
$$\kappa_{n_k}(k) := \begin{bmatrix} \kappa(k) & \cdots & \kappa(k-n_k) \end{bmatrix}^T$$

So, the criterion (E.11) can be written as

$$\hat{J} = \frac{1}{N_{tr}-m} \sum_{k=m+1}^{N_{tr}} \left|d_{tr}(k-n) - \hat{\mathbf{w}}^T\mathbf{y}_{n_w}(k) + \hat{\mathbf{k}}^T\mathbf{d}_{n_k}(k-n-1)\right|^2 + \mathrm{E}[|\hat{\mathbf{w}}^T\mathbf{d_f}_{n_w}(k)|^2] + \\ + \mathrm{E}[|\hat{\mathbf{k}}^T\kappa_{n_k}(k)|^2]$$

with $m := \max(n_w, n + n_k + 1)$. Further, define

$$Y := \begin{bmatrix} \mathbf{y}_{n_w}(m+1) & \cdots & \mathbf{y}_{n_w}(N_{tr}) \end{bmatrix}^T$$
$$D := \begin{bmatrix} \mathbf{d}_{n_k}(m-n) & \cdots & \mathbf{d}_{n_k}(N_{tr}-n-1) \end{bmatrix}^T$$
$$\mathbf{d} := \begin{bmatrix} d_{tr}(m-n+1) & \cdots & d_{tr}(N_{tr}-n) \end{bmatrix}^T$$

By constraining the gradient of $\hat{J}$ to be zero, we obtain the following set of normal equations for the filter coefficients

$$\left(\frac{1}{N_{tr}-m} \begin{bmatrix} Y^T \\ -D^T \end{bmatrix} \begin{bmatrix} Y^T \\ -D^T \end{bmatrix}^H + \begin{bmatrix} \mathrm{E}[\mathbf{d_f}_{n_w}(k)\mathbf{d_f}_{n_w}^H(k)] & 0 \\ 0 & \mathrm{E}[\kappa_{n_k}(k)\kappa_{n_k}^H(k)] \end{bmatrix} \right) \begin{bmatrix} \hat{\mathbf{w}}^* \\ \hat{\mathbf{k}}^* \end{bmatrix} =$$
$$= \frac{1}{N_{tr}-m} \begin{bmatrix} Y^T \\ -D^T \end{bmatrix} \mathbf{d}^* \tag{E.12}$$

with $(.)^H$ the Hermitian transpose operator and $\mathrm{E}[\kappa_{n_k}(k)\kappa_{n_k}^H(k)] = \sigma_\kappa^2 I_{n_k+1}$. $\mathrm{E}[\mathbf{d_f}_{n_w}(k)\mathbf{d_f}_{n_w}^H(k)]$ can be calculated using (E.10). Note that

$$\mathrm{E}[\mathbf{d_f}_{n_w}(k)\mathbf{d_f}_{n_w}^H(k)] = \begin{bmatrix} R_{d_f}(0) & \cdots & R_{d_f}(n_w) \\ \vdots & \ddots & \vdots \\ R_{d_f}^*(n_w) & \cdots & R_{d_f}(0) \end{bmatrix}$$

with $R_{d_f}(i) = \mathrm{E}[d_f(k)d_f^*(k-i)]$ which can be calculated using the inverse Fourier transform of the spectrum of $d_f(k)$ (E.10) $R_{d_f}(i) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{d_f}(\omega)e^{j\omega i}d\omega$. Now, (E.12) can be solved for the filter coefficients of the robust DFE stored in $\hat{\mathbf{w}}$ and $\hat{\mathbf{k}}$. Note, that Y and D are Toeplitz matrices, which structure might be exploited to obtain a fast implementation to calculate $\hat{\mathbf{w}}$ and $\hat{\mathbf{k}}$. The block diagonal covariance matrix with diagonal entries $\mathrm{E}[\mathbf{d_f}_{n_w}(k)\mathbf{d_f}_{n_w}^H(k)]$ and $\mathrm{E}[\kappa_{n_k}(k)\kappa_{n_k}^H(k)]$ in (E.12), can be considered as a regularization of the nominal DFE least squares problem [85]. In general $\Phi_{\Delta\mathcal{C}}(\omega)$ is a function of frequency and not a scalar, i.e. the uncertainty is frequency dependent. So, there exists $i \neq 0$ such that $R_{d_f}(i) \neq 0$ and the matrix $\mathrm{E}[\mathbf{d_f}_{n_w}(k)\mathbf{d_f}_{n_w}^H(k)]$, used for the regularization, is not just diagonal. Usually, for simplicity reasons, least squares problems are regularized with a scalar regularization parameter [70], to obtain a robust solution. Recently, [23] proposed a method to determine the optimal value of the scalar regularization parameter, given a bound on the uncertainty on the data matrix $\begin{bmatrix} Y & -D \end{bmatrix}$. However, regularization with a scalar only, does not take into account the frequency dependency of the uncertainty. The method described in this paper, can be seen as a generalization that explicitly takes the frequency dependency of the uncertainty into account.

E.3.3   Adaptive estimation of a robust DFE

For the FIR filter case, $\hat{\mathcal{W}} = \hat{W}$ and $\hat{\mathcal{K}} = \hat{K}$, the optimization problem (E.11) can also be minimized (approximately) in an adaptive way using e.g. the Least Mean Squares (LMS) method. The filter coefficients of the time varying filters $\hat{W}(k,q^{-1})$ and $\hat{K}(k,q^{-1})$ are contained in the vectors $\hat{\mathbf{w}}(k)$ and $\hat{\mathbf{k}}(k)$ respectively. The criterion $\hat{J}$ is replaced by

$$\hat{J}(k) = \left| d_{tr}(k-n) - \hat{\mathbf{w}}^T(k)\mathbf{y}_{nw}(k) + \hat{\mathbf{k}}^T(k)\mathbf{d}_{nk}(k-n-1) \right|^2$$
$$+ \mathrm{E}[|\hat{\mathbf{w}}^T(k)\mathbf{d_f}_{nw}(k)|^2] + \mathrm{E}[|\hat{\mathbf{k}}^T(k)\kappa_{nk}(k)|^2]$$

The filter coefficients are updated by

$$\hat{\mathbf{w}}(k+1) = \hat{\mathbf{w}}(k) - \mu_w \left(\frac{\partial \hat{J}(k)}{\partial \hat{\mathbf{w}}(k)}\right)^T$$

$$\hat{\mathbf{k}}(k+1) = \hat{\mathbf{k}}(k) - \mu_k \left(\frac{\partial \hat{J}(k)}{\partial \hat{\mathbf{k}}(k)}\right)^T$$

with μ_w and μ_k are user chosen step sizes and

$$\left(\frac{\partial \hat{J}(k)}{\partial \hat{\mathbf{w}}(k)}\right)^T = 2\big(\mathbf{y}_{nw}(k)d_{tr}^*(k-n) - \mathbf{y}_{nw}(k)\mathbf{y}_{nw}^H(k)\hat{\mathbf{w}}^*(k)+$$

$$+\mathbf{y}_{nw}(k)\mathbf{d}_{nk}^H(k-n-1)\hat{\mathbf{k}}^*(k) + \mathrm{E}[\mathbf{d}_{\mathbf{f}nw}(k)\mathbf{d}_{\mathbf{f}nw}^H(k)]\hat{\mathbf{w}}^*(k)\big)$$

$$\left(\frac{\partial \hat{J}(k)}{\partial \hat{\mathbf{k}}(k)}\right)^T = 2\big(-\mathbf{d}_{nk}(k-n-1)d_{tr}^*(k-n) + \mathbf{d}_{nk}(k-n-1)\mathbf{y}_{nw}^H(k)\hat{\mathbf{w}}^*(k)+$$

$$-\mathbf{d}_{nk}(k-n-1)\mathbf{d}_{nk}^H(k-n-1)\hat{\mathbf{k}}^*(k) - \mathrm{E}[\kappa_{nk}(k)\kappa_{nk}^H(k)]\hat{\mathbf{k}}^*(k)\big)$$

This yields a robust adaptive method to update the DFE filters.

E.4 Application to mobile radio channel equalization

E.4.1 Mobile radio channel

In mobile radio channels, like in GSM, the transmitted (base band) signal $d(k)$ is distorted due to multi path propagation of the radio waves and noise. Due to the mobility of the mobile receiver and changes in the environment, the multi path propagation is not constant in time. To compensate for the time variation in a burst, a robust DFE is estimated as described in Section E.3. We compare this robust DFE also with the robust DFE obtained by using the CW approach taken in [100, 165]. In the experiment, the transmitted data is a binary PAM signal, i.e. $d(k) \in \{-1, +1\}$. Each burst consists of 100 data symbols and 30 training symbols, the latter are located in the middle of the burst. The channel is Rayleigh fading with three independent real valued FIR coefficients [98] $C(k, q^{-1}) = c_0(k) + c_1(k)q^{-1} + c_2(k)q^{-2}$ with $\mathrm{E}[c_i(k)] = 0$ $(i = 0, 1, 2)$ and

$$\mathrm{E}[\mathbf{c}(k+l)\mathbf{c}^T(k)] = \begin{bmatrix} 1.00 & 0.00 & 0.00 \\ 0.00 & 0.66 & 0.00 \\ 0.00 & 0.00 & 0.33 \end{bmatrix} J_0(2\pi f_c \frac{v}{c_0} l) \qquad \text{(E.13)}$$

with $\mathbf{c}(k) = [c_0(k)\ c_1(k)\ c_2(k)]^T$, $J_0(.)$ the zero order Bessel function of the first kind, the carrier frequency $f_c = 1800\text{MHz}$. The velocity of the mobile receiver and the speed of light are denoted by v and c_0 respectively. Figure E.3 illustrates the time variation of the channel. It shows the variation of the coefficients over 10 bursts for a mobile receiver moving at 50km/h and at 300km/h, we see there is much more variation in the channel at high velocities. In the experiment, the noise is white ($\mathcal{M} = 1$) with standard deviation σ_v. The signal to noise ratio is defined as

$$SNR = 20 \log \frac{\sigma_d}{\sigma_v} \mathrm{E}||C(k, q^{-1})||_2 \qquad \text{(dB)}$$

with $\sigma_d = 1$ and $\text{E}||C(k, q^{-1})||_2$ the average 2-norm of the channel. Further, the time between two subsequent symbols is equal to the sample time $T_s = 3, 7\mu s$.

E.4.2   Modeling of uncertainty

In the CW approach [165], the estimated channel $\hat{C}_0$ is used as the average channel C_0. The time variation of the channel and the uncertainty in the identification are modeled by the uncertain polynomial $\Delta C(q^{-1})$, with covariance matrix [99]

$$R_{\Delta C} = R_{var} + R_{id}$$

R_{id} follows from the identification of C_0 and R_{var} can be calculated using the distribution of the channel coefficients. To determine R_{var}, we will follow [99]. The uncertainty in the channel during one burst is determined by R_{burst}

$$R_{burst} = \frac{1}{T_b} \int_{t=0}^{T_b} \big(\mathbf{c}(t) - \mathbf{c}(\frac{T_b}{2})\big)\big(\mathbf{c}(t) - \mathbf{c}(\frac{T_b}{2})\big)^T dt \qquad (E.14)$$

with $T_b = 130 \cdot Ts = 0,48ms$ the duration of the burst and (with some abuse of notation) $\mathbf{c}(t)$ the vector containing the coefficients of the channel at time t (relative to the beginning of the burst). Because $\mathbf{c}(t)$ is unknown, R_{burst} cannot be calculated explicitly. However, we can average R_{burst} over all bursts using the distribution of the channel (E.13)

$$R_{var} = 2R_c\Big(1 - \frac{2}{T_b} \int_0^{Tb/2} J_0(2\pi f_c \frac{v}{c_0}\tau) \ d\tau\Big) \qquad (E.15)$$

with $R_c = \text{E}[\mathbf{c}(k)\mathbf{c}^T(k)]$, which can be estimated by averaging over a number of bursts. Then, the robust DFE can be calculated by the CW approach using the estimated channel $\hat{C}_0$ and the covariance matrix $R_{\Delta C}$. When we directly estimate a robust DFE from the training data as described in Section E.3 we should use another covariance matrix $R_{\Delta C}$ for two reasons. First, we don't have to compensate for uncertainty in the identification of the channel. Second, the nominal DFE estimated from the training data is optimal for the training period, so it compensates already for variations in the training period. Because the variation
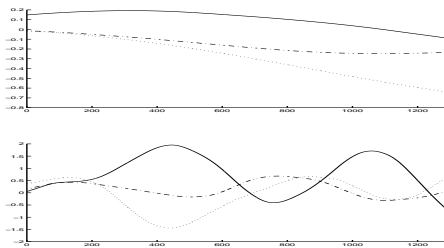

PSfrag replacements

first tap

Figure E.3: Fading channel taps (first: solid, second: dashed, third: dash-dotted) during the first 10 bursts for a mobile receiver moving with 50km/h (upper) and 300 km/h (lower).

of the channel coefficients is smooth, we may assume, that the nominal DFE is approximately optimal just outside the training period. Therefore, we can use a smaller covariance matrix, $R_{\Delta C}$, given by $R_{\Delta C} = \alpha R_{var}$, $(0 < \alpha < 1)$ with R_{var} given by (E.15). In the simulation example we used $\alpha = \frac{1}{N_{tr}-m} = 0.04$.

E.4.3 Simulation results

A series of simulations are done at various velocities of the mobile receiver (ranging from 50km/h to 300km/h) and different SNR's (ranging from 10dB to 30dB). In each simulation about 750 bursts were transmitted, the bursts were separated by 2 symbols, to account for the smoothing lag $n = 2$. We compared the average Bit Error Rate (BER) obtained by directly estimated nominal and robust DFE's and indirectly determined nominal and robust DFE's according to the CW approach. For the indirect approach the nominal DFE consists of a 2nd order feed forward filter and a 1st order feedback filter, which orders are optimal. The robust DFE consists of IIR filters, the feed forward filter and feedback filter have numerator polynomials of order 2 and 1 respectively, their denominator polynomials are equal and of order 1. For the direct approach, the nominal and robust DFE's consists of FIR filters which has the same order as for the indirectly estimated nominal DFE. Because of the FIR structure the robust DFE can be calculated as indicated in Section E.3.2. The standard deviation of $\kappa(k)$ was chosen $\sigma_\kappa = 0$, so we assumed the past decisioned data to be correct. However, additional robustness can be obtained by increasing σ_κ. Figure E.4 shows the average BER versus the SNR at various speeds of the mobile receiver by using the indirect CW method and the direct method. The performance of the indirectly estimated nominal and robust DFE's are indicated by regular solid and dashed curves respectively. The performance of the directly estimated nominal and robust DFE's are indicated by circle marked solid and dashed curves respectively. We infer, that in the direct and the indirect case the performance of the DFE can be improved by taking the uncertainty due to time variations in the channel into account, especially at high SNR's and high velocities. The performance of the directly estimated robust DFE approximates the performance of the indirectly estimated robust DFE (based on CW), designed using two coupled Diophantine equations and a spectral factorization. Further, note that the performance of the indirectly estimated nominal DFE is getting worse at high SNR's. This can be explained by the fact, that the nominal DFE's which are designed to be optimal at high SNR's are more sensitive to perturbations like variations in the channel. The Viterby algorithm has also be used to estimate the transmitted signal. However it's performance was much worse than the performance of the nominal and robust DFE's, because the Viterby algorithm is not optimal for time varying channels.

E.5 Conclusions

A method to estimate a robust Decision Feedback Equalizer (DFE) directly from data using regularized least squares was derived. The DFE is robust with respect to time variations in the channel, which are modeled as probabilistic uncertainty on the channel like in the Cautious Wiener approach. The robust DFE is a solution
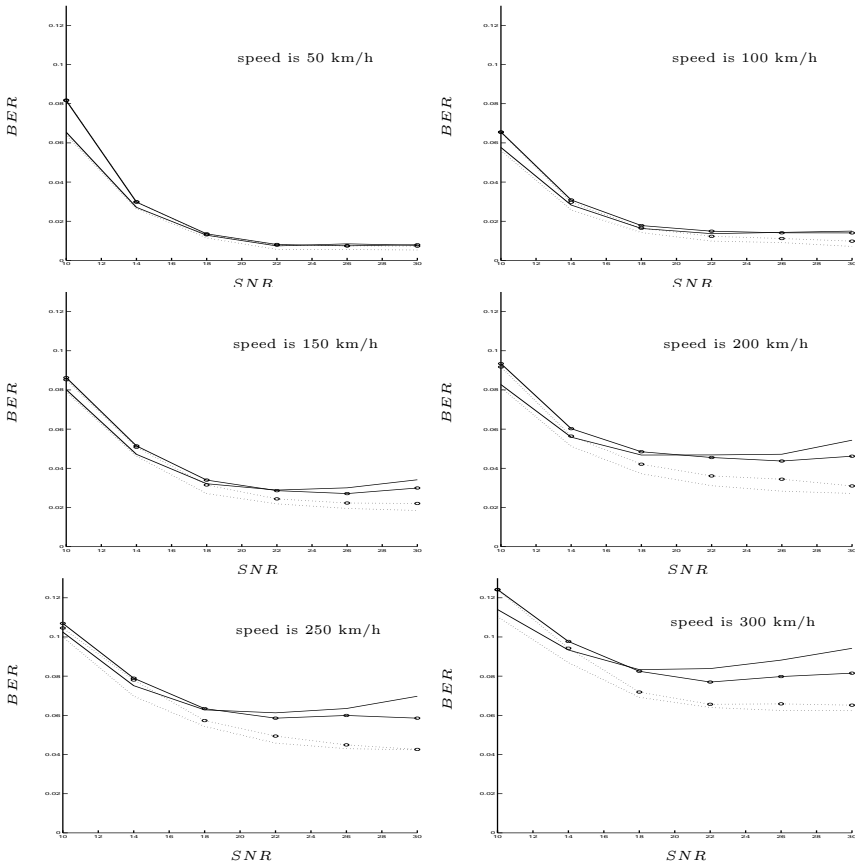
Figure E.4: BER versus SNR curves at 50-300km/h, obtained with the indirect nominal (solid) and robust (dashed) DFE and obtained with the direct nominal (solid with circles) and robust (dashed with circles) DFE.

to a least squares problem which is regularized using second order statistics of the uncertainty. This regularized least squares problem is obtained by minimizing the mean squared error (MSE) averaged over the distribution of the uncertainty in the channel. A simulation example with a fading mobile radio channel, showed that performance of the directly estimated robust DFE is comparable to the indirectly estimated robust DFE. The performance of the (directly and indirectly) nominal DFE could be improvewd by the robust DFE, especially when the variation in the channel and the signal to noise ratio are large. An advantage of the directly estimated robust DFE is, that no spectral factorization is necessary, and no model of the average channel and the noise color are necessary anymore. The structure in the least squares problem might be used to obtain fast implementations to calculate the DFE filters. Also adaptive estimation of a robust DFE was proposed, by minimizing the averaged MSE by the LMS-algorithm. Further analysis of this method and is left for further research.

Summary

Robust and fast schemes in broadband active noise and vibration control

This thesis presents robust and fast active control algorithms for the suppression of broadband noise and vibration disturbances. Noise disturbances, e.g., generated by engines in airplanes and cars or by airflow, can be reduced by means of *passive* or *active* methods. In passive methods, the undesired noise disturbances are isolated or damped with passive isolation or damping material. Since the required thickness of passive isolation material is determined by the largest wavelength of the disturbing noise, for low frequencies often thick and heavy materials are needed, which are undesired, e.g., in airplanes. The active methods are based on a different principle, namely on the principle of interference with a counteracting wave. This counteracting wave, generated by loudspeakers, is ideally exactly the opposite of the disturbing wave.

Similarly, the *active* control approach can also be used to suppress vibration disturbances in mechanical structures, e.g., to reduce the sound radiation of the structure due to the vibrations. The vibrations are suppressed by generating counteracting vibrations in the structure by means of, e.g., piezoceramic actuators or voice coil shakers. From the perspective of the design of the control algorithm, the active control of noise and vibration is very similar, and therefore we consider a generic control problem that covers both of these applications, and refer to it as the active noise and vibration control (ANVC) problem.

In ANVC the systems to be controlled are usually infinite dimensional systems, which can be approximated by finite dimensional systems in a limited bandwidth (typically 0-1kHz). However, since acoustical and mechanical systems often contain many resonance frequencies, these finite dimensional approximations need to be of high order. When the *delay* between the actuators and sensors can be arbitrarily reduced by collocation of actuator and sensor pairs, often simple (analogue) lead-, lag-, and proportional-integral-derivative controllers are effective solutions. However, in many situations the delay cannot be arbitrarily reduced and thus to retain closed loop stability the bandwidth should be significantly reduced. In this case model-based control methods can increase the bandwidth and improve the performance without leading to instability. However, since models always contain errors, the uncertainty in the models need to be accounted for in every model-based controller design.

This thesis considers the robust controller design for feedforward and feedback ANVC systems with stochastic broadband disturbances. The additional difficulty in comparison with harmonic disturbances is first that the disturbing frequencies are distributed over a particular bandwidth, which requires relatively complex con-

trollers with prescribed amplitude and phase behavior in this band. The second difficulty is, that the disturbance is stochastic and therefore the future values of the disturbance are not known precisely, which limits the freedom to design the controller.

Currently in ANVC applications the Filtered-X LMS adaptive algorithm is mainly used to compute the control signal. This algorithm has some attractive properties, such as its simplicity, its well understood convergence behavior and robustness properties. The latter are inherited from the H_∞ optimality of the LMS algorithm (c.f. [74]) and the passivity of the adaptive update rule if a particular strictly positive real condition relating the model and the true system is satisfied (c.f. [188]). Though the algorithm is well suited for applications with harmonic disturbances, for broadband applications its convergence rate is slow and the computational complexity increases considerably, especially in multichannel applications. This thesis proposes novel algorithms and offers insight to improve the convergence properties and to reduce the computational complexity of the Filtered-X LMS algorithm.

Chapter 2 lays the mathematical basis of the thesis and discusses the H_2 optimal feedforward and feedback control problems following a factorization approach, and clarifies their tight relation by means of the Youla parameterization. The effect of noise on the input signal (i.e., the reference signal in feedforward configurations) is analyzed, as well as the effect of control effort weighting. Both will reduce the gain of the optimal feedforward controller, or Youla parameter, in specific frequency bands.

Chapter 3 discusses the nominal controller problem and therefore considers first the accurate identification of models. To this end a comparison has been made between the identification of a prediction error model with the output error structure (PEM-OE) and the identification of a state-space model by a subspace model identification method, more precisely the PO-MOESP method [181]. It is shown that a fast implementation of the PO-MOESP algorithm outperforms the PEM-OE method, in identifying a 20th order vibrating plate system with 4 inputs and 4 outputs. The second part of the chapter proposes a method for the control-relevant estimation of the H_2 optimal controller for feedforward and feedback applications. The basic idea in control-relevant estimation of a controller (or a factor determining the controller), is that in the controller design the control cost function is explicitly minimized using measured data, rather than using a cost function that is fully determined by (uncertain) models of the system. It is shown, that uncertainty in the spectral factors of the reference signal and the disturbance signal can be partly compensated for, leading to better estimates of the H_2 optimal controller.

Chapter 4 considers the robust controller design problem by taking the uncertainty in the model explicitly into account by means of a probabilistic approach. The probabilistic approach was initially proposed for feedforward control and filtering problems by Sternad and Ahlén [164] in a polynomial model framework. In this thesis this approach has been adopted for probabilistic controller design and extended it to the state-space model framework, where the B and D or the C and D state-space model matrices are distorted by stochastic uncertainty with zero mean and known covariance. Additional frequency dependent weighting of

the uncertainty can also be included in the uncertainty model description. Using this uncertainty model a robust feedforward controller has been designed which optimizes the H_2 optimal performance averaged over the model uncertainty distribution. Similar to the control-relevant estimation of the nominal controller as discussed in Chapter 3, it is shown that the control-relevant estimation can be extended to estimate the robust feedforward controller. Furthermore, for the feedback application it has been shown, that using this robust feedforward controller in combination with Internal Model Control feedback compensation, the stability robustness of the closed loop has been improved.

Chapter 5 presents a theoretical spin-off of Chapter 2 with respect of the convergence analysis of the Filtered-U LMS algorithm, which is the infinite impulse response (IIR) equivalent of the Filtered-X LMS algorithm. The convergence analysis of adaptive algorithms which update the coefficients of an IIR filter is in literature well known to be difficult, since the cost function is usually a non-convex function of the adaptive filter coefficients. Recently, Wang and Ren [187] have derived conditions for the global convergence of the Filtered-U LMS algorithms, under the restrictive assumption that perfect cancellation of all disturbances (up to measurement noise) would be achievable. This assumption is often not satisfied in practice due to delays and non-minimum phase zeros in the system. This chapter shows that this assumption can be omitted, by analysis of the structure of the optimal controller and the remaining residual disturbance. Based on the structure of the optimal solution of the feedforward controller a preconditioning of the Filtered-U LMS algorithm has been proposed to increase its convergence rate. The preconditioning is analogous to the preconditioning of the Filtered-X LMS algorithm proposed by Elliott et al. [47].

Chapter 6 extends the probabilistic robust controller design of Chapter 4 to adaptive algorithms. The robustness of the Filtered-X LMS algorithm and the preconditioned Filtered-X LMS algorithm with respect to modeling errors is increased by taking model uncertainty explicitly into account. The new robust adaptive update algorithms are obtained by using LMS estimates of the gradient of the mean squared error averaged over the model uncertainty distribution. The robustness of the adaptive algorithm is effectively improved without losing too much performance as would be the case when using a leaky version of the Filtered-X LMS algorithm. However, the price paid is an increase of computational complexity since the number of columns of the regressor is augmented.

The last contribution of this thesis is contained in Chapter 7, which analyzes the feedforward active control problem in a Kalman-filter state-estimation context. This context may be especially attractive for applications where the (preconditioned) Filtered-X LMS does not provide a satisfactory convergence behavior. The active control problem is reformulated into a state estimation problem containing the adaptive filter coefficients and the state of the system. It has been shown that the algorithm reduces to the modified Filtered-X RLS algorithm, see e.g., [51], when the state of the system can be assumed to be perfectly known which is a simplification in practice. A fast array implementation of the Kalman filter algorithm has been derived to obtain an implementation which is more feasible in practical applications.

Concluding, this thesis proposes a set of algorithmic tools which may serve

the practical active controller design to improve robustness, to reduce convergence times and to lower computational complexity. The research leads also to new research questions which may yield further improvements of algorithms for ANVC, as outlined in Chapter 8.

Samenvatting

Robuuste en snelle ontwerpen in het breedbandig actief regelen van stoorgeluid en trillingen

Dit proefschrift presenteert robuuste en snelle actieve regel algoritmes voor de onderdrukking van breedbandig stoorgeluid en storende trillingen. Stoorgeluid, bijvoorbeeld gegenereerd door motoren in vliegtuigen of auto's of door luchtstroming, kan gereduceerd worden door *passieve* of *actieve* methodes. In passieve methodes worden de ongewenste stoorgeluiden geïsoleerd of gedempt door passief isolatie- of dempingsmateriaal. Omdat de benodigde dikte van het passieve isolatiemateriaal wordt bepaald door de grootste golflengte van het stoorgeluid, zijn voor lage frequenties vaak dikke en zware materialen nodig, hetgeen niet gewenst is in, bijvoorbeeld, vliegtuigen. De actieve methoden zijn gebaseerd op een ander principe, namelijk het principe van interferentie met een tegenwerkende golf. Deze tegenwerkende golf, gegenereerd door luidsprekers, is in het ideale geval exact tegengesteld aan de storende golf.

Op dezelfde manier kan de *actieve* regelbenadering gebruikt worden om storende trillingen in mechanische structuren te onderdrukken, bijvoorbeeld om de geluidsafstraling van deze structuren te reduceren. De trillingen worden onderdrukt door in de mechanische structuur tegengolven te genereren door middel van, bijvoorbeeld, piezokeramische actuatoren of voice-coil schudders. Vanuit het perspectief van het ontwerp van het regelalgoritme is het actief regelen van stoorgeluid en trillingen vrijwel aan elkaar gelijk, en daarom beschouwen we een generiek regelprobleem welke beide toepassingen omvat, en verwijzen hiernaar als het probleem van het actief regelen van stoorgeluid en trillingen (in engels: Active Noise and Vibration Control, ANVC).

In ANVC zijn de systemen, die moeten worden geregeld, gewoonlijk oneindigdimensionale systemen, die alleen kunnen worden benaderd door eindig dimensionale systemen in een beperkte bandbreedte (typisch 0-1kHz). Echter, omdat akoustische en mechanische systemen vaak veel resonantie-frequenties hebben, dienen deze eindig dimensionale benaderingen van hoge orde te zijn. Indien de *vertraging* tussen de actuatoren en de sensoren willekeurig verminderd kan worden door het gelijk positioneren van actuator-sensor paren, kan er volstaan worden met eenvoudige (analoge) lead-, lag- en proportional-integral-derivate regelaars. In veel gevallen kan de vertraging niet willekeurig worden verminderd en moet de bandbreedte aanzienlijk worden gereduceerd om de stabiliteit van de gesloten lus te behouden. In deze gevallen kunnen modelgebaseerde regelaars de bandbreedte vergroten en de prestatie verhogen zonder dat dit leidt tot instabiliteit. Echter, omdat modellen altijd fouten bevatten, moet onzekerheid in het model in rekening worden gebracht in het modelgebaseerd regelaarontwerp.

Dit proefschrift beschouwt het ontwerpen van robuuste regelaars voor feedforward en feedback ANVC systemen met stochastische breedbandige verstoringen. De extra moeilijkheid in vergelijking met harmonische verstoringen is allereerst dat de storende frequenties zijn verdeeld over een bepaalde bandbreedte, wat een relatief complexe regelaar vereist met voorgeschreven amplitude- en fase-gedrag in deze band. De tweede moeilijkheid is, dat de verstoring stochastisch is en daarom niet precies bekend is in de toekomst, hetgeen de vrijheid in het ontwerp van de regelaar beperkt.

Momenteel wordt in ANVC toepassingen voornamelijk het Filtered-X LMS algoritme gebruikt om het stuursignaal uit te rekenen. Dit algoritme heeft een aantal aantrekkelijke eigenschappen, zoals zijn eenvoud, zijn goed bekende convergentie-eigenschappen en zijn robuustheids-eigenschappen. Deze laatste worden verkregen vanwege de H_∞ optimaliteit van het LMS algoritme (v.g.l. [74]) en de passiviteit van het update algoritme indien voldaan wordt aan een bepaalde strikt positief reëel conditie welke het model en het echte systeem aan elkaar relateert (v.g.l. [188]). Ofschoon het algoritme goed geschikt is voor toepassingen met harmonische verstoringen, is voor breedbandige toepassingen de convergentie traag en kan de rekenkundige complexiteit aanzienlijk toenemen, vooral in multikanaals toepassingen. Dit proefschrift stelt nieuwe algoritmes voor en levert inzicht om de convergentie-eigenschappen van het Filtered-X LMS algoritme te verbeteren en de rekenkundige complexiteit te verminderen.

Hoofdstuk 2 legt de wiskundige basis van het proefschrift en beschouwt de H_2 optimale feedforward en feedback problemen overeenkomstig een factorisatie benadering, en verduidelijkt de nauwe relatie tussen feedforward en feedback problemen aan de hand van de Youla parameterizatie. Het effect van ruis op het input signaal (d.w.z., het referentie-signaal in feedforward configuraties) wordt geanalyseerd, evenals het effect van een weging op het niveau van het stuursignaal. Beide reduceren de versterking van de optimale feedforward regelaar, of de Youla parameter, in bepaalde frequentiebanden.

Hoofdstuk 3 gaat in op het nominale regelprobleem en beschouwt daarvoor allereerst het nauwkeurig identificeren van modellen. Hiervoor wordt een vergelijking gemaakt tussen de identificatie van een prediction error model met output error structuur (PEM-OE) en de identificatie van een state-space model door middel van een subspace model identificatie methode, meer precies de PO-MOESP methode [181]. Het blijkt dat een snelle implementatie van het PO-MOESP algoritme betere resultaten geeft dan de PEM-OE methode in het identificeren van een 20e orde trillende plaat systeem met 4 ingangen en 4 uitgangen. Het tweede deel van het hoofdstuk stelt een methode voor om de H_2 optimale regelaar voor feedforward en feedback toepassingen te schatten door middel van een control-relevante identificatie. De kern gedachte in control-relevant schatten van de regelaar (of een factor die de regelaar bepaalt), is dat de regelaar wordt ontworpen door expliciet de kostenfunctie voor het regelprobleem te minimaliseren gebruikmakend van gemeten data, in plaats van het minimaliseren van een kostenfunctie die volledig wordt bepaald door (onzekere) modellen van het systeem. Op deze manier kan onzekerheid in de spectrale factoren van het referentie- en het verstoringssignaal gedeeltelijk worden gecompenseerd, hetgeen leidt tot betere schattingen van de H_2 optimale regelaar.

Hoofdstuk 4 beschouwt het ontwerpen van een robuuste regelaar door de onzekerheid van het model expliciet in rekening te brengen door middel van een probabilistische benadering. De probabilistische benadering werd aanvankelijk door Sternad en Ahlén [164] voorgesteld voor feedforward- en filterproblemen in het kader van polynomiale-modellen. In dit proefschrift wordt deze probabilistische benadering overgenomen en uitgebreid voor state-space modellen, waarbij de B en D of de C en D matrices verstoord zijn door stochastische onzekerheid met een gemiddelde van nul en een bekende covariantie. Extra frequentie-afhankelijke weging van de onzekerheid kan worden meegenomen in de onzekerheidsmodel-beschrijving. Gebruikmakend van dit onzekerheidsmodel wordt een robuuste feedforward regelaar ontworpen welke de H_2 optimale prestatie, gemiddeld over de verdeling van de modelonzekerheid, optimaliseert. Evenals bij de control-relevante schatting van de nominale regelaar in Hoofdstuk 3, wordt er aangetoond dat de control-relevante schatting kan worden uitgebreid voor het schatten van de robuuste feedforward regelaar. Verder, blijkt dat voor feedback toepassingen gebruikmakend van deze robuuste feedforward regelaar in combinatie met Internal Model Control, de robuustheid van de stabiliteit van de gesloten lus wordt verbeterd.

Hoofdstuk 5 presenteert een theoretische spinoff van Hoofdstuk 2 met betrekking tot de convergentie analyse van het Filtered-U LMS algoritme, welke de infinite impulse response (IIR) uitbreiding is van het Filtered-X LMS algoritme. De analyse van de convergentie van adaptieve algoritmes welke de coëfficiënten van een IIR filter adapteren, is in de literatuur bekend als een moeilijk probleem, omdat de kostenfunctie in het algemeen een niet-convexe functie is van de adaptieve filtercoëfficiënten. Recentelijk, hebben Wang en Ren [187] voorwaarden afgeleid voor de globale convergentie van het Filtered-U LMS algoritme, onder de beperkende aanname dat perfecte uitdoving van alle stoorsignalen (op meetruis na) mogelijk is. Aan deze aanname wordt in de praktijk vaak niet voldaan vanwege vertraging en niet-minimumfase nulpunten in het systeem. In dit hoofdstuk wordt er aangetoont dat deze aanname overbodig is, door analyse van de structuur van de optimale regelaar en de resterende verstoring. Gebaseerd op de structuur van de optimale oplossing van de feedforward regelaar wordt er tevens een preconditionerings-methode voor het Filtered-U LMS algoritme voorgesteld, waarmee de convergentie snelheid wordt verbeterd. Deze preconditionering is analoog aan de preconditionering van het Filtered-X LMS algoritme zoals voorgesteld door Elliott e.a. [47].

Hoofdstuk 6 breidt het ontwerp van de probabilistische robuuste regelaar van Hoofdstuk 4 uit voor adaptieve algoritmes. De robuustheid van het Filtered-X LMS algoritme en het preconditioneerde Filtered-X LMS algoritme met betrekking tot modelfouten wordt vergroot door expliciet de modelonzekerheid in rekening te brengen. De nieuwe robuuste adaptieve algoritmes worden verkregen door LMS schattingen van de gradient van de gemiddelde kwadratische fout gemiddeld over de modelonzekerheids-verdeling. De robuustheid van het adaptieve algoritme wordt effectief vergroot, zonder teveel prestatie te verliezen, zoals het geval bij het gebruik van de leaky variant van het Filtered-X LMS algoritme. Echter, de prijs die wordt betaald is een toename in de rekenkundige complexiteit omdat het aantal kolommen van de regressor wordt uitgebreid.

De laatste bijdrage van dit proefschrift is beschreven in Hoofdstuk 7, waarin het probleem van feedforward actief regelen wordt geanalyseerd in een Kalman-

filter toestandsschattings probleem. Deze context is met name aantrekkelijk voor toepassingen waarin het (preconditioneerde) Filtered-X LMS algoritme niet het benodigde convergentiegedrag levert. Het probleem van het actief regelen wordt hergeformuleerd in een toestandsschattings probleem, waarbij de toestand bestaat uit de adaptieve filtercoëfficiënten en de toestand van het systeem. Er wordt aangetoond dat het algoritme vereenvoudigt tot het modified Filtered-X RLS algoritme, zie, bijvoorbeeld, [51], wanneer de toestand van het systeem perfect bekend kan worden verondersteld, hetgeen een vereenvoudiging is van de praktijk. Een fast-array implementatie van het Kalman-filter algoritme is afgeleid om een implementatie te verkrijgen die beter is geschikt voor praktische toepassingen.

Concluderend, dit proefschrift stelt een aantal algorithmische instrumenten voor die van dienst kunnen zijn bij het actief regelen in de praktijk om robuustheidseigenschappen te verbeteren en om convergentie-tijden en rekenkundige complexiteit te verminderen. Het onderzoek leidt ook tot nieuwe onderzoeksvragen, die tot verdere verbeteringen van algoritmes voor ANVC kunnen leiden, zoals in hoofdlijnen aangegeven in Hoofdstuk 8.

Bibliography

[1] M. Abdelghani, M. Verhaegen, P. Van Overschee, and B. De Moor. Comparison study of subspace identification methods applied to flexible structures. *Mechanical Systems and Signal Processing*, 12(5):679–692, 1998.

[2] A. Ahlén and M. Sternad. Wiener filter design using polynomial equations. *IEEE Trans. Signal Processing*, SP-39(11), 1991.

[3] B.D.O. Anderson and J.B. Moore. *Optimal Control – Linear Quadratic Methods*. Prentice-Hall, Englewood Cliffs, NJ, USA, 1989.

[4] T. Auspitzer, B. Rafaely, and S.J. Elliott. A fast adaptive feedback-controller for active noise control. In *Proc. of Internoise 96*, pages 1021–1024, Liverpool, UK, Aug. 1996.

[5] B. Bamieh, F. Paganini, and M.A. Dahleh. Distributed control of spatially invariant systems. *IEEE Trans. on Automat. Control*, AC-27(7):1091–1107, Jul. 2002.

[6] C.A.J. Beijers and A. de Boer. Hybrid isolation of structure-borne sound. In *Proc. of ACTIVE 2002*, pages 291–302, Southampton, UK, Jul. 2002.

[7] B. Berglund, T. Lindvall, and D.H. Schwela (ed.). Guidelines for community noise. Technical report, World Health Organization, Geneva, Switzerland, April 1999. `http://www.euro.who.int/Noise/Publications/20030528_2`.

[8] A.P. Berkhoff. Piezoelectric sensor configuration for active structural acoustic control. *Journal of Sound and Vibration*, 246(1):175–183, Sep. 2001.

[9] B. Bernhardsson. Robust performance optimization of open loop type problems using models from standard identification. *Systems & Control Letters*, 25(2):79–87, May 1995.

[10] D.S. Bernstein and W.M. Haddad. LQG control with an H_∞ performance bound: a Riccati equation approach. *Trans. Automat. Control*, AC-34(3):293–305, Mar. 1989.

[11] E. Bianchi, P. Gardonio, and S.J. Elliott. Smart panel with multiple decentralised units for the control of sound transmission. part iii: Control system implementation. In *Proc. of ACTIVE 2002*, pages 499–510, Southamption, UK, Jul. 2002.

[12] G. Billoud. Active control at lord corporation – a reality. Technical Report LL-6508, Lord Corporation, 2003. `http://www.lordmpd.com/inside/tech_papers.asp`.

[13] B. Bingham, M.J. Atalla, and N.W. Hagood. Comparison of structural-acoustic control designs on an vibrating composite panel. *Journal of Sound and Vibration*, 244(5):761–778, 2001.

[14] S. Bittanti and M. Lovera. Bootstrap-based estimates of uncertainty in subspace identification methods. *Automatica*, 36(11):1605–1615, November 2000.

[15] E. Bjarnason. Active noise cancellation using a modified form of the filtered-X LMS algorithm. In *Proc. 6th European Signal Processing Conf.*, volume 2, pages 1053–1056, Brussels, Belgium, Aug. 1992.

[16] M. Bodson, J.S. Jensen, and S.C. Douglas. Active noise control for periodic disturbances. *IEEE Trans. on Contr. Syst. Techn.*, 9(1):200–205, Jan. 2001.

[17] I.U. Borchers et al. Advanced study of active noise control in aircraft. In J.M.M. Hernandez, editor, *Advances in Acoustics Technology*, pages 1–81. Wiley, Chichester, UK, 1994.

[18] M. Bouchard. Multichannel affine and fast affine projection algorithms for active noise control and acoustic equalization systems. *IEEE Trans. Speech Audio Processing*, SAP-11(1):54–60, Jan. 2003.

[19] M. Bouchard and F. Albu. The multichannel Gauss-Seidel fast affine projection algorithm for active noise control. In *Proc. Seventh International Symposium on Signal Processing and Its Applications (ISSPA)*, volume 2, pages 579–582, Paris, France, Jul. 2003.

[20] J. Braakhuis. Linear subspace identification of car models from measurements. Master's thesis, Delft University of Technology, 1998.

[21] J.C. Burgess. Active adaptive sound control in a duct: A computer simulation. *J. Acoust. Soc. Am.*, 70(3):715–726, Sep. 1981.

[22] J.C. Carmona and V.M. Alvarado. Active noise control of a duct using robust control theory. *IEEE Trans. Contr. Syst. Techn.*, 8(6):930–938, Nov. 2000.

[23] S. Chandrasekaran, G.H. Golub, M. Gu, and A.H. Sayed. Parameter estimation in the presence of bounded data uncertainties. *SIAM Journal on Matrix Analysis and Applications*, 19(1):235–252, 1998.

[24] T. Chen and B. Francis. *Optimal Sampled-Data Control Systems*. Springer, 1996.

[25] T. Chen and B.A. Francis. State-space solutions to discrete-time and sampled-data H_2 control problems. In *Proc. of the 31st Conference on Decision and Control*, pages 1111–1116, Tucson, AZ, US, Dec. 1992.

[26] T. Chen and B.A. Francis. State-space solutions to discrete-time and sampled-data H_2 control problems. Technical Report 9205, Dept. of Elect. Eng., Univ. of Toronto, Toronto, Canada, Aug. 1992.

[27] N.L.C. Chui and J.M. Maciejowski. Realization of stable models with subspace methods. *Automatica*, 32(11):1587–1595, Nov. 1996.

[28] H. Coanda. Procédé de protection contre les bruits. French Patent FR 722.274, Oct. 1930.

[29] M.H. Costa, J.C.M. Bermudez, and N.J. Bershad. Stochastic analysis of the Filtered-X LMS algorithm in systems with nonlinear secondary paths. *IEEE Trans. Signal Processing*, SP-50(6):1327–1342, Jun. 2002.

[30] R.F. Curtain and H.J. Zwart. *An Introduction to Infinite-Dimensional Linear Systems Theory*. Springer-Verlag, New York, USA, 1991.

[31] R. D'Andrea and G.E. Dullerud. Distributed control design for spatially interconnected systems. *IEEE Trans. Automat. Contr.*, AC-48(9):1478–1495, Sep. 2003.

[32] M. de Boer. Control of a thermal deformable mirror — correction of a light beam with limited sensor information. Master's thesis, Delft University of Technology, Delft, The Netherlands, Oct. 2003.

[33] R. de Callafon. *Feedback Oriented identification for Enhanced and Robust Control – a fractional approach applied to a wafer stage*. PhD thesis, Delft University of Technology, 1998.

[34] P. De Fonseca, P. Sas, and H. Van Brussel. A comparative study of methods for optimising sensor and actuator locations in active control applications. *Journal of Sound and Vibration*, 221(4), 1999.

[35] W. Dehandschutter. *The reduction of structure-borne noise by active control of vibration*. PhD thesis, KU Leuven, Leuven, Belgium, 1997.

[36] P. Dewilde and A.-J. van der Veen. *Time-Varying Systems and Computations*. Kluwer, Boston, MA, USA, 1998.

[37] N.J. Doelman. Control strategies for active noise reduction. In *Proc. of Internoise 90*, pages 1273–1276, Gothenburg, Sweden, Aug. 1990.

[38] N.J. Doelman. *Design of Systems for Active Sound Control*. PhD thesis, Delft University of Technology, Delft, The Netherlands, 1993.

[39] N.J. Doelman. Minimum variance control of random sound. In *Active Control of Vibration and Noise, ASME 1994*, volume 75, pages 35–42, Nov. 1994.

[40] N.J. Doelman. Adaptive and robust systems for the active control of noise and vibration. In *Proc. Adaptronic*, 1999.

[41] S.C. Douglas. The fast affine projection algorithm for active noise control. In *Proc. 29th Asilomar Conf. Sign., Syst., Comp.*, volume 2, pages 1245–1249, Pacific Grove, CA, USA, Oct. 1995.

[42] J.C. Doyle. Guaranteed margins for lqg regulators. *IEEE Trans. Automat. Contr.*, AC-23(4):756–757, Aug. 1987.

[43] J.C. Doyle, K. Glover, P.P. Khargonekar, and B.A. Francis. State-space solutions to standard H_2 and H_∞ control problems. *IEEE Trans. Automat. Contr.*, AC-34(8):831–847, Aug. 1989.

[44] J.C. Doyle and G. Stein. Robustness with observers. *IEEE Trans. Automat. Contr.*, AC-25(4):607–611, Aug. 1979.

[45] S.J. Elliott. Optimal controllers and adaptive controllers for multichannel feedforward control of stochastic disturbances. *IEEE Trans. Signal Processing*, SP-48(4):1053–1060, Apr. 2000.

[46] S.J. Elliott. *Signal Processing for Active Control*. Academic Press, London, UK, 2001.

[47] S.J. Elliott and J.G. Cook. A preconditioned LMS algorithm for rapid adaptation of feedforward controllers. In *Proc. Int. Conf. Acoust., Speech and Sign. Proc. (ICASSP)*, pages 845–849, Istanbul, Turkey, Jun. 2000.

[48] S.J. Elliott and T.J. Sutton. Performance of Feedforward and Feedback Systems for Active Control. *IEEE Trans. Speech Audio Processing*, 4(3):214–223, May 1996.

[49] L.J. Eriksson, M. Allie, and R. Greiner. The selection and application of an IIR adaptive filter for use in active sound control. *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-35(4):433–437, Oct. 1987.

[50] P.L. Feintuch. An Adaptive Recursive LMS Filter. *Proc. IEEE*, 64(11):1622–1624, Nov. 1976.

[51] S.J. Flockton. Fast adaptation algorithms in active noise control. In *Second Conference on Recent Advances in Active Noise Control of Sound and Vibration*, pages 802–810, Virginia, Apr. 1993.

[52] G.D. Forney, Jr. The Viterbi algorithm. *Proc. IEEE*, 61:268–278, 1973.

[53] R. Fraanje, A.H. Sayed, M. Verhaegen, and N.J. Doelman. A Fast-Array Kalman Filter Solution to Active Noise Control. *submitted to the Int. J. Adaptive Control and Signal Processing*, Nov. 2003.

[54] R. Fraanje, M. Verhaegen, and N.J. Doelman. Blockwise Subspace Identification for Active Noise Control. In *Proc. 39th Conf. on Decision and Control*, pages 2397–2402, Sydney, Australia, Dec. 2000.

[55] R. Fraanje, M. Verhaegen, and N.J. Doelman. MIMO H_2 optimal controller design by Subspace Identification for Active Control Applications. Technical Report UT/TN/SCE-2001-11, University of Twente, Enschede, The Netherlands, Oct. 2001.

[56] R. Fraanje, M. Verhaegen, and N.J. Doelman. Robust decision feedback equalizer design via the solution of a regularized least squares problem. In *Proceedings of the European Control Conference*, Porto, Portugal, Sep. 4-7 2001.

[57] R. Fraanje, M. Verhaegen, and N.J. Doelman. Subspace identification for active noise control in a duct. In *Proc. Int. Con. and Exh. Noise Contr. Eng. (Inter-Noise)*, pages 2531–2534, The Hague, The Netherlands, Aug. 2001.

[58] R. Fraanje, M. Verhaegen, and N.J. Doelman. General convergence analysis of the Filtered-U LMS algorithm for Active Noise Control in case perfect cancelation is not possible. *Signal Processing, Elsevier.*, 83(6):1239 – 1254, Jun. 2003.

[59] R. Fraanje, M. Verhaegen, and N.J. Doelman. Increasing the Robustness of a Preconditioned Filtered-X LMS algorithm . *IEEE Signal Processing Lett.*, 11(2):285–288, Feb. 2004.

[60] R. Fraanje, M. Verhaegen, N.J. Doelman, and A.P. Berkhoff. MIMO H_2 and Robust Feedback Controller Estimation for a Vibrating Plate using Subspace Model Identification. In P. Gardonio and B. Rafaely, editors, *Proceedings of the 2002 International Symposium on Active Control of Sound and Vibration (ACTIVE)*, pages 1251–1262, Southampton, United Kingdom, Jul. 2002.

[61] R. Fraanje, M. Verhaegen, N.J. Doelman, and A.P. Berkhoff. MIMO H_2 Feedback Controller Estimation for a Vibrating Plate using Subspace Model Identification. In P. Sas and B. van Hal, editors, *Proceedings of the International Conference on Noise & Vibration Engineering (ISMA)*, pages 201–209, Leuven, Belgium, Sep. 16-19 2002.

[62] R. Fraanje, M. Verhaegen, N.J. Doelman, and A.P. Berkhoff. MIMO H_2 optimal and Robust Feedback Controller Estimation for a Vibrating Plate using Subspace Model Identification. *Accepted for publication in Control Engineering Practice*, 2004.

[63] R. Fraanje, M. Verhaegen, V. Verdult, and R. Pintelon. A frequency domain subspace algorithm for mixed causal, anti-causal lti systems. In *Proc. of the 13th IFAC Symposium on System Identification (SYSID)*, pages 893–898, Rotterdam, The Netherlands, Aug. 2003.

[64] C.R. Fuller, S.J. Elliott, and P.A. Nelson. *Active Control of Vibration*. Academic Press Limited, London, UK, 1996.

[65] M. Gad-el-Hak. *Flow Control*. Cambridge University Press, Cambridge, UK, 2000.

[66] P. Gardonio, E. Bianchi, and S.J. Elliott. Smart panel with multiple decentralised units for the control of sound transmission. part i: Theoretical predictions. In *Proc. of ACTIVE 2002*, pages 471–486, Southampton, UK, Jul. 2002.

[67] P. Gardonio, E. Bianchi, and S.J. Elliott. Smart panel with multiple decentralised units for the control of sound transmission. part ii: Design of the decentralised control units. In *Proc. of ACTIVE 2002*, pages 478–498, Southamption, UK, Jul. 2002.

[68] M. Gevers and G. Li. *Parametrizations in Control, Estimation and Filtering Problems : Accuracy Aspects*. Communication and Control Engineering Series. Springer Verlag, New York, 1993.

[69] M.R. Gevers. Towards a joint design of identification and control? In H.L. Trentelman and J.C. Willems, editors, *Essays on Control: Perspectives in the Theory and its Applications*, pages 111–151. Birkäuser, Boston, 1993.

[70] G.H. Golub and C.F. van Loan. *Matrix Computations*. The Johns Hopkins University Press, 1996.

[71] C.G. Goodwin, editor. *Model Identification and Adaptive Control — From Windsurfing to Telecommunications*. Springer-Verlag, London, UK, 2001. a collection of contributions in honour of Professor Brian Anderson's 60th birthday.

[72] G.C. Goodwin, M. Gevers, and B. Ninnes. Quantifying the Error in Estimated Transfer Functions with Application to Model Order Selection. *IEEE Trans. Automat. Contr.*, AC-37(7):913–928, Jul. 1992.

[73] G.C. Goodwin and D.E. Miller. Robust performance optimization based on stochastic model errors: the stable case. *International Journal of Robust and Nonlinear Control*, 12(14):1191–1208, Dec. 2002.

[74] B. Hassibi. *Indefinite Metric Spaces in Estimation, Control and Adaptive Filtering*. PhD thesis, Stanford University, Stanford, CA, USA, 1996.

[75] B. Hassibi, T. Kailath, and A.H. Sayed. Array algorithms for H_∞-estimation. *IEEE Trans. on Automat. Control*, AC-45(4):702–706, Apr. 2000.

[76] B. Hassibi, A. Sayed, and T. Kailath. H_∞ Optimality of the LMS Algorithm. *IEEE Trans. Signal Processing*, SP-44(2), Feb. 1996.

[77] B. Haverkamp and M. Verhaegen. *State Space Model Identification Software for Multivariable Dynamical Systems*. Delft University of Technology, Systems and Control Engineering Group, Delft, The Netherlands, 1997.

[78] L.R.J. Haverkamp. *State space identification – Theory and practise*. PhD thesis, Delft University of Technology, 2001.

[79] M.H. Hayes. *Statistical Digital Signal Processing and Modeling*. Wiley, New York, USA, 1996.

[80] S. Haykin. *Adaptive filter theory*. Prentice-Hall, Englewood Cliffs, NJ, USA, 1996.

[81] R. Hinchliffe, I. Scott, M. Purver, and I. Stothers. Tonal active control in production of a large turbo-prop aircraft. In *Proc. of ACTIVE 2002*, pages 369–376, Southampton, UK, Jul. 2002.

[82] H. Hjalmarsson, M. Gevers, S. Gunnarsson, and O. Lequin. Iterative feedback tuning: theory and applications. *IEEE Control Systems Magazine*, 18(4):26–41, Aug. 1998.

[83] T. Ishihara and H. Takeda. Loop transfer recovery techniques for discrete-time optimal regulators using prediction estimators. *IEEE Trans. Automat. Contr.*, AC-31(12):1149–1151, Dec. 1986.

[84] C.R. Johnson jr. and M.G. Larimore. Comments on and Additions to "An Adaptive Recursive LMS Filter". *Proc. IEEE*, 65(9):1399–1401, Sep. 1977.

[85] T. Kailath, A.H. Sayed, and B. Hassibi. *Linear Estimation.* Prentice-Hall, Englewood Cliffs, NJ, USA, 2000.

[86] O.E. Kaiser. *Active Control of Sound Transmission through a Double Wall Structure.* PhD thesis, Swiss Federal Institute of Technology (ETH), Zürich, 2001.

[87] O.E. Kaiser, F.J. Kraus, and M. Morari. Robust active noise control in a duct. In *Proc. 4th Internat. Conf. on Motion and Vibration Control, MOVIC '98*, pages 61–66, Zürich, Switzerland, Aug. 1998.

[88] M. Krutzen. Control in adaptive optics for atmospheric aberration compensation. Master's thesis, University of Twente, Enschede, The Netherlands, Aug. 2001.

[89] S.K. Kung. A new low-order approximation algorithm via singular value decomposition. In *Proceedings 12th Asilomar Conference on Circuits, Systems and Computers*, pages 705–714, 1978.

[90] S.M. Kuo and D.R. Morgan. *Active Noise Control Systems — Algorithms and DSP Implementations.* Wiley, New York, USA, 1996.

[91] V. Kučera. *Discrete Linear Control.* Wiley, New York, USA, 1979.

[92] H. Kwakernaak. Robust control and H_∞ optimization – tutorial paper. *Automatica*, 29(2):255–273, Mar. 1993.

[93] H. Kwakernaak and R. Sivan. *Modern Signals and Systems.* Prentice-Hall, Englewood Cliffs, NJ, USA, 1991.

[94] T. Laakso, V. Välimäki, M. Karjalainen, and U.K. Laine. Splitting the unit delay – tools for fractional delay filter design. *IEEE Signal Processing Mag.*, 13(1):30–60, Jan. 1996.

[95] T.L. Lagö. Industry overview of active control methods and applications. In *Proc. of Adaptronic Congress*, Apr. 2002.

[96] W.S. Lee, B.D.O. Anderson, R.L. Kosut, and I.M. Y. Mareels. A new approach to adaptive robust control. *International journal of adaptive control and signal processing*, 7:183–211, 1993.

[97] J.-Y. Lin and Z.-L. Luo. Internal Model-Based LQG/H_∞ Design of Robust Active Noise Controllers for an Acoustic Duct System. *IEEE Trans. on Control Systems Technology*, 8(5):864–872, Sep. 2000.

[98] L. Lindbom. *A Wiener Filtering Approach to the Design of Tracking Algorithms*. PhD thesis, Uppsala University, 1995.

[99] E. Lindskog. *Space-time processing and equalization for wireless communications*. PhD thesis, Uppsala University, 1999.

[100] E. Lindskog, M. Sternad, and A. Ahlén. Designing decision feedback equalizers to be robust with respect to channel time variations. In *Proceedings of Nordic Radio Symposium seminar*, Uppsala, Sweden, Nov. 1993.

[101] L. Ljung. Analysis of Recursive Stochastic Algorithms. *IEEE Trans. Automat. Contr.*, AC-22(4):551–575, Aug. 1977.

[102] L. Ljung. On Positive Real Transfer Functions and the Convergence of Some Recursive Schemes. *IEEE Trans. Automat. Contr.*, AC-22(4):539–551, Aug. 1977.

[103] L. Ljung. Model validation and model error modelling. In B. Wittenmark and A. Rantzer, editors, *The Åström Symposium on Control*, pages 15–42, Lund, Sweden, Aug. 1999.

[104] L. Ljung. *System Identification - Theory for the User*. Prentice-Hall, Upper Saddle River, NJ, USA, 1999.

[105] L. Ljung. *System Identification Toolbox — For Use with* MATLAB. The Mathworks, Natick, MA, USA, user's guide version 5 edition, 2002.

[106] L. Ljung and T. Söderström. *Theory and Practice of Recursive Identification*. The MIT Press, Cambridge, MA, USA, 1983.

[107] P.A.C. Lopes. The Kalman filter in active noise control. In *Proc. of ACTIVE 99*, pages 1111–1122, 1999.

[108] P. Lueg. Process of silencing sound oscillations. US Patent 2,043,416, Mar. 1934.

[109] J.M. Maciejowski. Asymptotic recovery for discrete-time systems. *IEEE Trans. Automat. Contr.*, AC-30(6):602–605, Jun. 1985.

[110] K. Maouche and D.T.M. Slock. Fast subsampled-updating stabilized fast transversal filter (FSU SFTF) RLS algorithm for adaptive filtering. *IEEE Trans. Signal Processing*, AC-48(8):2248–2257, Aug. 2000.

[111] J. Mari, P. Stoica, and T. McKelvey. Vector ARMA esitmation: A reliable subspace approach. *IEEE Trans. Signal Processing*, SP-48(7):2092–2104, Jul. 2000.

[112] T. McKelvey. Frequency domain system identification with instrumental variable based subspace algorithm. In *Proceedings of the 1997 ASME design engineering technical conferences*, Sacramento, CA, USA, Sep. 14-17 1997.

[113] T. McKelvey, H. Akçay, and L. Ljung. Subspace-based multivariable system identification from frequency response data. *IEEE Trans. Automat. Contr.*, AC-41(7):960–979, Jul. 1996.

[114] T. McKelvey, A. Fleming, and S.O.R. Moheimani. Subspace-based system identification for an acoustic enclosure. *Transactions of the ASME, Journal of Vibration & Acoustics*, 124(3):414–419, Jul. 2002.

[115] T. Meurers and S.M. Veres. Iterative design of vibration attenuation. *Int. Journal of Acoustics and Vibration*, 4(2):79–83, 1999.

[116] T. Meurers and S.M. Veres. Further results on the iterative design for noise attenuation. *Int. Journal of Acoustics and Vibration*, 6(4):219–225, 2001.

[117] T. Meurers and S.M. Veres. Implementation aspects for fsf-based feedback control with secondary path estimation. In *Proc. of ACTIVE 2002*, pages 1327–1338, Southampton, UK, Jul. 2002.

[118] T. Meurers, S.M. Veres, and S.J. Elliott. Frequency selective feedback for active noise control. *IEEE Control Systems Magazine*, 22(4):32–41, Aug. 2002.

[119] P. Monsen. Feedback equalization for fading dispersive channels. *IEEE Trans. Inf. Th.*, IT-17:56–64, 1971.

[120] M. Morari and E. Zafiriou. *Robust Process Control*. Prentice-Hall, Englewood Cliffs, NJ, USA, 1989.

[121] J. Mørkholt. *Active control of noise radiation from vibrating structures*. PhD thesis, Technical University of Denmark, Lyngby, Denmark, 1999.

[122] J. Mørkholt and S.J. Elliott. Active vibration control using state space lqg and internal model control methods. In *Proc. 4th Internat. Conf. on Motion and Vibration Control, MOVIC '98*, Zürich, Switzerland, Aug. 1998.

[123] C. Mosquera, J.A. Gómez, F. Pérez, M. Sobreira, and E. Alexandre. Adaptive IIR filters for active noise control. *Int. J. Acoust. Vib.*, 6(1):3–10, Mar. 2001.

[124] C. Mosquera and F. Pérez. Algebraic solution to the robust SPR problem for two polynomials. *Automatica*, 37, May 2001.

[125] C. Mosquera and F. Pérez-González. Convergence analysis of the multiple-channel Filtered-U recursive LMS algorithm for active noise control. *Signal Processing*, 80(5):849–856, May 2000.

[126] P.A. Nelson and S.J. Elliott. *Active control of sound*. Academinc Press, London, UK, 1992.

[127] G. Nijsse, M. Verhaegen, and N.J. Doelman. State Space Modeling of Acoustics for Active Control. In *Proc. Int. Con. and Exh. Noise Contr. Eng. (Inter-Noise)*, pages 2527–2530, The Hague, The Netherlands, Aug. 2001.

[128] G. Nijsse, D. Westwick, B. De Schutter, and M. Verhaegen. Simplification of multichannel active noise control systems: a feasability study. Internal report ITS/ET/RT/SCE:0.05, Faculty of Information Technology and Systems, Department of Electrical Engineering, Delft University of Technology, Jul. 1999.

[129] K. Öhrn. *Design of Multivariable Cautious Discrete-time Wiener Filters.* PhD thesis, Uppsala University, 1996.

[130] K. Öhrn, A. Ahlén, and M. Sternad. A probabilistic approach to multivariable robust filtering and open-loop control. *IEEE Trans. on Automat. Control*, AC-40(3):405–418, Mar. 1995.

[131] H.F. Olson. Electronic sound absorber. US Patent 2,983,790, Apr. 1953.

[132] F. Paganini. Convex methods for robust H_2 analysis of continuous-time systems. *IEEE Trans. Automat. Control*, AC-44(2):239–252, Feb. 1999.

[133] W. Passchier-Vermeer, H.M.E. Miedema, H. Vos, H.J.M. Steenbekkers, D. Houthuijs, and S.A. Reijneveld. Sleep disturbance and aircraft noise exposure. Technical Report TNO Inro 2002.028 / RIVM 441520019, TNO Inro / RIVM, TNO Inro, Delft / RIVM, Bilthoven, The Netherlands, Nov. 2002. (in Dutch).

[134] I.R. Petersen. Multivariable control of noise in an acoustic duct. In *Proc. of the European Control Conference*, pages 414–419, Porto, Portugal, Sept. 2001.

[135] I.R. Petersen and H.R. Pota. Minimax LQG optimal control of a flexible beam. *Control Engineering Practice*, 11:1273–1287, Nov. 2003.

[136] I.R. Petersen, V. Ugrinovskii, and A.V. Savkin. *Robust Control Design using H_∞ methods.* Springer-Verlag, Londen, UK, 2000.

[137] R. Pintelon. Frequency-domain subspace system identification using non-parametric noise models. *Automatica*, 38:1295–1311, Aug. 2002.

[138] R. Pintelon and J. Schoukens. *System Identification – A Frequency Domain Approach.* IEEE Press, 2001.

[139] S. Prajna, O.E. Kaiser, S.J. Pietrzko, and M. Morari. Robust active control of a vibrating plate. In *Proceedings of Noise-Con 2000*, Newport Beach, CA, USA, Dec. 2000.

[140] A. Preumont. *Vibration Control of Active Structures.* Kluwer Academic Publishers, 1997.

[141] B. Rafaely and S.J. Elliott. H_2/H_∞ Active Control of Sound in a Headrest: Design and Implementation. *IEEE Trans. Contr. Syst. Techn.*, 7(1):79–84, 1999.

[142] W. Ren and P.R. Kumar. Adaptive active noise control: Structures, algorithms and convergence analysis. In *Proc. Inter-Noise 89*, pages 435–440, Newport Beach, CA, USA, Dec. 1989.

[143] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, New York, USA, 1987.

[144] M. Rupp and A. Sayed. On the stability and convergence of Feintuch's algorithm for adaptive IIR filtering. In *Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing (ICASSP)*, pages 1388–1391, Detroit, MI, USA, May 1995.

[145] M. Rupp and A.H. Sayed. Robust FxLMS algorithms with improved convergence performance. *IEEE Trans. Speech Audio Processing*, 6(1):78–85, Jan. 1998.

[146] H. Sano, T. Yamashita, and M. Nakamura. Recent applications of active noise and vibration control to automobiles. In *Proc. of ACTIVE 2002*, pages 29–42, Southampton, UK, Jul. 2002.

[147] A. Sayed, H. Lev-Ari, and T. Kailath. Time-variant displacement structure and triangular arrays. *IEEE Trans. Signal Processing*, SP-42(5):1052–1062, May 1994.

[148] A.H. Sayed. A framework for state-space estimation with uncertain models. *IEEE Trans. Automat. Contr.*, AC-46(7):999–1013, Jul. 2001.

[149] A.H. Sayed. *Fundamentals of Adaptive Filtering*. Wiley, New York, USA, 2003.

[150] A.H. Sayed and T. Kailath. Extended Chandrasekhar recursions. *IEEE Trans. on Automat. Control*, AC-39(3):619–623, Mar. 1994.

[151] A.H. Sayed and T. Kailath. A state-space approach to adaptive RLS filtering. *IEEE Signal Processing Mag.*, pages 18–60, Jul. 1994.

[152] A.H. Sayed, V. Nascimento, and S. Chandrasekaran. Estimation and control with bounded data uncertainties. *Linear Algebra and its Applications*, 284:259–306, Nov. 1998.

[153] B. Sayyar-Rodsari. *Estimation-based adaptive filtering and control*. PhD thesis, Stanford University, Stanford, CA, USA, 1999.

[154] B. Sayyar-Rodsari, J.P. How, B. Hassibi, and A. Carrier. Estimation-based synthesis of H_∞-optimal adaptive FIR filters for filtered-LMS problems. *IEEE Trans. on Signal Processing*, SP-49(1):164–178, Januari 2001.

[155] C. Scherer. Mixed H_2/H_∞ control. In A. Isidori, editor, *Trends in Control, A European Perspective*, pages 173–216. Springer-Verlag, Berlin, 1995.

[156] R. Schirmacher, T. Auspitzer, and D. Guicking. Fast algorithms for active adaptive control of nonstationary sound fields. *NAG (Nederlands Akoestisch Genootschap) Journaal*, (120):41–55, Nov. 1993.

[157] J.J. Shynk. Adaptive IIR Filtering. *IEEE Acoust., Speech and Signal Processing Mag.*, 6(2):4–21, Apr. 1989.

[158] V. Sima. Cholesky or QR factorization for data compression in subspace-based identification. In *Proceedings of the Second NICONET Workshop on "Numerical Control Software: SLICOT, a Useful Tool in Industry"*, pages 75–80, INRIA Rocquencourt, France, Dec. 1999.

[159] SLICOT. Niconet international society: The SLICOT package. Secretariat: Mrs.Ida Tassens, Department of Electrical Engineering, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, 3001 Leuven, Belgium, 2002. http://www.win.tue.nl/niconet/.

[160] P.C.W. Sommen and J. Garas. Using phase information to decorrelate the Filtered-X algorithm. In *Proc. Int. Conf. Acoust., Speech and Sign. Proc. (ICASSP)*, pages 1397–1400, Seattle, WA, USA, May 1998.

[161] T.C. Sors and S.J. Elliott. Volume velocity estimation with accelerometer arrays for active structural acoustic control. *Journal of Sound and Vibration*, 258(1):867–883, Dec. 2002.

[162] M. Sternad and A. Ahlén. The Structure and Design of Realizable Decision Feedback Equalizers for IIR Channels with Colored Noise. *IEEE Trans. Inf. Th.*, (4):848–858, Jul. 1990.

[163] M. Sternad and A. Ahlén. A novel methodology for polynomial-lq controller design. *IEEE Trans. Automat. Contr.*, AC-38(1), Jan. 1993.

[164] M. Sternad and A. Ahlén. Robust Filtering and Feedforward Control Based on Probabilistic Descriptions of Model Errors. *Automatica*, 29(3):661–679, May 1993.

[165] M. Sternad, A. Ahlén, and E. Lindskog. Robust decision feedback equalizers. In *Proceedings of ICASSP 1993*, volume III, pages 555–558, Minneapolis, MN, USA, Apr. 1993.

[166] A.A. Stoorvogel. *The H_∞ control problem: a state space approach*. Prentice-Hall, Englewood Cliffs, NJ, USA, 1992.

[167] A. Subramanian and A.H. Sayed. Regularized Robust Estimators for Time Varying Uncertain Discrete-Time Systems. In *Proceedings of the 13th IFAC Symposium on System Identification (SYSID 2003)*, pages 1060–1065, Rotterdam, The Netherlands, Aug. 2003.

[168] X. Sun and G. Meng. Comments on "Convergence analysis of the filtered-u LMS algorithm for active noise control" by A.K. Wang and W. Ren. *submitted to Signal Processing*, 2003.

[169] C. Talotte. Aerodynamic noise: a critical survey. *Journal of Sound and Vibration*, 231(3):549–562, Mar. 2000.

[170] C.K.W. Tam. Supersonic jet noise. *Annu. Rev. Fluid Mech.*, 27:17–43, 1995.

[171] T. Tay, I. Mareels, and J. Moore. *High Performance Control*. Birkhäuser, 1998.

[172] D.J. Thompson and C.J.C. Jones. A review of the modelling of wheel/rail noise generation. *Journal of Sound and Vibration*, 231(3):519–536, Mar. 2000.

[173] M.O. Tokhi and S.M. Veres, editors. *Active Sound and Vibration Control*. The Institution of Electrical Engineers, 2002.

[174] R.K. Tyson. *Principles of Adaptive Optics*. Academic Press, Boston, USA, 1991.

[175] P.M.J. van den Hof and R.J.P. Schrama. Identification and control – closed-loop issues. *Automatica*, 31(12):1751–1770, 1995.

[176] S.W. van der Hoeven. Prefiltering of the FxLMS algorithm applied to a vibrating plate system. Master's thesis, University of Twente, Enschede, The Netherlands, Mar. 2002.

[177] P. Van Overschee and B. De Moor. *Subspace Identification for Linear Systems*. Kluwer, 1996.

[178] V. Verdult. *Nonlinear system identification – a state space approach*. PhD thesis, University of Twente, 2002.

[179] V. Verdult and R. Fraanje. Slicot determines models for active noise control. In *NICONET Newsletter*, number 8, pages 13–18. Working Group on Software WGS, Feb. 2002. `http://www.win.tue.nl/niconet`.

[180] M. Verhaegen. Subspace model identification part 3. Analysis of the ordinary output-error state-space model identification algorithm. *International Journal of Control*, 56(3):555–586, 1993.

[181] M. Verhaegen. Identification of the deterministic part of MIMO state space models given in innovations form from input-output data. *Automatica*, 30(1):61–74, Jan. 1994.

[182] M. Verhaegen. A subspace model identification solution to the identification of mixed causal anti-causal lti systems. *SIAM Journal on Matrix Analysis*, 17(2):332–347, Apr. 1996.

[183] M. Verhaegen and P. Dewilde. Subspace model identification part 1. The output-error state-space model identification class of algorithms. *International Journal of Control*, 56(5):1187–1210, 1992.

[184] M. Verhaegen and P. Dewilde. Subspace model identification part 2. Analysis of the elementary output-error state-space model identification algorithm. *International Journal of Control*, 56(5):1211–1241, 1992.

[185] M. Viberg. Subspace-based methods for the identification of linear time-invariant systems. *Automatica*, 31(12):1835–1851, 1995.

[186] M. Vidyasagar. *Control systems synthesis: a factorization approach.* The MIT Press, Cambridge, MA, USA, 1985.

[187] A.K. Wang and W. Ren. Convergence analysis of the Filtered-U algorithm for active noise control. *Signal Processing*, 73(3):255–266, Jan. 1999.

[188] A.K. Wang and W. Ren. Convergence Analysis of the Multi-Variable Filtered-X LMS Algorithm with Application to Active Noise Control. *IEEE Trans. Signal Processing*, SP-47(4):1166–1169, Apr. 1999.

[189] B. Widrow and M. Hoff. Adaptive switching circuits. In *Proc. IRE WESCON Convention Record, Part 4, Session 16*, pages 96–104, 1960.

[190] B. Widrow and J.M. McCool. Comments on "An Adaptive Recursive LMS Filter". *Proc. IEEE*, 65(9):1402–1404, Sep. 1977.

[191] B. Widrow, D. Schur, and S. Shaffer. On adaptive inverse control. In *Proc. 15th Asilomar Conf. Circuits, Systems and Computers*, pages 185–195, Pacific Grove, CA, USA, Nov. 1981.

[192] World Health Organization. Occupational and community noise. `http://www.who.int/inf-fs/en/fact258.html`, Feb. 2001.

[193] D.C. Youla, J.J. Bongiorno, and H.A. Jabr. Modern Wiener-Hopf design of optimal controllers, Pt.1, The single-input-single-output case. *IEEE Trans. Automat. Contr.*, AC-21(1):3–13, 1976.

[194] D.C. Youla, J.J. Bongiorno, and H.A. Jabr. Modern Wiener-Hopf design of optimal controllers, Pt.2, The multiple-input-multiple-output case. *IEEE Trans. Automat. Contr.*, AC-21(3):319–338, 1976.

[195] F. Yu and M. Bouchard. Multichannel active noise control algorithms using inverse filters. In *Proc. Int. Conf. Acoust., Speech and Sign. Proc. (ICASSP)*, pages 825–828, Istanbul, Turkey, Jun. 2000.

[196] H. Zhang. Numerical condition of polynomials in different forms. *Electronic Transactions on Numerical Analysis*, 12:66–87, 2001.

[197] M. Zhang, E. Shafai, and H.P. Geering. Robust adaptive control for active noise cancellation in a duct. In *Proc. of ACTIVE 99*, pages 1025–1034, Ft. Lauderdale, Fl, USA, Dec. 1999.

[198] K. Zhou and X. Chen. Design of optimal reduced order H_2 filters. *Systems & Control Lett.*, 38:135–138, 1999.

[199] K. Zhou, J.C. Doyle, and K. Glover. *Robust and optimal control.* Prentice-Hall, Englewood Cliffs, NJ, USA, 1996.

CURRICULUM VITAE

Rufus Fraanje was born on July 18, 1976 in Ten Boer, The Netherlands. He followed his secondary school education at the Gomarus College in Groningen (1988-1992) and at the Gereformeerde Scholengemeenschap Rotterdam (1992-1994). From 1994-1999 he studied Electrical Engineering at Delft University of Technology, The Netherlands, which he completed with an MSc. thesis on algorithms for fault detection and isolation in dynamical systems under advisory of Prof. Michel Verhaegen. In the scope of this thesis project, he worked for 3 weeks together with Prof. Michèle Basseville and Prof. Albert Benveniste at INRIA, Rennes, France. During his studies he worked at LMS International in Leuven, Belgium, for 4 months on identification of dynamical car models and the development of model validation tools. In 1999 he started his PhD. studies at the University of Twente in Enschede, The Netherlands, on control algorithms for robust active control of broadband noise and vibration, under advisory of Prof. Michel Verhaegen. His research was sponsored by the 'Knowledge Center Sound and Vibration', a joint initiative of TNO Institute of Applied Physics in Delft and the University of Twente. In 2001 he was posted on Delft University of Technology at the current Delft Center for Systems and Control research group. In 2003 he worked, within the scope of his PhD. project, for 3 months together with Prof. Ali H. Sayed at the University of California in Los Angeles, USA, on fast and robust adaptive algorithms for application in active control. Recently, he started as researcher at the Delft Center for Systems and Control, where his research is on distributed control and identification for application in large scale control systems.